M. TECH SYLLABUS - 2026

Semester I to IV

Artificial Intelligence and Data Science

Branch Code: PAD

Jyothi Engineering College (AUTONOMOUS)
Reaccredited with NAAC (Grade A) and
NBA Programmes* (*CE, CS, EC, EE, ME, MR)
Jyothi Hills, P. O. Vettikkattiri, Cheruthuruthy
Thrissur, Kerala, India, 679531
04884 259000 | info@jecc.ac.in | www.jecc.ac.in
A Centre of Excellence in Science and Technology by the Catholic Archdiocese of Trichur

SEMESTER I

Discipline: Computer Science and Engineering

Stream: Artificial Intelligence and Data Science

.

J1PCST101	ADVANCED MACHINE LEARNING	CATEGORY	L	T	P	CREDIT
		DISCIPLINE CORE 1	3	0	0	3

Preamble: This course introduces machine learning concepts and popular machine learning algorithms. It will cover the standard and most popular supervised learning algorithms including linear regression, logistic regression, decision trees, k-nearest neighbour, an introduction to Bayesian learning and the naive Bayes algorithm, support vector machines and kernels and basic clustering algorithms. Dimensionality reduction methods and some applications to real world problems will also be discussed. It helps the learners to develop application machine learning based solutions for real world applications.

Course Outcomes:

After the completion of the course the student will be able to: *

CO1	Analyse the Machine Learning concepts, classifications of Machine Learning algorithms and basic parameter estimation methods. **(Cognitive Knowledge Level: Analyse)**
CO2	Illustrate the concepts of regression and classification techniques**(Cognitive Knowledge Level: Apply)**
CO3	Describe unsupervised learning concepts and dimensionality reduction techniques. (**Cognitive Knowledge Level:Apply**)
CO4	Explain Support Vector Machine Concepts and graphical models. **(Cognitive Knowledge Level:Apply)**
CO5	Choose suitable model parameters for different machine learning techniques and to evaluate a model performance. **(Cognitive Knowledge Level: Apply)**
CO6	Design, implement and analyse machine learning solution for a real-world problem. **(Cognitive Knowledge Level: Create)**

Program Outcomes(PO)

Outcomes are the attributes that are to be demonstrated bya graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resourcesandstate-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An abilitytoengageinlife-longlearningforthedesignanddevelopment relatedtothestream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load managementskillsrelatedtoprojectmanagementand finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓		✓		✓	✓	
CO2	✓		✓	✓	✓	✓	
CO3	✓		✓	✓	✓	✓	
CO4	✓		✓	✓	✓	✓	
CO5	✓		✓	✓	✓	✓	
CO6	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Apply	60-80%
Analyse	20-40%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

ContinuousInternalEvaluationPattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

ContinuousInternalEvaluation :40marks

.

Micro project/Course based project: 20 marks

Course based task/Seminar/Quiz :10marks

Testpaper,1no. :10marks

Theprojectshallbedoneindividually.Groupprojectsnotpermitted. Test paper shall include minimum 80% of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the University.There will be two parts; Part A and Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (Such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students). Students shall answer all questions.

Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturityofthe students in a course, through long answer questions relating to theoretical/practicalknowledge, derivations, problemsolving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks.

Total duration of the examination will be 150minutes.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Suppose that X is a discrete random variable with the following probability mass function: where $0 \leq \theta \leq 1$ is a parameter. The following 10 independent observations were taken from such adistribution:***(3, 0, 2, 1, 3, 2, 1, 0, 2, 1)***.What is the maximum likelihood estimate of θ.

X	0	1	2	3
$P(X)$	$2\theta/3$	$\theta/3$	$2(1-\theta)/3$	$(1-\theta)/3$

2. What is the difference between Maximum Likelihood estimation(MLE)andMaximum a Posteriori (MAP) estimation?

3. A gamma distribution with parameters ***α, β*** has the following density function,where ***Γ(t)*** Is the gamma function.

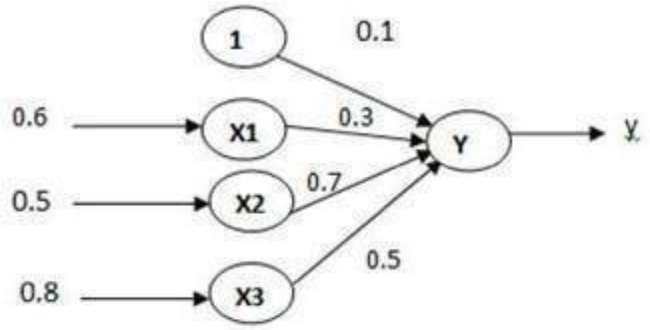
$$p(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

If the posterior distribution is in the same family as the prior distribution,then we say that the prior distribution is the conjugate prior for the likelihood function. Using the Gamma distribution as a prior, show that the Exponential distribution is a conjugate prior of the Gamma distribution. Also, find the maximum a posteriori estimator for the parameter of the Exponential distribution as a function of α and ***β***.

Course Outcome2(CO2)

1. How can we interpret the output of a two-class logistic regression classifier as a robability?

2. Calculate the output of the following neuronY if the activation function is a binary sigmoid.



3. Suppose you have a 3-dimensional input x = (x1, x2, x3) = (2, 2, 1) fully connected with weights (0.5,0.3, 0.2) to one neuron which is inthehiddenlayer withsigmoidactivation function.Calculate the output of the hidden layer neuron.

4. Consider the case of the XOR function in which the two points {(0,0),(1,1)}belongto one class, and the other two points {(1, 0),(0, 1)} belong to the other class. Design a multilayer perceptron for this binary classification problem.

5. Why does a single perceptron cannot simulate simple XOR function ? Explain howthis limitation is overcome?

6. Consider a naïve Bayes classifier with 3 boolean input variables, **X1**, **X2**and**X3**,and one boolean output, **Y**. How many parameters must be estimated to train such a naive Bayes classifier? How many parameters would have to be estimated to learn the above classifier if we do not make the naive Bayes conditional independence assumption?

Course Outcome3(CO3):

1. Describe the basic operation of k-means clustering.

2. A Poisson distribution is used to model data that consists of non-negative integers. Suppose you observe m integers in your training set. Your model assumption is that each integer is sampled from one of two different Gaussian distributions. You would like to

.

Learn this model using the EM algorithm.List all the parameters of the model.Derive the E-step and M-step for this model.

3. A uni-variate Gaussian distribution is used to model data that consists of non-negative integers. Suppose you observe m integers in your training set. Your model assumption is that each integer is sampled from one of two different Gaussian distributions. You would like to learn this model using the EM algorithm. List all the parameters of the model. Derive the E-step and M-step for the model.

4. Suppose you want to cluster the eight points shown below using **k**-means

	A_1	A_2
x_1	2	10
x_2	2	5
x_3	8	4
x_4	5	8
x_5	7	5
x_6	6	4
x_7	1	2
x_8	4	9

Assume that **k=3** and that initially the points are assigned to clusters as follows:

C1={x1,x2,x3},C2={x4,x5,x6},C3={x7,x8}.Apply the **k**-means algorithm until convergence, using the Manhattan distance.

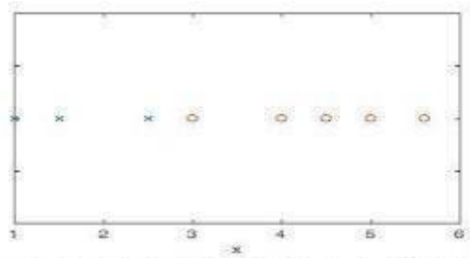
Course Outcome4(CO4):

1. Describe how Support Vector Machines can be extended to make use of kernels. Illustrate with reference to the Gaussian kernel $K(x, y) = e^{-y}$, where $y = (x-y)^2$.

2. Suppose that you have alinear support vector machine (SVM) binaryclassifier. Consider a point that is currently classified correctly, and is far away from the decision boundary. If you remove the point from the training set, and re-train the classifier, will the decision boundary change or stay the same? Justify your answer.

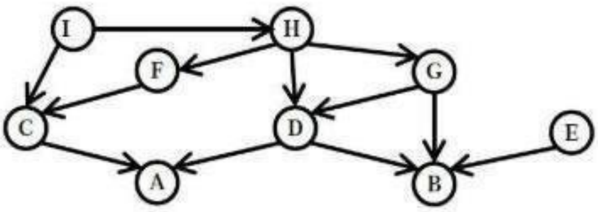
3. What is the primary motivation for using the kernel trick in machine learning algorithms?
4. Show that the Boolean function$(x_1 \wedge x_2) \vee (\neg x_1 \wedge \neg x_2)$is not linearly separable(i.e.there is no linear classifier **sign** $(w_1 x_1 + w_2 x_2 + b)$ that classifies all 4 possible input points correctly). Assume that —true‖ is represented by 1 and —false‖ is represented by −1. Show that there is a linear separator for this Boolean function when we use the kernel$K(x, y) = (x \cdot y)^2$($x.y$ denotes the ordinary inner product) . Give the weights and the value of ***b*** for one such separator.

5. Consider the following one-dimensional training dataset,‘x‘denote snegative examples and ‘o‘ positive examples. The exact data points and their labels are given in the table. Suppose aSVM is used to classifythis data. Indicate which are the support vectors and mark the decision boundary. Give the value of the cost function and of the model parameters after training.



x	1	1.5	2.5	3	4	4.5	5	5.6
y	-1	-1	-1	1	1	1	1	1

6. Write down the factored conditional Probability expression that corresponds to the graphical Bayesian Network shown below.



7. How do we learn the conditional probability tables(CPT) in Bayesian networks if information about some variables is missing? How are these variables called?

Course Outcome5(CO5):

1. Suppose 10000 patients get tested for flu; out of them, 9000 are actually healthy and 1000 are actually sick. For the sick people, a test was positive for 620 and negative for 380. For healthy people, the same test was positive for 180 and negative for 8820. Construct a confusion matrix for the data and compute the accuracy, precision and recall for the data.

2. Given th efollowing data,construct the ROC curve of the data. ComputetheAUC.

Thres hold	**TP**	**TN**	**FP**	**FN**
1	0	25	0	29
2	7	25	0	22
3	18	24	1	11

.

4	26	20	5	3
5	29	11	14	0
6	29	0	25	0
7	29	0	25	0

3. With an example classification problem, explain the following terms :a)Hyperparameters b)Training set c)Validation sets d)Bias e)Variance.

4. What is ensemble learning?Can ensemble learning using linear classifiers learn classification of linearly non-separable sets?

5. Describe boosting.What is there lation between boosting and ensemble learning?

6. Classifier Aattains 100%accuracy on the training set and70%accuracy on the test set. Classifier B attains 70% accuracy on the training set and 75%accuracyon the test set. Which one is a better classifier. Justify your answer.

7. What are ROC space and ROC curve in machine learning?In ROC space,which points correspond to perfect prediction, always positive prediction and always negative prediction? Why?

8. Suppose there are three classifiers A, B and C. The (FPR, TPR) measures ofthe three classifiersareas follows – A(0, 1), B(1,1), C(1,0.5).Which can be considered as a perfect classifier? Justify your answer.

9. What does it mean for a classifier to have a high precision but low recall?

.

ModelQuestionPaper

QPCODE:

Reg No: _________________

Name:____PAGES:4

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRST SEMESTER M. TECH DEGREE EXAMINATION, MONTH&YEAR

CourseCode: J1PCST101

Course Name: ADVANCED MACHINE LEARNING

Max.Marks:60 **Duration:2.5Hours**

PARTA

AnswerAllQuestions.EachQuestionCarries5 Marks

1. Explain the principle of the gradient descent algorithm.
2. In a two-class logistic regression model, the weight vector **w = [4,3,2,1,0]**. We apply it to some object that we would like to classify; the vectorized feature representation of this object is **x = [−2, 0, −3, 0.5, 3]**. What is the probability, according to the model, that this instance belongs to the positive class?
3. Expectation maximization(EM) is designed to find a maximum likelihood setting of the parameters of model when some of the data is missing. Does the algorithm converge? If so, do you obtain a locally or globally optimal set of parameters?
4. What is the basic idea of a Support Vector Machine?
5. What is the trade-off between bias and variance? **(5x5=25)**

PartB

(Answer any five questions. Each question carries 7 marks)

6. Suppose $x_1,\ldots,x_n$ are independent and identically distributed(iid)samples from a distribution with density **(7)**

$$f_X(x \mid \theta) = \begin{cases} \dfrac{\theta x^{\theta-1}}{3^\theta}, & 0 \le x \le 3 \\ 0, & \text{otherwise} \end{cases}$$

Find the maximum likelihood estimate(MLE)for θ.

7. Derive the gradient descent training rule assuming for the target function $o_d = w_0 + w_1x_1 + \ldots + w_nx_n$. Define explicitly the squared cost/error function E, assuming that a set of training examples D is provided, where each training example $d \in D$ is associated with the target output t_d. **(7)**

(7)

8. Cluster the following eight points representing locations into three clusters: ***A1(2,10), A2(2, 5), A3(8,4),A4(5, 8),A5(7,5),A6(6, 4),A7(1,2), A8(4,9)***.

Initial cluster centers are:***A1(2,10),A4(5, 8)***and***A7(1,2)***.

The distance function between two points ***a = (x1, y1)***and ***b =(x2, y2)***is defined as D***(a, b) = |x2 – x1| + |y2 – y1|***

Use **k**-Means Algorithm to find the three clustercenters after the second iteration.

9. Describe Principal Component Analysis. What criterion does the method minimize? What is the objective of the method? Give a way to compute the solution from a matrix ***X*** encoding the features. **(7)**

10 Consider a support vector machine whose input space is 2-D, and the inner products are computed by means of the kernel ***K (x, y) = (x.y + 1)²-1*** (***x.y***denotes the ordinary inner product). Show thatthe mappingto feature space That is implicitly defined by this kernel is the mapping to 5-D given by **(7)**

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \phi(\mathbf{x}) = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2}\, x_1 x_2 \\ \sqrt{2}\, x_1 \\ \sqrt{2}\, x_2 \end{bmatrix}.$$

11 How does random forest classifier work? Why is a random forest better than a decision tree? **(7)**

.

12 Consider a two-class classification problem of predicting whether a photograph contains a man or a woman. Suppose we have a test dataset of 10 records with expected outcomes and a set of predictions from our classification algorithm. Compute the confusion matrix, accuracy, precision, recall, sensitivity and specificity on the following data. **(7)**

Sl.No.	**Actual**	**Predicted**
1	man	woman
2	man	man
3	woman	woman
4	man	man
5	man	woman
6	woman	woman
7	woman	man
8	man	man
9	man	woman
10	woman	woman

Syllabus

Module-1(Parameter Estimation and Regression)8hours

Overview of machine learning: supervised, semi-supervised, unsupervised learning, reinforcement learning. Basics of parameter estimation: Maximum Likelihood Estimation (MLE), Maximum a Posteriori Estimation (MAP). Gradient Descent Algorithm, Batch Gradient Descent, Stochastic Gradient Descent. Regression algorithms: least squares linear regression, normal equations and closed form solution, Polynomial regression.

Module-2(Regularization techniques and Classification algorithms)9hours

Overfitting, Regularization techniques - LASSO and RIDGE. Classification algorithms: linear and non-linear algorithms, Perceptrons, Logistic regression, Naive Bayes, Decision trees. Neural networks: Concept of Artificial neuron, Feed-Forward Neural Network, Back propagation algorithm.

Module-3(Unsupervised learning)8hours

Unsupervised learning: clustering,k-means,Hierarchical clustering,Principal component analysis,

.

Density-based spatial clustering of applications with noise(DBSCAN). Gaussian mixture models: Expectation Maximization (EM) algorithm for Gaussian mixture model.

Module-4(Support Vector Machine and Graphical Models)7hours

Support vector machines and kernels: Max margin classification, Nonlinear SVM and the kernel trick, nonlinear decision boundaries, Kernel functions. Basics of graphical models - Bayesian networks, Hidden Markov model - Inference and estimation.

Module-5(Evaluation Metrics and Sampling Methods)8hours

Classification Performance Evaluation Metrics: Accuracy, Precision, Precision, Recall, Specificity, False Positive Rate (FPR), F1 Score, Receiver Operator Characteristic (ROC) Curve, AUC. Regression Performance Evaluation Metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R Squared/Coefficient of Determination. Clustering Performance Evaluation Metrics: Purity, Jaccard index, Normalized Mutual Information, Clustering Accuracy, Silhouette Coefficient, Dunn‘s Index. Boosting: AdaBoost, gradient boosting machines. Resampling methods: cross-validation, bootstrap. Ensemble methods: bagging, boosting, random forests Practical aspects in machine learning: data preprocessing, overfitting, accuracy estimation, parameter and model selection Bias-Variance tradeoff

CoursePlan

No	**Topics**	**No. of Lectures (40)**
1	**Module-1(Parameter Estimation and Regression)8hours**	
1.1	Overview of machine learning: supervised, semi-supervised, unsupervised learning, reinforcement learning.	1
1.2	Basics of parameter estimation: Maximum Likelihood Estimation(MLE)	1
1.3	Basics of parameter estimation:Maximum Likelihood Estimation(MLE) -Examples	1
1.4	Basics of parameter estimation:Maximuma Posteriori Estimation(MAP)	1
1.5	Basics of parameter estimation:Maximuma Posteriori Estimation(MAP) -Example	1
1.6	Gradient Descent Algorithm, BatchGradient Descent, Stochastic Gradient Descent	1
1.7	Regression algorithms:least squares linear regression,normal equations and closed form solution	1
1.8	Polynomial regression	1
2	**Module-2(Regularization techniques and Classification algorithms)9hours**	

.

2.1	Overfitting, Regularization techniques-LASSO and RIDGE	
2.2	Classification algorithms:linear and non-linear algorithms	
2.3	Perceptrons	
2.4	Logistic regression	
2.5	Naïve Bayes	
2.6	Decision trees	
2.7	Neural networks:Concept of Artificial neuron	
2.8	Feed-Forward Neural Network	
2.9	Back propagation algorithm	
3	**Module-3(Unsupervised learning)8hours**	
3.1	Unsupervised learning:clustering,k-means	
3.2	Hierarchical clustering	
3.3	Principal component analysis	
3.4	Density-based spatial clustering of applications with noise(DBSCAN)	
3.5	Gaussian mixture models:Expectation Maximization(EM) algorithm for Gaussian mixture model	
3.6	Gaussian mixture models:Expectation Maximization(EM)algorithm for Gaussian mixture model	
4	**Module-4(Support Vector Machine and Graphical Models)7hours**	
4.1	Support vector machines and kernels:Max margin classification	
4.2	Support vector machines:Max margin classification	
4.3	NonlinearSVMandthekerneltrick,nonlineardecisionboundaries	
4.3	Kernel functions	
4.5	Basics of graphical models-Bayesian networks	
4.6	Hidden Markov model- Inference and estimation	
4.7	Hidden Markov model- Inference and estimation	
4.8	Hidden Markov model- Inference and estimation	
5	**Module-5(Evaluation Metrics and Sampling Methods)8hours**	
5.1	Classification Performance Evaluation Metrics:Accuracy,Precision, Precision, Recall, Specificity, False Positive Rate (FPR), F1 Score, Receiver Operator Characteristic (ROC) Curve, AUC	
5.2	Regression Performance Evaluation Metrics:Mean Absolute Error	

.

	(MAE),Root Mean Squared Error (RMSE),RSquared/Coefficientof Determination	
5.3	Clustering Performance Evaluation Metrics:Purity,Jaccardindex, Normalized Mutual Information, Clustering Accuracy, Silhouette Coefficient, Dunn‘s Index	
5.4	Boosting:AdaBoost,gradient boosting machines.	
5.5	Resampling methods:cross-validation,bootstrap.	
5.6	Ensemble methods:bagging,boosting,random forests	
5.7	Practical aspects in machine learning:data preprocessing,overfitting, accuracy estimation, parameter and model selection	
5.8	Bias-Variance tradeoff	

Reference Books

1. Christopher Bishop.Neural Networks for Pattern Recognition,Oxford University Press, 1995.

2. KevinP. Murphy. Machine Learning:AProbabilisticPerspective,MITPress2012.

3. Trevor Hastie,Robert Tibshirani,JeromeFriedman,TheElementsOfStatisticalLearning, Second edition Springer 2007.

4. Ethem Alpaydin, Introduction to Machine Learning, 2nd edition,MITPress 2010.

5. TomMitch

.

CODE J1PADT102	MATHEMATICAL FOUNDATIONSFOR DATA SCIENCE	CATEGORY	L	T	P	CREDIT
		Program Core 1	**3**	**0**	**0**	**3**

Preamble:

This course is intended to provide the learners with an outlook on applying concepts in linear algebra in the fields of data science, machine learning, and artificial intelligence. This course helps the learners to acquire the skills to implement the concepts in MATLAB/Python and then apply linear algebra concepts to real datasets. Also, this course discusses the Challenges of applying the acquired knowledge in different Optimization and Linear Algebra concepts toward the inference and prediction stages of Data Analytics.

Course Outcomes: After the completion of the course,the student will be able to

CO1	Analyse the fundamentals of linear algebra and calculus,and other mathematical concepts for Artificial Intelligence,Machine Learning,and Data Science(**Cognitive knowledge level: Analyse**)
CO2	Apply the knowledge acquired in different optimization and linear algebra concepts Towards the inference and prediction stages of data analytics.(**Cognitive knowledge level: Apply)**
CO3	Implement linearalgebra concepts in scientific programming languages(MATLAB, Python) **(Cognitive knowledge level: Apply)**
CO4	Apply eigenvectors and SVD for Dimensionality reduction**(Cognitive knowledge level: Apply)**
CO5	Design,Develop,implement,and Present innovative Ideas on the application of linear Algebra for Data Science,Machine learning,and Artificial Intelligence**(Cognitive Knowledge Level: create)**

Program Outcomes(PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively and write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate mastery over the area as per the program's specialization.The mastery should be at a level higher than the requirements in the appropriate bachelor‘sprogram

PO4: An ability to apply stream knowledge to design or develop solutions forreal-world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tools to model, analyze and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream-related problems taking into consideration sustainability, societal, ethical, and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✔	✔	✔	✔	✔	✔	
CO2	✔	✔	✔	✔	✔	✔	
CO3	✔	✔	✔	✔	✔	✔	
CO4	✔	✔	✔	✔	✔	✔	
CO5	✔	✔	✔	✔	✔	✔	✔

AssessmentPattern

Bloom'sCategory	End Semester Examination
Understand	
Apply	60-80%
Analyze	20-40%
Evaluate	Can be done through Assignments
Create	Can be done through Assignments

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

The evaluation shall only be based on application, analysis or design-based questions (for both internal and end-semester examinations).

Continuous Internal Evaluation :40marks

Micro project/Course based project :20marks

Course based task/Seminar/Quiz :10marks

.

Test paper,1no. :10marks

The project shall be done individually.Group projects are not permitted.

The test paper shall include a minimum of 80% of the syllabus.

Course-based task/test paper questions shall be useful in testing the knowledge, skills, comprehension, application, analysis, synthesis, evaluation, and understanding of the students.

End Semester Examination Pattern:

The end-semester examination will be conducted by the University.There will be two parts; Part A and Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (Such questions shall be useful in testing knowledge, skills, comprehension, application, analysis, synthesis, evaluation, and understanding of the students). Students shall answer all questions.

Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturityofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations,problem-solvingandquantitativeevaluation),with a minimum one question from each module of which student should answer any five. Each question can carry 7 marks.

The total duration of the examination will be150minutes.

Course Level AssessmentQuestions

Course Outcome 1 (CO1):

1. Discuss Linear algebra handles large amounts of data,

2,List ten Powerful Applications of Linear Algebra in Data Science

3. ‗Support Vector Machine isanapplication of the concept of Vector Spaces in Linear Algebra‘- Investigate

4. Support Vector Machine Classification

Course Outcome2(CO2)

1. Write an algorithm for simple linear regression by gradient descent method
2. Implement Gradient Descent for multilinear regression from scratch
3. Elaborate regularized linear regression for model prediction and reduce errors

Course Outcome3(CO3):

.

1. Elaborate on the different ways to compute and conceptualize matrix multiplication with examples

2. Implement Vector Hadamard multiplication

3. w1= [1 35];

4. w2= [342];in MATLAB/Python

Course Outcome4(CO4):

1. Implement the determinant of a matrix product in Python

2. Elucidate the usefulness of Linear Algebra in Dimensionality Reduction

    i Principal Component Analysis(PCA)

    i Singular Value Decomposition(SVD)

3. Determine whether the following matrices have a null space.If so,provide basisvector(s)for that null space.

a $\begin{pmatrix} 4 & 3 \\ 1 & 1 \\ 0 & 5 \end{pmatrix}$ b. $\begin{pmatrix} 3 & 1 & 5 \\ 4 & 1 & 0 \end{pmatrix}$

Course Outcome5(CO5):

1. List the five steps of model fitting

2. Express the average or the mean of a set of numbers as a model and use least squares to fit that model

3. Elaborate on the usage of Linear Algebrain Machine Learning

    a. Loss functions

    b. Regularization

    c. Covariance Matrix

.

ModelQuestionPaper

QPCODE:

Reg No: ________________

Name: ____ **PAGES:4**

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRST SEMESTER M. TECH DEGREE EXAMINATION, MONTH&YEAR

CourseCode: J1PADT102

Course Name:Mathematical Foundation for DataScience

Max.Marks:60 **Duration:2.5Hours**

PARTA

AnswerAll Questions.Each Question Carries 5 Marks

1. Explain normal equation. Contrast normal equation and gradient descent. **5**

2. Explain overfitting and the method of resolving overfitting **5**

3. Make use of the parallel matrix multiplication method to find AB, where **5**

$$A = \begin{bmatrix} 2 & 1 & 5 \\ 0 & 7 & 1 \\ 9 & 2 & 4 \\ 3 & 6 & 7 \end{bmatrix} \quad B = \begin{bmatrix} 6 & 1 & 2 \\ 4 & 5 & 6 \\ 1 & 1 & -8 \\ 4 & 0 & -5 \end{bmatrix}$$

4. Let $v=\langle 1,3\rangle$ and $w=\langle -4,-2\rangle$.Write v as the sum of two orthogonal vectors,one of which Is the projection of $v\rightarrow$onto w **5**

5. Find the SVD for the matrix **5**

(5x5=25)

PartB

(Answer any five questions.Each question carries 7marks)

6. (a)Write an algorithm for simple linear regression bygradient descent method **(7)**

7. (a)Implement Gradient Descent for multilinear regression from scratch **(7)**

8. (a) Elaborate the different types of matrices with examples and implement them using MATLAB or NumPy **(7)**

i. Square matrix
ii. Rectangular matrix
iii. Symmetrix matrix
iv. Skey-symmetric matrices
v. Identity matrix
vi. Zero
vii. Diagonal matrix
viii. Triangular matrix
ix. Augmented
x. Complex

9. (a)Define rank and a maximum possible rank **(2)**

(b) Createa matrix of 10 X 10 with a rank of 4(use matrix multiplication) **(3)**

(c) Generalize the procedure to create any M x N matrix with a rank r (2) **(2)**

10 (a)Implement the determinant of a matrix product in Python **(3)**

(b)Determine whether the following matrices have a null space.If so,provide basis vector(s) for that null space. **(4)**

$$a = \begin{pmatrix} 4 & 3 \\ 1 & 1 \\ 0 & 5 \end{pmatrix} \qquad b = \begin{pmatrix} 3 & 1 & 5 \\ 4 & 1 & 0 \end{pmatrix}$$

11 (a)Write MATLAB or Python code to implement the following experiment: **(7)**

a. Generate a 2× 3 matrix of random numbers.
b. Compute its SVD.
c. Compute two eigen decompositions using the matrix and its transpose.
d. Confirm that the two sets of eigen values match,and check whether the eigenvalues match the singular values

Plot the eigen vectors and singular vectors in 2Dor3D(asappropriate)to confirm that SVD and transpose + eigen decomposition produce the same eigenspaces

12 (a)Determine th eeigenvalues and the corresponding eigenvectors of the following mat **(7)**

$$\text{matrix A} = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

.

Syllabus (Emphasis shall be on Problem-solving, implementing concepts,and application-Languages Python/MATLAB)

Module	Contents	hours
I	**GRADIENT DESCENT AND REGULARIZATION** Gradient Descent, Intuitions, Gradient Descent for a Regression algorithm, Multiple Features, Gradient Descent on Multiple Features, Practice onGradient Descent for Polynomial Regression, Normal Equation	8
II	**LINEAR ALGEBRA-VECTORS**: Vectors: geometry and algebra, Vector addition, and subtraction, Vector-scalar multiplication, Dot product geometry, Vector orthogonality, Cauchy-Schwarz inequality, Vector Hadamard multiplication, cross product, unit vectors, **VECTOR SPACES**: Dimensions, and fields in linear algebra, Subspaces, Subspaces vs. subsets, Span, Linear independence, Basis **MATRICES**: introduction: dimensionality, Matrix operations, Matrix-scalar multiplication, Implementation, Transpose, Complex matrices, Addition, equality, transpose, Diagonal, and trace,	7
III	**MATRICES-: MATRIX MULTIPLICATION-**Introduction, matrix multiplication by layering, Multiplication with diagonals, Matrix-vector multiplication, Symmetric matrices, multiply symmetric matrices Hadamard Multiplication, asymmetry Index, Code challenge, **RANK-**concepts, Maximum possible rank, Computing rank, Rank and scalar multiplication, Rank of added and multiplied matrices, Rank of A & $A^{T,}$ A^T A, $AA^{T,}$ random matrices, Boosting rank by shifting, rank difficulties, rank, and span, Code challenge: **MATRIX SPACES:** Column space and Row space of a matrix (A & AA^T), Null space of a matrix, orthogonal subspaces, Dimensions of column/row/null spaces, Example of the four subspaces, code challenge	7
IV	**DETERMINANTS, PROJECTIONS&ORTHOGONALIZATION-** **DETERMINANT-**Determinant, Determinant of a 2x2 matrix, Determinant of a 3x3 matrix, characteristic polynomial, the full procedure, determinant of triangles, determinant and row reduction, determinant and scalarmultiplication, theory vs practice, Code challenge **MATRIX INVERSE:** Concept and applications, Inverse of a Diagonal matrix, Inverse of a 2x2 matrix, The MCA algorithm to compute the inverse, Computing the inversevia row reduction, Left inverse and right inverse, Pseudo-inverse, Code challenge **PROJECTIONS, AND ORTHOGONALIZATIONS:** Projections in R^2, Projections in R^N, Orthogonal and parallel vector components, Orthogonal matrices, Gram-Schmidt procedure, QR decomposition, Inversevia QR Decomposition, Code challenge **LEAST SQUARES FOR MODEL- FITTING IN STATISTICS**: Introduction, Least squares via left inverse, Least squares via orthogonal projection, Least-squares via row-reduction, Model-predicted values, and residuals, Least-squares applications, Code challenge	10
V	**DIMENSIONALITY REDUCTION: EIGEN DECOMPOSITION-** Eigenvalues,eigenvectors,Eigendecomposition,Diagonalization,Matrix powersviadiagonalization,Distinctandrepeatedeigenvalues,symmetric	8

.

	matrices, Eigen layers of a matrix, Eigen decomposition of singular matrices, Matrix powers and Inverse, Generalized eigen decomposition, Codechallenges **SINGULAR VALUE DECOMPOSITION (SVD):** Singular value decomposition, Computing the SVD, singular values and eigenvalues, Symmetric Matrices, SVD and the four subspaces, SVD, and matrix rank, Spectral theory of matrices, SVD for low-rank approximations, Normalizing singular values, the Condition number of a matrix, SVD and Matrix Inverse, MP pseudo inverse, code challenges	

CoursePlan

S.NO	TOPIC	NO. OF LECTURES
MODULE1-GRADIENTDESCENT AND REGULARIZATION-8 hours		
1.1	Gradient Descent,Intuitions,	1
1.2	Gradient Descent for a Regression algorithm.	1
1.3	Gradient Descent for a Regression algorithm.	1
1.4	Multiple Features	1
1.5	Gradient Descenton Multiple Features,	1
1.6	Practiceon Gradient Descent	1
1.7	Gradient Descent for Polynomial Regression	1
1.8	Gradient Descent and Normal Equation	1
MODULE2 LINEAR ALGEBRA-7HOURS		
	LINEAR ALGEBRA- Vectors	1
2.1	**VECTORS:**geometry and algebra,Vector addition,and subtraction, Vector-scalar multiplication,	1
2.2	Dot productgeometry,Vector orthogonality,	1
2.3	Vector Hadamard multiplication,crossproduct,unit vectors,	
2.4	**VECTOR SPACES:** Dimensions, andfields inlinear algebra, Subspaces, Subspaces vs. subsets	1
2.5	Span,Linear independence,Basis	1
2.6	**MATRICES**: dimensionality, Matrix operations, Matrix-scalar multiplication,	1
2.7	Complex matrices,Addition,equality,transpose,Diagonal,andtrace,	1
MODULE3-MATRICES–7HOURS		
3.1	**MATRIX MULTIPLICATION:**matrix multiplication by layering,	1

.

	Multiplication with diagonals,Matrix-vectormultiplication,	
3.2	Symmetric matrices, multiply symmetric matrices Hadamard Multiplication, asymmetry Index, Code challenge,	1
3.3	**RANK-**concepts,Maximum possible rank,Computing rank,Rank and scalar multiplication, Rank of added and multiplied matrices	1
3.4	The rankofA&$A^{T,}A^{T}$A, $AA^{T,}$random matrices,	1
3.5	Boosting rank by shifting,rank difficulties,rank,and span,Code challenge:	1
3.6	**MATRIX SPACES**:Column space and Row space ofa matrix(A& AAT), Null space ofa matrix,	1
3.7	Orthogonal subspaces,Dimensions of column/row/null spaces,Example of the four subspaces, code challenge,	1
MODULE4-DETERMINANTS, & PROJECTIONS, AND ORTHOGONALIZATION-10 HOURS		
4.1	**DETERMINANT**-Determinant ofa 2x2 matrix, & 3x3 matrix, characteristic polynomial, the full procedure,	1
4.2	determinant of triangles, determinant and rowreduction, determinant and scalar multiplication, theory vs practice, Code challenge	1
4.3	**MATRIX INVERSE:** Concept and applications, Inverse of a Diagonal matrix, Inverse ofa 2x2 matrix, The MCA algorithm to compute the inverse,	1
4.4	Computing the inverse via row, Left inverse and right inverse, Pseudoinverse, Code challenge	1
PROJECTIONS AND ORTHOGONALIZATION:		
4.5	Projections in R^2, Projections in R^N, Orthogonal and parallel vector components,	1
4.6	Orthogonal matrices, Gram-Schmidt procedure, QR decomposition,	1
4.7	Inverse via QR Decomposition, Code challenge	1
LEAST SQUARES FOR MODEL-FITTING IN STATISTICS		
4.8	Introduction, least squares via left inverse, Least squares via orthogonal projection,	1
4.9	Least squares via row-reduction, Model-predicted values, and residuals	1
4.10	Least-squares applications, Code challenge	1
MODULE5-DIMENSIONALITY REDUCTION-8HOURS		
EIGEN DECOMPOSITION-,		
5.1	Eigenvalues,eigenvectors,Eigen decomposition,	1
5.2	Diagonalization,Matrix powers via diagonalization	1
5.3	Distinct and repeated eigenvalues,symmetric matrices,Eigen layers of a matrix,	1

.

5.4	Eigen decomposition of singular matrices,Matrix powers and Inverse, Generalized eigen decomposition, Code challenges	1
SINGULAR VALUE DECOMPOSITION (SVD)-		
5.5	Singular value decomposition,ComputingtheSVD,singular values and eigenvalues,	1
5.6	Symmetric Matrices,SVD and the four subspaces,SVD,and matrix rank,	1
5.7	Spectral theory of matrices,SVD for low-rank approximations, Normalizing singular values, the Condition number of a matrix,	1
5.8	SVD and Matrix Inverse, MP pseudo inverse,code challenges	1

TEXTBOOKS:

1. Mike X Cohen,LinearAlgebra:Theory,Intuition,Code[PrintReplica] Kindle Edition
2. GeneH. Golub,CharlesF.VanLoan,―MatrixComputations‖,JohnHopkins University Press.
3. Steven Cooper, Data Science from Scratch: The #1 Data Science Guide for Everything A Data Scientist Needs to Know: Python, Linear Algebra, Statistics,Coding,Applications,NeuralNetworks,andDecisionTreeKindle Edition
4. RandolfH.Reiss,B.S,―EigenValuesandEigenVectorsinDataDimension ReductionforRegression‖,SanMarcos,Texas.
5. GilbertStrang,―LinearAlgebraanditsApplications‖,ThomsonLearningInc.

REFERENCES:

1. Charu C.Aggarwal,―LinearAlgebraandOptimizationforMachineLearning‖, Springer 2020.
2. Singiresu S.Rao,―EngineeringOptimization:TheoryandPractice‖,FourthEdition 2009byJohnWiley& Sons,Inc.

.

J1PADT103	INTRODUCTIONTO AI AND NLP	CATEGORY	L	T	P	CREDIT
		ProgramCore2	3	0	0	3

Preamble:

This course introduces the concepts, tools, and techniques of machine learning for text data. The students will learn the elementary concepts as well as emerging trends in the field of NLP. This course helps the learners to extract information from unstructuredtext, identify linguistic structure of it, and to apply different the techniques for text analytics. The learners will able to implement and evaluate NLP applications using machine learning and deep learning methods.

Course Outcomes:

After the completion of the course,the student will be able to:*

CO1	Analyze the applications of AI in the domain of NLP**(Cognitive Knowledge Level:Apply)**
CO2	Transform text into an appropriate data structure(**Cognitive Knowledge Level:Apply)**
CO3	Apply Probability models,language models,and Markov models for Text processing **(Cognitive Knowledge Level: Apply)**
CO4	Build NLP applications using Machine Learning Methods**(Cognitive Knowledge Level:Apply)**
CO5	Design,Develop,Implementand Present innovative Ideas on NLP and AI **(Cognitive Knowledge Level:Create)**

Program Outcomes(PO)

Outcomes are the attributes that are to be demonstrated bya graduate after completing the course

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively,write and present technical reports on complex engineering activities by interacting with t he engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions forreal-world problems by following the standards

PO5: An ability to identify,select and apply appropriate techniques,resources and state-of-the-art tools to model, analyze and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmentalaspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓		✓		✓	✓	
CO2	✓		✓		✓	✓	
CO3	✓		✓		✓	✓	
CO4	✓	✓	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓	✓	✓

Assessment Pattern

Bloom'sCategory	EndSemester Examination
Apply	50-80%
Analyse	20-40%

.

Evaluate	Assess using Assignments/Project
Create	Assess using Assignments/Project

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

The evaluation shall only be based on application,analysis o rdesign-based questions (for both internal and end-semester examinations).

Continuous Internal Evaluation: 40 marks
Microproject/Coursebasedproject:20marks
Course based task/Seminar/Quiz : 10 marks
Test paper, 1 no. : 10 marks

The project shall be done individually. Group projects are not permitted.Test paper shall include a minimum of 80%ofthe syllabus.

Course-based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation, and understanding of the students.

End Semester Examination Pattern:

The end-semester examination will be conducted by the University. There will be two parts; Part A and Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (Such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation, and understanding of the students). Students shall answer all questions.

Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course,through long answer questions

.

Relating to theoretical/practical knowledge,derivations,problem-solving and quantitative evaluation), with a minimum one question from each module of which student should answer any five. Each question can carry 7 marks.

The total duration of the examination will be150minutes.

Course Level Assessment Questions

Course Outcome1(CO1):

1. Implement artificial intelligence-based solution in agriculture for optimization of irrigation and application of pesticides and herbicides

2. Develop Artificial Intelligence/Machine Learning based for Diabetes Care

3. Explain the role of Natural Language Processing Applications in Finance

Course Outcome2(CO2)

1. Which of the following techniques can be used for key word normalizationin NLP, the process of converting a keyword into its base form?

a. Lemmatization

b. Soundex

c. Cosine Similarity

d. N-grams

Interpret your answer.

2. Explain any two out of the eight methods that **Convert Text to Features**

a. Recipe1.One Hot encoding.

b. Recipe2.Count vectorizer.

c. Recipe3.N-grams.

d. Recipe4.Co-occurrence matrix.

e. Recipe5.Hash vectorizer.

f. Recipe6.Term Frequency-Inverse Document Frequency(TF-IDF)

.

g. Recipe7.Word embedding.

h. Recipe8.Implementing fast Text.

3. Explain the following with an example

a. Regular expression and its working

b. Properties of regular expression

c. Meta characters Big brackets[]and search[abcd]in _kasdfaiabcasdfaabc'and write the output

d. Let's look at the following sentence:—Iateanappleandplayedthepiano.‖ Generate the one-hot embedding matrix for this sentence

CourseOutcome3(CO3):

a. Implement TF-IDF froms cratch
b. Explain how you will implement NLP in other languages
c. Explain Markov Model for text classifier and build a text classifier using the Markov model

Example:Take two sets of poems by two different authors,Edgar Allan Poe and RobertFrost. Given a sentence, the system should be ableto predict the author

Course Outcome4(CO4):

1. Application :Latent Semantic Indexing for Search Engine Optimizationusing PCA/SVD

2. Implement ANN formulticlass classification

3. Implement SpamDetection using NaïveBayes or Logistic regression

CourseOutcome5(CO5):

1. Implement Text Summarization using python

2. How will you implement a sentiment analyzer in Python using logistic regression to predict sentiment on Amazon reviews?

3. Application: Topic Modeling Using Latent Dirichlet Allocation

.

ModelQuestionPaper

QPCODE:

Reg No: __________________

Name: ____ **PAGES:4**

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRST SEMESTER M. TECH DEGREE EXAMINATION,MONTH&YEAR

CourseCode: J1PADT103

Course Name:INTRODUCTION TO AI AND NLP

Max.Marks:60 **Duration:2.5Hours**

PARTA

Answer All Questions. Each Question Carries 5 Marks

1.	a.Contrast NLTK and Spacy b.Explain the different types of artificial Intelligence	**5**
2.	a.Analyse the importance of the concept of vectors in NLP b.Explain the simple functioning of the ELIZA chatbot	**5**
3.	Explain Markov Property and Markov model for NLP	**5**
4.	Discuss topic modeling and the intuition behind using LDA for topic modeling	**5**
5.	Describe implementing binary text classification using Tensor Flow	**5**

(5x5=25)

Part B

(Answer any five questions.Each question carries 7marks)

6. (a) Explain the applications of Artificial intelligence in Healthcare **(7)**

.

7. (a) Elaborate vectors,their usefulnessinNLP,tokenization,stopwords,stemming,and lemmatization with examples **(3)**

(b) Explain the limitations of the count vectorizer and the necessity of TF-IDF **(2)**

(c) In a corpus of N documents,one randomly chosen document contains a total of T terms,and the term—hello‖appears K times.What is the correct value for the Product of TF(termfrequency)and IDF(inverse-document-frequency),If the term —hello‖appears in approximately one-third of the total documents?

8. (a) Explain the regular expression and need for a regular expression in NLP **(3)**

(b) How do regular expressions work? **(2)**

(c) Explain the use of common Regex functions used in NLP **(2)**

9. (a) Explain the language model,applications of language modeling,method to Compute the probability of a sentence, and the curse of dimensionality **(2)**

(b) What Is Markov's assumption in language modeling and n-grams **(3)**

(c) Implement n-grams and update-function **(2)**

10. (a) Explain the spam detection problem and why you want to detect them(2) **(3)**

(b) Apply the Markov model n-gram approach to solve this problem and implement in Python **(4)**

11 (a) Explain RNN for text classification in TensorFlow **(2)**

(b) Explain Parts of Speech(PoS)Tagging in TensorFlow **(3)**

(c) Explain Named EntityRecognition in TensorFlow **(2)**

12. (a) Explain CNN fort ext classification **(7)**

.

Syllabus

Module	Contents	Total Lecture Hours (40hrs)
1	**INTRODUCTION TO ARTIFICIAL INTELLIGENCE**: Artificial Intelligence? History, AI on a conceptual level, Types of AI, Use Cases, importance and applications of AI, AI algorithms, types of machine learning, types of problems solved in AI, advantages, and disadvantages of AI, AI In Marketing, Banking, Finance, Agriculture, HealthCare, Gaming, Space Exploration., Autonomous Vehicles, Chatbots, Artificial Creativity, AI Tools & Frameworks, AI vs Machine Learning vs Deep Learning, an overview of python for AI, INTRODUCTION TO NLP: NLP in the Real World, NLP Tasks, Language? Its Building Blocks, Why Is NLP Challenging? Machine Learning, Deep Learning, and NLP: An Overview, Approaches to NLP, Heuristics-Based NLP, Machine Learning & Deep Learningfor NLP, NLP Pipeline, Applications of NLP-Machine translation, Speech recognition, Image Captioning, spam detection, text prediction- Introduction to Software Packages-Spacy, NLTK, Gensim, PyTorch, Regular Expression- importance, properties, working and python package (re), case study: working of Eliza chatbot	7
2	**REGULAR EXPRESSION & TEXT PROCESSING:** Common regex function, Meta Characters- Big brackets, cap, Backslash, Squared Brackets, Special Sequences, Asterisk, Plus, And Question mark, Curly Brackets **Understanding Pattern Objects-** Match Method Vs Search Method, Finditer Method, Logical Or, Beginning And End Patterns, Parenthesis **String Modification-** split method, sub- method, subn method, **Text Processing-**Words, Tokens, Counting words, vocabulary, corpus, tokenization in spacy- **Sentiment Classification-** (yelp) download a review dataset use data preparation using NumPy, pandas, counter, re- add tokens to vocabulary, build vocabulary from a data frame, from corpus, one hot encoding, encoding documents, train test splits, feature computation, confusion matrix, analysis. **Language Independent Tokenization**: Types of tokenization — Word, Character, and sub-word tokenization, problems with word tokenizer, drawbacks of a character- based tokenizer, problems with sub-word tokenization, Byte Pair Encoding, , **String Matching and Spelling Correction**-Minimum edit distance- table filling, dynamic programming,	9
	WORD EMBEDDING&PROBABILISTIC MODELS:Vector Models&Text Preprocessing: Vectors, Bag of Words, Count Vectorizer, Tokenization,Stopwords, Stemming, and Lemmatization, Stemming, and Lemmatization, Count Vectorizer, Vector Similarity. TF-IDF, Word-to-Index Mapping, Building TF-IDF, Neural Word Embeddings,Neural Word Embeddings.Vector Models Text Pre-Processing Summary,steps of NLP analysis,**Probabilistic Models-Language**	9

.

3	**Modelling:** importance, types of language modeling, the curse of dimensionality, Language Model Markov Assumption And N-Grams, Language Model Implementation – Setup, Ngrams Function, Update Counts Function, Probability Model Function,ReadingCorpus, LanguageModel Implementation Sampling Text, **Markov Models:** Markov Property, Markov Model, Probability Smoothing and Log-Probabilities,Buildinga TextClassifier, **Article Spinning** – Problem, N-Gram Approach, implementation, **Cipher Decryption with Language Modeling And Genetic Algorithm** Ciphers,substitutioncipher,bigrams,maximum likelihood,and log-likelihood, Language models, Genetic Algorithms,	
4	**NLP USING MACHINE LEARNING MODELS**-Spam Detection– Problem, Naive Bayes theorem, Intuition, spam detection using Naïve Bayes, classimbalance, ROC, AUC, AND F1 SCORE, Implementing spam detection in python, **Sentiment Analysis -**Problem, Logistic Regression Intuition, Multiclass Logistic Regression, Logistic Regression Training and Interpretation, sentiment analysis implementation in python, **Text Summarization-**Using Vectors, Text Rank Intuition,Text Rank in Python, Text Summarization in Python **Topic Modeling-** different topic modeling techniques,LatentDirichlet Allocation(LDA)–Essentials, Latent Dirichlet Allocation–Topic Modeling with Latent Dirichlet, **Latent Symmatc Modelling(Indexing)-**LSA / LSI Introduction, Singular Value Decomposition Intuition, LSA / LSI: Applying SVD to NLP, Latent Semantic Analysis / Latent Semantic Indexing in Python	7
5	**DEEP LEARNING-** word embeddings, nonlinear neural networks, **Neuron** –Intro, Fitting a Line, Classification Code Preparation, Text Classification in Tensorflow, The Neuron, How does a model learn?, **Feed Forward Neural Networks-** Ann-introduction, The Geometrical Picture, Activation Functions, Multiclass Classification, Text Classification ANN in Tensorflow, Text Preprocessing Code Preparation, Text Preprocessing in Tensorflow, Embeddings, CBOW(continuous bag of words), CBOW in Tensorflow, **Convolution Neural Networks-** Convolution, pattern matching, weight sharing, convolution in color images, CNNArchitecture,CNNforText,CNNfor NLPinTensorflow, **Recurrent Neural Networks-** Simple RNN / Elman Unit, RNNs: Paying Attention to Shapes, GRU, and LSTM. RNN for Text Classification in TensorFlow, Parts-of-Speech Tagging, andNamed Entity Recognition in TensorFlow	8

.

COURSE PLAN		
Sl.NO	**TOPIC**	**NO.OF LECTURES**
	Module1–IntroductionofAI&NLP(7hours)	
Module1:Introduction To Artificial Intelligence(7hours)		
1.1	ArtificialIntelligence?History,AI on a conceptual level,TypesofAI,Use Cases, importance and applications of AI,	**1**
1.2	AIalgorithms,types of machine learning,types of problems solvedinAI,advantages, and disadvantages of AI	**1**
1.3	AI In Marketing,Banking,Finance,Agriculture,HealthCare,Gaming,SpaceExploration, Autonomous Vehicles, Chatbots,	**1**
1.4	Artificial Creativity,AITools&Frameworks,AIvsMachineLearningvsDeep Learning, an overview of python for AI,	**1**
INTRODUCTIONTO NLP,		
1.5	NLP in the Real World, NLP Tasks, Language? Its Building Blocks, Why Is NLP Challenging?Machine Learning,Deep Learning,and NLP:An Overview,Approaches to NLP, Heuristics-Based NLP, Machine Learning & Deep Learning for NLP	**1**
1.6	NLP Pipeline,Applications ofNLP-Machine translation,Speech recognition,Image Captioning, spam detection, text prediction-	**1**
1.7	Introduction to Software Packages-Spacy, NLTK, Gensim, PyTorch, **Regular Expression-** importance, properties, working and python package (re), case study: working of Eliza chatbot	**1**
Module2-Regular Expression&Text Processing(9Hours)		
2.1	Common regex functionused in NLP,-	**1**
2.2	Meta Characters-Big brackets,cap,Backslash,	**1**
2.3	Squared Brackets,Special Sequences,Asterisk,Plus,AndQuestion mark,Curly Brackets	**1**
2.4	**Understanding Pattern Objects-**Match Method Vs Search Method,FinditerMethod, Logical Or, Beginning And End Patterns, Parenthesis	**1**
2.5	**String Modification-**split method,sub-method,subn method,	**1**
TEXT PROCESSING		
2.6	Words,Tokens,Counting words,vocabulary,corpus,tokenization in spacy-	**1**
2.7	**Sentiment Classification-** (yelp) download a review dataset use –data preparation using NumPy, pandas, counter, re-add tokens to vocabulary, build vocabulary from a data frame, from corpus, one hot encoding, encoding documents, train test splits, feature computation, confusion matrix, analysis.	**1**

.

2.8	**Language Independent Tokenization**: Types of tokenization — Word, Character, and sub-word tokenization, problems with word tokenizer, drawbacks of a character- based tokenizer, problems with sub-word tokenization, Byte Pair Encoding	**1**
2.9	**StringMatchingandSpellingCorrection-**Minimum edit distance-table filling, dynamic programming,	**1**
Module3-WordEmbedding&Probabilistic Models(9Hours)		
3.1	**Vector Models & Text Preprocessing:** Vectors,Bag of Words,CountVectorizer, Tokenization, Stop words,	**1**
3.2	Stemming and Lemmatization,CountVectorizer,VectorSimilarity.TF-IDF,	**1**
3.3	Word-to-IndexMapping,BuildingTF-IDF	**1**
3.4	Neural Word Embeddings,Neural Word EmbeddingsDemo.VectorModels&Text Pre-processing Summary, steps of a typical NLP analysis	**1**
	PROBABLISTIC MODELS	
3.5	**Language Modelling:** types of language modeling, the importance of language modeling, the curse of dimensionality, Language Model Markov Assumption And N-Grams,	**1**
3.6	Language Model Implementation–Setup,Ngrams Function,UpdateCountsFunction, Probability Model Function, Reading Corpus, Language Model Implementation Sampling Text,	**1**
3.7	**Markov Models:**Markov Property,Markov Model,Probability Smoothing and Log-Probabilities, Building a Text Classifier,	**1**
3.8	**ArticleSpinning**–Problem Description,N-Gram Approach,implementation in python,	**1**
3.9	**Cipher Decryption with Language Modeling And Genetic Algorithm-** Ciphers, substitution cipher, bigrams, maximum likelihood, and log-likelihood, Language models, Genetic Algorithms,	**1**
Module4–NLPUsing Machine Learning Models (7 Hours)		
4.1	**Spam Detection–** Problem, Naive Bayes theorem, Intuition, spam detection using Naïve Bayes, class imbalance, ROC, AUC, AND F1 SCORE, Implementing spam detection in python,	**1**
4.2	**SentimentAnalysis**-Problem,LogisticRegressionIntuition,MulticlassLogistic Regression,	**1**
4.3	LogisticRegressionTrainingandInterpretation,sentimentanalysisimplementationin python	**1**
4.4	**TextSummarization**-UsingVectors,TextRankIntuition	**1**
4.5	TextRankinPython,TextSummarizationinPython	**1**
4.6	**TopicModeling**-differenttopicmodelingtechniques,LatentDirichletAllocation (LDA)–Essentials,LatentDirichletAllocation–TopicModelingwithLatentDirichlet	**1**
4.7	**Latent Symmatc Modelling(Indexing)-**LSA / LSI Introduction, Singular Value Decomposition Intuition, LSA / LSI: Applying SVD to NLP, Latent Semantic Analysis / Latent Semantic Indexing in Python	**1**

.

Module5-Deep Learning (8Hours)		
5.1	Word embeddings,nonlinear neural networks	**1**
5.2	**Neuron**–Intro,FittingaLine,ClassificationCodePreparation,TextClassificationin Tensorflow, The Neuron,How does a model learn?,	**1**
5.3	**Feed Forward Neural Networks-** Ann-introduction, The Geometrical Picture, Activation Functions, Multiclass Classification, Text Classification ANN in Tensorflow,	**1**
5.4	Text Preprocessing Code Preparation, Text Preprocessing in Tensorflow, Embeddings, CBOW(continuous bag of words), CBOW in Tensorflow	**1**
5.5	CONVOLUTION NEURAL NETWORKS:-CNN-Introduction,Convolution,pattern matching, weight sharing, convolution in color images,	**1**
5.6	**Convolution Neural Networks-** Convolution, pattern matching, weight sharing, convolution in color images, CNN Architecture, CNN for Text, CNN for NLP in Tensorflow,	**1**
5.7	**Recurrent Neural Networks-**SimpleRNN/Elman Unit,RNNs:PayingAttentionto Shapes, , GRU, and LSTM. RNN for Text Classification in TensorFlow,	**1**
5.8	Parts-of-SpeechTagging,andNamed Entity Recognition in TensorFlow	**1**

Reference Books

1. Sowmya Vajjala,Bodhisattwa Majumder,Anuj Gupta,Harshit Surana,Practical Natural Language Processing, Shroff/O'Reilly, 2020.

2. Steven Bird,Ewan Klein,Edward Loper,Natural Language Processing with Python,O‘Reilly

3. Akshay Kulkarni,Adarsha Shivananda,Natural Language Processing Recipes Unlocking Text Data with Machine Learning and Deep Learning using Python, Apress

4. Taweh Beysolow II _Applied Natural Language Processing with Python- Implementing Machine Learning and Deep Learning Algorithms for Natural Language Processing,Apress

5. Palash Goyal,Sumit Pandey,Karan Jain_Deep Learning for Natural Language Processing Creating Neural Networks with Python,Apress

6. Hobson Lane.Cole Howard, Hannes Max Hapke_Natural Language Processing in Action, Understanding, analyzing,and generating text with Python‘,©2019byManning Publications Co

8. Wolfgang Ertel,Introduction to Artificial Intelligence,Springer,

.

J1PADT114	DATA ANALYTICS	CATEGORY	L	T	P	CREDIT
		PEC-2	3	0	0	3

Preamble:

This course enables the students to understand the concepts of Data Analytics.It covers Data and Relations, Correlation, Basic Data Analytics and visualization methods using R, Finite State Machines, Dimensionality reductions, Feature extraction, Clustering, Classification and Regression Techniques, and scalability through parallelization. It helps the learners to develop applications for real time data analysis.

Course Outcomes:
After the completion of the course the student will be able to

CO1	Identifydataerrorsanddependenciesamongattributesbymodellingthemassets& relations. **(Cognitive Knowledge Level: Apply)**
CO2	Apply statistical methods for evaluation hypothesis(**CognitiveKnowledgeLevel: Apply)**
CO3	Apply regression,classification,andclustering models on a given dataset**(Cognitive Knowledge Level: Apply)**
CO4	Apply correlation techniques to find the dependencies between the features. **(CognitiveKnowledgeLevel: Apply)**
CO5	Develop applications that uses the concepts in Data Analytics**(Cognitive Knowledge Level: Create)**

Program Outcomes(PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams.

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program.

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards.

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects.

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓		✓	✓	✓	✓	
CO2	✓		✓	✓	✓	✓	
CO3	✓		✓	✓	✓	✓	
CO4	✓		✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom'sCategory	EndSemester Examination
Apply	60-80%
Analyse	20-40%

.

Evaluate	
Create	

Assignments or course projects can be used for higher level assessment of course outcomes.

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

Continuous Internal Evaluation Pattern:

Evaluation shall only be based on application, analysis or design-based questions (for both internal and end semester examinations).

Continuous Internal Evaluation:40marks

i. Preparing a review article based on peer reviewed original publications (minimum10 publications shall be referred) :15marks

i. Course base dtask/Seminar/Data collection and interpretation :15marks

i. Test paper(1number) :10 marks

Test paper shall include minimum80%of the syllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

End Semester Examination Pattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part Awill contain5 numerical/short answer questions with 1question from each module, having 5marks for each question.Students should answer all questions.Part B will contain7

.

questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry7 marks

Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. How to model stochastic and deterministic errors. Explain with examples.
2. What are the ways in which various errors can be handled?

CourseOutcome2(CO2):

1. What R commands would you use to remove null values from a dataset.
2. Which function R can be used to a fit nonlinear line to the data.

CourseOutcome3(CO3):

1. Consider the data sets for two classes X1 = {(0,0)} and X2 = {(1,0), (0,1)}. Which classification probabilities will a naive Bayes classifier produce for the feature vector (0,0)?
2. Explain SVM classifier with an example.

Course Outcome4(CO4):

1. For the datasetX={(1,0),(2,0),(3,1),(4,1),(5,1),(6,1),(7,0),(8,0)}compute chi-Square test statistic for 4 bins.
2. Explain the difference between correlation and causality.

CourseOutcome5(CO5):

.

1. Develop a small application for a manufacturing industry to maintain their works using the concepts in data analytics.
2. Develop a small application for improving Transportation System using the concepts in data analytics.

ModelQuestion Paper

QPCODE:

Reg No: ________________

Name: ____ PAGES:2

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRST SEMESTER M. TECH DEGREE EXAMINATION,MONTH&YEAR

Course Code: J1PADT114
Course Name:DATA ANALYTICS

Max.Marks:60 **Duration:2.5 Hours**

PARTA

Answer All Questions .Each Question Carries5 Marks

1. Compute the output of (a) an asymmetric moving mean filter,q=3, (b) an asymmetric moving median filter, q=3, (c) an exponential filter, y0=0, Correction term n=0.5 for the time series (0,0,0,1,0,0,0).Which filter result do you like best?
2. Differentiate betweenType I and Type II errors.
3. Construct the data tuples for an autoregressive forecasting model with a time horizon of m = 2 for the time series x = (1,2,3,5,8).
4. Explain prototype-based clustering.
5. Explain Rectified linear activation unit. **(5x5=25)**

.

PartB

(Answer any five questions.Each question carries7 marks)

6.	Explain stochastic and deterministic errors with examples. Using 2-sigma rule and m-sigma rule how a value is classified as outliers?	(7)
7.	Consider the datasets for two classesX1={(0,0)} and X2={(1,0),(0,1)}. Which classification probabilities will a naive Bayes classifier produce for the Feature vector(0,0)?	(7)
8.	Illustrate the importance of visualizing data before analysis.	(7)
9.	What is feature scaling?Explain the feature scaling techniques.	(7)
10	Justify how the computational complexity of the nearest neighbor is reduced by the LVQ approach.	(7)
11	For the data set X={(1,0),(2,0),(3,1),(4,1),(5,1),(6,1),(7,0),(8,0)} Compute chi-square test statistic for 4 bins.	(7)
12	Consider the two-dimensional patterns(2,1),(3,5),(4,3),(5,6),(6,7),(7,8). Compute the principal component using PCA Algorithm.Use PCA Algorithm to transform the pattern (2, 1) onto the Eigen vector.	(7)

.

Syllabus: Error Handling, Correlation, Models,Clustering,Data and Process Parallelization, Batch processing frameworks.

	Syllabus	
Module	**Content**	**Hours**
1	**Data and Relations** - Data scales, Set and Matrix representations, Relations, Similarity and dissimilarity measures, Sequence relations. Data pre-processing - Error types, error handling, filtering, transformation, merging. Correlation - Linear, Causality, Chi-Square tests. Cross validation and feature selection.	**7**
2	**BasicDataAnalyticsMethodsUsingR-**DescriptiveStatistics, Statistical methods for evaluation, Hypothesis Testing, ANOVA. **VisualizationmethodsusingR-Exploratory**DataAnalysis, visualizing single Variable, Examining Multiple Variables	**9**
3	**Models-**Finitestatemachines,Recurrentmodels,Autoregressivemodels, Moving Average Models.	**6**
4	**Clustering -** Cluster partitions, Sequential clustering, Prototype based clustering, Relational clustering, Cluster tendency assessment, Cluster validity, Self-organising map (SOP). Use Cases **Regression-**LinearRegression,Logisticregression,UseCases	**7**
5	**Classification Methods-** Naive Bayes classifier, Decision Trees, LDA, SVM, Learning Vector Quantization. **Scalability through parallelization** - Data parallelization, Process parallelization, Scaling using feature engineering, Dimensionality Reduction, Cascading, Feature reduction through spatial transforms. Global Average Pooling, Data Augmentation, **CaseStudies:**ReLUnonlinearity,MLP,ConvolutionalLayer.	**11**

.

CoursePlan

No	Topic	No.ofLectures
1	**Data and Relations,Correlation**	
1.1	Data scales,Set and Matrix representations	1
1.2	Relations,Similarity and dissimilarity measures	1
1.3	Sequence relations.	1
1.4	Data pre-processing-Error types,error handling,Filtering	1
1.5	Transformation,merging	1
1.6	Correlation-Linear,Causality Chi-Square tests	1
1.7	Cross validation and feature selection	1
2	**Basic Data Analytics Methods Using R**	
2.1	Introduction to R and GUI	1
2.2	Attributes and Data Types	1
2.3	Descriptive Statistics	1
2.4	Statistical methods for evaluation and Hypothesis	1
2.5	Hypothesis Testing-Difference of Means	1
2.6	Type–I,Type–II Errors,Problems	1
2.7	ANNOVA	1
2.8	Visualization–Single variable.	1
2.9	Examining multiple variables	1

.

3	**Models**	
3.1	Finite state machines	1
3.2	Recurrent models	1
3.3	Autoregressive models	1
3.4	Autoregressive models	1
3.5	Moving AverageModels-ARIMA	1
3.6	Moving Average Models-ARMA	1
4	**Clustering**	
4.1	Cluste rpartitions	1

4.2	Sequential clustering,Prototype based clustering	1
4.3	Relational clustering,	1
4.4	Cluster tendency assessment,Cluster validity	1
4.5	Self-organisingmap(SOP).	1
4.6	Regression: Linear regression	1
4.7	LogisticRegression,Use cases	1
5	**Classification**	
5.1	NaiveBayesclassifier	1
5.2	LDA	1
5.3	SVM	1

.

5.4	LearningVectorQuantization	1
5.5	Scalabilitythroughparallelization	1
5.6	Dataparallelization,Processparallelization	1
5.7	Scalingusingfeatureengineering	1
5.8	Cascading,Featurereductionthroughspatialtransforms	1
5.9	ReLUnonlinearity	1
5.10	DataAugmentation,MLP,ConvolutionalLayer	1
5.11	GlobalAveragePooling,DimensionalityReduction	1

Reference Books

1. Thomas A. Runkler, —DataAnalytics-ModelsandAlgorithmsforIntelligentData Analysis‖, Springer 2012.
2. Stefanos Vrochidis, BenoitHuet, EdwardChang, IoannisKompatsiaris, —BigDataAnalysis for Large-Scale Multimedia Search‖, Wiley 2019.
3. J. O. Moreira, Andre Carvalho, Tomas Horvath, —AGeneralIntroductiontoData Analytics‖, Wiley 2019.
4. —Data Science and Big Data Analytics, Discovering, Analyzing, Visualizingand Presenting Data‖ EMC EducationService.
5. GouzhuDongandHuanLiu, —FeatureEngineeringForMachineLearningandData Analytics —CRCPress.

.

J1PADT124	R FOR DATA SCIENCE	CATEGORY	L	T	P	CREDIT
		Program Elective1	3	0	0	3

Preamble:

The course introduces the R programming environment and its use in Data Science. It covers data loading and organization, data exploration and cleaning, building machine learning models, statistical models, and documentation and data visualization using R. The course enables learners to use R programming in data analytics related tasks for making predictions.

CourseOutcomes:

Afterthecompletionofthecoursethestudentwillbeableto

CO 1	Organize,explore,cleanandanalysedatatofindrelativepatternsindata.(**Cognitive Knowledge Level: Apply**)
CO 2	Designdifferentmachinelearningmodelstomakepredictionsfromdata.(**Cognitive Knowledge Level: Apply**)
CO 3	Demonstratethepatternsindatausingvariousdatavisualizationpackages.(**Cognitive Knowledge Level: Apply**)
CO4	Buildpredictionmodelsfordifferenttypesofdata,evaluateandvalidatethem.(**Cognitive Knowledge Level: Apply**)
CO5	ApplyRprogrammingskillstosolvereal-lifedataanalyticsproblems.(**Cognitive Knowledge Level: Apply**)

ProgramOutcomes(PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams.

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program.

PO4: Anabilitytoapplystreamknowledgetodesignor developsolutionsforrealworldproblemsby following the standards.

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects.

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Apply	60
Analyse	40
Evaluate	-
Create	-

Assignmentsorcourseprojectscanbeusedforhigherlevelassessmentofcourseoutcomes.

Markdistribution

Total Marks	CIE	ESE	ESEDuration
100	40	60	2.5 hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication, analysisordesignbasedquestions (forboth internal and end semester examinations).

ContinuousInternalEvaluation:40marks

.

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications (minimum10 publications shall be referred) : 15 marks

i. Coursebasedtask /Seminar/Datacollectionandinterpretation :15 marks

i. Testpaper(1number) :10 marks

Testpapershallincludeminimum80%ofthesyllabus.

Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

Theendsemester examinationwillbeconductedbythe respectiveCollege.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marksfor eachquestion.Studentsshouldanswerall questions.Part B willcontain7 questions (such questions shall be useful in the testing of overallachievement and maturity ofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks.

Totaldurationofthe examinationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. Givenavectorofvalues, demonstratehowwouldyouconvertitintoatimeseries object?
2. Ifx=c (1,2,3,3,5,3,2,4, NA), what arethelevelsoffactor(x)?
3. Writeacustomfunctionwhichwillreplaceallthemissingvaluesinavectorwiththemeanof values.

CourseOutcome2(CO2)

.

1. Demonstratetheuseofanyfivegenericfunctionsforextractingmodelinformation.

2. ElaborateonhowAnovamodelcanbeusedfordataanalysis.

3. Describehowtofitanon-linearregressionmodelinR.

CourseOutcome3(CO3):

1. Withthehelpofanexample, showhowtocreatescatter plotusingRlibraries.

2. Analysethetypeofcharttobeusedwhentryingtodemonstratetherelationshipbetween variables/parameters.

3. Suggestscenarioswhereyouwoulduseabarchartandahistogram. Justify.

CourseOutcome4(CO4):

1. Suppose, givenadataset (x1, y1), (x2, y2), (x3, y3) …. (xn, yn) ofnobservationfroman experiment fit a straight line y=a+bx to the given data.

2. Compare the situation where you want to compare your data distribution with another. How would you accomplish using R, the scenario where you need to check if a sample follows a normal distribution or not or if two samples are drawn from the same distribution?

3. Given a dataset, elaborate on how you would choose a suitable model for prediction. Justify your answer.

CourseOutcome5(CO5):

1. Write an R program to read a given a dataset, clean the data, organize the data and build a suitable machine learning model to make predictions. Also, evaluate and validate the model.

2. Suppose you are asked to build a model for spam detection, discuss about the method you would follow if you would prefer using: a) supervised learning method b) an unsupervised learningapproach. Analyse which one would be better.

3. Choose any real-life scenario and apply R programming skills to develop a good machine learning model for making predictions. Also, evaluate the performance of your model.

.

ModelQuestionPaper

QPCODE:

Reg No: __________________

Name: ____ PAGES:4

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRSTSEMESTERM.TECHDEGREEEXAMINATION, MONTH&YEAR

Course Code: J1PADT124

CourseName: RforDataScience

Max.Marks: 60 **Duration:2.5Hours**

PARTA

AnswerAllQuestions. EachQuestionCarries5 Marks

1. Demonstratetheworkingofwhichfunctioninvectorswiththehelpofsuitable examples.

Compareandcontrastlistanddataframes.

23. Elaboratetheconceptofgeneralizedlinearmodelandtheuseofglm () function.

4. Withthehelpofasuitableexample, illustratehowaccuracyofamodelis evaluated.

5. ShowhowthedifferentplotfunctionsinRcouldbeusedfordatavisualization. **(5x5=25)**

PartB

(Answeranyfivequestions.Eachquestioncarries7 marks)

6. Illustrateanyfivevectormethodswithappropriateexamples. **(7)**

7. Write an R code to generate an uppertriangular matrix in R. Convert the same into a lower triangular matrix. **(7)**

8. Compare linear regression and logistic regression. Explain how to create a linear regression model in R. **(7)**

9. The number of awards earned by students at one high school. Predictors of the number of awards earned include the type of program in which the student was enrolled (e.g., vocational, general or academic) and the score on their final exam in math. Select a suitable regression model to perform the analysis. Also write thecode for the same. **(7)**

10 DescribethecommonprobabilitydistributionfunctionsinRandtheirapplications. **(7)**

11 Justifytheuseofscatterplotindataanalyticsbyquotingsuitableapplication. **(7)**

.

12	Given is the _diamonds‘ dataset in R whichis part of theggplot2 library. It contains prices of approximately 50000 round cut diamonds. How would you use an approach to plot a histogram that will display a type of diamonds based on the quality of cut (Ideal, Premium, Very Good, Good and Fair)?	**(7)**

Syllabus

	Syllabus	
Module	**Content**	**Hours**
1	**Introduction -** Reading and getting data into R, Vectors and assignment, Logical and Index vectors, Generating regular sequences, Missing values, Ordered and Unordered Factors, The function tapply() and ragged arrays, Ordered factors.	**9**
2	**Exploring and cleaning data for analysis -** Reading data from files, Data organization, Arrays and Matrices, Basics of Arrays in R, Matrixoperations, Advanced Matrix operations, Additional Matrix facilities, Lists and Data frames.	**11**
3	**Buildingmachinelearningmodels-**Buildinglinearmodels,Generalized linear models, Nonlinear least squares and maximum likelihood models.	**8**
4	**Evaluatingand Validatingmodels-** Evaluating and Validatingmodels, Probability distributions in R, Statistical models in R.	**5**
5	**Data Visualization -** Documentation, Graphical analysis, plot() function, Displaying multivariate data, Using graphics parameters, Matrix plots, Exporting graphs, ggplot package.	**7**

CoursePlan

No	**Topic**	**No. ofLectures**
1	**IntroductiontoR**	
1.1	IntroductiontoRenvironment, InstallationofRenvironmentandR studio.	1
1.2	VariablesanddatatypesinR	1
1.3	ReadingandgettingdataintoR	1
1.4	VectorsandAssignment	1
1.5	LogicalandIndex vectors	1
1.6	Generatingregularsequences	1
1.7	Missingvalues,OrderedandUnorderedFactors	1
1.8	Thefunctiontapply()andraggedarrays	1
1.9	Orderedfactors	1

.

2	**Exploringandcleaningdataforanalysis**	
2.1	Readingdatafromfiles	1
2.2	Dataorganization	1
2.3	ArraysandMatrices	1
2.4	BasicsofArraysinR	1
2.5	Matrixoperations	1
2.6	AdvancedMatrixoperations	1
2.7	AdditionalMatrixfacilities	1
2.8	IntroductiontoLists	1
2.9	IntroductiontoLists	1
2.10	Dataframes	1
2.11	Dataframes	1
3	**Buildingmachinelearningmodels**	
3.1	Buildinglinearmodels	1
3.2	Buildinglinearmodels	1
3.3	Generalizedlinearmodels	1
3.4	Generalizedlinearmodels	1
3.5	Nonlinearleastsquares	1
3.6	Nonlinearleastsquares	1
3.7	Maximumlikelihoodmodels	1
3.8	Maximumlikelihoodmodels	1
4	**EvaluatingandValidatingmodels**	
4.1	EvaluatingandValidatingmodels	1
4.2	Probability distributionsinR	1
4.3	ProbabilitydistributionsinR	1
4.4	StatisticalmodelsinR	1
4.5	StatisticalmodelsinR	1
5	**DataVisualization**	
5.1	Graphicalanalysis,plot()	1
5.2	Displayingmultivariatedata	1
5.3	Usinggraphicsparameters	1
5.4	Matrixplots	1
5.5	Exportinggraphs	1
5.6	ggplotpackage	1
5.7	Documentation	1

References:

1. RogerD.Peng, ―RProgrammingforDataScience‖, LeanPublishing,2015.

2. NinaZumel, JohnMount"PracticalDataSciencewithR.ManningPublications.2014

3. NathanYau,"VisualizeThis: TheFlowingdataGuidetoDesign, VisualizationandStatistics", Wiley, 2011.

4. Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman, "Mining of Massive Datasets‖, Cambridge University Press, 2014.

.

5. Tilman M. Davies, _TheBook ofR-AFirst Coursein R Programming and Statistics‘, No Starch Press, 2016.

6. TonyOjeda, SeanPatrickMurphy, BenjarninBengfort.AbhijitDasgupta."PracticalData Science Cookbook", Packt Publishing Limited, 2014.

7. W.N. Venables, D.M. SmithandtheRCoreTeam,"AnIntroductiontoR",2013

.

J1PADT134	Data visualization with python	CATEGORY	L	T	P	CREDIT
		Program Elective1	3	0	0	3

Preamble: This course is intended to provide basic concepts of Data Visualization. This course helps students learn visualization libraries in python and apply them to obtain and understand the underlying information. This course helps students to implement the visualization techniques to build an interactive dashboard

CourseOutcomes:

Afterthecompletionofthecourse, thestudentwillbeableto

CO1	AnalyzetheneedfordataVisualizationinDataAnalytics(**Cognitiveknowledgelevel: Apply)**
CO2	IdentifytherighttoolforDataVisualizationbasedontheDataAnalyticsProblem (**Cognitive knowledge level: Apply)**
CO3	Utilizetherightvisualizationtechniquetoobtaininformation, knowledge, andinsight from a particular Dataset. (**Cognitive knowledge level: Apply)**
CO4	ImplementVisualizationtasksusingPython(**Cognitiveknowledgelevel:Apply)**
CO5	ApplyPlotlytocreateplotslikeBarCharts,LineCharts,ScatterPlots,andHeatMaps, and create and deploy an interactive dashboard (**Cognitive knowledge level: create)**

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1: Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams

PO2: An ability to communicate effectively and write and present technical reports on complex engineeringactivitiesbyinteractingwiththeengineeringfraternityandwithsocietyat large.

PO3: Anabilitytodemonstrateadegreeofmasteryover theareaasperthespecializationofthe program. Themastery should be at a level higher than the requirements inthe appropriate bachelor‘s program

PO4: Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforreal-worldproblems by following the standards

PO5: Anabilitytoidentify, selectandapplyappropriatetechniques, resourcesandstate-of-the-art tools to model, analyze and solve practical engineering problems.

PO6: Anabilitytoengageinlifelonglearningforthedesignanddevelopmentrelatedtothestream- related problems taking into consideration sustainability, societal, ethical, and environmental aspects

PO7: Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagementand finance which focus on Entrepreneurship and Industry relevance.

.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO 1	✓	✓	✓	✓	✓	✓	
CO 2	✓	✓	✓		✓	✓	
CO 3		✓		✓		✓	✓
CO 4		✓	✓		✓	✓	
CO 5	✓	✓		✓		✓	✓

AssessmentPattern

Bloom'sCategory	**EndSemesterExamination**
Apply	60-80%
Analyse	20-40%
Evaluate	Assignments/Project
Create	Assignments/Project

Markdistribution

Total Marks	**CIE**	**ESE**	**ESEDuration**
100	40	60	2.5 hours

ContinuousInternalEvaluationPattern:

Theevaluationshallonlybebasedonapplication, analysis, ordesign-basedquestions (forboth internal and end-semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeer-reviewedoriginalpublications (minimum10 publications shall be referred) : 15 marks
i. Coursebasedtask/Seminar/Datacollectionand interpretation :15 marks
i. Testpaper(1number) :10 marks

Testpapershallincludeaminimumof80%ofthesyllabus.

Course-basedtask/testpaperquestionsshallbeusefulinthetestingofknowledge, skills, comprehension, application, analysis, synthesis, evaluation, and understanding of the students.

EndSemesterExaminationPattern:

TheendsemesterexaminationwillbeconductedbytherespectiveCollege. There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answerquestionswith 1 question from each module, having 5 marksforeachquestion. Studentsshouldanswerallquestions. PartBwillcontain7questions (such

.

questions shall be useful in the testing of overallachievement and maturity ofthe students in a course, through long

answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. a. Distinguishcounthistogram, relativefrequencyhistogram, cumulativefrequency histogram, and density histogram

   b. ForwhattypeofdataisaHistogramplotusuallyused?

   c.ExplainthevariousfeaturesthatyoucanfindfromHistogramplots
2. a. ExplainthesignificanceofBoxplots

   b. Listthe informationyou couldgainfromabox plot.
3. a. Whatisascatterplot? Forwhattypeofdataisascatterplotusuallyused?

   b. Listthe featuresthatmightbevisibleinscatterplots

   c.Choosethetypeofplotyouwouldliketouseifyouneedtodemonstrate ―the relationship‖ between variables/parameters

CourseOutcome2(CO2)

1. ExplainthedifferenttypesofDataStructuresavailableinpandasandtheadvantages of pandas
2. Explainthemethodsofchoosingtheright visualizationtool
3. Explainthedifferenttypesofdataandlevelsofmeasurementandhowtheyinfluence a data analyst to choose a chart for the data he would like to visualize

CourseOutcome3(CO3):

1. Discussthedifferentdatatypes, NumPycansupportanddiscoverchanging the type of the variable affects the data it stores.
2. Explainthegroupingandaggregationmethodsinpandas
3. Lookat thefollowing data

```
countrycontinentyearlifeExppop          gdpPercap
0Afghanistan    Asia    195228.8018425333779.445314
```

.

```
1Afghanistan    Asia     1957 30.3329240934820.853030
2Afghanistan    Asia     1962  31.99710267083853.100710
3Afghanistan    Asia     1967 34.02011537966836.197138
4Afghanistan    Asia   1972  36.08813079460739.981106
5Afghanistan    Asia   1977  38.43814880372786.113360
6Afghanistan    Asia   1982   39.85412881816978.011439
7Afghanistan    Asia   1987   40.82213867957852.395945
8Afghanistan    Asia   1992   41.67416317921649.341395
9Afghanistan    Asia   1997   41.76322227415635.341351
```

Foreachyearinourdata, estimatetheaveragelifeexpectancy. Also, appraiseabout population and GDP.

CourseOutcome4(CO4):

1. a. Explaintheseabornlibrary.Doesseabornrequirematplotlib
   b. Explaintheloaddataset () methodwithanexample
   c.Discusscreatinglineplot, violinplot, andhistograminseaborn
2. DiscussseveralattributesofthePandasdataframeobjectthatyouwillfrequently need while cleaning, pre-processing, or analyzing a data set.
3. a. Explaintheseabornlibrary.Doesseabornrequirematplotlib
   b. Explaintheloaddataset () method withanexample
   c.Discusscreating lineplot, violinplot, andhistograminseaborn

CourseOutcome5(CO5):

1. Useplotlytocreateinteractiveplots,
2. Createmaintypesofplotswithplotlyandpython
3. workontheactualDASHlibraryfromplotlytobeginservingcomponentsandseveral plots as a web app in our browser. So that is going to be a little different than just a singular plotly plot. Instead, combine multiple things and have a full-service dashboard that is basically a Web service or Web app in your browser.
4. Createdataandexploreadvancedandcomplexfeaturesofthedash-multipleinputs and outputs, interactive components, controlling callbacks with state, and linking together grouped Plotly plots
5. Using Dash dashboard and python, create a basic Web app that will automatically lookupandservestocktickerdataforyouintwosetsoftimestampsthatyougetto choose.

.

Model Question Paper
QPCODE:
Reg No: ________________
Name: ______________________________ **PAGES:4**
APJABDULKALAMTECHNOLOGICALUNIVERSITY
FIRSTSEMESTERM.TECHDEGREEEXAMINATION, MONTH&YEAR
CourseCode:J1PADT134
CourseName: Data visualization with python
Max.Marks:60 **Duration:2.5Hours**
PARTA(5x5)
AnswerAllQuestions. EachQuestionCarries5 Marks

1.	Explain the importance of DataVisualization	**5**
2.	WriteaPandasprogramtogetthepowersofarrayvalueselement-wise. Note: First array elements raised to power the form the second array Sampledata: {'X‘: [78,85,96,80,86],'Y‘: [84,94,89,83,86],'Z‘: [86,97,96,72,83] Writethesampleoutput	**5**
3.	ExplainthemethodstoconvertStringtodate.	**5**
4.	a. Createascatterplot of1000randomdatapoints. b. ExplainthestepstomakeaCSVfileupdateautomatically	**5**
5.	a. Compareplotlyandmatplotlib b. Defineunivariatedistributionsandgraphsforvisualizingunivariate distributions	**5**

PartB

(Answeranyfivequestions.Eachquestioncarries7 marks)

6.	(a)	ExplainthedifferenttypesofDataVisualizationcharts.	**(7)**
7.	(a)	a.ExplainthedataframeinPandas	**(3)**
	(b)	Explainthefollowingdataframemethodsandcomputationswith examples i. MinandMax ii. SumandCount iii. Mean, MedianandMode iv. DescribewithNumericvalues v. DescribewithObject(with)Textvalues	**(4)**
8.	(a)	WriteaPandasprogramtogetthefirst3rowsofagivenDataFrame. Sample Data Frame: exam_data= {'name‘: ['Anastasia','Dima','Katherine','James','Emily','Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'], 'score‘: [12.5,9,16.5, np.nan,9,20,14.5, np.nan,8, 19], 'attempts‘: [1, 3,2,3,2,3,1,1,2,1], 'qualify‘: ['yes','no','yes','no','no','yes','yes','no','no','yes']} labels=['a','b','c','d','e','f','g','h','i','j'] Write the sample output	**(3)**

.

	(b)	Whatisthenameofpandaslibrarytoolsusedtocreateascatterplot matrix?	**(1)**
	(c)	WriteaPandasprogramtomergetwogivendataframeswithdifferentcolumnsand give a sample output TestData: data1: key1key2PQ 0 K0K0P0Q0 1 K0K1P1Q1 2 K1K0P2Q2 3 K2K1P3Q3 data2: key1key2RS 0 K0K0R0S0 1 K1K0R1S1 2 K1K0R2S2 3 K2K0R3S3	**(3)**
9.	(a)	ElucidatethemethodofcreatingaScatterPlotwithseveralcolorsinMatplotlibwith thehelpofanexample	**(2.5)**
	(b)	ExplainthemethodofaddingalegendtoascatterplotinMatplotlib	**(2)**
	(c)	ExplainthemethodofincreasingthesizeofscatterpointsinMatplotlib	**(2.5)**
10	(a)	ExplainTimeSeriesinpandas	**(2)**
	(b)	HowwillyoucreateaseriesfromdictinPython?	**(2)**
	(c)	ExplainthefollowingoperationsonSeriesinpandas. BuildadashboardusingdatafromaCSVfile i. Concatenatingseries ii. Concatenatingseriesbyindexing iii. Concatenatingseriesbycolumn iv. DataframeMerge()method	**(3)**
11	(a)	BuildadashboardusingdatafromaCSVfile	**(7)**
12	(a)	Whatisaninteractivedashboard?	**(1)**
	(b)	ExplainDashComponents: HTMLcomponents, corecomponents, andMarkdown, UsingHelp()withDash	**(3)**
	(c)	ExplainSingleCallbackforInteractivity	**(3)**

.

Syllabus

Mod	Content	Hrs
1	**PYTHONVISUALIZATIONCHARTS&INTRODUCTIONTONUMPYAND PANDAS:** Data visualization, importance, advantages, Categories and tools, design principles, listing libraries of python, JavaScript, and R, Dimensions and measures, types of data, visualizing charts using python, general theory, creation, interpretation, conditions, and disadvantages- Bar chart, Pie chart, stacked areachart, Line chart, Histogram, scatter plot, regression plot, Combining Bar and Line chart, **Numpy** –Array, NaNand INF, Statistical Operations, Shape, Reshape, Ravel, Flatten, Sequence, Repetitions, and Random Numbers, Where, File Read and Write, Concatenate and Sorting, Dates, **Pandas-**Data Frame and Series, File Reading and Writing, Info, Shape, Duplicated, and Drop, Columns, NaN and Null Values, Imputation, Lambda Function,	8
2	**PANDAS-** Statistical functions, **Data Visualisation With Pandas-** Line Plot, Bar Plot, Stacked Plot, Histogram, Box Plot, Area and Scatter Plot, Hex and Pie Plot, Scatter Matrix and Subplots, **Series And Columns-**Selecting A Single and multiple Column, Series Methods, The powerful value_counts() method, Using plot() to visualize **Indexing And Sorting-** Set_Index Basics, set_index: The WorldHappiness Index Dataset, setting index with read_csv, sort_values intro, sorting by multiple columns, sorting text columns, sort_index, Sorting and Plotting!, loc, iloc, loc& iloc with Series, **Filtering Data Frames-** Filtering data frames with a Boolean series, Filtering With Comparison Operators, The Between Method, The isin() Method, Combining Conditions Using AND (&). Combining Conditions Using OR (|), Bitwise Negation, isna() and notna() Methods, Creating and Adding, dropping, And Removing Columns and rows	8
3	**PANDAS & MATPLOTLIB: UPDATING VALUES-**Renaming Columns and Index Labels, The replace () method, Updating Multiple Values Using loc[],Updates With loc[] and Boolean Masks, **Working With Types:-** Casting Types With astype(), Introducing the Category Type, Casting With pd.to_numeric(), dropna() and isna(), fillna(), Working With Dates And Times **Matplotlib** –LinePlot, Label, Scatter, Bar, and Hist Plots, Box Plot, Subplot, Pie Plot Text Color, Nesting,&Labeling, Bar Chart, Line plot &Scatter plot on Polar Axis, Animation Plot, **Pandas Plotting** –Changing Plot Styles, Adding Labels and Titles, rename (),Multiple plots on The Same Axes, Automatic Subplots, Manual Subplots With Pandas,Exporting Figures With savefig(), Grouping And Aggregating	8
4	**PANDAS & SEABORN: PANDAS:** Hierarchical Indexing in pandas, working with Text in pandas, Pandas- Apply, Map and Applymap-, Combining Series and Dataframes- **Seaborn: -** The Helpful load_dataset () method, Seaborn Scatterplots, Line plots. The relplot () Method, Resizing Seaborn Plots: Aspect & Height, Histograms, KDE Plots, Bivariate Distribution Plots, Rugplots, The Amazing displot () Method. **Seaborn Categorical Plots-** Countplot, Strip & Swarm Plots, Boxplots, Boxenplots, Violinplots, Barplots, The Big Boy Catplot Method, **Controlling Seaborn Aesthetics-** Changing Seaborn Themes, Customizing Styles with set_style(), Altering Spines With despine(), Changing Color Palettes	6
5	**PYTHONDASHBOARDSWITHPLOTLYANDDASH:**PLOTLYBASICS-	10

.

	Plots, Line Charts, Bar Charts, Bubble Plots, Box Plots, Histograms, Distplots,, Heatmaps, DASH BASICS –dash layouts, - styling, Converting Simple Plotly Plot to Dashboard with Dash, create a simple dashboard, **Dash Components,** HTML Components, Core Components, Markdown with Dash, Using Help() with Dash INTERACTIVE COMPONENTS- Single Callbacks for Interactivity, Dash Callbacks for Graphs, Multiple Inputs &Outputs, Controlling Callbacks with Dash State, INTERACTING WITH VISUALISATION -Hover Over Data, Click Data, Selection Data, Updating Graphs on Interactions- Project: Building an Interactive dashboard with Plotly and Dash	

CoursePlan

S.NO	TOPIC	NO. OF LECTURES
MODULE 1-PYTHONVISUALIZATIONCHARTS&INTRODUCTIONTONUMPYAND PANDAS:8hours		
1.1	Data visualization, importance, advantages, Categories and tools, design principles, listing libraries of python, JavaScript, and R, Dimensions and measures, types of data,	1
1.2	visualizingchartsusingpython,generaltheory,creation,interpretation,conditions, and disadvantages- Bar chart, Pie chart,	1
1.3	Stackedareachart,Linechart	1
1.4	Histogram,scatterplot	1
1.5	Regressionplot,CombiningBarandLinechart,	1
1.6	**NUMPY** –Array, NaN and INF, Statistical Operations, Shape, Reshape, Ravel, Flatten, Sequence, Repetitions, and Random Numbers, Where, File Read and Write, Concatenate and Sorting, Dates,	1
1.7	**PANDAS-**DataFrameandSeries,FileReadingandWriting,Info,Shape, Duplicated, and Drop, Columns,	1
1.8	NaNandNullValues,Imputation,LambdaFunction,	1
MODULE2-PANDAS-8hours		
2.1	**PANDAS-**Statisticalfunctions	1
2.2	**DataVisualisationWithPandas**-LinePlot,BarPlot,StackedPlot,Histogram, Box Plot, Area and Scatter Plot, Hex and Pie Plot, Scatter Matrix and Subplots,	1
2.3	**SeriesAndColumns**-SelectingASingleandmultipleColumn,SeriesMethods, The powerful value_counts() method, Using plot() to visualize	1
2.4	SelectingMultipleColumns,Thepowerfulvalue_counts()method,Usingplot() to visualize!, EXERCISE: Series & Plotting	1
2.5	**Indexing And Sorting-** Set_Index Basics, set_index: The World Happiness Index Dataset, settingindex withread_csv, sort_values intro,sorting bymultiple columns, sorting text columns, sort_index, Sorting and Plotting, loc, iloc, loc & iloc with Series	1

.

2.6	**FilteringDataFrames-**FilteringdataframeswithaBooleanseries,Filtering With Comparison Operators, The Between Method, The isin() Method,	1
2.7	CombiningConditionsUsing AND (&).CombiningConditionsUsingOR (|), Bitwise Negation, isna() and notna() Methods,	1
2.8	CreatingandAdding,dropping,AndRemovingColumnsandrows	1
MODULE3-PANDAS&MATPLOTLIB-8hours		
3.1	**UpdatingValues-**RenamingColumnsandIndexLabels,Thereplace()method, Updating Multiple Values Using loc[], Updates With loc[] and Boolean Masks	1
3.2	**WorkingWithTypes:-**CastingTypesWithastype(),IntroducingtheCategory Type, Casting With pd.to_numeric(), dropna() and isna(), fillna(),	1
3.3	WorkingWithDatesAndTimes	1
3.4	**Matplotlib**–LinePlot,Label,Scatter,Bar,andHist Plots,Box Plot, Subplot, Pie Plot Text Color, Nesting,&Labeling,	1
3.5	BarChart,Lineplot&ScatterplotonPolarAxis,AnimationPlot,	1
3.6	**PandasPlotting**–ChangingPlotStyles,AddingLabelsandTitles,rename(), Multiple plots on The Same Axes, Automatic Subplots,	1
3.7	ManualSubplotsWithPandas,ExportingFiguresWithsavefig(),	1
3.8	GroupingAndAggregating	1
MODULE4-PANDAS&SEABORN-6hours.		
4.1	**PANDAS:**HierarchicalIndexinginpandas	1
4.2	WorkingWithTextinpandas	1
4.3	Pandas-Apply,MapAndApplymap	1
4.4	CombiningSeriesAndDataframes-	1
4.5	**SEABORN: -** The Helpful load_dataset () method, Seaborn Scatterplots, Line plots. The relplot() Method, Resizing Seaborn Plots: Aspect & Height, Histograms, KDE Plots, Bivariate Distribution Plots, Rugplots, The Amazing displot() Method	1
4.6	**SEABORN CATEGORICAL PLOTS-** Countplot, Strip & Swarm Plots, Boxplots, Boxenplots, Violinplots, Barplots, The Big Boy Catplot Method, **CONTROLLING SEABORN AESTHETICS-** Changing Seaborn Themes, Customizing Styles with set_style(), Altering Spines With despine(), Changing Color Palettes	1
MODULE5-PYTHONDASHBOARDSWITHPLOTLYANDDASH=10hours		
5.1	**PLOTLYBASICS-**Plots,LineCharts	1
5.2	BarCharts,BubblePlots,BoxPlots	1
5.3	Histograms,Distplots,,Heatmaps,	1
5.4	**DashBasics**–dashlayouts,-styling,ConvertingSimplePlotlyPlotto DashboardwithDash, creatingasimpledashboard,**DashComponents,**HTML	1

.

	Components,CoreComponents,MarkdownwithDash,UsingHelp()withDash	
5.5	**InteractiveComponents-**SingleCallbacksforInteractivity,DashCallbacksfor Graphs	1
5.6	MultipleInputs&Outputs,ControllingCallbackswithDashState,	1
5.7	INTERACTINGWITHVISUALISATION-HoverOverData,ClickData	1
5.8	SelectionData,UpdatingGraphsonInteractions	1
5.9	Project:BuildinganInteractivedashboardwithPlotlyandDash	1
5.10	Project:BuildinganInteractivedashboardwithPlotlyandDash	1

ReferenceBooks

TEXTBOOKS:

1. Dr Ossama Embarak, ‗Data Analysis and Visualisation Using python- Analyze Data to Create Visualizations for BI Systems -Apress
2. KirthiRaman, ―MasteringPythonDataVisualization‖, PacketPublishingLtd, UK
3. Elias Dabbas, Interactive Dashboards and Data Apps with Plotly and Dash: Harness the power of a fully-fledged frontend web framework in Python ̊ no JavaScript required, ISBN: 9781800568914
4. FabioNelli, ―PythonDataAnalyticswithpandas, NumPy, Matplotlib‖, Second Edition, Apress
5. MattHarrisonandMichaelPrentiss, ―LearningthePandasLibrary-Pythontools forDataMunging, Data Analysis, and Visualisation‖
6. DanielY.Chen, ―PandasforEveryone: PythonDataAnalysis,1ePaperback‖

.

J1PADT144	REINFORCEMENT LEARNING	CATEGORY	L	T	P	CREDIT
		Program Elective1	3	0	0	3

Preamble:

This course provides the basic concepts and advanced techniques in reinforcement learning. This course covers policies and value functions, Q-learning, function approximation and policy optimization methods. This course helps the learners to acquire key skills required to develop applications in the exciting area of reinforcement learning.

CourseOutcomes:

Afterthecompletionofthecoursethestudentwillbeableto

CO 1	Utilizethekeyfeaturesofreinforcementlearningdevelopingmachine learning applications. (**Cognitive knowledge level: Apply**)
CO 2	ExploresuitablelearningtaskstowhichRLtechniquescanbeapplied. (**Cognitiveknowledgelevel: Apply**)
CO 3	ApplycoreprinciplesbehindtheRL, includingpolicies, valuefunctions, deriving Bellman equations. (**Cognitive knowledge level: Apply**)
CO 4	Implementandanalyzeapproximatesolutions(**Cognitiveknowledgelevel:Analyze**)
CO 5	AnalyzecurrentadvancedtechniquesandapplicationsinRL(**Cognitiveknowledge level: Analyze**)
CO 6	Practiceusingpopularopen-sourcelibraryforimplementingRLalgorithms. (**Cognitive knowledge level: Apply**)

ProgramOutcomes(POs)

Outcomesaretheattributesthataretobedemonstratedbyagraduateafter completing thecourse.

PO1: Anabilitytoindependentlycarryoutresearch/investigationand development work inengineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complexengineeringactivitiesbyinteractingwiththeengineeringfraternityand with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specializationoftheprogram. Themasteryshouldbeatalevelhigherthanthe requirements in the appropriatebachelor program

PO4: Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5: Anabilitytoidentify, selectandapplyappropriatetechniques, resourcesandstate- of-the-arttool to model, analyse and solve practical engineering problems.

PO6: An ability toengagein life-longlearning forthedesign anddevelopmentrelatedto thestreamrelatedproblemstakingintoconsiderationsustainability, societal, ethical and environmentalaspects

PO7: Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoproject management andfinance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO 1			✓		✓	✓	
CO 2			✓		✓	✓	
CO 3			✓		✓	✓	
CO 4			✓	✓	✓	✓	
CO 5			✓		✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom'sCategory	**EndSemesterExamination**
Apply	60-80

.

Analyse	20-40
Evaluate	Can be evaluated using Miniprojects/assignments
Create	Can be evaluated using Miniprojects/assignments

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternal Evaluation Pattern:

Evaluationshallonlybebasedonapplication, analysisordesign-basedquestions (for both internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewed originalpublications (minimum 10 publications shall be referred): 15 marks

i. Coursebasedtask/Seminar/Datacollectionandinterpretation:15marks

i. Testpaper(1number):10marks

Testpapershallincludeminimum80%ofthe syllabus.

Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

TheendsemesterexaminationwillbeconductedbytherespectiveCollege.

Therewillbetwoparts; PartAandPartB.

.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

Forexampleiftheaverageendsemestermark%foracorecourseis40, thenthemaximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. Listtheelementsofreinforcementlearning.

2. Explainmulti-armedbanditproblems.

3. Explainaction-valuemethods.

CourseOutcome2(CO2)

1. GiveexamplesofMarkovianandnon-Markovianenvironments.

2. Whataretheadvantagesanddisadvantagesofvalue methodsvspolicymethods?

3. Imaginethattherewardsareatmost1everywhere.Whatisthemaximum value that the discounted return can attain? Why?

CourseOutcome3(CO3):

1. DescribeMonteCarlo prediction, estimationandcontrol.

2. ForQ-learningtoconvergeweneedtocorrectlymanagetheexplorationvs. exploitationtradeoff. What property needs to be hold for the exploration

.

strategy?

3. WithrespecttotheexpectedSarsaalgorithm, isexplorationrequiredasitisin the normalSarsa and Q-learning algorithms? Justify.

CourseOutcome4(CO4):

1. Describevaluefunctionapproximation.

2. Justifytheuseofdeepneuralnetworksinnonlinearfunctionapproximation.

3. Inthecontextofλ-return, differentiatebetweensimpleupdateandcompoundupdate.

CourseOutcome5(CO5):

1. Statethe policygradienttheorem.

2. Comparetwoversionsofactor-critic algorithms.

3. WhatarethebasicelementsofOpenAIGym environment.

.

ModelQuestionPaper

QPCODE:

Reg No: ___________________

Name: ____ PAGES:4

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRSTSEMESTERM.TECHDEGREEEXAMINATION, MONTH&YEAR

ModelQuestionPaper

Course Code: J1PADT144

CourseName: Reinforcement

Learning

Max.Marks:60 **Duration:2.5Hours**

PARTA

AnswerAllQuestions.EachQuestionCarries5 Marks

1. Describebrieflytheelementsofreinforcementlearning.

2. WritedowntheBellmanexpectationequationforstate-valuefunctions.

3. Drawthebackupdiagramfor2-stepQ-learning. Writethe corresponding learning rule for 2-step Q-learning.

4. Compareandcontrastanytwolinearmethodsusedforfunctionapproximation.

5. Statethepolicygradienttheoremanditsapplications. **(5x5=25)**

PartB

(Answeranyfivequestions.Eachquestioncarries7 marks)

.

(7)

6. Describetheprinciplesbehindincrementalimplementationofcomputations for estimated action values. Give a simple algorithm for the same.

7. Distinguishbetweenpolicyiterationandvalueiteration.Giverelevant algorithms. (7)

8. WhyisQ-learningconsideredanoff-policycontrolmethod? (7)

9. Demonstratehowpolynomialfunctionapproximationcanbehelpina reinforcement learning problem having 3 dimensional states. (7)

10. DerivetheREINFORCEpolicy-gradientlearning algorithm. (7)

11. Provethattheapproximationfortheλ-returnbecomesexactiftheapproximate value function does not change. (7)

12. Providepseudocodeoftheactor-criticalgorithmthatuseseligibilitytraces. (7)

Syllabus

Module1:Introductiontoreinforcementlearning(8hours)
IntroductiontoRL,Examples,ElementsofRL,Multi-armedbanditproblems,Action- value methods, The ten-armed testbed, Incremental implementation
Module2:Policiesandvaluefunctions(8hours)
MarkovDecisionProcess,Goalsandrewards,Returnsandepisodes,Policiesand value functions, Policy evaluation, Policy improvement, Policy iteration, Value iteration

Module3:Qlearning(8hours)

.

MonteCarloprediction,estimationandcontrol,TDprediction,Sarsa,Q-learning,n-step TD prediction, n-step Sarsa, n-step Off-policy learning, Dyna

Module4:Functionapproximation(8hours)
Value functionapproximation, Stochastic gradient methods, Linear methods, Non-linear functionapproximation,Episodicsemi-gradientcontrol,Semi-gradientn-stepSarsa,The λ return, TD(λ)
Module5:Policyapproximation(8hours)
Policy approximation, Policy gradient theorem, REINFORCE algorithm, Actor-Critic methods,Trust-RegionPolicyOptimization,ProximalPolicyOptimization,Introduction to OpenAI Gym

CoursePlan

No	Topic	**NumberofHours (39 Hours)**
1	**Introductiontoreinforcementlearning(7hours)**	
1.1	Introductionto RL.	1
1.2	Examples	1
1.3	ElementsofRL	1
1.4	Multi-armedbanditproblems	1
1.5	Action-valuemethods	1
1.6	Theten-armedtestbed	1
1.7	Incrementalimplementation	1
2	**Policiesandvaluefunctions(8hours)**	

.

2.1	MarkovDecisionProcess	1
2.2	Goalsand rewards	1
2.3	Returnsandepisodes	1
2.4	Policiesandvaluefunctions	1
2.5	Policyevaluation	1
2.6	Policyimprovement	1
2.7	Policyiteration	1
2.8	Valueiteration	1
3	**Qlearning(8hours)**	
3.1	MonteCarloprediction,estimationandcontrol	1
3.2	TD prediction	1
3.3	Sarsa	1
3.4	Q-learning	1
3.5	n-stepTD prediction	1
3.6	n-stepSarsa	1
3.7	n-stepOff-policylearning	1
3.8	Dyna	1
4	**Functionapproximation(8hours)**	

.

4.1	Valuefunctionapproximation	1
4.2	Stochasticgradient methods	1
4.3	Linear methods	1
4.4	Non-linearfunctionapproximation	1
4.5	Episodicsemi-gradientcontrol	1
4.6	Semi-gradientn-stepSarsa	1
4.7	Theλ-return	1
4.8	TD(λ)	1
5	**Policymethods(8hours)**	
5.1	Policyapproximation	1
5.2	Policygradienttheorem	1
5.3	REINFORCEalgorithm	1
5.4	Actor-Criticmethods	1
5.5	Trust-RegionPolicyOptimization	1
5.6	ProximalPolicyOptimization	1
5.7	IntroductiontoOpenAI Gym	1
5.8	IntroductiontoOpenAI Gym	1

.

ReferenceBooks

1. Reinforcementlearning: anintroduction, RichardS.SuttonandAndrewG. Barto, Secondedition, MIT Press, 2018.

2. AlgorithmsforReinforcementLearning.C. Szepesvari.MorganandClaypool Publishers,2010

3. ReinforcementLearning: State-of-the-Art.M. WieringandM.vanOtterlo. Springer, 2012

ReferencePapers

1. JohnSchulman, SergeyLevine, PhilippMoritz, MichaelJordan, andPieterAbbeel. 2015.Trust region policy optimization. In Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37 (ICML'15). JMLR.org, 1889– 1897.

2. Schulman, John, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg limov. "Proximalpolicyoptimizationalgorithms."*arXivprepritarXiv:1707.06347*(2017

.

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PADT154	Computational Linguistics	Program Elective1	3	0	0	3

Preamble:

This course introduces the fundamentals of Language processing froma computational viewpoint. This course covers Language models, Computational Phonology and Morphology Unification, Semantics and knowledge representation and Pragmatics. It helps the student to apply NLP tasks such as POST,WSD, and modeling of languages.

ProgramOutcomes

Graduatesofthisprogramwillbeabletodemonstratethefollowingattributes.

PO1:Anabilitytoindependentlycarryoutresearch/investigationanddevelopment work in engineering and allied streams.

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specializationof the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program.

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards.

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesand state-of-the-art tools to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects.

PO7: An ability to develop cognitive load management skills related to project managementandfinancewhichfocusonEntrepreneurshipandIndustryrelevance.

.

CourseOutcomes: TheCOsshownareonlyindicative.Foreachcourse,therecan be 4 to 6 COs.

Afterthecompletionofthecoursethestudentwillbeableto

CO 1	ApplyProbabilisticModelsofPronunciationandSpelling(Cognitive Knowledge Level: Apply)
CO 2	ApplythedifferentmethodsforParsingwithContext-FreeGrammarsfor English (Cognitive Knowledge Level: Apply)
CO 3	ApplybasicconceptsforProbabilisticContext-FreeGrammars(Cognitive Knowledge Level: Apply)
CO 4	DescribeUnificationofFeatureStructures(CognitiveKnowledgeLevel: Understand)
CO 5	ApplythekeyconceptsWordSenseDisambiguationandInformation Retrieval (Cognitive Knowledge Level: Apply)
CO 6	DevelopanapplicationthatusesNaturalLanguageGenerationconcepts (Cognitive Knowledge Level: Apply)

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO 1	✓		✓		✓	✓	
CO 2	✓		✓		✓	✓	
CO 3	✓		✓		✓	✓	
CO 4			✓		✓	✓	
CO 5	✓		✓		✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

.

Assessment Pattern

Bloom'sCategory	EndSemester Examination
Apply	50-60
Analyse	30-40
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Test1	Test2	Assignments	Total
15	15	10	40

InternalExaminationPattern:

Each of the two internal examinations has to be conducted out of 50 marks. The first series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing theremaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each

question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed

.

modulesand1questionfromthepartlycompletedmodule), eachwith7marks.Out of the 7 questions, a student should answer any 5.

End SemesterExaminationPattern:

There will be two parts; Part A and Part B. Part A contains 5 questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B contains 7 fullquestionsfromeachmoduleofwhichstudentsshouldansweranyfullquestions.Eachquestion can have a maximum 2 subdivisions and carries 7 marks.

<table>
<tr><td colspan="3">APJABDULKALAMTECHNOLOGICALUNIVERSITY</td></tr>
<tr><td>Q.P.Code:</td><td></td><td>Name:
....
Reg.No:
....</td></tr>
<tr><td colspan="3">FIRSTSEMESTERM.TECH.DEGREEEXAMINATION</td></tr>
<tr><td colspan="2">Branch:ComputerScienceandEngineering</td><td></td></tr>
<tr><td colspan="3">J1PADT154 Computational Linguistics</td></tr>
<tr><td>Time:2.5hours</td><td colspan="2">Max.Marks:60</td></tr>
<tr><td colspan="3">Answerall5questions.</td></tr>
</table>

Q.No.	PartA **Answerall5questions**.	Marks
1.	Compareinflectionalandderivationalmorphology.	**5**
2.	Whatarethedifferenttypesofsingle-errormisspellingsgiveexamples for each	**5**

.

3.	S→NPVP NP→DetN VP→VNP N→flight|meal V→ includes Det→ the | a Parsethesentence***“theflightincludesameal”***UsingCYKalgorithm	**5**
4.	Discussthedifferentfeaturestructuresassociatedwithgrammar. Explain with examples	**5**
5.	Explainpolysemywithexample.	**5**
Q.No.	**PartB** **Answerany5 questions**	**Marks**
6.	WriteChomskyandHallenotationforthefollowingrules i.Keep the first letter of the name,and drop all occurrences of a,e,i,o,u,w,yii.Replaceanysequenceofidenticalnumberswitha single number (ie., 333→ 3)	**7**
7.	Writeaprogram(pythonissufficient)tocomputeunsmoothedbigrams count	**7**
8.	Writeashort noteonHumanparsing	**7**

9.	DrawtheDAGscorrespondingtotheAVMsgivenin Examples $$\begin{bmatrix} \text{AGREEMENT} & \begin{bmatrix} \text{NUMBER} & \text{SG} \end{bmatrix} \\ \text{SUBJECT} & \begin{bmatrix} \text{AGREEMENT} & \begin{bmatrix} \text{NUMBER} & \text{SG} \end{bmatrix} \end{bmatrix} \end{bmatrix} \sqcup \begin{bmatrix} \text{SUBJECT} & \begin{bmatrix} \text{AGREEMENT} & \begin{bmatrix} \text{PERSON} & 3 \\ \text{NUMBER} & \text{SG} \end{bmatrix} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \text{AGREEMENT} & \begin{bmatrix} \text{NUMBER} & \text{SG} \end{bmatrix} \\ \text{SUBJECT} & \begin{bmatrix} \text{AGREEMENT} & \begin{bmatrix} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \end{bmatrix} \end{bmatrix} \end{bmatrix}$$	**7**
10.	JustifytheneedofWordSenseDisambiguation?Explainsupervised method of WSD in detail	**7**
11.	Explaintheselectionalrestriction-baseddisambiguationanditslimitations	**7**
12.	Describethedifferentfeaturestructuresassociatedwithgrammar. Explain with examples	**7**

SyllabusandCorsePlan(For3creditcourses,thecontentcanbefor40hrsand for 2 credit courses, the content can be for 26 hrs. The audit course in third semester can have content for 30 hours).

No	Topic	No. of Lectures
1	**Introduction(8Hours)**	
1.1	Words-RegularExpressions	1
1.2	Automata	1
1.3	Morphology	1
1.4	Finite-State Transducers	1
1.5	ComputationalPhonology	1

.

1.6	Pronunciation Modeling	1
1.7	ProbabilisticModelsofPronunciation	1
1.8	ProbabilisticModelsofSpelling	1
2	**Syntax(6Hours)**	
2.1	N-grammodels	1
2.2	N-grammodelsofSyntax-WordClasses	
2.3	Part-of-SpeechTagging	1
2.4	Context-FreeGrammarsforEnglish	1
2.5	Parsing	1
2.6	ParsingwithContext-FreeGrammars	1
3	**Probabilistic Context-Free Grammars(7Hours)**	
3.1	ProbabilisticCYK	1
3.2	ParsingofPCFGs	1
3.3	LearningPCFGProbabilities	1
3.4	ProblemswithPCFGs	1
3.5	ProbabilisticLexicalizedCFGs	1
3.6	DependencyGrammars	1
3.7	HumanParsing	1
4	**UnificationofFeatureStructures(6Hours)**	
4.1	FeatureStructuresintheGrammar	1

.

4.2	Agreement-HeadFeatures	1
4.3	Subcategorization	1
4.4	Long DistanceDependencies	1
4.5	Implementing Unification	1
4.6	UnificationData Structures-	1
5	**SemanticsandPragmatics(7Hours)**	
5.1	RepresentingMeaning	1
5.2	SemanticAnalysis	1
5.3	LexicalSemantics	1
5.4	WordSenseDisambiguation	1
5.5	InformationRetrieval	1
5.6	NaturalLanguageGeneration	1
5.7	MachineTranslation	1

ReferenceBooks

1. Jurafsky,D.andJ.H.Martin,Speechandlanguageprocessing:
2. AnIntroductiontoNaturalLanguageProcessing,ComputationalLinguistics, and Speech Recognition,Prentice-Hall, 2000.
3. Charniak,E.:StatisticalLanguageLearning.TheMITPress.
4. J.Allen:NaturalLanguageUnderstanding.Benjamin/Cummins.

.

J1PCST154	COMPUTATIONAL INTELLIGENCE	CATEGORY	L	T	P	CREDIT
		Program Elective1	**3**	**0**	**0**	**3**

Preamble: The aim of this course is to provide the students with the knowledge and skills required to design and implement effective and efficient Computational Intelligence solutions to problems for which a direct solution isimpracticalorunknown. This course covers concepts of fuzzy logic, genetic algorithms, and swarm optimization techniques. The learners will be able to provide Fuzzy and AI –based solutions to real world problems.

CourseOutcomes: Afterthecompletionofthecoursethestudentwillbeableto

CO1	Applyfuzzylogictohandleuncertaintyandsolveengineeringproblems.**(Cognitive Knowledge Level: Apply)**
CO2	ApplyFuzzyLogicInferencemethodsinbuildingintelligentmachines.**(Cognitive Knowledge Level: Apply)**
CO3	Designgeneticalgorithmsforoptimizedsolutionsinengineeringproblems. **(CognitiveKnowledgeLevel: Analyze)**
CO4	AnalyzetheproblemscenariosandapplyAntcolonysystemtosolvereal optimizationproblems.**(CognitiveKnowledgeLevel: Analyze)**
CO5	ApplyPSOalgorithmtosolverealworldproblems.**(CognitiveKnowledgeLevel: Apply)**
CO6	Design,developandimplementsolutionsbasedoncomputationalintelligence conceptsandtechniques.**(CognitiveKnowledgeLevel: Create)**

ProgramOutcomes(PO)

Outcomes arethe attributes thatareto be demonstrated bya graduateafter completing the course.

PO1: An ability to independently carry out research/investigation and developmentwork in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO 1				✓		✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	
100	40	60	2.5 hours

ContinuousInternal EvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesignbasedquestions(forboth internal and end semester examinations).

ContinuousInternalEvaluation:40marks

.

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

i. Coursebasedtask/Seminar/Datacollectionand interpretation :15 marks

i. Testpaper(1number) :10 marks

Testpapershallincludeminimum80%ofthe syllabus.
Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

Theendsemester examinationwillbeconductedbythe respectiveCollege.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marksfor eachquestion.Studentsshouldanswerall questions.Part B willcontain7 questions (such questions shall be useful in the testing of overallachievement and maturity ofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks
Totaldurationofthe examinationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. Let V = {A, B, C, D} be the set of four kinds of vitamins, F = {f_1, f_2, f_3} be three kinds of fruits containing the vitamins to various extents, and D = {d_1, d_2, d_3} be the set of three diseases that are caused by deficiency of these vitamins. Vitamin contents of the fruits are expressed with the help of the fuzzy relation R over F x V, and the extent of whichdiseases arecaused the deficiency ofthese vitamins is givenbythe fuzzy relationS over V x D. Relations R and S are given below

.

R=[0.5 0.2 0.20.7 0.4 0.4 0.1 0.1 0.4 0.3 0.8 0.1]S
= [0.3 0.5 0.1 0.8 0.7 0.4 0.9 0.1 0.5 0.5 0.2 0.3]

Find the correlation between the amount of certain fruit that should be taken while suffering from a disease.

CourseOutcome2(CO2):

1. In mechanics, the energy of a moving body is called kinetic energy. Suppose we model mass and velocity asinputs toa moving body and energy as output. Observe the system for a while and the following rule is deduced.

   IFxissmallandyishigh

   THEN z is medium

   The graphical representation of rule is given below. Let the inputs given are 0.35kg and 55m/s. What willtheoutputusingMamdaniinference?Anydefuzzificationmethodcanbe used to obtain the crisp single output.

CourseOutcome3(CO3):

1. Describe how Roulette wheel is used for selection. Draw the Roulette wheel for six chromosomes corresponding to the table given below.

Chromosome#	***Fitness***
1	10
2	5
3	25
4	15
5	30
6	20

.

CourseOutcome4(CO4):

1. Consider an Ant Colony System based on Ant Quantity model for solving the following Travelling Salesman Problem.Compute the pheromone content at each of the edges after 4 steps(1 iteration). Assume pheromone decay factor ρ=0.1, Q = 120. Assume initial pheromone of 50 units at each of the edges and that three ants k1, k2 and k3 follow the paths given below in the first iteration.

   k1=abcda;k2=acbda;k3=adcba

2. SixjobsgofirstonmachineA,thenonmachineB,andfinallyonmachineC.Theorder ofthe completion of the jobs in the three machines is given in Table

Jobs	Processingtime(hr)		
	MachineA	MachineB	MachineC
1	8	3	8
2	3	4	7
	7	5	6
4	2	2	9
5	5	1	10
6	1	6	9

Findthesequenceofjobsthatminimizesthetimerequiredtocompletethejobsusingthe ACS model.

.

CourseOutcome5(CO5):

1. Consideraparticle swarmoptimizationsystemcomposed ofthree particlesand maximum velocity 10. Assume that both the random numbers r1 and r2 used for computing the movement of the particle towards the individual best position and social best position are 0.5. Also assume that the space of solutions is the two-dimensional real valued space and the current state of swarm is as follows:

Position of particles:x1 = (4,4); x2 = (8,3); x3 = (6,7)

Individual best positions:x14,4) = *); x2*=(7,3);x3*=(5,6)

Velocities: v1 = (2,2); v2 = (3,3);v3 = (4,4)

Whatwouldbethenextpositionofeachparticle afteroneiterationofthe PSOalgorithm iftheinertiaparameterω thatisused alongwithcurrentvelocityupdate formulais0.8 ?

CourseOutcome6(CO6):

1. Implementtravellingsalesmanproblemusingappropriateoptimizationtechnique.

.

ModelQuestionPaper

QPCODE:

Reg No: ________________

Name: ____ **PAGES:5**

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRSTSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

CourseCode: J1PCST154

CourseName:ComputationalIntelligence

Max.Marks:60 **Duration:2.5Hours**

PARTA

AnswerAllQuestions.EachQuestionCarries5 Marks

1. ConsiderthesetofColoursA={Blue,Red,Orange,Yellow,Green},AttributesB = {Bright, Warmth, Dullness}, Feelings C= {Unpleasant, happiness, Angry}. Given R and S where R is the relationship between colours and their attributes andSistherelationshipbetweencolourattributesandfeelingscreated.Findthe relationship Q between colours and feelings created **(5)**

2. Developamembershipfunctionfor—Tall‖.Basedonthatdevisemembership functionfor—VeryTall‖.Explainhowitisdone **(5)**

3. Mentionthe importanceofobjective(fitness)functioningeneticalgorithm **(5)**

4. Describehowpheromoneisupdated.Whatiselitist/elasticants?Arethey useful in this scenario? **(5)**

5. Whatisthesignificanceofpbestandgbestparticlesinsolvingproblemswith particle swarm optimization? **(5)**

PartB

(Answeranyfivequestions.Eachquestioncarries7 marks)

6. (a)ConsiderthesetoffruitsF={Apple,Orange,Lemon,Strawberry,Pineapple}. **(3)**

LetsweetfruitsB={ + }and

SourFruitsF={ + }

FindFruitsthatareSweetorSour,Sweet butnotSour,Sweet andSour

(b)ConsidertwofuzzySetsgivenby **(4)**

FindthefuzzyrelationfortheCartesianproductofPandQi.e,R=PxQ.

IntroduceafuzzysetTgivenby

$$T = \left\{\frac{0.9}{short} + \frac{0.3}{medium} + \frac{0.6}{tall}\right\}$$

andFindToRusing max-mincomposition

.

7. Consider a Fuzzy Inference System for checking climate comfortability of humanbeingsforlongtimeliving.Thesystemacceptstwoinputs– temperature and humidity. The rules and membership functions of FIS is given below. Using Mamdani inference and center of sum, calculate output when the temperature is 50Fahrenheit and humidity is 50%. **(7)**

Rule 1: IF temperature is cool and humidity is low, THEN climate is comfortable.

Rule2:IFtemperatureiscoolandhumidityishigh,THENclimateis livable

Thefuzzysets—EasyQuestionPaper‖andtheircorresponding—StudentPerformance‖ are given below

$$\text{Easy\_QP}=\left\{\frac{0.8}{1}+\frac{0.2}{}+\frac{0.6}{2}+\frac{0.7}{34}\right\}$$

$$\text{Stud\_Perf}=\left\{\frac{0.3}{a}+\frac{0.4}{b}+\frac{0.8}{c}+\frac{0.9}{d}\frac{0.8}{1}+\frac{0.2}{2}+\frac{0.6}{3}+\frac{0.8}{4}+\frac{0.7}{5}\right\}$$

Findtheperformanceofstudentscanddforthequestionpaper—Somewhat Easy‖

$$\text{Somewhat\_Easy}=\{\qquad\qquad\qquad\qquad\}$$

8. Explainanyproceduretomapasolutiontothecorrespondingchromosome and vice versa in genetic algorithms. Also illustrate it with an example **(7)**

9. Describetwomethods usedtoselectindividualsfroma populationfor the mating pool in Genetic Algorithms **(7)**

10 (a)ConsidertheTSP withthefollowingedgecosts.Giventheevaporationfactor ρ=0.02andinitialpheromoneatalledgesT_{ij}=100 **(1)**

Whatisthecostofbest tour?

(b) Usingtheequation$T_{ij}(t+1)=(1-\rho)T_{ij}(t)+\Delta T_{ij}(t,t+1)$,computethe$T_{ij}$ofthe edge<a,c> when 10 ants uses the edges <a, c>, using the following models: **(6)**

i AntDensityModel(ConstantQ=10)

i Ant Quantity Model(Constant Q=100)

whereQistheconstantrelatedtothepheromoneupdation.

11 DescribeAntColonySystem.WhatarethedifferenttypesofAntsystems? **(7)**

.

12 Consider a particle swarm optimization system composed of three particles andmaximumvelocity 10.Assume thatboth the random numbersr1andr2 used for computing the movement of the particle towards the individual best position and social best position are 0.5. Also assume that the space of solutions is the two-dimensional real valued space and the current state of swarm is as follows: **(7)**

Position of particles:x1 = (4,4); x2 = (8,3); x3 = (6,7)

Individual best positions:x1* = (4,4); x2*=(7,3);x3*=(5,6)

Velocities: v1 = (2,2);v2 = (3,3);v3 = (4,4)

What would be the next position of each particle after one iteration of thePSO algorithm if the inertia parameter ω that is used along with current velocity update formula is 0.8?

Syllabus

Module1:Fuzzy Logic

Crisp sets vs fuzzy sets- Operations and properties of Fuzzy sets. Membership functions - Linguistic variables. Operations on fuzzy sets- Fuzzy laws- Operations on fuzzy relations, Fuzzy composition- Max- min**,** Max – product. Alpha-cut representation.

Module2:FuzzySystems

Fuzzy Reasoning – GMP and GMT. Fuzzy Inference System: Defuzzification methods - Fuzzy Controllers -Mamdani FIS, Larsen Model

Module3:GeneticAlgorithms

Introduction to Genetic Algorithms – Theoretical foundation - GA encoding, decoding - GA operations – Elitism – GA parameters – Convergence. Multi-objective Genetic Algorithm – Pareto Ranking.

Module4:AntColonySystems

.

Swarm intelligent systems - Background Ant colony systems – Biological systems- Development of the ant colony system- - Working - Pheromone updating- Types of ant systems- ACO algorithms for TSP

Module5:ParticleSwarmOptimization

Basic Model - Global Best PSO- Local Best PSO- Comparison of _gbest‘ to _lbest‘- PSO Algorithm Parameters- Problem Formulation of PSO algorithm- Working. Rate of convergence improvements -Velocity clamping- Inertia weight- Constriction Coefficient- Boundary Conditions- Guaranteed Convergence PSO- Initialization, Stopping Criteria, Iteration Terms and Function Evaluation.

CoursePlan

No	Topic	No.of Lectures(40)
1	**Module1:FuzzyLogic**	**9**
1.1	Crispsetsvsfuzzysets,OperationsandpropertiesofFuzzysets	1
1.2	Membershipfunctions	1
1.3	LinguisticVariables	1
1.4	Operationsonfuzzysets	1
1.5	Fuzzylaws	1
1.6	Operationsonfuzzyrelations	1
1.7	FuzzyComposition-Max-min	1
1.8	FuzzyComposition–Max-Product	1
1.9	Alpha-cutrepresentation	1
2	**Module2:FuzzySystems**	**7**
2.1	FuzzyReasoning –GMP	1
2.2	FuzzyReasoning–GMT	1
2.3	FuzzyInferenceSystem	1
2.4	Defuzzificationmethods	1
2.5	FuzzyControllers	1
2.6	MamdaniModel	1
2.7	LarsenModel	1

.

3	**Module3:GeneticAlgorithms**	**7**
3.1	IntroductiontoGeneticalgorithm	1
3.2	Theoreticalfoundation	1
3.3	GAencoding-decoding	1
3.4	GAoperations	1
3.5	Elitism,GAparameters,ConvergenceofGA	1
3.6	Multi–objectiveGeneticAlgorithm	1
3.7	ParetoRanking	1
4	**Module4:AntColonySystems**	**8**
4.1	Swarmintelligentsystems	1
4.2	Background	1
4.3	Antcolonysystems–biologicalsystems	1
4.4	Developmentoftheantcolonysystem	1
4.5	Working	1
4.6	Pheromoneupdating	1
4.7	Typesofantsystems	1
4.8	ACOalgorithmsforTSP	1
5	**Module5:ParticleSwarmOptimization**	**9**
5.1	BasicModel	1
5.2	GlobalBestPSO	1
5.3	LocalBestPSO,Comparisonof_gbest‘to _lbest‘	1
5.4	PSOAlgorithmParameters	1
5.5	ProblemFormulation	1
5.6	Working	1
5.7	Rateofconvergenceimprovements–velocityclamping	1
5.8	Inertia-weight-ConstrictionCoefficient-BoundaryConditions	1
5.9	Initialization,StoppingCriteria,IterationTermsandFunction Evaluation	1

References

1. SamirRoy,UditChakraborty,IntroductiontoSoftComputingNeuro-FuzzyGenetic Algorithms, Pearson, 2013

.

2. N.P.Padhy,ArtificialIntelligenceandIntelligentsystems,OxfordPress,NewDelhi, 2005.
3. Xin-SheYangSchoolofScienceandTechnology,MiddlesexUniversityLondon,Nature-Inspired Optimization Algorithms, Elsevier, First edition, 2014
4. SatyobrotoTalukder,BlekingeInstituteofTechnology,MathematicalModellingand Applications of Particle Swarm Optimization, February 2011
5. MitchellMelanie,AnIntroductiontoGeneticAlgorithm,PrenticeHall,1998
6. AndriesEngelbrecht,ComputationalIntelligence:AnIntroduction,Wiley,2007
7. MarcoDorigoand Thomas Stutzle, ―AntColonyoptimization‖,Prentice HallofIndia, New Delhi 2005

.

CODE	ADVANCED DATABASE	CATEGORY	I			CREDIT
J1PADT115		PROGRAM ELECTIVE2	3	0	0	3

Preamble:

his course provides an exposure to the concepts and techniques in advanced database management. The topics covered in this course includes Relational Model –Conceptual Model and Schema Design, Strategies regarding query processing and optimization, Distributedsystemarchitecture,Semi-structureddatahandlingandmodern datamanagement techniques.Thiscoursehelpsthelearnerstodevelopapplicationsthatmanagedataefficiently with the help of suitable data models and techniques.

CourseOutcomes: Afterthecompletionofthecoursethestudentwillbeableto

CO1	Makeuseoftheconceptsinrelationaldatabasesystemsincluding:datamodelsrelational algebra,ER features, andthe different normalization techniques to relational models. **(Cognitive Knowledge Level: Apply)**
CO2	Illustratethebasicdatabasestorage,fileorganization,databaseaccessingandindexing **(CognitiveKnowledgeLevel: Apply)**
CO3	Identifyvariousmeasuresofqueryprocessingandoptimization.**(Cognitive Knowledge Level: Apply)**
CO4	Analyzeimplementationaspectsofdistributedsystemondatabase architecture. **(CognitiveKnowledgeLevel: Analyze)**
CO5	Makeuseofsemistructured data,XMLandXMLqueriesfordatamanagement. **(CognitiveKnowledgeLevel: Apply)**
CO6	Design,Develop,andImplementinnovativeideasonadvanceddatabaseconcepts and techniques. **(Cognitive Knowledge Level: Create)**

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstrated by a graduate after completing the course.

PO1: Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentwork in engineering and allied streams

.

PO2: An ability to communicate effectively, write and present technical reports on complexengineeringactivitiesbyinteractingwiththeengineeringfraternityand with society at large.

PO3: Anabilitytodemonstrateadegreeofmasteryovertheareaasperthespecializationof the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworldproblems by following the standards

PO5: Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate- of-the-art tool to model, analyse and solve practical engineering problems.

PO6: Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedtothe stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project managementandfinancewhichfocusonEntrepreneurshipandIndustryrelevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓		✓	✓	✓	✓	
CO2	✓		✓	✓		✓	
CO3	✓		✓	✓		✓	
CO4	✓		✓	✓		✓	
CO5	✓		✓	✓	✓	✓	
CO6	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom'sCategory	EndSemesterExamination

.

Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

ContinuousInternalEvaluationPattern:

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15marks

iii. Testpaper(1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.

Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

Theendsemester examinationwillbeconductedbytherespectiveCollege.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the studentsinacourse,throughlonganswerquestionsrelatingtotheoretical/practical

.

knowledge, derivations, problem solving and quantitative evaluation), with minimum one question fromeach module of which student should answer any five. Each question can carry 7 marks

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

i. How does a query tree represent a relational algebra expression? Draw the initialquery tree, apply heuristic rules and obtain an optimized query expression for the following SQL query.

SELECTS.SName,C.CName, G.grade
FROMStudentS,CoursesC,FacultyF,GradesG
WHERES.Sid=G.SidANDC.Cid= G.CIdANDF.Fid =G.Fid ANDG.grade>=7.0
ANDF.DName= 'CSE'

i. Consider the ER diagram shown below. Identifythe minimum set ofrelations requiredtomaptoarelationalmodel.Identifyforeignkeysandprimarykeys.Draw a schema diagram showing all relations

i.

.

CourseOutcome2(CO2):

1. Differentiatebetweenfixedlengthrecordsandvariable lengthrecords.
2. A file has r=36000 STUDENT records of fixed length. Each record has the following fields: Name (25 bytes), SSN (8 bytes), Address(35 bytes), Phone(10 bytes), Date of Birth (8 bytes), Sex(1 bytes), Class code (3 bytes).
   - i. Calculatetherecordsize ofR
   - ii. Calculate the blocking factor bfr and number of file blocks b, assuming an unspanned organization
   - iii. Calculate the average time it takes to find a record doing a linear search on the file (assume blocks are stored continuously)
   - iv. Assume the file ordered by SSN, bydoing binary search, calculate time it takes.
   - v. Assume a primary index is created with keyas SSN and the data pointer needs 8 bytes, find the number blocks required to keep theprimary index. Also find the average time required to find a record using the index.

CourseOutcome3(CO3):

- **i.** Let *s* be selection cardinality and *bfr* be the blocking factor. Compare the cost function for SELECT operation in the following cases i) when a clustering index is available ii)When a secondary index is available.
- ii. Consider a STUDENT file with 20,000 records stored in a disk with fixed length blocks of size 1024 bytes. Each record is of40 bytes.Assume that in theSTUDENTfile,there existsasecondaryindexonkeyfiled,Sid,withX_{SID}=3. There is another file, COURSE_REG, with attributes StudID, CourseID,CourseName and Date of Registration. There are 40,000 records in COURSE_REG file, stored as 4000 blocks. A secondary index on non key key field StudID with X_{STUDID}=4 is available. Let the join selectivity be 1/8 and6 output records be stored in a block.Find the number of block accesses required for nested join and single loop join for the following query:

STUDENT $_{SID=STUDID}$ **COURSE_REG**

CourseOutcome4(CO4):

1. There are four sites S1, S2, S3 and S4 in a distributed database systemwith weights2, 3, 4 and 3 respectively. Assume read quorum value is 6. If a data item x is replicated across these sites and quorum consensus protocol is followed:
   - i. findtheminimumpossiblevalueofwritequorum.
   - ii. Minimumnumberofsiteslocked toperformaread operation

.

- i. Minimumnumber ofsiteslocked toperformawriteoperation

2. Explain2phasecommitprotocolinadistributedenvironment.Whatactionswouldbetaken when a site recovers from failure?

CourseOutcome5(CO5)

1.It is required to represent a University database in XML form. A University has one ormore departments. Each department hasa name, a specialization, Head of theDepartment. Several faculty membersare working on each department and severalcourses are run by a department. Each department is uniquely identified by a number(attribute).Name, Area of Specialization each faculty needs to be stored.

Information such as CourseId, CourseName, Duration and Credits are to be kept about each course.Design a DTD for this University structure

ModelQuestionPaper

QPCODE:

Reg No: __________________

Name: ____ PAGES:4

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRSTSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

Course Code: J1PADT115
CourseName:ADVANCEDDATABASE

Max.Marks: 60 **Duration:2.5Hours**

PARTA

AnswerAllQuestions.EachQuestionCarries5 Marks

1. GivenarelationschemaR(ABCD)andsetofdependenciesG={$A \rightarrow B$,$BC \rightarrow D$, $A \rightarrow C$} **(5)**

1. Identifythekey.
2. Identifythenormalform.
3. DecomposeintoBCNF.

.

2. A file has r=36000 STUDENT records of fixed length. Each record has thefollowing fields: Name (25 bytes), SSN (8 bytes), Address (35 bytes), Phone (10 bytes), Date of Birth (8 bytes), Sex(1 bytes), Class code (3 bytes). **(5)**

i. Calculatetherecordsize ofR

ii. Calculatetheblockingfactor bfrandnumberoffileblocksb,assumingan unspanned organization

iii. Calculate the average time it takes to find a record doing a linear searchon the file (assume blocks are stored continuously)

iv. Assume a primary index is created with key as SSN and the data pointer needs 8 bytes, find the number blocks required to keep theprimaryindex. Also find the average time required to find a record using the index.

3. Discusstherulesfortransformationofquerytreesandidentifywheneachrule should be applied during optimization **(5)**

4. Explain2phasecommitprotocolinadistributed environment. **(5)**

5. DesignanXMLdocumentforstoringhostelmessfood details(mealstakensuchas breakfast, lunch, dinner) with their charges for the month of June 2022. Charges may vary depending on the food taken. Students can opt not to take any meals on certain days. **(5)**

i WriteasampleXMLfor2studentsfor2days.

ii WriteaXQuerytoreturnthelunchdetailsofall.

iii CreateanXSDforthesame.

PartB

(Answeranyfivequestions.Eachquestioncarries7 marks)

6. (a) SupposeyouaregivenwitharelationschemaR(ABCD).Eachofthefollowing FDs, assuming they are the dependencies hold over R, state whether or not proposeddecompositionofRintosmallerrelationisagooddecomposition. Explain Why? **(5)**

1. AB→BC→A,C→D,decomposeintoACDandBC
2. A→BC,C→AD,A→C,decomposeintoBCDandAC

(b) WhatMinimalCover. Illustratewithanexample **(2)**

.

7. Notown Recordshas decided tostoreinformationabout musicians whoperformon its albums(aswellasothercompanydata)inadatabase.Thecompanyhaswiselychosen tohireyouasadatabasedesigner(atyourusualconsultingfeeof$2500/day).Designa conceptual schema for Notown and draw an ER diagram for your schema. **(7)**

a) EachmusicianthatrecordsatNotownhasanSSN,aname,anaddress,anda phone number.

b) EachinstrumentusedinsongsrecordedatNotownhasauniqueidentification number, a nameand a musical key.

c) Eachalbumrecorded ontheNotownlabelhasauniqueidentificationnumber, atitle,acopyright date, aformat,andanalbumidentifier.Eachsongrecorded at Notown has a title and an author.

d) Eachmusicianmayplayseveralinstruments,andagiveninstrumentmaybe played by several musicians.

e) Eachalbumhasanumberofsongsonit,butnosongmayappearonmorethan one album.

f) Eachsongisperformedbyoneormoremusicians,andamusicianmay perform a number of songs.

g) Eachalbumhasexactlyonemusicianwhoactsasitsproducer.Amusician may produce several albums, of course.

8. Consider the following statistics about a relational table, STUDENT(Sid, SName, Branch, CNo).There are 16000 records in 4000 blockswith ablocking factor of4.There is a secondary index on non keyattribute CNo with X_{CNO} = 3. Assume, there are only 100 different courses. We have another relation, COURSES (CId, CName, Credit, Type).There are 100 rows in this table,storedin20diskblocks.There existsaprimaryindexonCIdwithX_{CID}= 1. Assumetheselectioncardinalityforthejoinattributeis160. **(7)**

Estimate the cost of join operation ($STUDENT_{CNo=CId}COURSES$) by the following type of join operation (avoid the cost incurred for the storage of resultant records) .

i) nestedloopjoin

ii) singleloopjoin.

iii) nestedloopjoinwithabufferspaceavailabilityof12blocks

.

(7)

9. Consider the bitmap representation ofthe free-space map, where for each block in the file, two bits are maintained in the bitmap. If the block is between 0 and 30 percent full the bits are 00, between 30 and 60 percent the bits are 01, between 60 and 90 percent the bits are 10, and above 90 percent the bits are 11. Such bitmaps can be kept in memory even for quite large files.

i Outlinetwobenefitsandonedrawbacktousingtwobitsforablock,instead of one byte as described earlier in this chapter.

ii Describehowtokeepthebitmapuptodateonrecordinsertionsand deletions.

iii Outline thebenefitofthebitmaptechniqueoverfreelistsinsearchingfor free space and in updating free space information.

10(a)Letr(A,B,C)withtuples{(1,2,3),(4,5,6),(1,2,4),(5,3,2),(8,9,7)}and s(C,D,E)withtuples{(3,4,5),(3,6,8),(2,3,2),(1,4,1),(1,2,3)}aretwo relationinstances.Compute*rsemijoins* **(4)**

(b)Whatactionswouldbetakenwhenasiterecoversfrom failure? **(3)**

11 (a)Assume that a Movie database in XML form is available and title, director, year ofrelease,costofproductionastheinformationstoredinit.LetMOVIE, TITLE, DIRECTOR, YEAR, COST are the XML elements, and the element MOVIE has an attribute CATEGORY which indicates the type of movies(*Horror, Comedy, Thriller)*. Similarly, the TITLE has an attribute LAN which indicates the language (*Malayalam, English, Hindi*). A movie can have more than one director. **(4)**

WriteXPATHqueriesforthefollowing

I.ListallEnglishMovies

i. Listallmovieswherelanguageisnot specified

i. Listallmovieshavingtwo directors

iv. Listall*Comedy*type moviesinthedatabase

v. Listallmovieswhosecostproductionisbelow10million.

(b)Explaintheterms: i)*WellFormedXMLii)ValidXML* **(3)**

.

12 ConsiderthefollowingXMLTree **(7)**

WriteanXMLschemafortheabove,andalsoprovideanXQueryexpressionto get the books published in the year 1992.

SyllabusandCoursePlan
J1PADT115-ADVANCEDDATABASE

Module1(RelationalDatabases–RelationalModel,Normalization)11Hours

Relational Model Introduction - Structure of Relational Database, database Schema, Keys

The Relational Algebra: Fundamental Operations, The Entity-Relationship model: Entity Set, Relationship Set, Attributes, Constraints: : Mapping cardinalities-E-R Diagrams, Real world Scenarios – ER diagrams. Normalization - The Need for Normalization, Process, Rules for Functional Dependencies, First Normal Form, Second Normal Form, Third Normal Forms, Boyce/Codd Normal Form, Functional Dependencies- Minimal cover, Equivalence, Properties of Relational Decomposition Relational Databases

Module2:QueryProcessingandOptimization(8Hours)

Placing file records on disk- Record types, Record blocking and spanned versus Unspanned records, Hashing techniques –Internal, External hashing for disk files, IndexingandHashing:BasicconceptOrderedIndices,B+treeIndexFiles:Structureofa B+-Tree (structure only, algorithms not needed), Indexes on Multiple keys,Hash Indexes, Bitmap indices, Indexing spatial data

Module3:IntroductiontoQueryProcessingandOptimization(6hours)

.

Measuresofquerycost,AlgorithmsforSelectionwithcostanalysis,AlgorithmsforJoin with cost analysis, Evaluation of expressions, Heuristics in Query Optimization, Optimization of Relational Algebra expressions.

Module4:DistributedSystemArchitecture(6Hours)

IntroductiontoDistributedSystemarchitecture,Distributedstorage&Distributedfile systems

DistributedRDBdesign&itsTransparency,DistributedTransactions,CommitProtocols & Concurrency Control, Distributed Query Processing,

Module5:XML,XPath,Non-relationalDatabases---9Hours

IntroductiontoSemi-structuredDataandXMLDatabases,XMLDataModel–XSD, XML: DTD and XML Schema, XML presentation,XPath Queries ,XQuery, Next Generation Databases: Distributed Relational Databases - MongoDB Shading and Replication,Object Relational Systems

CoursePlan

No	Topic	No.of Lectures(40 hours)
1	**Module1(RelationalDatabases–RelationalModel,Normalization):11hours**	
1.1	RelationalModelIntroduction-StructureofRelationalDatabase, database Schema, Keys	1
1.2	TheRelationalAlgebra:FundamentalOperations	1
1.3	The Entity-Relationship model: EntitySet, Relationship Set, Attributes,Constraints::Mappingcardinalities-E-RDiagrams	1
1.4	Realworld Scenarios –ERdiagrams	1
1.5	Normalization-TheNeedforNormalization,Process	1
1.6	RulesforFunctionalDependencies	1
1.7	FirstNormalForm,SecondNormalForm,ThirdNormalForms	1
1.8	Boyce/CoddNormalForm	1
1.9	FunctionalDependencies-Minimalcover, Equivalence	1
1.10	PropertiesofRelationalDecomposition	1
1.11	AlgorithmsforRelationalDatabaseDesign	1

.

2	**Module2:QueryProcessingandOptimization(8Hours)**	
2.1	Placingfilerecordsondisk-Recordtypes,Recordblockingand spanned versus Unspanned records	1
2.2	Hashingtechniques–Internal,Externalhashingfordisk files	1
2.3	IndexingandHashing:BasicconceptOrderedIndices	1
2.4	B+treeIndexFiles:	1
2.5	StructureofaB+- Tree(structureonly,algorithmsnotneeded)	1
2.6	IndexesonMultiple keys	1
2.7	HashIndexes,Bitmap indices	
2.8	Indexingspatialdata	
3	**Module3:IntroductiontoQueryProcessingandOptimization(6hours)**	
3.1	Measuresofquerycost	1
3.2	AlgorithmsforSelectionwithcostanalysis	1
3.3	AlgorithmsforJoinwithcost analysis	1
3.4	Evaluationofexpressions	1
3.5	HeuristicsinQueryOptimization	1
3.6	OptimizationofRelationalAlgebraexpressions	1
4	**Module4:DistributedSystemArchitecture(6Hours)**	
4.1	IntroductiontoDistributedSystemarchitecture	1
4.2	Distributedstorage&Distributed filesystems	1
4.3	DistributedRDBdesign& itsTransparency	1
4.4	DistributedTransactions	1
4.5	CommitProtocols&ConcurrencyControl	1

.

4.6	DistributedQueryProcessing	1
5	**Module5:XML,XPath,Non-relationalDatabases---9Hours**	
5.1	IntroductiontoSemi-structuredDataandXMLDatabases	1
5.2	XMLDataModel–XSD	1
5.3	XML:DTDandXMLSchema,XML presentation	1
5.4	XPathQueries	1
5.5	XQuery	
5.6	NextGenerationDatabases:DistributedRelationalDatabases-	1
5.7	NonrelationalDistributedDatabases	1
5.8	MongoDBShadingandReplication	1
5.9	ObjectRelationalSystems	1

ReferenceBooks

1. RamezElmasri,ShamkantB.Navathe,―FundamentalsofDatabaseSystems―, PearsonEducation,6thEdition,2007.(Module1-Chapter7.1to7.7,14.1to14.5,15.1 to15.3,Module2:16.4,16.8,17.1to17.3,Module:18.1to18.3,18.7to18.9)
2. AbrahamSilberschatz,HenryF.Korth,S.Sudarshan,‖DatabaseSystemConcepts‖, McGrawHillEducation,6thEdition,2011.(Module4)
3. GuyHarrison,―NextGenerationDatabases:NoSQL,NewSQL,andBigData‖, Apress,1stEdition,14December2015.
4. Rob,PeterandCarlosCoronel,―DatabasePrinciples:FundamentalsofDesign, ImplementationandManagement‖,9thEdition,2011.
5. ThomasMConnollyandCarolynEBegg,―Databasesystems-APracticalApproach toDesign,ImplementationandManagement‖,PearsonEducation,4thEdition(2014).
6. AshutoshKumarDubay,―DatabaseManagementConcepts‖,S.K.Kataria&Sons,1st Edition (2012).
7. RaghuRamakrishnanandJohannesGehrke,―DatabaseManagementSystems‖,McGraw Hill, 3rd Edition (2014).

.

J1PADT125	CONCEPTS IN CLOUD COMPUTING	CATEGORY	L	T	P	CREDIT
		Program Elective2	3	0	0	3

Preamble: Cloud computing is the delivery of computing services over theInternet. The syllabus is prepared with a view to equip the students to learn basic concepts in cloud computing - compute, storage, networking. They should gain basic understanding of orchestration, HA and failover. After learning this course computation services can offer faster innovation, flexible resources, and economies of scale.

CourseOutcomes: After thecompletionofthecoursethestudentwillbeableto

CO 1	MakeuseoftheconceptsincloudcomputingandOpenStacklogicalarchitectureto develop applications (Apply)
CO 2	ExploreOpenStack*cloudcontrollerandcommonservices(Apply)*
CO 3	ComparedifferentOpenStackcomputeservicecomponentsandstoragetypes(Apply)
CO 4	Analyse*theOpenStackNetworking-Connectiontypesandnetworkingservices(Analyse)*
CO 5	Analysethe*orchestration,HAandfailoverinOpenStack(Analyse)*
CO 6	Design,develop,implementandpresentinnovativeideasoncloudcomputing(Analyse)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateafter completing the course.

PO1: Anabilitytoindependentlycarryoutresearch/investigationanddevelopment work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

.

PO5: Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate- of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoproject management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7
CO 1	√		√			√	
CO 2	√		√	√	√	√	
CO 3	√		√	√		√	
CO 4	√		√		√	√	
CO 5	√		√			√	
CO 6	√	√	√	√	√	√	√

AssessmentPattern

Bloom'sCategory	**EndSemester Examination**
Apply	50-80
Analyse	20-40
Evaluate	Canbedonethrough Assignments/projects
Create	Canbedonethrough Assignments/projects

Markdistribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

.

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesign-basedquestions(forboth internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) :15marks

i. Coursebasedtask/Seminar/Datacollectionandinterpretation :15marks

i. Testpaper(1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.

Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern

Theendsemester examinationwillbeconductedbythe respectiveCollege.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question fromeach module of which student should answer any five. Each question can carry 7 marks

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

.

ModelQuestionPaper

QPCODE:

Reg No: ________________

Name: ____ **PAGES:4**

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRSTSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

CourseCode: J1PADT125

CourseName:CONCEPTSINCLOUDCOMPUTING

Max.Marks: 60 **Duration:2.5Hours**

PARTA

AnswerAllQuestions.EachQuestionCarries5 Marks

1.	Differentiatebetweenprivatecloudandpubliccloud?Illustratethedesignof OpenStack logical architecture	**(5)**
2.	*Illustrate* asymmetric clustering and symmetric clustering. *Categorize* the functionalities handled by the cloud controller	**(5)**
3.	Designthesystems ofdockercontainers.	**(5)**
4.	Derivetheprocedureforconnectingtwonetworksusingavirtualrouter.	**(5)**
5.	Howhigh-performance,open-sourceloadbalancerandreverseproxyforTCP andHTTPapplicationsareobtained.ListtheHAlevelsinOpenStack.	**(5)**

PartB

(Answeranyfivequestions.Eachquestioncarries7 marks)

6.	(a) HowtheprovisioningofVMinOpenStackisorganised.Designthesystemanddraw a diagrammatic representation	**(4)**
	(b) IdentifythebestpracticesusedinPhysicalmodedesign	**(3)**
7.	DesignamethodwithstepsforrunningOpenStackplaybooks	**(7)**
8.	Explainthedeployingswiftservices.	**(7)**
9.	Designthearchitectureofneutronandexplain.	**(7)**
10	Identifythestepsinvolvedinsettinga databasewithhighavailability	**(4)**

.

11	CompareobjectstoragewithNAS/SANbasedstorage	**(7)**
12	Identifythecategorizationofneutronvirtualnetworksandjustifythecategories.	**(7)**

Syllabus-CONCEPTSINCLOUDCOMPUTING

Module1:OverviewofOpenStack 6 Hours)

Introduction to cloud computing, private cloud, public cloud, hybrid cloud architecture. Cloud Services - Infrastructure as a Service, Platform as a Service, Storage as a Service. Designing OpenStack Cloud Architectural Consideration - OpenStack - The new data center paradigm - OpenStack logical architecture. Nova - Compute service . Neutron - Networking services . Gathering the pieces and building a picture. A sample architecture setup.

Module 2: OpenStack cluster - Controller and common services (5 Hours) OpenStack Cluster – The Cloud Controller and Common Services, Asymmetric clustering, Symmetric clustering. The cloud controller - The keystone service. The nova-conductor service, Thenova- scheduler service, TheAPIservices, Imagemanagement. The network service, The horizon dashboard, The telemetry services.

Module 3: OpenStack compute and Storage (11 Hours)Thecomputeservicecomponents-DecidingonthehypervisorOpenStackMagnum project.SegregatingthecomputecloudOvercommitmentconsiderations.Storinginstances' alternatives.Understandinginstancebooting.Planningforservicerecovery.OpenStackStorage - Block, Object, and File Share-Understanding the storage types. A spotlight on swift. Deploying swift service. Using Block Storage Service Cinder.

Module4:OpenStackNetworking (9 Hours)

The architecture of Neutron. Implementing virtual networks - VLAN, Tunnel based. Virtual Switches, The ML2 Plugin. Neutron Subnets Connecting virtual networks with routers - Configuringtheroutingservice Connectingnetworksusingavirtualrouter,Connectingtothe external world Connectivity fromthe external world,Associatinga floatingIP Implementing network security in OpenStack.

.

Module 5: OpenStack Orchestration, HA and Failover (9 Hours) Orchestration in OpenStack, Heat and its Components. Stacking in OpenStack. OpenStackOrchestrationwithTerraform.ScopeofHAinOpenStack.HAinthedatabase.HAin the Queue, Implementing HA on RabbitMQ

CoursePlan

No	Topic	No.of Lectures(40 hrs)
1	**OverviewofOpenStack**	**(6Hours)**
1.1	Introductiontocloudcomputing,privatecloud,publiccloud,hybrid cloudarchitecture.	1
1.2	CloudServices-InfrastructureasaService,PlatformasaService, Storageasa Service	1
1.3	DesigningOpenStackCloudArchitecturalConsideration-OpenStack- Thenewdatacenterparadigm-OpenStacklogicalarchitecture	1
1.4	Nova-Computeservice.Neutron-Networkingservices.	1
1.5	Gatheringthepiecesandbuildingapicture	1
1.6	Asamplearchitecturesetup	1
2	**OpenStackcluster-Controllerandcommonservices**	**(5Hours)**
2.1	OpenStackCluster–TheCloudControllerandCommonServices Asymmetricclustering,Symmetricclustering	1
2.2	Thecloudcontroller-Thekeystoneservice.Thenova-conductor service	1
2.3	Thenova-schedulerservice,TheAPIservices,Imagemanagement.	1
2.4	Thenetworkservice	1
2.5	Thehorizondashboard,Thetelemetryservices	1
3	**OpenStackcomputeandStorage**	**(11 Hours)**

.

3.1	Thecomputeservicecomponents-Decidingonthe hypervisorOpenStackMagnumproject.	1
3.2	Segregatingthecomputecloud	1
3.3	Overcommitmentconsiderations.	1
3.4	Storinginstances'alternatives.Understandinginstancebooting.	1
3.5	Planningforservicerecovery.	1
3.6	OpenStackStorage–	1
3.7	Block,Object,andFileShare	1
3.8	Understandingthestoragetypes.	1
3.9	Aspotlightonswift.	1
3.10	Deployingswiftservice.	1
3.11	UsingBlockStorageServiceCinder	1
4	**OpenStackNetworking**	**(9Hours)**
4.1	ThearchitectureofNeutron.	1
4.2	Implementingvirtualnetworks-VLAN,Tunnelbased.	1
4.3	VirtualSwitches,TheML2Plugin	1
4.4	NeutronSubnetsConnectingvirtualnetworkswithrouters	1
4.5	ConfiguringtheroutingserviceConnectingnetworksusingavirtual router	1
4.6	Connectingtotheexternalworld	1
4.7	Connectivityfromtheexternalworld,	1
4.8	AssociatingafloatingIPImplementingnetworksecurityinOpenStack-Part1	1
4.9	AssociatingafloatingIPImplementingnetworksecurityinOpenStack-Part2	1
5	**OpenStackOrchestration,HAandFailover**	**(9Hours)**
5.1	OrchestrationinOpenStack,	1
5.2	HeatanditsComponents.	1

.

5.3	StackinginOpenStack	1
5.4	OpenStackOrchestrationwithTerraform.	1
5.5	ScopeofHA	1
5.6	ScopeofHAinOpenStack	1
5.7	HAinthedatabase.	1
5.8	HAintheQueue	1
5.9	ImplementingHAonRabbitMQ	1

TextBook

1. OmarKhedher,ChandanDattaChowdhury,MasteringOpenStack,2ndEdition,Packt Publishing, 2017

ReferenceBooks

1. TomFifield,DianeFleming,AnneGentle,LorinHochstein,JonathanProulx,Everett Toews, and Joe Topjian, OpenStack Operations Guide, O'REILY,1/e, 2014.

2. UchitVyas,AppliedOpenStackDesignPatterns,Apress,1/e,2016.

3. V.K.CodyBumgardner,OpenStackinaction,Manning,2016.

4. AmarKapadia,SreedharVarma,KrisRajana,ImplementingCloudStoragewith OpenStack Swift, Packt Publishing, 2014.

5. https://docs.openstack.org/wallaby/?_ga=2.231002015.1428061357.1620834394-

1139122985.1620834394

.

CODE J1PADT135	STATISTICS FOR DATA SCIENTISTS	CATEGORY	L	T	P	CREDIT
		PEC-2	3	0	0	3

Preamble: This course is intended to systematically master the core concepts in statistics & probability, descriptive statistics, hypothesis testing, regression analysis, analysis of variance, and some advanced regression/machine learning methods such as logistics regressions, polynomial regressions and decision trees. This course helps the students to work with different types of data and implement the techniques and make data-driven decisions

CourseOutcomes:

Afterthecompletionofthecourse,thestudentwillbeableto

CO 1	Apply the fundamentals of statistics, from bar plots to ANOVAs, regression to k-means,andt-testtonon-parametricpermutationtestingformachinelearning,AI,and data science. (**Cognitive knowledge level: Apply)**
CO 2	Visualizethedataindifferentdescriptive,inferential,andpredictiveconceptsfor relevant stages of data analytics. (**Cognitive knowledge level: Apply)**
CO 3	Analysethedatabymakinguseofconceptssuchasmean,median,andmode,plus range and IQR and box-and-whisker plots (**Cognitive knowledge level: Apply)**
CO 4	Applytherightstatisticaltechniqueatappropriate stageofadataanalyticsproject
CO 5	ImplementstatisticalconceptsinPython/MATLAB(**Cognitiveknowledgelevel: Apply)**
CO 6	Drawinferencesfromthedatafordifferentmachinelearningmodelsthrough hypothesis testing. (**Cognitive knowledge level: Apply)**

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1: Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineeringactivitiesbyinteractingwiththeengineeringfraternityandwithsocietyat large.

PO3: Anabilitytodemonstrateadegreeofmasteryover theareaasperthespecializationofthe program. Themastery should be at a level higher than the requirements inthe appropriate bachelor program

PO4: Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforreal-worldproblems by following the standards

PO5: Anabilitytoidentify,selectandapplyappropriatetechniques,resources,andstate-of-the-art tools to model, analyse and solve practical engineering problems.

PO6: Anabilityto engageinlifelonglearningforthedesignanddevelopmentrelatedtothestream- related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO 1	✓		✓	✓	✓	✓	
CO 2	✓		✓	✓	✓	✓	
CO 3	✓		✓	✓	✓	✓	
CO 4	✓		✓	✓	✓	✓	
CO 5	✓		✓	✓	✓	✓	
CO 6	✓	✓	✓	✓	✓	✓	

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Apply	60-80%
Analyse	20-40%
Evaluate	Attain through Project/Assignments
Create	Attain through Project/Assignments

.

Markdistribution

Total Marks	CIE	ESE	ESEDuration
100	40	60	2.5 hours

ContinuousInternalEvaluationPattern:

Theevaluationshallonlybebasedonapplication,analysis,ordesign-basedquestions(forboth internal end-semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeer-reviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

i. Coursebasedtask/Seminar/Datacollectionand interpretation :15 marks

i. Testpaper(1number) :10 marks

Testpapershallincludeaminimumof80%ofthesyllabus.

Course-basedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation, and understanding of the students.

EndSemesterExaminationPattern:

The endsemester examinationwillbeconductedbythe respectiveCollege.

There will be two parts: Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem-solving and quantitative evaluation), with a minimum one question from each module of which student should answer any five. Each question can carry7 marks

Totaldurationofthe examinationwillbe150minutes.

.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. Identifythevariablesinthefollowingdatadescriptionandclassifythevariablesas categorical or quantitative. If the variable is quantitative, list the units.

   —The Indianapolis 500 is a car race that‗s been taking place since 1911 and is often scheduled to take place over Memorial Day weekend. The race takes place at the Indianapolis Motor Speedway and a driver needs to complete 200 laps that cover a distance of 500 miles. Race results are reported by driver number, the driver‗s name, the type of car the driver uses, and the time to the nearest ten-thousandth of a second. If a driver doesn‗t finish the race, their number of laps completed is recorded instead of the time to complete the race.‖

2. Compare the scatterplots. The second graph includes extra data starting in 1880. How doesthiscomparetotheplotthatonlyshows1910to2010? Explaintrendsinthedata, and how the regression line changes by adding in these extra points. Which trend line would be best for predicting the temperature in 2050?

.

3. Calculateandinterpret thecorrelationcoefficientforthedataset

CourseOutcome2(CO2)

1. ExplainDescriptiveanalytics,Diagnosticanalytics,PredictiveAnalytics,andPrescriptive analytics
2. Sarah's monthly sales to date are shown in the ogive. Signify the meaning of circled point

.

3. Bethanystartedasit-upprogramsothatshecando200sit-upsinaday. Attheend ofweek 6she‘ll have completed 1,685sit-ups. Createan ogiveof the data.

Week	Numberofsit-ups
Week1	350
Week2	455
Week3	600
Week4	540
Week5	1,275
Week6	1,685

CourseOutcome3(CO3):

1. a.ElaborateNormaldistributionsandz-scores
   b. Let‘s say the mean finishing time for male speed skaters in the winter Olympics on the 500-meter track is 70.42 seconds, with a standard deviation of 0.34 seconds(thedata isnormallydistributed).What isthemaximumtimeaskatercanpost ifhe wants to skate faster than 95% of his competitors?
2. a.Illustratetheadvantagesofthebarchartoverthe piechart
   b.ListthechartsforcategoricalvariablesandquantitativevariablesandIllustrate

3. Create the total-relative frequency table for the data, and then answer the question: Carl is in charge of creating an activity for the students in his college dorm. If Carl wants the highest possible turnout, which activity should he choose? Why?

	Movie	**Bowling**	**PizzaParty**
Male	20	40	55
Female	35	50	62

CourseOutcome4(CO4):

1. Explainthetechniquesyouwillusefor yourdataanalysisbeforeyoucollect anydata

2. The first stage in any analysis should be to describe your data and the population from whichitisdrawn.Thestatisticsappropriateforthisactivityfallintothreebroadgroupsand depend on the type of data you have.

.

Answerthefollowing

a. Whatdoyouwanttohavetolookatthedistribution,todescribethecentral tendency, and describe the spread

b. Withwhattypeofdata

c. Appropriatetechniques

3. ExplaintheappropriateStatisticaltechniquestofindtheDifferencesbetweengroupsand variables and relationship between the variables

CourseOutcome5(CO5):

1. ExplaintheCentralLimitTheoremandImplementitinPython/ MATLAB

2. ExplainKNNandPCAandImplementtheminPython/ MATLAB

CourseOutcome6(CO5):

1. Explainthesignificanceofhypothesistestingformachinelearningandtheparametersof hypothesis testing

2. Afast-foodrestaurantisimplementingnewworkplacepolicieswiththegoalofincreasing employee satisfaction by 2 points on a scale of 1 to 10. The restaurant surveys 10 employees, asking them both before and after the policies are enacted to rate their workplace satisfaction on the 1 – 10 scale and records the results in the table below.

Cantherestaurantsayat5%significancethatthepoliciesincreasedemployeesatisfaction by2 points?

.

ModelQuestionPaper

QPCODE:

Reg No: ________________

Name: ____ PAGES:4

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRSTSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

CourseCode: J1PADT135

CourseName:STATISTICSFORDATASCIENTISTS

Max.Marks: 60 **Duration:2.5Hours**

PARTA

AnswerAllQuestions.EachQuestionCarries5 Marks

1. Acompanyisanalyzingtheresultsfromarecentsurveyaboutwhypeopleleft theiremployment. Theresultsareshowninthedatatablebelow. Ingeneral, isa bar graph or a pie chart a better choice to display the data? Why? **5**

Reasonsforleavinga job	
Reducedjobduties	30%
Companyrestructuring	15%
Toomuchtraveltime	12%
Lookingformoreopportunity	11%
Needmorepersonaltime	9%
Poorexpected companygrowth	8%
Thejobwasacontractorshortterm	8%
Needmoreofachallenge	5%
Other	2%

2. Johnstopsatthelocalgasstationanddecidestobuylotterytickets.Eachticket hasa20%chanceofbeingawinner.Hewillbuyalotteryticketandcheckto seeifit‘sawinner.He‘llcollecthismoneyandbedoneifit'sawinner.Ifit‘s notawinner,he‘llbuyanother.He‘llrepeatthisuntilhegetsawinningticket. Butifhehasn‘twonbyhisfifthticket,hewon‘tbuyanymoretickets.LetLbe thenumberoflotteryticketsJohnwillbuy, thenfind E(L). **5**

3. Estimatetheeffectofbias onsampling. **5**

4. Comparecorrelationand regression **5**

5. ExplainSignaldetectiontheory **(5x5=25)**

PartB

(Answeranyfivequestions.Eachquestioncarries7 marks)

.

6. (a)Appraisethehappeningstoourmeasuresofcentraltendencyandspread whenwe make changes toour data set(mean, median, mode, range, and IQR) **(3)**

i. Changingtheentiredataset
i. Addingorremoving adatapointfromthe set

(b) Illustrate the appropriate visualization technique to represent a data set whenwewanttoshowthemedianandspreadofthedataatthesametime **(2)**

(c) DescribetheimportanceofBox-and-whisker plots **(2)**

7. (a) Illustratemeasuresofcentraltendency **(3.5)**

(b) Illustratethemeasuresofspread(Range,interquartilerange(IQR), variance, standard deviation) **(3.5)**

8. (a)Two factories A and B produce heaters for car seats. A customer received a defective car seat heater and the manager at factory B would like to knowif it came from her factory. Use the table below to determine the probability that the heater came from factory B. **(2.5)**

(b) ExplainPoissonprocess **(2)**

(c) There are 30 students in a Kindergarten class and each one of them has a4% chance of forgetting their lunch on any given day. What is the probability that exactly 5 of them will forget their lunch today? **(2.5)**

9. (a)You are planning a day at the beach, but the morning is cloudy. 64% of all rainy days start off cloudy, but cloudy mornings are common (55% of days start cloudy). This month is usually a dry month and only 18% of the days tend to be rainy. What‘s the chance that it will rain during your day at the beach? **(3)**

(b) The time it takes students to complete multiple choice questions on an AP Statistics Exam has a mean of 55 seconds with a standard deviation of 12 seconds. Ifthe examconsistsof40 multiple choicequestions, find themean total time to finish the exam. Then find the standard deviation in the total time.

W
h
at
a

. ssumption must be made? **(2)**

(c) We toss a fair coin 15 times and record the number of tails. Is this experi mentmodeledbyabinomialrandomvariable?Ifitisn‘t,explain why.Ifitis,determineitsparametersnandpandexpressthebinomial random variable as $X \sim B(n, p)$. **(2)**

.

10 (a) Suppose we want to determine whether the mean height of men issignificantly higher than the mean height of women in a certain city, so we randomly sample 100 men and 100 women. Given the mean and standard deviation of both samples below, use the critical value approach to say whether men are significantly taller than women at a 1% level of significance. **(7)**

11 (a)Three types of batteries were tested for battery life. See the battery lives in thetablebelow.InconstructingtheANOVAtable,whatwillbethevalues of factor and error degrees of freedom? **(3)** **(3.5)**

C1	C2	C3
Battery1	Battery 2	Battery 3
90	102	102
85	98	82
92	97	85
95	103	75
88	107	92

(b) Explainconfidenceintervalanditsimporance **(5)**

12 (a) Explainthestatisticalpoweranditsimportance **(2)**

(b) Use the Average Global Sea Surface Temperatures data shown in the tabletocreatealineofbestfitforthedata.Consider1910asyear10.Use the equation to predict the average global sea surface temperature in the year 2050. **(5)**

Year	**Temperature,F**
1910	-1.11277
1920	-0.71965
1930	-0.58358
1940	-0.17977
1950	-0.55318
1960	-0.30358
1970	-0.30863
1980	0.077197
1990	0.274842
2000	0.232502
2010	0.612718

.

Syllabus

Mod	Contents	hrs
I	DESCRIPTIVE STATISTICS: DATA,Types of data,data,Samplevs.populationdata, Samples (N=1 & N>1 studies), Visualizing Data: Bar plots. Box-and-whisker plots, Boxplots of normal and uniform noise, Histograms, Histogram proportion, Pie charts, Implementation, descriptive vs. inferential statistics, Accuracy, precision, and resolution. Data distributions, histograms of distributions, Measures of central tendency, central tendencies with outliers, Measures of dispersion Interquartile range, QQ plots, Statistical "moments", Histograms: Violin plots, Shannon entropy, entropy, and number of bins, Implementation in Python	7
II	**DATA NORMALISATION AND PROBABILITY:** Z-score standardization, Min-max scaling, Removing outliers: z-score method, modified z-score method, z vs. modified-z, Multivariate outlier detection, Euclidean distance for outlier removal, Removing outliers by data trimming, Non-parametric solutions to outliers, Implementation Probability: Computing probabilities, Probability mass vs. density, PDF, CDF, creating sample estimate distributions, Monte Carlo sampling, Sampling variability, noise, and other annoyances, Expected value, Conditional probability, and Tree diagrams, The Law of Large Numbers, The Central Limit Theorem,, Implementation Random Variables: Discrete RV, Binomial Poisson, Bernoulli, and Geometric random variables	8
III	**SAMPLING & HYPOTHESIS TESTING: Sampling:** Types of studies, Sampling, and bias, Sampling distribution of the sample mean, Conditions for inference with the SDSM, Sampling distribution of the sample proportion, Conditions for inference with the SDSP, **Hypothesis Testing:** Independent and Dependent Variables, models, residuals, Sample distributions under null and alternative hypotheses, P-values: definition, tails, and misinterpretations, P-z combinations that you should memorize, Degrees of freedom, Type 1 and Type 2 errors, Parametric vs. non-parametric tests, Multiple comparisons and Bonferroni correction, Statistical vs. theoretical vs. clinical significance, Cross-validation, Statistical significance vs. classification accuracy, **The T-Test Family:** Purpose and interpretation, One-sample t- test, The role of variance, Two-samples t-test, Importance of N for t-test, Wilcoxon signed-rank (nonparametric t-test), Mann-Whitney U test (nonparametric t-test), U test, Permutation testing for t-test, significance, Python Implementation	8
IV	**CONFIDENCE INTERVALS & ANOVA: Confidence Intervals on Parameters:** Computing confidence intervals via formula, Confidence intervals via bootstrapping (resampling), Misconceptions about confidence intervals. CORRELATION: Motivation and description of correlation, Covariance, and correlation: formulas, Correlation matrix, correlation to the covariance matrix, Partial correlation, The problem with Pearson, Nonparametric correlation: Spearman rank, Fisher-Z transformation for correlations, Spearman correlation, the confidence interval on the correlation, Kendall's correlation for ordinal data, The subgroups correlation paradox, Cosine similarity, **Analysis Of Variance:** Sum of squares, The F-test and the ANOVA table, The omnibus F-test and posthoccomparisons,Thetwo-wayANOVA,One-wayANOVAexample,Two-way ANOVAexample,**Regression:**IntroductiontoGLM/regression,Least-squaressolution	8

.

	totheGLM,Evaluatingregressionmodels:R2andF,Simpleregression,implementation	
V	**REGRESSION, CLUSTERING, AND PCA: Regression:** Multiple regression, Standardizing regression coefficients, Polynomial regression models, Logistic regression, Under andover-fitting,comparing"nested"models,missingdata, **Statistical Power, And Sample Tests:** Importance of statistical power, Estimating statistical power and sample size, Compute power and sample size using G*Power, **Clustering And Dimension-Reduction-**K-means clustering, K-means, and normalization, K-means on a Gauss blur, Clustering via dbscan, dbscan vs. k-means, K-nearest neighbor classification, Principal components analysis, K-means on PC data, independent components analysis, **Signal Detection Theory:** The two perspectives of the world, d-prime, Response bias, F-score, Receiver operating characteristics (ROC), Python Implementation	9

S.NO	TOPIC	NO. OF LECTURES (40hrs)
	Module1-DescriptiveStatistics-7Hours	
1.1	Typesofdata,Samplevs.populationdata,Samples,	1
1.2	**VisualizingData**:Barplots.Box-and-whiskerplots,Boxplotsofnormaland uniform noise, Histograms,	1
1.3	Histogramproportion,Piecharts,implementation	1
1.4	descriptivevs.inferentialstatistics,Accuracy,precision,andresolution.Data distributions, histograms of distributions,	1
1.5	Measuresofcentraltendency,centraltendencieswithoutliers	1
1.6	MeasuresofdispersionInterquartilerange,QQplots,Statistical"moments	1
1.7	Histograms:Violinplots,Shannonentropy,entropy,andnumberofbins,code	1
	Module2-DataNormalisationandProbability-8Hrs	
2.1	**DATA NORMALISATION AND PROBABILITY:** Z-score standardization, min-max scaling, removing outliers: z-score method, The modified z-score method, z vs. modified-z,	1
2.2	Multivariateoutlierdetection, Euclideandistancefor outlierremoval, Removing outliers by data trimming, non-parametric solutions to outliers,	1
2.3	**Probability:**Computingprobabilities,	1
2.4	Probabilitymassvs.density,PDF,CDF.CDFsforvariousdistributions,	1
2.5	Creating sample estimate distributions, Monte Carlo sampling, Sampling variability, noise, and other annoyances, Expected value, Conditional probability, and Tree diagrams,	1
2.6	TheLawofLargeNumbers,TheCentralLimitTheorem,Implementation	1
2.7	**RandomVariables:**BinomialandPoissonRV,	1

.

2.8	Bernoulli,andGeometricrandomvariables	1
	Module3-SamplingandHypothesisTesting-7Hrs	
3.1	**Sampling:**Typesofstudies,Sampling,andbias,Samplingdistributionofthe sample mean,	1
3.2	ConditionsforinferencewiththeSDSMSamplingdistributionofthesample proportion, Conditions for inference with the SDSP,	1
3.3	HypothesisTesting:IndependentandDependentVariables,models,residuals, Sample distributions under null and alternative hypotheses,	1
3.4	P-values:definition,tails,andmisinterpretations,P-zcombinationsthatyou should memorize, Degrees of freedom, Type 1 and Type 2 errors,	1
3.5	Parametric vs. non-parametric tests, Multiple comparisons, and Bonferroni correction, Statistical vs. theoretical vs. clinical significance, Cross-validation, Statistical significance vs. classification accuracy,	1
3.6	**TheT-TestFamily:**Purposeandinterpretation,One-samplet-test,Theroleof variance, Two-samples t-test,	1
3.7	ImportanceofNfort-test,Wilcoxonsigned-rank(nonparametrict-test),	1
3.8	Mann-WhitneyUtest(nonparametrict-test),Utest,Permutationtestingfort-test, significance, Python Implementation	1
Module4-ConfidenceIntervals,CorrelationandAnova-8Hrs		
4.1	Confidence Intervals on Parameters: Computing confidence intervals via formula, Confidence intervals via bootstrapping (resampling), Misconceptions about confidence intervals	1
4.2	**Correlation:**Motivationanddescriptionofcorrelation,Covariance,and correlation: formulas, Correlation matrix,	1
4.3	correlation to the covariance matrix, Partial correlation, The problem withPearson, Nonparametric correlation: Spearman rank, Fisher-Z transformation for correlations,	1
4.4	Spearmancorrelation,theconfidenceintervalonthecorrelation,Kendall'scorrelationfor ordinaldata,Thesubgroupscorrelationparadox,Cosinesimilarity,	1
4.5	**Analysisof Variance:**Sum ofsquares,TheF-testandtheANOVAtable,The omnibus F-test and posthoc comparisons,	1
4.6	Thetwo-wayANOVA,One-wayANOVAexample,	1
4.7	Two-wayANOVAexample,	1
4.8	**Regression:**IntroductiontoGLM/regression,Least-squaressolutiontotheGLM, Evaluating regression models: R2 and F, Simple regression, code	1
	Module5-Regression,ClusteringandPCA-9Hrs	
5.1	**Regression:**Multipleregression,Standardizingregressioncoefficients	1
5.2	Polynomialregressionmodels,Logisticregression,	1
5.3	Underandover-fitting,comparing"nested"models,missingdata,	1
5.4	**StatisticalPowerandSampleTests**Estimatingstatisticalpowerandsamplesize, Compute power and sample size using G*Power,	1

.

5.5	**ClusteringAndDimension-Reduction-**K-meansclustering,K-means,and normalization, K-means on a Gauss blur, Clustering via dbscan	1
5.6	dbscanvs.k-means,K-nearestneighborclassification,Principalcomponents analysis,	1
5.7	K-meansonPCdata,independentcomponentsanalysis,	1
5.8	**SignalDetectionTheory:**d-prime,Responsebias	1
5.9	F-score,Receiveroperatingcharacteristics(ROC),PythonImplementation	1

TEXTBOOKS:

1. A.Abebe,J.Daniels,J.W.Makean,―StatisticsandDataAnalytics‖,Statistical Computation Lab, Western Michigan University, Kalamazoo.2001
2. PeterGoosandDavidMeintrup,‗StatisticswithJMP:Graphs,Descriptive Statistics, and Probability, WILEY 2015
3. PeterGoosandDavidMeintrup,‗StatisticswithJMP:HypothesisTests,Anova, andRegression‗WILEY2016
4. Bruce Ratner, ‗Statistical and Machine-Learning Data Mining-Techniques for BetterPredictiveModelingandAnalysisofBigData,ThirdEdition,CRCPress, Tailor and Francis group, 2017
5. CharlesWheelan,‗NakedStatistics_StrippingtheDreadfromtheData‗,W.W. Norton Company, New York, 2014

REFERENCES:

1. JimAlbert,―BayesianComputationwithR‖,2nd Edition,Springer2009
2. TrevorHasti,RobertTibshirani,JeromeFriedman,―DataMining,Inferenceand Statistics‖,2nd Edition,SpringerSeriesinStatistics2008

.

CODE J1PADT145	ETHICS FOR DATA SCIENTISTS	CATEGORY	L	T	P	CREDIT
		PROGRAM ELECTIVE2	3	0	0	3

Preamble:

This course is intended to provide an introduction to critical and ethical issues using data and its implications in the society. This course helps the learners to understandthe benefits and drawbacks of using data while using them for making predictions by understanding the structureofethics, law,and societalvalues.Also,thiscourseblends social and historical perspectives on data with ethics, policy, and case examples to help students develop a workable understanding of current ethical issues in data science.

CourseOutcomes:

Afterthecompletionofthecourse,thestudentwill beableto

CO 1	Applyingtheconceptofethics,andthenecessityandthebenefitofadoptingshared ethical principles in Data Science (**Cognitive knowledge level: Apply)**
CO 2	AnalyzetheroleofanIRBinahumansubject‘sstudyandthedifferencebetween datacollected forbusinesspurposesversusresearchpurposesandrelatehowthis changesthe requirement for IRB approval. (**Cognitiveknowledgelevel: Apply)**
CO 3	Distinguishbetweenthethreemaincategoriesofintellectualpropertyandidentify the data owner (**Cognitive knowledge level: Apply)**
CO 4	Describe the reasonable expectation of privacy relates to data collection and recognizethevoluntarylimitsontheuseofdatathatariseoutofsocialconsensus. (**Cognitive knowledge level: Apply)**
CO 5	ApplydataethicswhilepracticingDataScience.(**Cognitiveknowledgelevel: Apply)**
CO 6	Identifyinvaliddata,incorrectmodels,biasinalgorithms,andbadanalysis conducted on good data (**Cognitive knowledge level: Apply)**

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompleting the course.

PO1: Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complexengineeringactivitiesbyinteractingwiththeengineeringfraternityand with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5: Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1			✓			✓	
CO2			✓			✓	
CO3			✓			✓	
CO4			✓			✓	

CO5			✓			✓	✓
CO6			✓			✓	

AssessmentPattern

Bloom'sCategory	**EndSemesterExamination**
Apply	50-80%
Analyse	20-40%
Evaluate	Assignments/Projects
Create	Assignments/Projects

Markdistribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5hours

Continuous InternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesign-basedquestions(for both internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeer-reviewedoriginalpublications(minimum 10publications shall be referred) : 15 marks

i. Coursebasedtask/Seminar/Datacollectionandinterpretation:15marks

i. Testpaper(1 number):10marks

.

Testpapershallincludea minimum80%ofthesyllabus.

Course-basedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation, and understanding of the students.

EndSemester ExaminationPattern:

Theendsemester examinationwillbeconductedbythe respectiveCollege.

There will be two parts: Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. PartB will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem-solving and quantitative evaluation), with a minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20%over the average ESE mark % for the core courses. ESE marks awarded to a studentfor each elective course shall be normalized accordingly.

Forexample,iftheaverageendsemestermark%foracorecourseis40,thenthe maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions
Course Outcome 1 (CO1):

1. Illustrateanexampleofasituation,inwhichweall,asasociety,arebetteroff, because we agree to behave ethically.

2. Company X has learned about Facebook‘s mood manipulation experiment and believes that a happy person is much more likely to buy than a grumpy one. Therefore, it has designed its website to tell heart-warming stories in callout boxes on every page. These stories, at best, are tangentially related to the products being sold on the page. They A/B test this website before launch to see if the story boxes do have the intended effect. They find that the boxes do have the desired effect of increasing sales. They then adopt the new website design with the story boxes, and they write an article describing their findings in a Marketing Journal.

.

Discuss:

Does Company X need to inform its customers about this effort? To what extent? Does it need to obtain consent? If so, for what? If you answered YES to theconsent question above, what is the smallest change to the scenario describedabove that would make you change your answer to NO?

3. You go to the bus stop, and everyone is patiently in line waiting for the bus. Rather thanwait in line, you just jump onto the bus when it arrives. Discuss the ethics about yourbehavior- whether it is legal/ethical/ unethical/ unethical and legal/ ethical and illegal etc

4. You conduct research on user interface design. You wish to evaluate a new layout you have developed for presenting the results of a web search. For this purpose,you need to get the opinions of several users. Even though the users of your new interface have no possibility of suffering any harm, and furthermore your test is no more intrusive than the A/B testing performed by so many web companies, Discuss- is it true that you are nevertheless required to obtain IRB clearance?

CourseOutcome2(CO2)

1. Creative Commons has a set of standard copyright licenses that are used widely. Thiscourseasawholeisreleased CC-BY-NC,whichmeansitcanbereproduced with attribution (BY) for non-commercial use (NC). Individual components are released CC BY-NC-ND, which means they can be reproduced with attribution (BY) for non-commercial use (NC) without making any changes (ND = no derivatives). Discuss whether it is OK to reuse, with attribution, a single video from this course in your own (non-commercial) presentation.

2. You conduct research on user interface design. You wish to evaluate a new layout youhave developed for presenting the results of a web search. For this purpose, you need to get the opinions of several users. Even though the users of your new interface have nopossibilityof suffering any harm, and furthermore your testisno more intrusive than theA/B testing performed by so many web companies, is ittrue that you are neverthelessrequired to obtain IRB clearance?

3. Analyse- You have designed a human subjects experiment with appropriate provisionforinformedconsent,andyousubmitthistotheappropriateIRB.The IRB will automatically

approveyourexperimentsinceyouhavedemonstratedyouwillproperlyobtain informed consent.

.

CourseOutcome3(CO3):

1. If teenagers knew that their parents could have access to all their social media posts,thenthe teenagers wouldlikelybeverycarefulaboutwhat theypost-Asses

2. Iagreetoposeforsomephotographsyoutakewiththepromisethatyouwillkeep these photos private. Some years later, you change your mind and publish these photos. Since you own these photos, are you within your rights?, Evaluate the action.

CourseOutcome4(CO4):

Intermsofundesireduseofdata,therearethreedistinctsteps.Justifyyouranswer foreach of the following three questions about whether it violates privacy.

1. Undesiredcollectionofpersonaldataviolatesprivacy.

2. Undesiredanalysisofpreviouslycollectedpersonaldataviolates privacy.

3. Undesireddisseminationofpreviouslycollectedpersonaldataviolates privacy.

4. A major shortcoming of all-or-nothing data access policies (e.g., when an app wants access to your location data) is that it defeats privacy because you have no control over any data you choose to share with the app- Discuss

CourseOutcome5(CO5):

1. It turns out that the government funding for public health departments is computed on a formula that is heavily dependent on the number of cases of flu. Forefficiency, the government decides to adopt Google flu numbers forthis parameter. If you run a public health department, you seek to maximize your funding byasking the public in your county to perform searches for flu. Will this work? Evaluate.

2. The university in the preceding question conducts some additional investigation, and determines that both the mean and the median scores obtained by minority applicants on the standardized test are substantially lower than the corresponding mean and median forother students. Based on this fact, in conjunction with the facts from the preceding question, can we conclude that the test is unfair to a minority applicant? Justfy

.

3. A university uses performance on a standardized test as the only scoring mechanism usedto admit applicants. The university observes that it is admittingfar fewer minority studentsthan their proportion in the population at large. Based on only these facts, can we concludethat the test is unfair? Justify your answer

CourseOutcome6(CO6):

1. Yourcityhasdecidedtomakepropertytaxpaymentdatasemi-public:youjust have toenter your property identifier to get that information.

Yourneighborhas asmall businessthat you have invested in, and arefeeling nervous about. You enter your neighbor‘s property ID into the city system to check on your neighbor‘s tax payments. You find that he has missed paying the last two quarters, after many years of paying ontime. You suspect acash flow crunch in his business and ask for your loan back.

Your neighbor is forced to sell some business assets to pay you back. Another investor sees this sale of business assets, and also decided to liquidate her investment.Inthismanner,problemssnowball,untilyourneighborisdrivenoutof business.

Whosefaultisit,ifany?Identifyspecificstepswheresomeethicalrulewasviolated.

2. Youwork for a major cellphone service provider and have access to large volumes of detailed location data for your customers. One day, you are able to correlate location with building footprints and hence determine whether the cell phone user (your customer) isindoors or outdoors. On this basis, you obtain analytical results that lead to a new signal amplification algorithm that is amazingly effective in improving call quality. You surprise yourself, your boss, and your company, with these results. Does this analysis violate the—Do Not Surprise‖ rule? Evaluate

3. Seeking to expand their business and improve their product, suppose that Amazon sends a survey to all Kindle owners asking them what they like and dislike about their Kindle. What validity concerns would you have about the survey results obtained? If the primarygoal is to grow Kindle sales, what could Amazon do toget more valid data?

.

ModelQuestionPaper		
QPCODE:		
Reg No: ________________		
Name: ____ **PAGES:4**		
APJABDULKALAMTECHNOLOGICALUNIVERSITY		
FIRSTSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR		
CourseCode: J1PADT145		
CourseName:ETHICSFORDATASCIENTISTS		
Max.Marks:60Duration:2.5Hours		
PARTA		
AnswerAllQuestions. EachQuestionCarries5 Marks		
1.	BrieflyexplaintheimportanceofInstitutionalReviewBoardsandIndependent Ethics Committees	**5**
2.	DiscussthevarioustypesofIntellectualPropertyRights	**5**

.

3.	a. I am a fan of art called mauve pottery. I painstakingly create a directory of mauve potters around the world. I publish this directory, copyrighted, withallrightsreserved. Is it OKforyouto makeacopyofthisdirectoryforuse in your (non-commercial) class? b. In the example of the preceding question, is it OK for you to publish a new version of this directory showing the locations of all mauve potters on amap? c. Many psychology experiments are conducted on the university campus by academic researchers. The human subjects recruited tend to be college students, who are generally younger and smarter than the population as a whole. This is clearly not a representative sample of the generalpopulation. To show the universal validity of an important effect, researchers went off-campus to a nearby city and recruited volunteersubjects by offering a small cash incentive for their time. i. Thesecond experimentisagoodrandomsampleofthe population **i.** Thesecondexperimentisnotagoodrandomsampleeither,andsois pointless. **i.** Thesecondexperimentisnotagoodrandomsamplebutisstillvaluable. Explain and justify your answer.	**5**
4.	a. In machine learning, it is common practice to use k-fold cross-validation, where the data set is divided into k parts. k-1 of these are used for training and the remaining part is used for testing. And this can be repeated k times,leaving out a different part each time. Appraise is this a good way to measurehow well the learned model predicts the test data (labels, values, or whatever is being predicted)? b. Based on face recognition technology applied to records from in-store video cameras, Fancy Store is immediately able to identify you when you enter their store if you have shopped there before. If you are identified as a high-value shopper, from your previous purchasing history in the store, a personal shop assistant is immediately assigned to stay with you during your visit and helpyou choose and locate the items you want. Ordinary shoppers, not identified as high-value shoppers, do not get the same service. Is this unfair? Why orwhy not?	**5**

5.		a. A travel website has empirically determined that Mac users are more willing to pay for higher-priced hotel rooms. Therefore, the website modifies the default order in which hotels are shown, with higher-priced hotels ranking slightly higher for Mac users than for other PC users. Analyze: Is this reordering—discrimination‖ ethical? b. b.A leading algorithm-based employment agency determines, based on data analysis,thatcandidateswithstraighthairmakemorereliableemployeesthan candidates with curly hair. They use this as a criterion (one among many, but with significant weight to this one) in choosing which candidates tointerview, using submitted photographs with the application as their basis to determine whether hair is straight or curly. They do not tell prospective candidates whatcriteria they are using. This is unethical. c.	**5** **(5x5=25)**
		PartB **(Answeranyfivequestions.Eachquestioncarries7 marks)**	
6.	(a)	Analyzetheimpactofnotfollowingthedataethicsinbusiness	**(7)**
7.	(a)	Discusstheimportanceofethics-centricityindata projects	**(7)**
8.	(a)	AppraiseModernPrivacyRisksandProtectionStrategiesinData Analytics	**(7)**
9.	(a)	Discussthecasestudyoftargetedadvertisement,whetheritishelpfulor annoying. If annoying, is there any method digitally to avoid that?	**(3)**
	(b)	DiscusstheCaseStudyofdataprivacybreach:SneakyMobileApps	**(4)**
10.	(a)	a.Explainchoosingattributesthatwewillusetoachieveourobjectivesand methods of measuring them.	**(4)**
	(b)	b.Explaintheerrorsinthedataprocessing	**(3)**
11.	(a)	aExplaintheneed foralgorithmfairness	**(2)**
	(b)	Discussremovingbiasfromhiringalgorithms	**(5)**

.

12.	(a)	Discuss the societal consequences of Data Science that we should be concerned about even if there are no issues with fairness, validity, anonymity,privacy, ownership or human subjects research	**(3.5)**
	(b)	DiscussSocialcreditScoresanditssocietalconsequences	**(3.5)**

Syllabus

MODULE NO	**CONTENT**	**HOURS**
I	**Introduction:**Ethics,Definition,whatisDataEthics?Needof Data Science Ethics, Examples, Case Study and Discussion. HumanSubject Research and Informed Consent, Cases, US InstitutionalReviewBoard(IRB),IRBinIndia,Limitationsof Informed Consent,Case Study and Discussion.	**8**
II	**Data Ownership and Privacy:** Data ownership, Limits of RecordingData and Using Data, Intellectual Property rights, Privacy: Introduction, History, Degrees, Privacy Risks, Case Studies- TargetedAds, The Naked Mile. Sneaky Mobile Apps, andDiscussion,Anonymity:De-Identification,CaseStudies,and Discussion	**8**
III	**Data Validity:** Validity: Introduction, Choices of Attribute and Pressures, Errors in Data Processing, Errors in Model Design, Managing Change, Case Study- Three Blind Mice, AlgorithmsandRace, Algorithms intheOffice, GermanWings Crash,Google Flu,and Discussion	**8**
IV	**AlgorithmicFairness:**AlgorithmFairness:Introduction,Correct andMisleading results, Hiring algorithms, P Hacking, Case Study-HighThroughputBiology,Geopricing,YourSafetyIsMy LostIncomeandDiscussion-fairnessofapplyingfacerecognition to givepreference to the often visiting customers in business	**8**
V	**Societal Consequences:** Introduction, Societal Impact, Ossification, Surveillance, Code of ethics, Wrap Up, CaseStudies-SocialCreditScores,PredictivePolicing,and Discussion	**8**

.

CoursePlan

No	Topic	No. of Lectures()
	Module1(IntroductiontoEthics)-8hours	
1.1	Ethics,Definition,whatisDataEthics?	1
1.2	NeedofDataScienceEthics.	1
1.3	Examples,CaseStudyand Discussion	1
1.4	HumanSubjectResearchandInformedConsent,Cases,US Institutional Review Board (IRB), IRB in India,	1
1.5	LimitationsofInformedConsent,	1
1.6	CaseStudyand Discussion	1
1.7	CaseStudyand Discussion	1
1.8	CaseStudyand Discussion	1
	Module2(DataOwnershipandPrivacy)-8hours	
2.1	DataOwnershipandPrivacy	1
2.2	DataOwnershipandPrivacy	1
2.3	IntellectualPropertyrights,	1
2.4	Privacy:Introduction,History,Degrees.	1
2.5	PrivacyRisks,CaseStudies,andDiscussion,	1
2.6	TargetedAds,TheNakedMile.SneakyMobileApps	1
2.7	Anonymity:Introduction,De-Identification,	1

.

2.8	CaseStudiesandDiscussion	1
Module3(DataValidity)-8hours		
3.1	Validity:Introduction,ChoicesofAttributeandPressures,	1
3.2	ErrorsinData Processing,	1
3.3	ErrorsinModelDesign,	1
3.4	ManagingChange	1
3.5	ManagingChange	1
3.6	CaseStudy-ThreeBlindMice,AlgorithmsandRace	1
3.7	CaseStudy- AlgorithmsintheOffice	1
3.8	CaseStudy-GermanWingsCrash,Google Flu,andDiscussion	1
Module4(AlgorithmicFairness)-8hours		
4.1	AlgorithmFairness:Introduction,	1
4.2	CorrectandMisleadingresults.	1
4.3	CorrectandMisleadingresults.	1
4.4	PHacking,	1
4.5	PHacking,	1
4.6	CaseStudy-HighThroughputBiology, Geopricing	1
4.7	CaseStudy- YourSafetyIsMyLostIncome	1
4.8	CaseStudy-Discussion-fairnessofapplyingfacerecognitionto give preference to the often-visiting customers in business	1

.

	Module5(SocietalConsequences)	
5.1	Introduction,	1
5.2	SocietalImpact,	1
5.3	Ossification,	1
5.4	Surveillance	1
5.5	Codeofethics,	1
5.6	Codeofethics,WrapUp,	1
5.7	CaseStudies-SocialCreditScores,andDiscussion	1
5.8	CaseStudies-PredictivePolicing,andDiscussion	1

ReferenceBooks

Books:

1. DJPatil,HilaryMason,MikeLoukides,_EthicsandDataScience,',O'Reilly Media,Inc.

2. ShannonVallor,Ph.D.WilliamJ.Rewak,S.J.Professorof Philosophy,_AnIntroductiontoDataEthics'SantaClara University

3. KordDavis_Ethicsofbigdata'ISBN:9789350238806,9350238802

4. SamikshaShukla,JossyP.George,KapilTiwari, Joseph Varghese Kureethara,_DataEthicsandChallenges'Springer, ISBN:978-981-19-0752-4

5. JohnHavens_ArtificialIntelligence:Embracing OurHumanityto MaximizeMachines'TarcherPerigee.ISBN-10:0399171711

.

URLs

MOOCS:https://courses.edx.org/courses/coursev1:

MichiganX+DS101x+1T2018/courseware/94ac457869964552a69a3f37ba579954/67
1f46458 36145eea658edbf9298be64/

.

J1PADT155	SPEECH PROCESSING	CATEGORY	L	T	P	CREDIT
		PROGRAM ELECTIVE-2	3	0	0	3

Preamble

The course aims to introduce the student to the various aspects of speech processing including modelling of human speech. The topics covered in the course includes Computational Phonology, Models of Spelling and Pronunciation, speech synthesis and speech recognition. It helps the learners to develop application involving speech processing.

CourseOutcomes: Afterthecompletionofthecoursethestudentwillbeableto

CO1	Analysethedifferentaspectsofproductionofspeechinhumans(**Cognitive Knowledge Level: Analyse**)
CO2	Makeofusethemethodsusedforspellingerrordetectionandcorrection (**Cognitive Knowledge Level: Apply**)
CO3	Illustratethevariousmodelsforspeechpronunciationvariations(**Cognitive Knowledge Level: Apply**)
CO4	Makeuseofthedifferentmodelsforrecognizinghumanspeechandconverting into equivalent text (**Cognitive Knowledge Level: Apply**)
CO5	Comprehendtheprocessesinvolvedintheacousticprocessingofhumanspeech (**Cognitive Knowledge Level: Apply**)
CO6	Design,Develop,andImplementinnovativeideasonspeechprocessingconcepts and techniques. **(Cognitive Knowledge Level: Create)**

ProgramOutcomes

Graduatesofthisprogramwillbeabletodemonstratethefollowingattributes.

PO1:Anabilitytoindependentlycarryoutresearch/investigationand development work inengineering and allied streams.

.

PO2:Anabilitytocommunicateeffectively,writeandpresenttechnicalreports on complexengineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialisation ofthe program. The masteryshould be at a level higherthan the requirements in the appropriatebachelor program.

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsfor real worldproblems by following the standards.

PO5: An ability to identify, select and apply appropriate techniques, resources andstate-ofthe-arttoolstomodel,analyseandsolvepracticalengineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects.

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoproject managementand finance whichfocusonEntrepreneurshipand Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1			✓			✓	
CO2			✓		✓	✓	
CO3			✓		✓	✓	
CO4			✓			✓	
CO5			✓			✓	
CO6	✓	✓	✓	✓	✓	✓	✓

.

AssessmentPattern

Bloom'sCategory	**EndSemester Examination**
Apply	50-80%
Analyse	20-40%
Evaluate	-
Create	-

Markdistribution

TotalMarks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesignbased questions (for both internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum 10 publications shall be referred) : 15 marks

i. Coursebasedtask/Seminar/Datacollectionandinterpretation:15marks

i. Testpaper(1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.

Course based task/test paper questions shall be useful in the testing of knowledge,skills,comprehension,application,analysis,synthesis,evaluation and understanding of the students.

.

EndSemesterExaminationPattern:

Theend semester examinationwillbeconducted bytherespectivecollege.

There will be two parts; Part A and Part B.

PartAwillcontain5 numerical/short answer questions with1 question fromeach module,having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity ofthe students in a course, through longanswer questions relating to theoretical/practical knowledge, derivations,problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry7 marks.

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over theaverage ESE mark % for the core courses. ESE marks awarded toa student for each electivecourse shall be normalized accordingly.

Forexampleiftheaverageendsemestermark% foracorecourseis40,thenthe maximumeligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. Explainthevariousvocalorgansandhowtheyproduce speech phones. 2. Draw and explain the transducer for z-devoicing rule.

CourseOutcome2 (CO2):

1. Writebriefnotesonthenoisychannelmodelofpronunciationandspelling.

2. Whatismeantbytheminimumeditdistancebetweentwostrings. Calculate theminimum edit distance between the words 'kitten' and 'sitting'.

.

CourseOutcome3 (CO3):

1. Illustrate the concept of pronunciation dictionaries with anexample.2.Discussthevariousphonologicalaspectsof prosody in speech.

CourseOutcome4 (CO4):

1. Give a brief overview of the Bayesian model of speech recognition.2.Listandbrieflyexplainthevariousparameters needed to define an HMM.

CourseOutcome5 (CO5):

1. Highlightthestepsinvolvedinextractionofspectralfeaturesfrom sound wave. 2. List the different approaches for calculating probabilities of acoustic feature vectors.

.

ModelQuestionpaper

TotalPages:2

Reg No.:_______________ Name:_________________________

___

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRSTSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

Course Code: J1PADT155

CourseName:SpeechProcessing

Max.Marks:60 Duration:2.5Hours

Branch:ComputationalLinguistics

PARTA

AnswerallquestionsinPARTA.Eachquestioncarries5 marks

Q. No		**Marks**
1	WritethelexicalentryforthepronunciationoftheEnglishpasttense suffix -d, and the two-level rules that express the difference in its pronunciation.	(5)
2	Statetheimportantthreeconcernswhileperformingpronunciation dictionary lookup.	(5)
3	Defineweightedautomata.Giveweightedautomatafortheword _tomato‘.	(5)

.

4 WhatarehiddenMarkovmodels?Statetheparametersneededtodefine . (5)

.

5		Summarizetheprocessoffeatureextractionofsoundwaves.	(5)

PARTB

AnsweranyfivefullquestionsinPARTB.Eachfullquestioncarries7marks

6		Designatransducerfori-insertionrulebywritingappropriate phonological rules.	(7)
7		Writenotesontheplace andmannerofarticulationofconsonants.	(7)
8	a)	WritenotesontheBayesianmodelforspelling correction.	(4)
	b)	Statetheforwardalgorithm.Also,explainhowitisusefulin computinglikelihoods.	(3)
9		Usingtheminimumeditdistancemethod,findwhethertheword _drive‘iscloserto_brief‘ orto _divers‘.	(7)
10		Explainthe useoftheViterbialgorithmin speechrecognitionprocess.	(7)
11		Writesnotesonspeechrecognitionarchitecture.	(7)
12		Discusshowaspeechrecognitionsystemcanbe trained.	(7)

****

.

Syllabus

Module	Syllabus Content	No.of Lecture Hours(38)
I	**Computational Phonology** - Articulatory Phonetics – Production andClassificationofSpeechSounds–VocalOrgans-Consonants–PlaceofArticulation;Consonants –MannerofArticulation;Vowels -PhonemeandPhonologicalRules;PhonologicalRulesand Transducers	8
II	**Speech Synthesis** - Mapping Text to Phonemes for TTS PronunciationDictionaries-TextAnalysis-FSTbasedpronunciation lexicon - Prosody in TTS – Phonological aspects of Prosody – Phonetic aspects of Prosody – Prosody in speech synthesis	8
III	**ModelsofSpellingandPronunciation**-Spellingerrors-Spelling ErrorPatterns-DetectingNonwordErrors-Probabilisticmodelsof spelling - Bayesian method to spelling – Minimum Edit Distance - The Bayesian Method for Pronunciation-Decision Tree Models of Pronunciation Variation – Weighted Automata and Segmentation	10
IV	**SpeechRecognition**-SpeechRecognitionArchitecture–Bayesian Model of Speech Recognition - Hidden Markov Models - Viterbi Algorithm – Advanced Methods for Decoding - A* Decoding	6
V	**Acoustic processing of speech** - Sound Waves - Interpreting a Waveform – Spectra – Feature Extraction - Computing Acoustic Probabilities-GaussianModels-NeuralNetModels-Traininga Recognizer	6

.

CoursePlan

No	Topic	No. of Lectures
1	**ComputationalPhonology(8Hours)**	
1.1	ArticulatoryPhonetics-Introduction	1
1.2	ProductionandClassificationofSpeechSounds	1
1.3	VocalOrgans	1
1.4	Consonants–PlaceofArticulation	1
1.5	Consonants–Mannerof Articulation	1
1.6	Articulationofvowels	1
1.7	PhonemeandPhonologicalRules	1
1.8	PhonologicalRulesandTransducers	1
2	**SpeechSynthesis(8Hours)**	
2.1	MappingTexttoPhonemesforTTS	1
2.2	PronunciationDictionaries	1
2.3	Text Analysis	1
2.4	FSTbasedPronunciationLexicon	1
2.5	Prosodyin TTS	1
2.6	PhonologicalAspectsofProsody	1

.

2.7	PhoneticAspectsofProsody	1
2.8	ProsodyinSpeechSynthesis	1
3	**ModelsofSpellingand Pronunciation(10Hours)**	
3.1	Spellingerrors-SpellingErrorPatterns	1
3.2	DetectingNonwordErrors	1
3.3	ProbabilisticModelsofSpelling	1
3.4	BayesianMethodtoSpelling	1
3.5	MinimumEditDistance	1
3.6	TheBayesianMethodforPronunciation	1
3.7	DecisionTreeModelsofPronunciationVariation	1
3.8	WeightedAutomata	1
3.9	ComputingLikelihoodsfromWeightedAutomata-TheForward Algorithm	1
3.10	Segmentation	1
4	**SpeechRecognition(6Hours)**	
4.1	SpeechRecognition Architecture	1
4.2	BayesianModelofSpeech Recognition	1
4.3	HiddenMarkovModels	1
4.4	ViterbiAlgorithm	1
4.5	AdvancedMethodsforDecoding	1

.

4.6	A*Algorithm	1
5	**Acousticprocessingofspeech(6Hours)**	
5.1	SoundWaves- InterpretingaWaveform	1
5.2	SpectraAnalysis	1
5.3	FeatureExtractionfromWaveformsand Spectra	1
5.4	ComputingAcousticProbabilities-GaussianModels	1
5.5	ComputingAcousticProbabilities-NeuralNetModels	1
5.6	TrainingaRecognizer	1

ReferenceBooks

1. Jurafsky, D. and J. H. Martin, Speech and language processing: An IntroductiontoNaturalLanguageProcessing,ComputationalLinguistics,and Speech Recognition, Upper SaddleRiver, NJ: Prentice-Hall, 2000

2. ClaudioBecchettiandLucioPrinaRicotti,―SpeechRecognition‖,John Wiley and Sons,1999

3. FrederickJelinek,―StatisticalMethodsofSpeechRecognition‖,MITPress,1997

4. Ben Gold and Nelson Morgan, ―Speech and Audio Signal Processing, ProcessingandPerceptionofSpeechandMusic‖,Wiley-IndiaEdition, 2006

.

J1PADT165	INFORMATION THEORY	CATEGORY	L	T	P	CREDIT
		PROGRAM ELECTIVE 2	3	0	0	3

Preamble: The course introduces the mathematical and fundamental notions of information theory that play a significant role in building modern communication systems. It covers entropy, mutual information, source coding, channel coding, continuous sources and channels and rate distortion theory. The course enables the learners to effectively choose appropriate source codes and channelcodes according to different applications.

CourseOutcomes:

After thecompletionofthecoursethestudentwillbeableto

CO 1	Comparedifferenttypesofentropyandusetheconceptofmutualinformation.(**Cognitive Knowledge Level: Apply**)
CO 2	DesignsourcecodesandappreciatetheuseofShannon‘ssourcecodingtheorem. (**Cognitive Knowledge Level: Apply**)
CO 3	DesignchannelcodesandappreciatetheuseofShannon‘schannelcoding theorem. (**Cognitive Knowledge Level: Apply**)
CO 4	Demonstratethenotionsofcontinuoussourcesandchannels.(**CognitiveKnowledge Level:Apply**)
CO 5	Elaborateonthevariousaspectsoftheratedistortiontheorem.(**Cognitive Knowledge Level: Apply**)
CO 6	Design,Develop,andImplementapplicationsusingInformationtheory concepts and techniques. **(Cognitive Knowledge Level: Create)**

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduate aftercompleting the course.

PO1: Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams.

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards.

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO 1	✔		✔		✔		
CO 2	✔		✔	✔	✔		
CO 3	✔		✔	✔	✔		
CO 4	✔		✔		✔		
CO 5	✔						
CO 6	✔	✔	✔	✔	✔	✔	✔

.

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Apply	50-80%
Analyse	20-40%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

Continuous InternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesign-basedquestions(for both internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

i. Coursebasedtask/Seminar/Datacollectionandinterpretation: 15marks

i. Testpaper(1number) : 10marks

Testpapershallincludeminimum80%ofthesyllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension,application,analysis,synthesis,evaluationandunderstandingofthestudents.

.

EndSemester ExaminationPattern:

Theendsemesterexaminationwillbeconductedbytherespectivecollege. There will

be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having5marksforeachquestion.Studentsshouldanswerallquestions.PartBwillcontain7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module

ofwhichstudentshouldansweranyfive.Eachquestioncancarry 7 marks. Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

Forexample,iftheaverageendsemestermark%foracorecourseis 40,thenthemaximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. Asourceproduces4symbolswithprobabilities1/2,1/2,1/8,and1/8.Findtheinformation contentof each symbol.

2. Azeromemorysourcehasasourcealphabet,S={s1,s2,s3}withP={0.5,0.3, 0.2}. Findthe entropy of the source.

3. Given a binary source with two symbols x1 and x2. Given x2 is twice as long as x1 andhalf as probable. The durationofx1 is 0.3 seconds. Calculate the informationrate of thesource.

CourseOutcome2(CO2)

1. Considera[7,4]linearblockcodewithparitycheckmatrix

.

$$H = \begin{matrix} 1\,0\,1\,1\,1\,0\,0 \\ 1\,1\,0\,1\,0\,1\,0 \\ 0\,1\,1\,1\,0\,0\,1 \end{matrix}$$

a) Constructcodewordsforthe[7,4]code.

b) ShowthatthiscodeisaHammingcode.

2. Consider a source with 8 alphabets, a to h with respective probabilities 0.2, 0.2, 0.18, 0.15, 0.12, 0.08, 0.05 and 0.02. Construct a minimum redundancy code and determine the codeefficiency.

3. Theparitymatrixfora(6,3)systematiclinearblockcodeisgivenby

(i) Findallcodewords.(ii) Findgeneratorandparitycheckmatrix.

CourseOutcome3(CO3):

1. StateandproveShannon‘schannelcodingtheorem.

2. Derivethe capacityofbinarysymmetricchannel.

3. Stateand provesourcechanneltheorem.

CourseOutcome4(CO4):

1. Derivethedifferentialentropyofanormaldistribution.

2. Derivethechainrulefordifferentialentropy.3.

CourseOutcome5(CO5):

.

1. Stateand provetheconverseto theratedistortion theorem.

2. Evaluatetheratedistortionfunctionforabinary source.

ModelQuestionPaper

QPCODE:

Reg No: __________________

Name: ____ PAGES:2

APJABDULKALAMTECHNOLOGICALUNIVERSITY

FIRSTSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

Course Code: J1PADT165

CourseName:InformationTheory

Max.Marks:60 **Duration:2.5Hours**

PARTA

AnswerAllQuestions.EachQuestionCarries5 Marks

1. Derivetherelationbetweenentropyandmutualinformation.

2. StateKraft-McMillaninequality.

3. Explainchannelcodingtheorem.

4. DiscussaboutGaussianchannels.

5. Defineratedistortiontheorem. **(5x5 =25)**

.

PartB

(Answeranyfivequestions. Eachquestioncarries7 marks)

6. Differentiatebetweenjointentropyandconditionalentropy. **(7)**

7. Suppose one has ncoins, among whichthere mayor maynot be onecounterfeitcoin.If there is a counterfeit coin, it may be either heavier or lighter than the othercoins. The coins are to be weighed by a balance. **(7)**

a) Find an upper bound on the number of coins n so that k weighings will findthe counterfeitcoin(ifany)andcorrectlydeclareittobeheavierorlighter.b)Whatisthe coin weighing strategy for k = 3 weighings and 12 coins?

8. FindabinaryHuffmancode forthe source emitting symbolswith probabilities0.49, 0.14, 0.14, 0.07, 0.04, 0.02, 0.02, 0.01. Also find the code efficiency andredundancy. **(7)**

9. Explainwiththehelpofaneatdiagram,discretememorylesschannelwithfeedback. **(7)**

10. Considertherandomvariable. **(7)**

a) FindabinaryHuffmancodefor X.

b) Findtheexpectedcodelengthforthis encoding.

c) Finda ternaryHuffmancodefor X.

11. ConsiderthediscretememorylesschannelY=X+Z(mod11),where **(7)**

andX∈{0,1,2,…,10}.AssumethatZisindependentofX.

a) Findthecapacity.

b) Whatisthemaximizingp*(x)?

(7)

12. Consider asourceXuniformlydistributedontheset {1,2,...,m}.Findtherate distortion function for this source with Hamming distortion, i.e.,

$$d(x, \hat{x}) = \begin{cases} 0 & \text{if } x = \hat{x}\,, \\ 1 & \text{if } x \neq \hat{x}\,. \end{cases}$$

Syllabus

Entropy,losslesssourcecoding,Huffmancode,Shannon‘ssourcecodingtheorem,Shannon‘s channelcodingtheorem,continuoussourcesandchannels,ratedistortiontheory.

	Syllabus	
Module	**Content**	**Hours**
1	**Introduction to Entropy**: Entropy- Memoryless sources - Markov sources Entropy of a discrete random variable - joint, conditional and relativeentropy-mutualInformationandconditionalmutualinformation -Chainrelationforentropy,relativeentropyandmutualinformation	12
2	**Losslesssourcecoding**-Uniquelydecodablecodes-Instantaneouscodes Kraft's inequality - Optimal codes - Huffman code- Shannon's Source Coding Theorem.	7
3	**Channelcoding**-Shannon'sChannelCodingTheoremanditsconverse-Channels with feedback - Joint source channel coding Theorem.	7
4	**Continuous Sources andChannels**:ContinuousSourcesandChannels-DifferentialEntropy-Joint,relativeandconditionaldifferentialentropy–Mutual information- Waveform channels- Gaussian channels.	7
5	**Rate Distortion Theory**: Introduction - Rate Distortion Function - Properties - Continuous Sources and Rate Distortion measure - Rate DistortionTheorem–Converse–InformationTransmissionTheorem- Rate Distortion Optimization.	7

.

CoursePlan

No	Topic	No. of Lectures (40 hours)
1	**Module1(Introductiontoentropy)(12hrs)**	
1.1	Entropy-Memorylesssources	2
1.2	Markovsources	2
1.3	Entropyofadiscreterandomvariable	2
1.4	Joint,conditionalandrelativeentropy	2
1.5	Mutualinformationandconditionalmutualinformation	2
1.6	Chainrelationforentropy,relativeentropyandmutualinformation	2
2	**Module2(Losslesssourcecoding)(7hrs)**	
2.1	Uniquelydecodablecodes	1
2.2	Instantaneouscodes	1
2.3	Kraft'sinequality	1
2.4	Optimalcodes-Huffmancode	2
2.5	Shannon'sSourceCodingTheorem	2
3	**Module3(Channelcoding)(7 hrs)**	
3.1	Introductiontochannelcoding	1
3.2	Shannon'sChannelCodingTheoremanditsconverse	2
3.3	Channelswithfeedback	2
3.4	JointsourcechannelcodingTheorem	2
4	**Module4(ContinuousSourcesandChannels)(7hrs)**	
4.1	ContinuousSourcesandChannels	2

.

4.2	DifferentialEntropy	1
4.3	Joint,relativeandconditionaldifferentialentropy	1
4.4	Mutualinformation	1
4.5	Waveformchannels	1
4.6	Gaussianchannels	1
5	**Module5(RateDistortionTheory)(7hrs)**	
5.1	Introductiontoratedistortiontheory	1
5.2	RateDistortionFunction–Properties	1
5.3	ContinuousSourcesandRateDistortionmeasure	1
5.4	RateDistortionTheorem	1
5.5	Converseofratedistortiontheorem	1
5.6	InformationTransmissionTheorem	1
5.7	RateDistortionOptimization.	1

References

1. T.CoverandThomas,ElementsofInformationTheory,JohnWiley&Sons.

2. RobertGallager,InformationTheoryandReliableCommunication,JohnWiley & Sons.

.

3. R.J.McEliece,Thetheoryofinformation&coding,AddisonWesley
Publishing Co.

4. T.Bergu,RateDistortionTheoryaMathematicalBasisforDataCompression PH
Inc.

5. R.W.Hamming.CodingandInformationTheory.PrenticeHallInc.

.

J1PADL107	ADVANCED MACHINE LEARNING LAB	CATEGORY	L	T	P	Credit
		Laboratory1	0	0	2	2

Preamble: Studyof thecourseenables the learners to make use of themachine learning concepts and algorithms to derive data insights. The course provides exposure to the design and implementation aspects of machine learning algorithms such as decision trees, regression, naive bayes algorithm, clustering algorithms and artificial neural network. This helps the students to develop machine learning based solutions to real world problems.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO#	CourseOutcomes
CO1	Applymodernmachinelearningnotionsinpredictivedataanalysis(**Cognitive KnowledgeLevel:Apply**)
CO2	Analyzetherangeofmachinelearningalgorithmsalongwiththeirstrengthsandweaknesses (**Cognitive Knowledge Level: Analyze**)
CO3	Designanddevelopappropriatemachinelearningmodelstosolverealworldproblems. (**Cognitive Knowledge Level: Analyze**)
CO4	Buildpredictivemodelsfromdataandanalyzetheirperformance(**CognitiveKnowledge Level: Create**)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompleting the course.

PO1: Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineeringactivitiesbyinteractingwiththeengineeringfraternityandwithsocietyatlarge.

PO3: Anabilitytodemonstrateadegreeofmasteryovertheareaasperthespecializationofthe program.Themastery shouldbeatalevel higherthantherequirements intheappropriate bachelor program

PO4: Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworldproblemsby following the standards

PO5: Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedtothestream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagementand finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓	✓	✓	✓	✓	✓	
CO2	✓	✓	✓	✓	✓	✓	
CO3	✓	✓	✓	✓	✓	✓	
CO4	✓	✓	✓	✓	✓	✓	

ContinuousInternalEvaluationPattern:

ThelaboratorycourseswillbehavingonlyContinuousInternalEvaluationandcarries100 marks.

Finalassessmentshallbedonebytwoexaminers;oneexaminerwillbeaseniorfacultyfrom the same department.

Continuous Evaluation : 60 marks

Finalinternalassessment:40marks

.

LabReport:

All the students attending the Lab should have a Fair Report. The report should contains details of experiment such as Objective, Algorithm/Design, Description, Implementation, Analysis, Results, and Outcome.The report should contain a print out of the respective code with inputs addressing all the aspects of the algorithm described and corresponding outputs. All the experiments noted in the fair report should be verified by the faculty regularly. The fair report, properly certified by the faculty, should be produced during the time of the final assessment.

Syllabus

Decision tree (ID3), Naïve bayesian classifier , Bayesian network, Expectation Maximization (EM) algorithm,K-means algorithm, K-nearest neighbor, Regression, Cross validation, Support Vector Machine (SVM), Artificial neural network, Backpropagation algorithm, Recurrent Neural Networks (RNN), Long Short Term Memory (LSTM), Google colab.

PracticeQuestions

1. Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.
2. Write a program to implement the naïve bayesian classifier for a sample training dataset stored as a .CSV file. Compute the accuracy of the classifier, considering few testdata sets.
3. Assuming a setof documents that need tobe classified,use the naïvebayesian Classifier model to perform this task. Calculate the accuracy, precision, and recall for your dataset.
4. Write a program to construct a Bayesian network considering medical data. Use this modeltodemonstrate the diagnosis of heart patientsusing standard Heart Disease Data Set. You can use Python ML library classes/API.
5. Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data setfor clusteringusingk-Means algorithm. Comparetheresultsof thesetwoalgorithms and comment on the quality of clustering. You can add Python ML library classes/API in the program.
6. Write a program to implement k-Nearest Neighbour algorithm to classify the iris dataset. Print both correct and wrong predictions. Python ML library classes can be used for this problem.
7. Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs.
8. Write a program to implement 5-foldcross validation on a given dataset. Compare the accuracy, precision, recall, and F-score for your data set for different folds.

.

9. ImplementSVM/SoftmaxclassifierforCIFAR-10dataset:(i)usingKNN,(ii)using3layer neural network.
10. BuildanArtificialNeuralNetworkbyimplementingtheBackpropagationalgorithmand test the same using appropriate data sets.
11. ImageCaptioningwithVanillaRNNs.
12. ImageCaptioningwithLSTMs.
13. FamiliarisationofcloudbasedcomputinglikeGoogle colab.

References:

1. JiaweiHan,MichelineKamber,JianPei.DataMiningConceptsandTechniques,Third Edition. Morgan Kaufmann.

2. ChristopherM.Bishop.Patternrecognitionandmachinelearning.Springer2006.

3. EthemAlpaydin,IntroductiontoMachineLearning,2ndedition,MITPress2010.

4. MohammedJ.ZakiandWagnerMeira,DataMiningandAnalysis:Fundamental Concepts and Algorithms, Cambridge University Press, First South Asia edition, 2016.

5. Goodfellow,I.,Bengio,Y.,andCourville,A.,DeepLearning,MITPress,2016.

6. NeuralNetworksandDeepLearning,Aggarwal,CharuC.,cSpringerInternational Publishing AG, part of Springer Nature 2018

.

:

SEMESTERII

Discipline:COMPUTERSCIENCEANDENGINEERING

Stream :CS2

.

:

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PCST201	ADVANCED DATA STRUCTURES AND ALGORITHMS	DISCIPLINE CORE 2	3	0	0	3

Preamble:Thecourseintroducesadvanceddatastructuresandalgorithmsindifferent domains. The goal of this course is to provide a solid background in the design and analysis of themajorclassesofalgorithms.Thecoursehelpsthelearnerstodeveloptheirownversions for a given computational task and to compare and contrast their performance.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto:*

CO1	Analysetherelevanceofamortizedanalysisandapplications.(CognitiveLevel: Apply)
CO2	Illustratestringmatchingalgorithms.(CognitiveLevel:Apply)
CO3	IllustrateadvanceddatastructureslikeBinomialheap,Fibonacciheap,Disjointset andstringmatchingalgorithms.(CognitiveLevel:Apply)
CO4	Illustratenetworkflowalgorithmsandapplications.(CognitiveLevel:Apply)
CO5	Makeuseofprobabilisticalgorithmsandapproximationalgorithmsincomputing. (CognitiveLevel:Apply)
CO6	Design,developandimplementsoftwareusingadvanceddatastructuresand algorithms.(CognitiveLevel:Create)

* The COs shown are only indicative. For each course, there can be 4 to 6 COs.

Program Outcomes (PO)

Outcomes are the attributes that are tobe demonstrated byagraduateaftercompletingthe course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3:Anabilitytodemonstrateadegreeofmasteryovertheareaasperthespecializationof the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of- the-art tool to model, analyse and solve practical engineering problems.

PO6:Anabilitytoengageinlife-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✔		✔	✔		✔	
CO2	✔		✔	✔		✔	
CO3	✔		✔	✔		✔	
CO4	✔		✔	✔		✔	
CO5	✔		✔	✔		✔	
CO6	✔	✔	✔	✔	✔	✔	✔

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	80%
Analyse	20%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
00	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesignbasedquestions(forboth internal and end semester examinations).

ContinuousInternalEvaluation :40marks

Micro project/Course based project: 20 marks

Course based task/Seminar/Quiz :10marks

Testpaper,1no. :10marks

The project shall be done individually. Group projects not permitted.

Test paper shall include minimum 80% of the syllabus.

Course based task/paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the University. There will be two parts; Part A and Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students). Students shall answer all questions.

Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks.

Total duration of the examination will be 150 minutes.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Explain how the accounting method of amortized analysis can be applied to stack operations.

2. Suppose we perform a sequence of n operations on a data structure in which the i^{th} operation costs i if i is an exact power of 2, and 1 otherwise. Use aggregate analysis to determine the amortized cost per operation.

3. What is the total cost of executing n of the stack operations PUSH, POP, and MULTIPOP, assuming that the stack begins with s_0 objects and finishes with s_n objects? Use potential method.

CourseOutcome2(CO2)

1. Use an aggregate analysis to show that the running time of KMP-MATCHER is $\theta(n)$.

2. Working modulo q=11, how many spurious hits does the Rabin-Karp matcher encounter in the text T = 3141592653589793 when looking for the pattern P = 26?

3. Compute the prefix function π for the pattern ababbabbabbababbabb

CourseOutcome3(CO3):

1. Analyse the time complexity of decrease-key operation of Fibonacci heap.

2. Illustrate extract-min operation of Binomial heap.

.

3. Explaintheheuristicsusedindisjo:iCn Ot set data structure to improve the running time.

CourseOutcome4(CO4):

1. ShowtheexecutionoftheEdmonds-Karpalgorithmonthegivenflownetwork(source:A and sink: J).

2. In the following figure, compute flow across the cut ({s, v_2, v_4}, {v_1, v_3, t}). What is the capacity of this cut?

Stateandprovemaxflowmincuttheorem.

Illustrate Miller-Rabin primalitytesting method.

Explain probabilistic selection algorithm.

Explaintheapproximationalgorithmforsubsetsumproblem.

Consider each of the following words as a set of letters: {*arid, dash, drain, heard, lost, nose,shun,slate,snare,thread}*.ShowwhichsetcoverGREEDY-SET-COVERproduces when we break ties in favor of the word that appears first in the dictionary.

.

ModelQuestionPaper

QPCODE:

RegNo:___________________

Name: ____PAGES: 4

APJABDULKALAMTECHNOLOGICALUNIVERSITY

SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

CourseCode: J1PCST201
CourseName:ADVANCEDDATASTRUCTURESANDALGORITHMS

Max.Marks:60 Duration:2.5Hours

PARTA

AnswerAllQuestions.EachQuestionCarries5Marks

1. Explainaccountingmethodofamortizedanalysiswithasuitableexample.
2. Explainthealgorithmforunitingtwobinomialheapsandanalysetherunning time.
3. MaximummatchinginabipartitegraphGcorrespondstoamaximumflowinits corresponding flow network G‘. Comment on this statement. Explain how maximum flow problem can be used to solve maximum bipartite matching problem.
4. Explainthbeprobabilisticalgorithmforverifyingmatrixmultiplicationproblem.
5. Explaintheapproximationalgorithmfortravelingsalespersonproblem.

PartB

(5x5=25)

(Answeranyfivequestions.Eachquestioncarries7marks)

6. DescribeKnuth-Morris-PrattalgorithmandillustrateusinggiventextT= AABAACAADAABAABA and pattern P = AABA.

7. (a)Using potential method, compute the amortized cost of incrementing a binary counter. (7)

   (b) Suppose we perform a sequence of n operations ona data structure in whichthe i^{th} operation costs i if i is an exact power of 2, and 1 otherwise. Use accounting method of amortized analysis to determine the amortized cost per operation. (3)

   (4)

8. (a)Explainhowdisjointsetdatastructureisusedtofindconnectedcomponents on an undirected graph.

(3)

.

(b) Showthebinomialheapthatresultswhenanodewithkey11isdeletedfrom the binomial heap shown in figure. (4)

9. (a)ExplainthestructureofFibonacciheap. (2)

(b) ApplyextractminimumoperationontheFibonacciheapshowninfigureand show the result. (5)

10 DescribeFord-Fulkersonalgorithmandapplyonthefollowingnetwork.Also obtain minimum cut across the network. (7)

11 (a)ApplyMiller-Rabinalgorithmtotestwhetherthenumber341isprimeornot. (4)

(b) Explainprobabilisticquicksortalgorithm. (3)

12 (a)Describepolynomial-timeapproximationschemeandfullypolynomial-time approximation scheme. (3)

(b) Give an example of a graph for which APPROX-VERTEX-COVER always yields a optimal solution. (4)

s

u

b

.

Syllabus

Module–1(AmortizedanalysisandStringmatching)

Overviewofasymptoticnotationsandcomplexityanalysis,Amortizedanalysis –aggregate analysis, accounting method, potential method.
Stringmatching–introduction,Rabin-Karpalgorithm,Knuth-Morris-Prattalgorithm.

Module–2(Advanceddatastructures)

Overview of binary heap operations, Binomial tree and heap, Binomial heap operations, Fibonacci heap structure, Fibonacci heap operations, Disjoint set – overview, linked list representation, disjoint set forests.

Module–3(Networkflow)

Network flow properties, examples, residual network, augmenting path, cut of network, maxflow-mincuttheorem,Ford-Fulkersonalgorithm,Edmonds-Karpalgorithm,maximum bipartite matching.

Module–4(Probabilisticalgorithms)

Introduction,typesofprobabilisticalgorithms,Numericalalgorithms–Numerical integration, Probabilistic counting, Monte-Carlo algorithms – Verifying matrix multiplication.
Numbertheoryfundamentals–modulararithmetic,modularexponentiation,Euler'sTheorem and Fermat's Theorem, Primality testing – Miller-Rabin test.
LasVegasalgorithms–Probabilisticselectionandquicksort.

Module–5(Approximationalgorithms)

Introduction,Vertex-coverproblem,Traveling-salesmanproblem,Set-coveringproblem, Subset-sum problem.

CoursePlan

No	Topic	No.of Lectures (37)
1	Module–1(AmortizedanalysisandStringmatching)	
1.1	Overviewofasymptoticnotationsandcomplexityanalysis	1
1.2	Amortizedanalysis–aggregateanalysis	1
1.3	accountingmethod	1
1.4	potentialmethod	1
1.5	Stringmatching–introduction	1
1.6	Rabin-Karpalgorithm	1
1.7	Knuth-Morris-Prattalgorithm(1)	1
1.8	Knuth-Morris-Prattalgorithm(2)	1
2	Module–2(Advanceddatastructures)	
2.1	Overviewofbinaryheapoperations	1

.

2.2	Binomialtreeandheap	1
2.3	Binomialheapoperations(1):COMPUTERSCIENCEANDENGINEE	1
2.4	Binomialheapoperations(2)	1
2.5	Fibonacciheapstructure	1
2.6	Fibonacciheapoperations(1)	1
2.7	Fibonacciheapoperations(2)	1
2.8	Disjointset–overview,linkedlistrepresentation	1
2.9	disjointsetforests	1
3	Module–3(Networkflow)	
3.1	Networkflowproperties,examples	1
3.2	residualnetwork,augmentingpath,cutofnetwork	1
3.3	maxflow-mincuttheorem	1
3.4	Ford-Fulkersonalgorithm	1
3.5	Edmonds-Karpalgorithm	1
3.6	Maximumbipartitematching	1
4	Module–4(Probabilisticalgorithms)	
4.1	Introduction,typesofprobabilisticalgorithms	1
4.2	Numericalalgorithms–Numericalintegration,Probabilisticcounting	1
4.3	Monte-Carloalgorithms–Verifyingmatrixmultiplication	1
4.4	Numbertheoryfundamentals–modulararithmetic,modular exponentiation	1
4.5	Euler'sTheoremandFermat'sTheorem	1
4.6	Primalitytesting–Miller-Rabintest(1)	1
4.7	Primalitytesting–Miller-Rabintest(2)	1
4.8	LasVegasalgorithms–Probabilisticselectionandquicksort	1
5	Module–5(Approximationalgorithms)	
5.1	Introduction	1
5.2	Vertex-coverproblem	1
5.3	Traveling-salesmanproblem	1
5.4	Set-coveringproblem	1
5.5	Subset-sumproblem(1)	1
5.6	Subset-sumproblem(2)	1

ReferenceBooks

1. T.H.Cormen,C.E.Leiserson,R.L.RivestandC.Stein,—Introductionto Algorithms‖, MITPress,3rdedition,2009.

2. GillesBrassardandPaulBratley,—Fundamentalsofalgorithms‖,Prentice-hallofIndia Private Limited, 2001.

3. RajeevMotwani,PrabhakarRaghavan,—RandomizedAlgorithms‖,CambridgeUniversity Press,2000.

4. DexterC.Kozen,—TheDesignandAnalysisofAlgorithms‖,Springer.
5. JonKleinbergandEvaTardos,—AlgorithmDesign‖,PearsonEducation,2006.

.

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PADT202	DEEP LEARNING	PROGRAM CORE3	3	0	0	3

Preamble: This course introduces the core concepts of deep learning and also provides aninsight into recent developments in the field.The concepts covered in the course includeNeural Network Optimization techniques, Regularization, Convolutional Neural networks, Recurrent Neural Networks, Word Embedding and Transformers. This course helps thestudents to develop solutions to real world applications using deep learning techniques.

Course Outcomes: After the completion of the course the student will be able to:*

CO1	Analyse the challenges in learning ofNeural Networksand develop solutions to overcome the issues (Cognitive knowledge level: Apply)
CO2	Constructconvolutionalneuralnetworksfordeeplearningapplications(Cognitive knowledge level: Apply)
CO3	Make use of recurrent neural network and its variants in relevant application areas (Cognitive knowledge level: Analyze)
CO4	Apply the deep learning techniques in natural language based applications (Cognitive knowledge level: Analyze)
CO5	Distinguishthetransformerarchitecturefromearlierarchitectures(Cognitive knowledge level: Analyze)
CO6	Design,developandimplementsolutionsbasedonDeepLearningconceptsand techniques(Cognitive knowledge level: Create)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1:Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams

PO2:An ability to communicate effectively, write and present technical reports on complex engineeringactivitiesbyinteractingwiththeengineeringfraternityandwithsocietyatlarge.

PO3:Anability to demonstrate a degree of mastery over thearea as per the specialization of theprogram.Themasteryshouldbeatalevelhigherthantherequirementsintheappropriate bachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anability toidentify,selectandapplyappropriatetechniques,resourcesandstate-of-the-art tool to model, analyse and solve practical engineering problems.

PO6:Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedtothe stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓		✓	✓	✓	✓	
CO2	✓		✓	✓	✓	✓	
CO3	✓		✓	✓	✓	✓	
CO4	✓		✓	✓	✓	✓	
CO5	✓		✓	✓	✓	✓	
CO6	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	40
Analyse	20
Evaluate	Can be evaluated using miniprojects/assignments
Create	Can be evaluated using miniprojects/assignments

Markdistribution

Total Marks	CIE	ESE	ESEDuration
100	40	60	2.5hours

.

ContinuousInternalEvaluationPattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

ContinuousInternalEvaluation :40marks

Micro project/Course based project :20marks

Coursebasedtask/Seminar/Quiz :10marks

Testpaper,1no. :10marks

Theprojectshallbedoneindividually.Groupprojectsnotpermitted. Test

paper shall include minimum 80% of the syllabus.

Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the University.There will be two parts; Part A and Part B. Part A contain 5 numerical questions with 1 question from each module, having 5 marks for each question. (such questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students). Students shall answer all questions.

Part B will contain 7 questions (such questions shall be useful in the testing of overall achievementandmaturityofthestudentsinacourse,throughlonganswerquestionsrelating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), withminimumonequestionfromeachmoduleofwhichstudentshouldansweranyfive. Each question can carry 7 marks.

Totaldurationoftheexaminationwillbe150minutes.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. ApplytheAdamoptimizationalgorithmtoanexampleofyourchoiceandimplementit using appropriate tools..

2. Explorethevalidityofthestatement:—ComputationalcomplexityduetoBatch normalization is proportional to the size of the mini batch‖.

.

CourseOutcome2(CO2)

1. DesignandsketchaCNNforanyapplicationofyourchoice.

2. DifferentiatebetweenaCNNandanautoencoderintermsofarchitectureandapplications.

3. Demonstrate the use of transfer learning in natural language processing.

Course Outcome 3(CO3):

1. IllustratetheworkflowinsideanRNNbyusinganexample.

2. IncomparisontoanRNN,howmanyordersofexcesscomputationarethereinanLSTM?

3. ImplementanysequencetosequencebasedapplicationusingRNNandLSTM.Analyse which one of these two is more relevant for your application.

CourseOutcome4(CO4):

1. Provideschematicdiagramsshowinginput,outputandhiddenlayersalongwiththeir connectivity for the word2vec embedding method.

2. AnalyzehowgoodBERTisintheapplicationareaofautomatedquestionanswering.

3. Derive the loss functions used for training of a GAN.

Course Outcome 5 (CO5):

1. Differentiatebetweenself-attentionandmulti-headattention.

2. Justifythestatement:—Visiontransformerisanadaptationofthetransformerarchitecture forcomputervisionapplications".

3. Analyzethecomputationalcomplexitiesinvolvedinlargescalepretrainingoftransformers.

CourseOutcome6(CO6):

1. Developadeeplearningmodelasasolutiontoarealworldproblemandanalyseits performance.

.

ModelQuestionPaper			
QPCODE:			
Reg No: ___________________			
Name:____PAGES:4			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode: J1PADT202			
CourseName:DEEPLEARNING			
Max.Marks:60 Duration:2.5Hours			
PARTA			
AnswerAllQuestions.EachQuestionCarries5Marks			
1.		Consider a simple perception f such that f:$R^4 \rightarrow R$ which uses sigmoid as itsactivation function. The inputs areX=[0.1 ,0.2, 0.3, 0.4] and thecorresponding weights W=[−0.2, 0.1, 0.2, 0.3]. Compute the derivative ofsigmoidofz(σ'(z)),wherezisalinearcombinationofweightsvectorWandinput vector X. Ignore the bias term.	
2.		Illustrateconvolutionandpoolingoperationwithanexample.	
3.		ExplaintherelevanceofLSTMinthecontextofrecurrentneuralnetworks..	
4.		Describetheword2vectechniqueofwordembedding.	
5.		Illustratetheconceptofself-attentionintransformers.	(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			

6.	a	Explaintheconceptsbehindi)Earlystoppingii)dropoutiii)weightdecay	(4)
	b	UsingAdagrad-basedgradientdescent,findthenewvalueofparameter θ_{t+1},given that the old value θ_t=0.7, aggregated gradient $\Delta\theta_t$=0.9, gradient accumulationr_{t-1}=0.6,learningrateα=0.1andsmallconstantδ=10.	(3)
7.	a	TheinputtoaCNNarchitectureisacolorimageofsize112x112x3.Thefirst convolution layer comprises 64 kernels of size 5x5 applied with a stride of 2 and padding 0. What will be the number of parameters?	(3)
	b.	ExplainhowtheVanishingandExplodinggradientproblemisaddressedin ResNet.	(4)
8.		ExplaintheworkingofRNNanddiscusshowbackpropagationthroughtime is used in recurrent networks.	(7)
9.		IllustratetheGloVetechniqueofwordembedding.	(7)
10.		DemonstratetheworkflowoftheBERTtechniqueofwordembedding.	(7)
11.		Justifytheneedforself-attentioninthetransformerarchitecture.	(7)
12.		IllustratethearchitectureofaVisiontransformer.	(7)

Syllabus

Module1:IntroductiontoDeepLearning	(8hours)
DeepLearningvstraditionalmachinelearning,GradientDescent,AdamOptimization,Weight initialization strategies,Batch Normalization, Regularization techniques,Cross entropy loss function	
Module2:ConvolutionalNeuralNetworks	(9hours)
Convolutionoperation,CNNlayers,BuildingaCNNmodel,TrainingaCNN,DeepAutoencoders,	
Module3:RecurrentNeuralNetworks	(8hours)
Basic architecture and variants, Backpropagation through time, LSTM, Deep recurrent neuralnetworks, Machine Translation, Encoder Decoder architecture, Sequence to sequence learning, Beam search	
Module4:Textprocessingusingdeeplearning	(8hours)

Wordembeddings,word2vec,GloVe,Subwordembedding,BERT
Module5:Transformers (8hours)
Bahdanauattention,Multi-headattention,Self-attentionandpositionalencoding,Transformerarchitecture, Transformers for vision, Large-scale pretraining with transformers

CoursePlan

No	Topic	No.ofLectures
1	IntroductiontoDeepLearning	
1.1	DeepLearningvstraditionalmachinelearning	1
1.2	GradientDescent	1
1.3	ChallengesinNeuralNetworkOptimization-Illconditioning, local Minim, Saddle Points, Exploding Gradients	1
1.4	Adagrad,RMSProp,AdamOptimization	1
1.5	Weightinitializationstrategies	1
1.6	BatchNormalization	1
1.7	Regularization-ParameterNormPenaltiesL2,L1Regularization	1
1.8	Regularization-DatasetAugmentation,NoiseRobustness,Early Stopping, Dropouts	1
2	ConvolutionalNeuralNetworks	
2.1	Convolutionoperation	1
2.2	Convolution-SparseInteraction,Parametersharing,Equi-varient Representation	1
2.3	Pooling,ConvolutionalNeuralNetworks,Parameters	1
2.4	Variantsofbasicconvolutionfunction-dilatedconvolution, Transpose Convolution, 1x1 convolution, 3DConvolution	1
2.5	BackpropagationinConvolutionalLayers	1
2.6	CrossEntropyLossFunction,AlexNet	1
2.7	GoogleNetArchitecture,InceptionModule,	1
2.8	Autoencoders-sparseautoencoders	1
2.9	Denoisingautoencoders	1
3	RecurrentNeuralNetworks	
3.1	UnfoldingGraphs	1

.

3.2	RecurrentNeuralNetworks,ComputingGradients	1
3.3	ModelingSequencesConditionedonContextwithRNNs	1
3.4	Encoder-DecoderSequence-to-SequenceArchitecture	1
	Transformers	

.

3.5	Deeprecurrentneuralnetworks,RecursiveNeuralNetworks	1
3.6	ExplodingandVanishingGradients	1
3.7	LSTM	
3.8	Beamsearch	1
4	Wordembedding,BERT	
4.1	Wordembeddings	1
4.2	word2vec-Skip-gramandContinuousBagofwords	1
4.3	WordEmbeddingwithGlobalVectors(GloVe)	1
4.4	Subwordembedding	1
4.5	HierarchicalSoftmax	1
4.6	BidirectionalEncoderRepresentationsfromTransformers (BERT)	1
4.7	NaturalLanguageProcessing:withpretrainingusingword embedding	1
4.8	Casestudies-SentimentAnalysis	1
5		
5.1	Bahdanauattention	1
5.2	Multi-headattention	1
5.3	Self-attentionandpositionalencoding	1
5.4	Transformerarchitecture	1
5.5	Transformersforvision(Lecture1)	1
5.6	Transformersforvision(Lecture2)	1
5.7	Large-scalepretrainingwithtransformers(Lecture1)	1
5.8	Large-scalepretrainingwithtransformers(Lecture2)	1

ReferenceBooks

1. IanGoodfellow,YoshuaBengioandAaronCourville,DeepLearning,Secondedition, MITPress, 2016.

2. M.Gopal,DeepLearning,Pearson,2022

3. AstonZhang,ZacharyC.Lipton,MuLi,andAlexanderJ.Smola,DiveintoDeepLearning, available online at d2l.ai,

.

COURSE CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PADP206	MINIPROJECT	PROJECT	0	0	4	2

Mini project can help to strengthen the understanding of student's fundamentals throughapplicationoftheoreticalconceptsandtoboosttheirskillsandwidenthe horizonoftheirthinking.Theultimateaimofanengineeringstudentistoresolve a problem by applying theoretical knowledge. Doing more projects increases problem solving skills.

The introduction of mini projects ensures preparedness of students to undertake dissertation. Students should identify a topic of interest in consultation with PG Programme Coordinator that should lead to their dissertation/research project. Demonstrate the novelty of the project through the results and outputs.The progress of the mini project is evaluated based on three reviews, two interim reviews and a final review. A report is required at the end of the semester.

EvaluationCommittee-ProgrammeCoordinator,OneSeniorProfessorandGuide.

Sl.No	Typeofevaluations	Mark	Evaluationcriteria
1	Interimevaluation1	20	
2	Interimevaluation2	20	
3	Finalevaluationby a Committee	35	Will be evaluating the level of completion and demonstrationoffunctionality/ specifications, clarity of presentation, oral examination, work knowledge andinvolvement
4	Report	15	the committeewillbeevaluating for the technical content, adequacy of references, templatesfollowedandpermitted plagiarism level(not more than 25%)
5	Supervisor/Guide	10	
	TotalMarks	100	

.

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PADL207	DEEPLEARNINGLAB	LABORATORY 2	0	0	2	1

Preamble: This course provides a practical introduction to deep learning algorithms in Python.This course includes programming exercises in computer vision, time series, natural language processing and generative modelling. Upon completing the course, the student will acquire the skills necessary to develop applications using deep learning frameworks.

CourseOutcomes:Afterthecompletionofthecoursethe studentwillbeableto

CO#	CourseOutcomes
CO1	Implementdeeplearningtechniquestosolveproblemsin computervision (Cognitive Knowledge Level: Apply)
CO2	Implement deep learningtechniques to solve problems involving time series data (Cognitive Knowledge Level: Apply)
CO3	Implementdeeplearningtechniquestosolveproblemsintextprocessing (Cognitive Knowledge Level: Apply)
CO4	Implementdeeplearningtechniquestodevelopgenerativemodeling(Cognitive Knowledge Level: Apply)

ProgrammeOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthecourse.

PO1:Anabilitytoindependentlycarryoutresearch/investigationanddevelopment workinengineering and allied streams

PO2:Anabilitytocommunicateeffectively,writeandpresenttechnicalreportsoncomplexengineering activities by interacting with the engineering fraternity and with societyat large. PO3:An abilitytodemonstrate a degree of mastery over thearea asperthe specializationof the program.Themasteryshouldbeatalevelhigherthantherequirementsintheappropriate bachelorprogram

PO4:An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5:An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6:An ability to engage in life-long learning for the design and development related to thestream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourse outcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1		✓			✓	✓	✓
CO2		✓			✓	✓	✓
CO3		✓			✓	✓	✓
CO4		✓			✓	✓	✓

ContinuousInternalEvaluationPattern:

Thelaboratorycourseswillbehavingonly ContinuousInternalEvaluationandcarries100marks. Final assessment shall be done by two examiners; one examiner will be a senior faculty from the same department.
Continuous Evaluation : 60 marks
Finalinternalassessment:40marks

LabReport:

All the students attending the Lab should have a Fair Report. The report should contains details of experiment such as Objective, Algorithm/Design, Description, Implementation, Analysis, Results, and Outcome.The report should contain a print out of the respectivecode with inputs addressing all the aspects of the algorithm described and corresponding outputs. All the experiments noted in the fair report should be verified by the faculty regularly. The fair report, properly certified by the faculty, should be produced during the time of the final assessment.

.

Syllabus

1) Basicdeeplearningforcomputervision
2) Advanceddeeplearningforcomputervision
3) Deeplearningfortimeseries
4) Deeplearningfortext
5) Generativedeeplearning
6) Deeplearningusingtransformers

PracticeQuestions

Syllabus

1. Basicimageprocessingoperations:Histogramequalization,thresholding,edgedetection, data augmentation, morphological operations
2. ImplementSVM/SoftmaxclassifierforCIFAR-10dataset:(i)usingKNN,(ii)using3 layer neural network
3. Studytheeffectofbatchnormalizationanddropoutinneuralnetworkclassifier
4. Familiarizationofimagelabellingtoolsforobjectdetection,segmentation
5. ImagesegmentationusingMaskRCNN,UNet,SegNet
6. Objectdetectionwithsingle-stageandtwo-stagedetectors(Yolo,SSD,FRCNN,etc.)
7. ImageCaptioningwithVanillaRNNs,ImageCaptioningwithLSTMs.
8. Chatbotusingbi-directionalLSTMs
9. Implementtimeseriesforecastingusingsuitabledatasets.
10. Implementsequencetosequencelearning.

ReferenceBooks:

1. DeepLearningwithPython,byFrançoisChollet,Manning,2021

.

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PCST243	COMPUTER VISION	PROGRAMME ELECTIVE3	3	0	0	3

Preamble:Thiscourseprovidesthebasicconceptsandadvancedtechniquesincomputervision. The areas comprising the syllabus include modern CNN architectures, modern object detection architectures, and most recent architectures using transformers. On completion of this course, the studentwouldhave aninsightinto the latestdeeplearning techniques usedincomputer vision.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Expressthecomputervisionpipleline(Cognitiveknowledgelevel:Apply)
CO2	ExploreModernconvolutionalneuralnetworkarchitectures(Cognitive knowledge level:Apply)
CO3	ApplyModernobjectdetectionarchitectures(Cognitiveknowledgelevel:Apply)
CO4	ExploreGenerativeAdversarialNetworksforcomputervision(Cognitive knowledge level:Apply)
CO5	Explore recent architectures for computer vision (Cognitive knowledge level: Apply)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1:Anabilitytoindependentlycarryoutresearch/investigationanddevelopment work inengineering and allied streams

PO2:Anabilitytocommunicateeffectively,writeandpresenttechnicalreportson complexengineeringactivitiesbyinteractingwiththeengineeringfraternityand with society at large.

PO3:Anabilitytodemonstrateadegreeofmasteryovertheareaasperthe specialization of the program. The mastery should be at a level higher than the requirements in the appropriatebachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate- of-the-arttool to model, analyse and solve practical engineering problems.

PO6:Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedto thestreamrelatedproblemstakingintoconsiderationsustainability,societal,ethical and environmental aspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoproject managementandfinancewhichfocusonEntrepreneurshipandIndustryrelevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO 4	PO5	PO6	PO7
CO1			✓			✓	
CO2			✓			✓	
CO3	✓	✓	✓	✓	✓	✓	✓
CO4			✓			✓	
CO5			✓	✓			

Bloom'sCategory	EndSemesterExamination
Apply	30
Analyse	30
Evaluate	Can be evaluated through assignmentsandminiproject
Create	Can be evaluated through assignmentsandminiproject

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

Continuous Internal Evaluation Pattern:

Evaluationshallonlybebasedonapplication,analysisordesignbasedquestions(forboth internal and end semester examinations).

.

Continu ousInternalEvaluation:40:

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks
ii. Coursebasedtask/Seminar/Datacollectionandinterpretation:15marks
iii. Testpaper(1number):10marks
Testpapershallincludeminimum80%ofthesyllabus.
Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module,having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity ofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldansweranyfive.Eachquestioncancarry 7 marks Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. Listtheelementsofthecomputervisionpipeline.

2. Giveacasestudyofanytwoapplicationsofcomputervision.

3. Illustratetherelevanceofeachstageofthecomputervisionpipeline.

CourseOutcome2(CO2)

- IllustratethearchitectureofaCNN.

- ExploretheVGGNetarchitecture.

.

- ComparetheResNetarchitecturewithGoogLeNetarchitecture. :COMPUTER SCIENCE A

CourseOutcome3(CO3):

1. Outlinethegeneralobjectdetectionframework.

2. Comparevariousevaluationmetricsforobjectdetection.

3.Show theworkflow oftheYoloalgorithm.

.

ModelQuestionPaper		
QPCODE:		
Reg No: ___________________		
Name: ___________________ PAGES:4		
APJABDULKALAMTECHNOLOGICALUNIVERSITY		
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR		
CourseCode: J1PCST243		
CourseName:ComputerVision		
Max.Marks:60Duration:2.5Hours		
PARTA		
AnswerAllQuestions.EachQuestionCarries5Marks		
1.	Classifythevariouselementsofcomputervision.	
2.	ExaminethepeculiaritiesoftheMobileNetarchitecture.	
3.	Justifytheelementsinageneralobjectdetectionframework.	
4.	IllustrateanyoneapplicationofGANincomputervision.	

ModelQuestionPaper		
5.	SketchtheinnovationsintheYolov7techniqueincomparisonwitholder versions.	(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)		
6.	YouaregivenaCNNhavingatotalof10layers.Deviseawaytoapply transfer learning to this network.	(7)

7.	IllustratetheinnerarchitectureoftheresidualblockinResNetarchitecture.	(7)

8.	ComparecontrastheGoogLeNet and ResNet architectures.	(7)
9.	Usingadiagram,demonstratethedataflowintheSingleShotDetector architecture.	(7)
10.	ExpressindetailtheworkflowinaYoloarchitecturewhena13x13gridis applied on the input image.	(7)
11.	Statetheequationfortotallossinaneuralstyletransferarchitectureand justify each loss term with respect to its function in the architecture.	(7)
12.	IntheVisionTransformerarchitecture,justifytheneedforatrainable linearprojection and show the computational steps prior to this projection.	(7)

Module1IntroductiontoComputerVision(7hours)
Introduction to computer vision, Computer vision, Applications of computer vision, Computervisionpipeline,Classification,Transferlearningandvariousapproaches,Open source dataset
Module2AdvancedCNNarchitectures(9hours)
CNNarchitectureandcomponents,ImageclassificationusingCNNs,AlexNet,VGGNet, InceptionandGoogLeNet, ResNet, MobileNet
Module3Objectdetection(10hours)
Generalobjectdetectionframework,Object-detectorevaluationmetrics,Region-based convolutional neural networks, Faster R-CNN, Single-shot detector (SSD), YOLO, YOLOv3
Module4GANapplications(5hours)

.

DCGANImage-to-imagetranslation(Pix2PixGAN),Imagesuper-resolutionGAN (SRGAN),Neural style transfer, Visual embeddings
Module 5 Advanced architectures (9 hours) YoloV7,VisionTransformer,TransGAN, GPV-1,Mobile-former

.

Course Plan (For 3 credit courses, the content can be for 40 hrs and for 2 credit courses,thecontentcanbefor26hrs.Theauditcourseinthirdsemestercanhave content for 30 hours).

No	Topic	Hours
1	Introductiontocomputervision	
1.1	Computervision	1
1.2	Applicationsofcomputervision	1
1.3	Applicationsofcomputervision	1
1.4	Computervisionpipeline	1
1.5	Classification	1
1.6	Transferlearningandvariousapproaches	1
1.7	Opensourcedatasets	1
2	AdvancedCNNarchitectures	
2.1	CNNarchitectureandcomponents,	1
2.2	ImageclassificationusingCNNs	1
2.3	ImageclassificationusingCNNs	1
2.4	AlexNet	1
2.5	VGGNet	1
2.6	InceptionandGoogLeNet	1

.

2.7	InceptionandGoogLeNet	1
2.7	ResNet	
2.8	MobileNet	1
3	Objectdetection	
3.1	Generalobjectdetectionframework	1
3.2	Object-detectorevaluationmetrics	1
3.3	Region-basedconvolutionalneuralnetworks	1
3.4	FasterR-CNN	1
3.5	Single-shotdetector(SSD)	1
3.6	Single-shotdetector(SSD)	1
3.7	YOLO	1
3.8	YOLO	1
3.9	YOLOv3	1
3.10	YOLOv3	1
4	GANapplicationsincomputervision	
4.1	DCGAN	1
4.2	Image-to-imagetranslation(Pix2PixGAN)	1
4.3	Imagesuper-resolutionGAN(SRGAN)	1

4.4	Neuralstyletransfer	1
4.5	Visualembeddings	1

5	Advancedarchitectures	
5.1	YoloV7	1
5.2	YoloV7	1
5.3	VisionTransformer	1
5.4	VisionTransformer	1
5.5	TransGAN	1
5.6	TransGAN	1
5.7	GPV-1	1
5.8	Mobile-former	1
5.9	Mobile-former	1

ReferenceBooks

1. DeepLearningforVisionSystems,MohamedElgendy,Manning,
2020 Reference Papers

2. Chien-YaoWang,AlexeyBochkovskiy,Hong-YuanMarkLiao,—YOLOv7:Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,‖ https://doi.org/10.48550/arXiv.2207.02696, 2022.

3. AlexeyDosovitskiyetal,—AnImageisWorth16x16Words:TransformersforImage Recognition at Scale‖, https://doi.org/10.48550/arXiv.2010.11929, 2020.

4. YifanJiang,ShiyuChang,ZhangyangWang,—TransGAN:TwoPureTransformersCan Make One Strong GAN, and That Can Scale Up‖, https://doi.org/10.48550/arXiv.2102.07074, 2021.

5. TanmayGupta,AmitaKamath,AniruddhaKembhavi,DerekHoiem,—TowardsGeneral

.

Purpose Vision Systems‖, https://doi.org/10.48550/arXiv.2104.00743, 2021.

6. Yinpeng Chen, Xiyang Dai, Dongdong Chen, Mengchen Liu, Xiaoyi Dong, Lu Yuan, ZichengLiu;—Mobile-Former:BridgingMobileNetandTransformer,‖Proceedingsofthe IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 5270-5279.
https://openaccess.thecvf.com/content/CVPR2022/papers/Chen_MobileFormer_Bridging_MobileNet_and_Transformer_CVPR_2022_paper.pdf

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PADT213	BLOCKCHAIN TECHNOLOGIES	PROGRAMME ELECTIVE3	3	0	0	3

Preamble:The purpose of this course is to create awareness and understanding among studentson the foundation of blockchain technology. The course introduces the cryptographic principles behind blockchain and helps the students understand concepts like consensus, crypto-currency, smart contracts, use cases etc. The course enables students to develop simple decentralized applications using blockchain networks such as Ethereum.

Prerequisite:Basicknowledgeindatastructuresandoperatingsystems.

CourseOutcomes:Afterthecompletionofthecoursethe studentwillbeableto

CO1	Illustrate and implement the cryptographicbuilding blocks of blockchain technology. (Cognitive Knowledge Level: Apply)
CO2	Makeuseoftheconceptsofblockchaintechnology.(CognitiveKnowledgeLevel: Apply)
CO3	Summarizetheclassificationofconsensusalgorithms.(CognitiveKnowledgeLevel: Understand)
CO4	Illustrate theconceptsoffirstdecentralizedcryptocurrencybitcoin.(CognitiveKnowledge Level: Apply)
CO5	Implementsmartcontractsanditsuse cases.(Cognitive Knowledge Level:Apply)
CO6	DevelopsimpleapplicationsusingSoliditylanguageonEthereum platform. (Cognitive Knowledge Level: Apply)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthecourse.

PO1:Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams

PO2:Anabilitytocommunicateeffectively,writeandpresenttechnicalreportsoncomplexengineering activities by interacting with the engineering fraternity and with societyat large. PO3:An abilitytodemonstrate a degree of mastery over thearea asperthe specializationof the program.Themasteryshouldbeatalevelhigherthantherequirementsintheappropriate bachelorprogram

PO4:An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5:An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6:An ability to engage in life-long learning for the design and development related to thestream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:An ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourse outcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓		✓	✓		✓	
CO2	✓		✓	✓		✓	
CO3	✓		✓	✓		✓	
CO4	✓		✓	✓		✓	
CO5	✓		✓	✓		✓	
CO6	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	80%
Analyze	20%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesign-basedquestions(forbothinternal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15 marks

iii. Testpaper(1number) :10 marks

Testpapershallincludeminimum80%ofthesyllabus.
Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:
TheendsemesterexaminationwillbeconductedbytherespectiveCollege.
There will be two parts; Part A and Part B.
PartAwill contain5numerical/short answerquestions with1questionfromeach module,having5 marks for each question. Students should answer all questions. Part B will contain 7 questions (shallbeusefulinthetestingofoverallachievementandmaturityofthestudentsinacourse,through long answer questions relating totheoretical/practical knowledge, derivations, problem solving and quantitativeevaluation), with minimum one question from each module of which studentshould answer any five. Each question can carry 7 marks
Totaldurationoftheexaminationwillbe150minutes.
Note: Themarks obtained forthe ESEforan elective courseshall not exceed 20%overthe average ESE mark %forthecore courses. ESE marks awarded to astudent foreach elective courseshall be normalized accordingly.
For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions Course Outcome 1 (CO1):

1. DistinguishbetweenSymmetriccryptographyandasymmetriccryptography.
2. ExplaintheworkingofAES algorithm.

CourseOutcome2(CO2):

1. Categorizeconsensusmechanismusedinblockchain.
2. DefineBlockchain.Explainhowdecentralizationofcomputingorprocessingpoweris achieved by a blockchain.

CourseOutcome3(CO3):

1. IllustratehowProofofStakecanachieveconsensusamongpeers.
2. ExplaintheworkingofRaftprotocol.

CourseOutcome4(CO4):

1. Describetheuseofgenesisblock.
2. Implementtheminingalgorithmusedinbitcoin.

CourseOutcome5(CO5):

1. Illustratehowblockchaintechnologycanbeusedinsupplychainmanagement.
2. Whatareoraclesinablockchainecosystem?Explainthegenericdataflowfromasmart contract to an oracle.

CourseOutcome6(CO6):

1. Developasmartcontractforvotingprocess.Inthisapplication,delegatedvotingisallowe dandthecountingisautomaticandcompletelytransparentatthesametime.
2. Developasmartcontractforauctionprocess.Thecontractshouldbeablindauctionwhere it is not possible to see the actual bid until the bidding period ends.

ModelQuestionPaper			
QPCODE:			
Reg No: _________________			
Name:____ PAGES:4			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode :J1PADT213			
CourseName: BLOCKCHAIN TECHNOLOGIES			
Max.Marks:60 Duration:2.5Hours			
PARTA			
AnswerAllQuestions. EachQuestionCarries5 Marks			
1.	ExplainhowhashfunctionsareusedtobuildMerkletreesinblockchain.		
2.	Explainthebenefits,featuresandlimitationsofblockchain.		
3.	WhatistheroleofaBitcoinminer?ExplaintheminingalgorithmusedinBitcoin with the help of a flowchart.		
4.	Explainthedesignprocessofdecentralizedapplicationswithdiagrams.		
5.	Defineblockdifficulty.ExplainhowblockdifficultyisadjustedinEthereum blockchain network.		(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			
6.	(a)	ExplainthedesignofSHA-256anditscompressionfunctionusinga diagram.	(7)
7.	(a)	ExplaintheconceptofGasinEthereum.Explainhowtransactioncostcanbe calculated in an Ethereum blockchain network.	(7)

ModelQuestionPaper			
8.	(a)	Explainconsensusmechanismsusedinblockchain.Listoutanysix consensus algorithms used in the context of blockchain.	(7)
9.	(a)	ShowhowPracticalByzantineFaultTolerancecanachieveconsensusinthe presence of Byzantine faults.	(7)
10	(a)	Illustratehowblockchaintechnologycanbeimplementedinfinancesector.	(7)
11	(a)	UsingSoliditylanguage,createasimplebankcontractthatallowsauserto deposit, withdraw and view balance.	(7)

Syllabus

Module – 1 (Fundamentals of Cryptography)

Introduction to Cryptography, Symmetric cryptography – AES. Asymmetric cryptography – RSA. Elliptic curve cryptography, Digital signatures – RSA digital signature algorithms. Secure Hash Algorithms – SHA-256. Applications of cryptographic hash functions – Merkle trees, Distributed hash tables.

Module–2(FundamentalsofBlockchainTechnology)

Blockchain – Definition, architecture, elements of blockchain, benefits and limitations, types of blockchain. Consensus – definition, types, consensus in blockchain.

Decentralization – Decentralization using blockchain, Methods of decentralization, Routes to decentralization, Blockchain and full ecosystem decentralization.

Module-3(ConsensusAlgorithmsandBitcoin)

Consensus Algorithms, Crash fault-tolerance (CFT) algorithms – Paxos, Raft. Byzantine fault-tolerance(BFT)algorithms–PracticalByzantineFault Tolerance(PBFT), Proofofwork(PoW), Proof of stake (PoS), Types of PoS.

Bitcoin–Definition,Cryptographickeys–Privatekeys,publickeys,addresses.Transactions–Lifecycle,coinbasetransactions,transactionvalidation.Blockchain–Thegenesisblock.

Mining – Tasks of miners, mining algorithm, hash rate. Wallets – Types of wallets.

Module-4(SmartContractsandUsecases)

Smart Contracts – Definition, Smart contract templates, Oracles, Types of oracles, Deploying smart contracts. Decentralization terminology – Decentralized applications, Decentralized Autonomous Organizations.

Use cases of Blockchain technology – Government, Health care, Finance, Supply chain management.

Blockchain and allied technologies – Blockchain and Cloud Computing, Blockchain and Artificial Intelligence.

Module-5(EthereumandSolidity)

Ethereum – The Ethereum network. Components of the Ethereum ecosystem – Keys and addresses, Accounts, Transactions and messages. The Ethereum Virtual Machine, Blocks and blockchain.

The Solidity language – The layout of a Solidity source code, Structure of a smart contract, variables, data types, control structures, events, inheritance, libraries, functions, error handling.

SmartcontractsCasestudy:Voting,Auction.

TextBook

1. Imran Bashir, Mastering Blockchain: A deep dive into distributed ledgers, consensus protocols, smart contracts, DApps, cryptocurrencies, Ethereum, and more, Packt Publishing, Third edition, 2020.

References

2. Ritesh Modi, Solidity Programming Essentials: A beginner's guide to build smart contracts for Ethereum and blockchain, Packt Publishing, First edition, 2018.
3. Kumar Saurabh, Ashutosh Saxena, Blockchain Technology: Concepts and Applications,First Edition, Wiley Publications, First edition, 2020.
4. Chandramouli Subramanian, Asha A George, et al, Blockchain Technology, Universities Press (India) Pvt. Ltd, First edition, August 2020.
5. Lorne Lantz, Daniel Cawrey, Mastering Blockchain: Unlocking the Power of Cryptocurrencies, Smart Contracts, and Decentralized Applications, O'Reilly Media, First edition, 2020.
6. Andreas M. Antonopoulos, Gavin Wood, Mastering Ethereum: Building Smart Contractsand DApps, O'Reilly Media, First edition, 2018.

TeachingPlan

No	Contents	No. of Lecture Hours (35hours)
Module-1(FundamentalsofCryptography)(7hours)		
1.1	Introductiontocryptography	1hour
1.2	Symmetriccryptography,AES	1hour
1.3	Asymmetriccryptography,RSA	1hour
1.4	Ellipticcurvecryptography	1hour
1.5	Digitalsignatures–RSAdigitalsignaturealgorithm	1hour
1.6	SecureHashAlgorithms–SHA-256	1hour
1.7	Applicationsofcryptographichashfunctions–Merkletrees,Distributed hash tables	1hour
Module-2(FundamentalsofBlockchainTechnology)(6hours)		
2.1	Blockchain–definitionandarchitecture	1hour
2.2	Elementsofblockchain.	1hour
2.3	Blockchain–benefitsandlimitations,types.	1hour
2.4	Consensus–definition,types,consensusinblockchain	1hour
2.5	Decentralizationusingblockchain,Methodsofdecentralization	1hour
2.6	Routestodecentralization,Blockchainandfullecosystemdecentralization	1hour
Module-3(ConsensusAlgorithmsandBitcoin)(7hours)		
3.1	ConsensusAlgorithms–Crashfault-tolerance(CFT)algorithms–Paxos, Raft (working is expected).	1hour
3.2	Byzantinefault-tolerance(BFT)algorithms–PracticalByzantineFault Tolerance (PBFT) (working is expected).	1hour
3.3	Proofofwork(PoW),Proofofstake(PoS),TypesofPoS	1hour

3.4	Bitcoin–Definition,Cryptographickeys–Privatekeys,publickeys, addresses.	1hour
3.5	Transactions–Lifecycle,coinbasetransactions,transactionvalidation	1hour
3.6	Blockchain–Thegenesisblock.Mining–Tasksofminers	1hour
3.7	Mining–miningalgorithm,hashrate.Wallets–Typesofwallets.	1hour
Module-4(SmartContractsandUsecases) (6hours)		
4.1	SmartContracts–Definition,Smartcontracttemplates	1hour
4.2	Oracles,Typesoforacles,Deployingsmartcontracts.	1hour
4.3	Decentralizationterminology–Decentralizedapplications,Decentralized Autonomous Organizations.	1hour
4.4	UsecasesofBlockchaintechnology–Government,Healthcare.	1hour
4.5	UsecasesofBlockchaintechnology–Finance,Supplychainmanagement.	1hour
4.6	BlockchainandAlliedTechnologies–BlockchainandCloudComputing, BlockchainandArtificialIntelligence.	1hour
Module-5(EthereumandSolidity)(9hours)		
5.1	Ethereum-TheEthereumnetwork, ComponentsoftheEthereumecosystem–Keysandaddresses,Accounts	1hour
5.2	ComponentsoftheEthereumecosystem–Transactionsandmessages	1hour
5.3	TheEthereumVirtualMachine	1hour
5.4	EthereumBlocksandblockchain	1hour
5.5	TheSoliditylanguage–ThelayoutofaSoliditysourcecode,Structureofa smart contract, variables, data types	1hour
5.6	TheSoliditylanguage–controlstructures,events,inheritance,libraries	1hour
5.7	TheSoliditylanguage–functions,errorhandling.	1hour
5.8	SmartcontractsCasestudy:Voting.	1hour
5.9	SmartcontractsCasestudy:Auction.	1hour

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PADT223	IMAGE AND VIDEO ANALYTICS	PROGRAMME ELECTIVE3	3	0	0	3

Preamble: This course introduces the salient features and techniques used in soft computing. Thiscoursecoversthevariousprinciplesbehindprocessingofimageandvideo,their analytics and their applications. On completion of the course, students will be able to acquire key skills required for solving real-life problems on image and video analytics.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Interpretthemathematicalprinciplesindigitalimageenhancementandapplythemin spatialdomainandfrequencydomain(Cognitiveknowledgelevel:Apply)
CO2	Applyvariousmethodsforimagefilteringandsegmentation(Cognitiveknowledge level:Apply)
CO3	Make use of various video enhancement and noise reduction techniques (Cognitive knowledge level: Apply)
CO4	Applyvariousfeatureextractionandpatternclassificationtechniqueson images.(Cognitiveknowledgelevel:Apply)
CO5	Analysevariousobjectdetectionandrecognitiontechniquesonimageand video.(Cognitiveknowledgelevel:Analyse)
CO6	Design, develop, implement and present solutions to simple real-life problems with popularopen source library using image and video analytical techniques (Cognitive knowledgelevel:Apply)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1:Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams

PO2:Anability tocommunicateeffectively,writeandpresenttechnicalreportsoncomplex engineering activities by interacting with the engineering fraternity andwith society at large.

PO3:An ability to demonstrate a degree of mastery overthe areaas per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anability toidentify,selectandapply appropriate techniques,resourcesandstate-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project managementand finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO 4	PO5	PO6	PO7
CO1	✓		✓	✓		✓	
CO2	✓		✓	✓		✓	
CO3	✓		✓	✓		✓	
CO4	✓		✓	✓	✓	✓	
CO5	✓		✓	✓	✓	✓	
CO6	✓	✓	✓	✓	✓	✓	

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	40
Analyse	20
Evaluate	Canbeevaluatedusing Course based task/Seminar/Data collectionandinterpretation/Assignments
Create	Canbeevaluatedusing Course based task/Seminar/Data collectionandinterpretation/Assignments

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15marks

iii. Testpaper(1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.
Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldanswerany five.Eachquestioncancarry 7 marks
Totaldurationoftheexaminationwillbe150minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. ApplyopeningandclosingoperationontheimagesampleAgivenbelowwith

structuring element B

A= B= [111]

1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

2. WhatisGammacorrectionandillustrateitsneed.

3. VerifywhethertheDFTmatrixisunitaryornotforN=4.

CourseOutcome2(CO2)

1. DifferentiateAveragingFilterandWeightedAveragingFilter,withanexample.

2. Illustratewithanexamplethattheinitialthresholdinthebasicglobalthreshoding algorithm should lie between minimum and maximum values in the image.

3. IllustrateUnsharpMaskingandHighBoostFiltering.

CourseOutcome3(CO3):

1. For the same line number per frame, what is the relation between the maximum temporalfrequencythataprogressiverastercanhaveandthatofaninterlacedraster which divides each frame into two fields? Also, give the relation between the maximum vertical frequencies.
2. Compareandcontrastthefeaturesofcomponentandcompositeformat?
3. Illustrate the Video Production, Transmission, and Reception in an Analog Color TV systems, with a block diagram.

CourseOutcome4(CO4):

1. Illustratesolvingthefacerecognitionproblemusingk-NNclassifier.

2. Illustratethestepsinrulebasedimageunderstanding,withanexample.

3. GivecasestudiesontheapplicationsofContentBasedImageRetrieval(CBIR).

CourseOutcome5(CO5):

1. Illustratetexturebasedobjectrecognitionperformed?

1. IllustratetheExhaustiveSearchBlockMatchingAlgorithm(EBMA)withanexample.

CourseOutcome6(CO6):

1. Preparereviewarticlebasedonpeerreviewedoriginalpublications(referringminimum 10 publications) about the recent image noise filtering techniques.

2. CourseprojecttodesignanddevelopaBirdsSpeciesClassificationSystemusingSupport Vector Machine Classifier.

ModelQuestionPaper		
QPCODE:		
Reg No: ___________________		
Name: ___________________ PAGES:4		
APJABDULKALAMTECHNOLOGICALUNIVERSITY		
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR		
CourseCode: J1PADT223		
CourseName:IMAGEANDVIDEOANALYTICS		
Max.Marks:60 Duration:2.5 Hours		
PARTA		
AnswerAllQuestions.EachQuestionCarries5Marks		
1.	DifferentiateBit-planeslicingandContraststretching.	
2.	WhatistheConvolutionpropertyof2DDFT?	
3.	DifferentiateProgressiveandInterlacedscaninrastersystems.	
4.	WhatistheroleofPCAinapatternrecognitionproblem?	
5.	Comparepixelbasedandblockbased2Dmotionrepresentationmethods.	(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)		

6.	(a)	Findthe4orderHadamardTransformforthefollowingimagesegment: 1 1 2 2 2 1 1 3 3 2 1 2 1 2 2 3	(4)
	(b)	Demonstratethestepsofgeneratingdigitalimagesfromthesenseddata. Findoutthenumberofbitsneededtorepresent512Graylevels.	(3)
7.	(a)	DerivetwopointDFT.	(4)
	(b)	Illustratethesignificanceofnegativeandlogarithmictransformations.	(3)
8.	(a)	Fortheimagegivenbelow,applyhistogramequalizationtoachieveimage enhancement. 4 4 4 3 4 f(x,y)= 3 4 5 4 3 3 5 5 5 3 3 4 5 3 3 4 4 4 4 4	(4)
	(b)	Whatarethevarioussmoothingnon-linearfiltersinthespatialdomain? Which is the best among them and why?	(3)

9.	(a)	a) For the image given below, apply Prewitt filter both in X and Y direction and find the resultant image. Also, find the magnitude and angle of gradientat the two underlined pixel positions. $f(x,y)= \begin{matrix} 2 & 0 & 1 & 3 & 2 \\ 1 & \underline{4} & 7 & 0 & 3 \\ 5 & 4 & \underline{3} & 5 & 7 \\ 1 & 2 & 0 & 4 & 4 \end{matrix}$	(4)
		Illustratehowsharpeningcanbedoneinthefrequencydomainusing Butterworth high-pass filters.	(3)
10.	(a)	Illustratetheprocessforformingacompositecolorvideosignal.How should youselectthecolorsub-carrierfrequencyandaudiosub-carrier frequency ?	(4)
	(b)	Whatistheperceivedcolorifyoumixred,green,andbluedyesinequal proportion?Whatistheresultifyoumixredandgreendyesonly ?	(3)
11.	(a)	IllustratethestepsofSupportVectorMachinesinsolvingatwo-class object recognition problem.	(5)
		ComparethefeaturesofSIFTandSURFfeatureextractiontechniques	(2)
12.	(a)	IllustratethemainstepsinAutomaticTrafficMonitoringsystemusing object detection and recognition techniques.	(4)
		HowisPixel-BasedMotionEstimationperformedusingMultipoint Neighborhood?	(3)

	Syllabus	
Module	Content	Hours
I	FundamentalsofImageProcessing&ImageEnhancement:StepsinImageProcessingSystems,Digitalimagerepresentation-, Sampling and Quantization, Pixel Relationships, Image Operations – Arithmetic, Geometric & Morphological operations, Colour Models, ImageEnhancementinSpatialDomain–Transformations–Negative,Logarithmic,Gamma,ContrastStretching,Greylevel& Bit Plane Slicing. Image Transforms - DFT, DCT, Hadamard Transforms	7
II	HistogramProcessing,Filtering:HistogramProcessing- Histogram Equalisation. Spatial correlation and convolution, Spatial filtering- Smoothing and Sharpening spatial filters,Basics of filteringinfrequencydomain-Smoothingandsharpeninginfrequency domain. Image Segmentation-Fundamentals, Thresholding. Edge Detection-Point,LineandEdgeDetection,EdgeDetection operators.	7
III	Video Processing : Video Formation, Perception and Representation: Color Perception and Specification-Human Perception of Color, The Trichromatic Theory of Color Mixture, Color Specification by Tristimulus Values, Color Specification by Luminance and Chrominance Attributes. Video Capture and Display-Principles ofColor Video Imaging, Video Cameras, Video Display, Composite versus Component Video, Gamma Correction, Analog Video Raster- Progressive and Interlaced Scan, Characterization of a Video Raster. Analog Color Television Systems- Spatial and Temporal Resolution, Color Coordinate, Signal Bandwidth, Multiplexing of Luminance, Chrominance, and Audio, Analog Video Recording. Digital video-ITU-R BT.601 Digital Video, Other Digital Video Formats and Applications.	8
IV	Image Analytics : Feature Extraction - Binary object feature,Histogram based (Statistical) Features, PCA - SIFT – SURF. Visual Pattern Recognition-Patterns and Pattern Classes- Statistical Pattern Classification Techniques- k-Nearest Neighbours Classifier, Support VectorMachines.ImageUnderstanding,ContentBasedImage	7

	Retrieval.	
V	VideoAnalytics&ImageandVideoAnalyticsApplications:Object detection and recognition in video-Texture models, Video classification models- Object tracking in Video – Applications.. Two Dimensional Motion Estimation: Pixel Based Motion Estimation-BlockMatchingAlgorithm.Implementationexamplesofsimple Image and Video processing problems using TensorFlow and Keras. Case Study : Face Detection and Recognition, Automatic Traffic Monitoring.	7

CoursePlan

No	Topic	No. of Lectures (36)
1	Module1(FundamentalsofImageProcessing&ImageEnhancement)7Hours	
1.1	StepsinImageProcessingSystems,Digitalimagerepresentation	1
1.2	ImageOperations–Arithmetic,Geometric,Morphologicaloperations	1
1.3	ColourModels	1
1.4	ImageEnhancementinSpatialDomain–Transformations–Negative, Logarithmic	1
1.5	Gamma,.ContrastStretching,Greylevel&BitPlaneSlicing	1
1.6	ImageTransforms-DFT	1
1.7	DCT,HadamardTransforms	1
2	Module2(HistogramProcessing,Filtering)7Hours	
2.1	HistogramProcessing-HistogramEqualisation.	1
2.2	Spatialcorrelationandconvolution,	1
2.3	Spatialfiltering-Smoothing,Sharpeningspatialfilters	1
2.4	Basicsoffilteringinfrequencydomain,Smoothinginfrequencydomain	1
2.5	Sharpeninginfrequencydomain	1
2.6	ImageSegmentation-Fundamentals,Thresholding	1
2.7	EdgeDetection-Point,LineandEdgeDetection,EdgeDetection operators.	1
3	Module3(Video Processing)8Hours	
3.1	VideoFormation,PerceptionandRepresentation:ColorPerceptionand Specification-HumanPerceptionofColor,TheTrichromaticTheory of	1

	ColorMixture	
3.2	ColorSpecificationbyTristimulusValues,ColorSpecificationby Luminance and Chrominance Attributes.	1
3.3	VideoCaptureandDisplay-PrinciplesofColorVideoImaging,Video Cameras, Video Display	1
3.4	CompositeversusComponentVideo,GammaCorrection,	1
3.5	AnalogVideoRaster-ProgressiveandInterlacedScan,Characterization of a Video Raster.	1
3.6	AnalogColorTelevisionSystems-SpatialandTemporalResolution, Color Coordinate, Signal Bandwidth, Multiplexing of Luminance, Chrominance, and Audio	1
3.7	AnalogVideoRecording.	1
3.8	Digitalvideo-ITU-RBT.601DigitalVideo,OtherDigitalVideoFormats and Applications.	1
4	Module 4(ImageAnalytics)7Hours	
4.1	FeatureExtraction,Binaryobjectfeature	1
4.2	Histogrambased(Statistical)Features,PCA	1
4.3	SIFT,SURF	1
4.4	VisualPatternRecognition,PatternsandPatternClasses	1
4.5	StatisticalPatternClassificationTechniques-k-NearestNeighbours Classifier	1
4.6	SupportVectorMachines	1
4.7	ImageUnderstanding,ContentBasedImageRetrieval	1
5	Module5(VideoAnalytics&ImageandVideoAnalyticsApplications)7 Hours	
5.1	Objectdetectionandrecognitioninvideo,Texturemodels	1
5.2	Videoclassificationmodels,ObjecttrackinginVideo,Applications	1
5.3	TwoDimensionalMotionEstimation	1

5.4	PixelBasedMotionEstimation	1
5.5	BlockMatchingAlgorithm	1
5.6	ImplementationexamplesofsimpleImageandVideoprocessing problems using TensorFlow and Keras.	1
5.7	CaseStudy:FaceDetectionandRecognition,AutomaticTraffic Monitoring	1

ReferenceBooks

1. RafaelC.GonzalezandRichardE.Woods,—DigitalImageProcessing‖,ThirdEdition, PearsonEducation,2009.
2. AnilK.Jain,—FundamentalsofDigitalImageProcessing‖,PearsonEducation,2003
3. MilanSonka,VaclavHlavacandRogerBoyle,—ImageProcessing,Analysisand MachineVision‖,ThirdEdition,CengageLearning,2007.
4. Yao Wang, Jorn Ostermann and Ya-Qin Zhang, —Video Processing and Communications‖,PrenticeHall,2001.
5. SimonJ.D.Prince,—ComputerVision:Models,Learning,andInference‖,Cambridge UniversityPress
6. CaifengShan,FatihPorikli,TaoXiang,ShaogangGong,—VideoAnalyticsfor BusinessIntelligence‖,Springer,2012.
7. HimanshuSingh,—PracticalMachineLearningandImageProcessing‖,APress,2019
8. VesnaZeljkovic,—VideoSurveillanceTechniquesandTechnologies‖,IGIGlobal

.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PADT233	INTERNET OF THINGS	PROGRAMME ELECTIVE3	3	0	0	3

Preamble:Thiscourse intendstoprovideinsightintonewinnovationsthatwillbuild novel type of interactions among things and humans, and enables the realization of smart cities, infrastructures, and services for enhancing the quality of life and utilization of resources. Knowledge about IOT and its related concepts, different IOT architectures and their components, emerging paradigms such as Fog computing, Platforms and solutions supporting developmentanddeploymentofIOTapplications,messagepassingmechanismssuchas RPC,REST,andCoAP,dataandknowledgemanagement,dataconfidentiality,data integrity, and operationcontrol issues faced byIOT are included in the course.This coursehelps the students to develop IOT based applications.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	MakeuseoftheconceptsandfeaturesoftheIOTparadigm.(CognitiveKnowledge Level:Comprehension)
CO2	AnalyseFogcomputing,TinyOS-nesCandprogrammingframeworksfor IOT(CognitiveKnowledgeLevel:Analze)
CO3	AnalysethedatamanagementtechniquesappliedtotheIOTenvironment.(Cognitive Knowledge Level:Analyze)
CO4	Analysesecurity,andprivacyinIOTenvironments.(CognitiveKnowledgeLevel: Apply)
CO5	AnalysekeyenablersandsolutionstoenablepracticalIoTsystems.(Cognitive Knowledge Level:Apply)
CO6	DesignofmicroprojectsforSmartGasLeakageDetector/NightPatrolathome (CognitiveKnowledgeLevel: Apply)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1:Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization ofthe program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anabilitytoidentify,selectandapplyappropriate techniques,resources and state-of-the-arttooltomodel,analyseandsolvepracticalengineeringproblems.

PO6:Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedtothe stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO 4	PO5	PO6	PO7
CO1	√		√			√	
CO2	√		√	√	√	√	
CO3	√		√			√	
CO4	√	√		√		√	
CO5	√		√		√	√	√
CO6	√		√			√	

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	50%
Analyse	50%

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

i. Preparingareviewarticlebasedonpeer-reviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15marks

iii. Testpaper(1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.

Coursebasedtask/testpaperquestions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldanswerany five.Eachquestioncancarry 7 marks

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. —IOTistheframeworkthatenablescollaborationbetweensmartmobiledevicesand cloud.‖ Justify.

2, Explain the state diagram of the open IOT services life cycle with a case study.

Course Outcome 2 (CO2)

1. Justify the need on the four broad requirements that motivate the design of TinyOS.

Course Outcome 3(CO3):

1. CompareStreamManagementSystem(DSMS)andComplexEventProcessing(CEP).

CourseOutcome4(CO4):

1. IllustratetheerrordetectiontechniqueswhichareapplicableinanIoTenvironment.

CourseOutcome5(CO5):

1. WhatsTWI?HowtoConfiguretheTWIforI2CCommunication?

.

ModelQuestionPaper :COMPUTERSCIENCEANDENGINEERING-CS2			
QPCODE:			
Reg No: ___________________			
Name: _____________________ PAGES:4			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode: J1PADT233			
CourseName::INTERNETOFTHINGS			
Max.Marks:60 Duration:2.5 Hours			
PARTA			
AnswerAllQuestions.EachQuestionCarries5Marks			
1.	—IOTistheframeworkthatenablescollaborationbetweensmartmobile devices and cloud.‖ Justify.		
2.	ExplainthemajorchallengesfacedintheFogparadigm.		
3.	Howareanomalydetectionandcategorizationofanomaliesindata performed?		
4.	Discusstheadvantagesofobfuscationanddiversificationtechniques.		
5.	WhatsTWI?HowtoConfiguretheTWIforI2CCommunication?		(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			
6.		ExplainthestatediagramoftheopenIOTserviceslifecycle.	(7)
7.	(a)	Justifytheneedonthefourbroadrequirementsthatmotivatethedesignof TinyOS.	(4)
	(b)	ExplainthedesigndecisionsfornesC.	(3)
8.		CompareStreamManagementSystem(DSMS)andComplexEventProcessing (CEP).	(7)
9.		Explaintheerrordetectiontechniqueswhichareapplicableinan IoT environment.	(7)

.

ModelQuestionPaper		:COMPUTERSCIENCEANDENGINEERING-CS2	
10.		DeveloptheStation-to-Stationprotocol(STS)andIdentifythetwomain shortcomings of STS.	(7)
11.		Suggestthesensorsrequiredtobuildtheenvironmental-sensingIoT gateway device for weather monitoring.	(7)
12.		Developthestepsforthedevelopmentofasensorproject.	(7)

Module1:Introduction(9Hours) Overview of Internet of Things: Open-source semantic web infrastructure for managing IOT resourcesintheCloud-Device/CloudCollaborationframeworkforintelligenceapplications.
Module2:Programmingframeworks(11Hours) IntroductiontoFogComputing:principles,architectures,andapplications.TinyOS–NesC, Programming frameworks for Internet of Things
Module3:Datamanagementtechniques(8Hours) StreamprocessinginIoT:foundations,state-of-the-art,andfuturedirections-Aframeworkfor distributed data analysis for IoT
Module4:Securityandprivacy (9Hours) Security and privacy in theInternet of Things-Internet of Things -robustness and reliability. TinyTO:two-wayauthenticationforconstraineddevicesintheInternetofThings -Obfuscationand diversification for securing the Internet of Things
Module 5: IoT Implementation(8 Hours) Creatinga simple IoT project - PreparingRaspberryPi – Interfacing the hardware - Internal representation of sensor values- Persisting data - Creating the actuator project - Creating a controller.

.

CoursePlan :COMPUTERSCIENCEANDENGINEERING-CS2

No	Topic	No.of Lectures
1	Introduction(7Hours)	
1.1	Internetofthings-definition,evolution.Applications-Smart homeapplications,Healthcare,Eldercare,Trafficsurveillance.	1
1.2	SOA-BasedArchitecture,APIorientedArchitecture,Resource Management.ComputationalOffloading,	1
1.3	Identification and Resource/Service Discovery, IOT Data Management and Analytics, IOT and the CLOUD	1
1.4	Open IOT architecture for IOT/Cloud convergence, Sensor middleware, Cloud computing infrastructure, Directory service, GlobalScheduler,LocalSchedulercomponent,Servicedelivery andutilitymanagerWorkflowofopenIOTplatform	1
1.5	SchedulingprocessandIOTServices,lifecycle,Statediagramof theOpenIOTServiceslifecyclewithintheschedulermodule.Scheduling and resource management	1
1.6	Resource optimization schemes, Caching techniqueServicecreation flowchart, Comparison of cost - with cache server and public cloud data-score	1
1.7	Runtimeadaptationengine,Device/cloudcollaborationframework applicationsofdevice	1
1.8	cloudcollaboration,SemanticQAcache	
2	Programmingframeworks(9Hours)	
2.1	IntroductiontoFogComputing:principles,architectures,and ApplicationsMotivatingscenarioforFogComputing,	1
2.2	AdvantagesofFogComputing,ReferencearchitectureofFog ComputingSoftware-DefinedResourcemanagementlayer,	1
2.3	Services of Software-Defined Resource management layer, ApplicationsofFogComputing.	1
2.4	HistoryofTinyOS,Implementation,Requirementsmotivatingthe design of TinyOS, Component Model, Interfaces. TinyOS computational conceptsOverview of TinyOS Execution Model	1
2.5	Concurrency,TinyOSTheoryofExecution:Events&Tasks,TinyOS Architecture. TinyOS-Programming Model.	1
2.6	nesCdesign,ComponentImplementation,DesignDecisionsfor nesC,ModuleComponents,ConfigurationComponentsWhole-ProgramAnalysis	1
2.7	DetectingRaceConditions,DealingwithRaceConditions,Issues for nesC. Overview of Embedded Programming Languages- nesC, Keil C, Dynamic C, B#	1

2.8	MessagePassinginDevices -Remote Procedure Call (RPC), LightweightRPC(LRPC)Representationalstatetransfer(REST), Computational REST (CREST), Constrained Application Protocol(CoAP), Comparison of HTTP and CoAP	1
2.9	Advantages of CoAP Coordination Languages- Orchestration, Choreography, Linda and eLinda, Orc, Features of Orc, Java OrchestrationLanguageInterpreterEngine(Jolie),Polyglot Programming	1
3	Datamanagementtechniques(6Hours)	
3.1	Stream,StreamProcessing,DataStreamManagementSystem (DSMS)	1
3.2	differencesbetweentwousecasesofStreamProcessing:Complex EventProcessing(CEP)	1
3.3	DSMSandCEPThecharacteristicsofstreamdatainIOT	1
3.4	general architecture of a stream-processing system in IOT Continuous logic processing system, challenges in streamprocessingsystems.Anomalydetection,problemstatementand definitions.	1
3.5	Hyperellipsoidalanomalydetection.	1
3.6	Distributedanomalydetection.	1
4	Securityandprivacy(8Hours)	
4.1	IOT security threats, IOT security requirements, security frameworksforIOT,IOTsecurityoverview,IOTgatewaysand security	1
4.2	IOTroutingattacks.SecurityframeworksforIOT-Lightweight cryptography,asymmetricLWCalgorithms,privacyinIOTnetworks	1
4.3	IOTcharacteristicsandreliability issues,reliabilitychallenges.	1
4.4	Addressingreliability,securityaspectsandsolutions	1
4.5	TinyTO:Two-wayauthenticationforconstraineddevicesinthe InternetofThings.	1
4.6	TinyTOprotocol.BCKwithpre-sharedkeysforTinyTO, handshakeimplementation	1
4.4	IOTnetworkstackandaccessprotocols,Obfuscationand diversificationtechniques.	1
4.5	Enhancing the security in IOT using obfuscation and diversificationtechniques,motivationsandlimitations,different use-casescenariosonsoftwarediversificationandobfuscation	1
5	IoTImplementation(7Hours)	
5.1	ThreekeycomponentstoanIOTarchitecture,Sensortogateway communication	1
5.2	wiredgatewayinterfaces,wirelessgatewayinterfaces.Sensors-sensorsrequiredtobuildtheenvironmental	1
5.3	sensingIOTgatewaydeviceforweathermonitoring.Gateway,	1

	Gatewayhardware,Gatewaysoftware.	
5.4	Datatransmission-advancedmessagequeuingprotocol,backend processing,toCLOUDornottocloud	1
5.5	Creatingasimplesensorproject-PreparingRaspberryPi–Claysterlibraries.Hardware,Interfacingthehardware	1
5.6	Internalrepresentationofsensorvalues-Persistingdata.External representationofsensorvalues,Exportingsensordata.	1

TextBook

1. RajkumarBuyya;AmirVahidDastjerdi,―InternetofThings‖,MorganKaufmann, 2016

ReferenceBooks

1. PeterWaher,―LearningInternetofThings‖,PacktPublishing,2015

2. S. SitharamaIyengar; Nandan Parameswaran; Vir V. Phoha; N. Balakrishnan; Chuka Okoye,―FundamentalsofSensorNetworkProgramming:ApplicationsandTechnology‖, Wiley, December 14, 2010

3. RobertStackowiak,ArtLicht,VenuMantha,LouisNagode,―BigDataandTheInternetof Things: Enterprise Information Architecture for A New Age‖, Apress, 2015

WebResources

1. https://www.coursera.org/specializations/internet-of-things

2. http://web.mit.edu/professional/digital-programs/courses/IoT

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PADT243	STATISTICAL NLP	PROGRAMME ELECTIVE3	3	0	0	3

Preamble:Thiscoursehelpstofamiliarizewiththebasicsofstatisticaltechniquesfor processinghumanlanguagesandlargetextcorpora.ItcoversWordSenseDisambiguation, Part-of-Speech Tagging, Machine Translation etc.At the end of the course, the learner will be able to develop natural language-based applications and frameworks.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Apply the concepts of Probability theory and linguistic essentials for processing large text corpora. (Cognitive Knowledge Level: Apply)
CO2	ApplytheconceptsofWordSenseDisambiguationforprocessinghumanlanguages (CognitiveKnowledge Level:Apply)
CO3	Apply the concepts of Part of speech tagging and Probabilistic Context free Grammarforprocessingnaturallanguages(CognitiveKnowledgeLevel:Apply)
CO4	AnalysetheconceptsofMachineTranslationforhumanlanguages(Cognitive Knowledge Level:Analyse)
CO5	Apply the concepts of Clustering and Text Categorization for natural language processing to solve real-life problems.(CognitiveKnowledgeLevel:Apply)

ProgramOutcomes(PO)

Studentswillbeabletodemonstratethefollowingattributesaftercompletingthecourse.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3:Anabilitytodemonstrateadegreeofmasteryovertheareaasperthespecializationof the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedtothe stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: Ability to develop cognitive load management skills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1				✔	✔	✔	
CO2	✔			✔	✔	✔	
CO3	✔		✔	✔	✔	✔	
CO4	✔		✔	✔	✔	✔	✔
CO5	✔			✔	✔	✔	

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	40
Analyse	20
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesignbasedquestions(forboth internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred): 15 marks

ii. Course-basedtask/Seminar/Datacollectionandinterpretation :15marks

iii. Testpaper(1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.

Course-basedtask/testpaperquestionsshallbeusefulinthetestingofknowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module,having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity ofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldansweranyfive.Eachquestioncancarry 7 marks Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. It is determined that in a text 25 of every 100 verbs and 600 of every 1,000 nouns are inflected. If the number of nouns on a particular page is four times more than that of verbs, calculate the probability of

   a.Awordwithoutinflectionbeingrandomlyselected.

   b,Anounwithinflectionbeingrandomlyselected.

   C.Averbwithoutinflectionisrandomlyselected.

2. Analysethedierentissuesinvolvedinprocessingacorpus?Describe toresolve them.

.

CourseOutcome2(CO2)

1. Constructphrasestructuregrammarofthefollowingsentences.

   a) Iprefermorningflight.

   b) Showmethelastflighttoleave.

2. Explain ip opalgorithmfordisambiguationandimplementthepartitionbasedon the given data.

   Translation{t1,...,tm}:{take,make,rise,speak}Indicators{x1,

   .... ,xn}:{mesure,note,exemple,decision,parole}

CourseOutcome3(CO3):

1. IsitpossibletouseHMMforPOStagging?Justifyyouranswer.

2. Illustratethetaggingandpartialparsinginquestion-answeringsystemwith appropriate examples.

CourseOutcome4(CO4):

1. Describethenoisychannelmodelinmachinetranslation.

2. Compareandcontrastdifferentmachinetranslationstrategies.

CourseOutcome5(CO5):

1. Whenisthemaximumentropyframeworkappropriateasaclassificationmethod?

2. IllustratehowtheVectorSpacemodelhelpsinad-hocretrieval?

.

ModelQuestionPaper :COMPUTER SCIENCE AND ENGINEERING-CS2			
QPCODE:			
Reg No: ___________________			
Name: _____________________ PAGES:4			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode:.. J1PADT243			
CourseName: STATISTICAL NLP			
Max.Marks:60 Duration:2.5 Hours			
PARTA			
AnswerAllQuestions.EachQuestionCarries5Marks			
1.	Discusstheproblemofwordsensedisambiguationasanoisychannelmodel		
2.	ExplainWitten-Bellsmoothing.		
3.	IllustratetheusageofPennTreebankinProbabilisticParsing?		
4.	Comparesyntacticandsemantictransferapproachesformachinetranslation.		
5.	Describetheadvantagesanddisadvantagesof LatentSemanticIndexing.		(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			
6.	(a)	Atextcontains10nounsand5verbs.Awordisselectedandreplacedbythe other category and then a second word is selected. Calculate : 1. Theprobabilitythatthesecondwordisaverb. 2. Theprobabilitythatbothwordsselectedbelongtothesamecategory. 3. Theprobabilitythatbothwordsselectedareofdierentcategory.	(4)
	(b)	Computethet-valueofwords*newcompanies*inacorpuswithfollowing countsC(new)=30,000,C(companies)=9,000,C(newcompanies)=20and corpus size N=15,000,000	(3)
7.	(a)	WhatarethelimitationsofMaximumLikelihoodEstimation(MLE)?Give solutions to overcome these limitations.	(4)

	(b)	ExplainDictionary-BasedDisambiguationwithexample	(3)

8.		Considerasimplethree-stateMarkovmodeloftheweather.Anygivenday, theweathercanbedescribedasbeingprecipitation(rainorsnow),cloudyor sunny.Giventhattheweatherondayt=1issunny,whatistheprobabilitythat theweatherforthenext7dayswillbe—sun,sun,rain,rain,sun,clouds,sun‖ ?	(7)
9.		Findtheprobabilityofthestring—Astronomerssawstarswithears‖withthe helpofinsideprobabilityforthegivenPCFGgrammar. S→ NP VP (1.0) P → with (1.0) NP→ NPPP(0.4) V→saw(1.0) PP→ PNP(1.0) NP→ ears(0.18) VP→ VNP(0.7) NP→astronomers(0.1) VP→ VPPP(0.3) NP→ stars(0.18) NP→saw(0.04) NP→telescopes(0.1)	(7)
10		Measuringlengthinnumberofcharactersispreferablebecausethe variance in number of words is greater. Do you agree that word-based length is more variable? Why?	(7)
11	(a)	Explainthefollowing: a) Perceptrons b) KNNclassification	(3)

.

	(b)	We have a document collection D with 5 documents. Given aquery q, we knowthatthreedocumentsarerelevanttoq.Aretrievalalgorithmproduces the ranking (of all documents in D) shown in the following table. i) Calculateprecision,recallandF-Scoreforeachdocument. ii) Calculateaverageprecision. iii) Isthereanybreakevenpoint?	(4)

Ranki	+/-		p(i)	r(i)	F-Score
1	+				
2	+				
3	-				
4	+				
5	-				

LettherandomvariableX hasthreepossible outcomes{a, b, c}.Consider two distributions on this random variable.

Symbol	p(x)	q(x)
a	1/2	1/3
b	1/4	1/3
c	1/4	1/3

a) CalculateH(p),H(q)

b) VerifythatinthiscaseD(p||q)!=D(q|| p)

Syllabus

	Syllabus	
Module	Content	Hours
I	Introduction-Mathematical Foundations: Elementary Probabilitytheory-Probability spaces, Conditional probability and independence, Bayes'theorem,Randomvariables,Jointandconditional distributions, Bayesian statistics, Information Theory- Entropy, Joint entropy and conditional entropy, The noisy channel model, Relative entropy or Kullback-Leibler divergence, The relation to language:Crossentropy,LinguisticEssentials-PartsofSpeechand Morphology, Phrase Structure, Semantics and Pragmatics, Corpus-Based Work- Corpora, tokenization, marked-up data, Zipf's laws, Collocations, Concordances	8
II	Statistical Inference: n-gram Models over Sparse Data, n-gram, Buildingn-grammodels,StatisticalEstimators,Combining Estimators, WordSenseDisambiguation-SupervisedDisambiguation, Dictionary-Based Disambiguation, Lexical Acquisition	8

III	Grammar-Markov Models, Hidden Markov Models - Implementations, properties and variants, Part-of-Speech Tagging-Information Sources in Tagging, Markov Model Taggers, Hidden Markov Model Taggers, Transformation-Based Learning of Tags, Probabilistic Context free Grammar-features, Probability of a String, Problems with the Inside-Outside Algorithm	9
IV	Probabilistic parsing-Parsing for disambiguation, Parsing models vs. language models, Tree probabilities and derivational probabilities, Phrase structure grammars and dependency grammars, Statistical Alignment -Text alignment, word alignment, Statistical Machine Translation	8
V	Clustering-Hierarchical Clustering, Non-Hierarchical Clustering, Text Categorization-Decision Trees, Maximum Entropy Modelling, Perceptrons, k- Nearest Neighbour Classification, Topics inInformation Retrieval	7

CoursePlan

No	Topic	No. of Lectures (40)
1	Module1(LinguisticEssentials)	
1.1	Introduction,MathematicalFoundations:ElementaryProbabilitytheory-Probability spaces, Conditional probability and independence	1
1.2	Bayes'theorem,Randomvariables,Jointandconditionaldistributions, Bayesian statistics	1
1.3	InformationTheory-Entropy,Jointentropyandconditionalentropy,The noisychannelmodel	1
1.4	RelativeentropyorKullback-Leiblerdivergence,Therelationto language:Crossentropy	1
1.5	LinguisticEssentials-PartsofSpeechandMorphology,	1
1.6	PhraseStructure,SemanticsandPragmatics	1
1.7	Corpus-BasedWork-Corpora,tokenization,marked-updata	1
1.8	Zipf'slaws,Collocations,Concordances	1
2	Module2(WordSenseDisambiguation)	
2.1	StatisticalInference:n-gramModelsoverSparseData,n-gram	1
2.2	Buildingn-grammodels	1
2.3	StatisticalEstimators	1

2.4	CombiningEstimators	1
2.5	WordSenseDisambiguation	1
2.6	SupervisedDisambiguation :COMPUTERSCIENCEANDENGINEER	ING-C1S2
2.7	Dictionary-BasedDisambiguation	1
2.8	LexicalAcquisition	1
3	Module3(Part-of-SpeechTagging)	
3.1	Grammar-MarkovModels	1
3.2	HiddenMarkovModels-Implementations,propertiesandvariants,	1
3.3	Part-of-SpeechTagging-InformationSourcesinTagging	1
3.4	MarkovModelTaggers	1
3.5	HiddenMarkovModelTaggers	1
3.6	Transformation-BasedLearningofTags	1
3.7	ProbabilisticContextfreeGrammar-features	1
3.8	ProbabilityofaString	1
3.9	ProblemswiththeInside-OutsideAlgorithm	1
4	Module4(StatisticalAlignmentandMachineTranslation)	
4.1	Probabilisticparsing	1
4.2	Parsingfordisambiguation	1
4.3	Parsingmodelsvs.languagemodels	1
4.4	Treeprobabilitiesandderivationalprobabilities	1
4.5	Phrasestructuregrammarsanddependencygrammars	1
4.6	StatisticalAlignment-Textalignment	1
4.7	Wordalignment	1
4.8	StatisticalMachineTranslation	1
5	Module5(ClusteringandTextCategorization)	
5.1	Clustering-HierarchicalClustering	1
5.2	Non-HierarchicalClustering	1
5.3	TextCategorization-DecisionTrees	1
5.4	MaximumEntropyModelling	1
5.5	Perceptrons	1
5.6	k-NearestNeighbourClassification	1
5.7	TopicsinInformationRetrieval	1

ReferenceBooks

1. C.D.ManningandH.Schubert:FoundationsofStatisticalNaturalLanguageProcessing: MIT Press, 2001

2. Jurafsky,D.andJ.H.Martin,Speechandlanguageprocessing:AnIntroductiontoNatural Language Processing, Computational Linguistics, and Speech Recognition, Prentice-Hall, 2000.

3. Charniak,E.:StatisticalLanguageLearning.TheMITPress

.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PADT253	KNOWLEDGE BASED SYSTEM DESIGN	PROGRAMME ELECTIVE3	3	0	0	3

Preamble: This course helps to familiarize with the architecture and design principles of knowledge-basedsystems.ItcoversKnowledgeAcquisition,KnowledgeRepresentationand the development of Knowledge-Based Systems. On successful completion of the course, the studentswillbeabletodistinguishbetweenhumanandartificialexperts,theirproblem- solving strategies, and their knowledge representation methods.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Understand the fundamentals of knowledge engineering (Cognitive Knowledge Level: Understand)
CO2	Apply the concepts of problem solving for building Expert System (Cognitive Knowledge Level: Apply)
CO3	Apply the concepts of Knowledge based system for developing applications. (Cognitive Knowledge Level: Apply)
CO4	Apply knowledge representation methods for developing Knowledge based system (Cognitive Knowledge Level: Apply)
CO5	Make use of Semantics of Expert Systems for building applications. (Cognitive Knowledge Level: Apply)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3:Anabilitytodemonstrateadegreeofmasteryovertheareaasperthespecializationof the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of- the-art tool to model, analyse and solve practical engineering problems.

PO6:Anabilitytoengageinlife-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmentalaspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO 4	PO5	PO6	PO7
CO1	✓						
CO2	✓		✓	✓	✓	✓	
CO3	✓		✓		✓		
CO4			✓	✓			
CO5			✓			✓	✓

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	40
Analyse	20
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesignbasedquestions(forboth internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation:15marks

.

iii. Testpaper(1number) : 10 marks

Testpapershallincludeminimum80%ofthesyllabus.

Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module,having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity ofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldansweranyfive.Eachquestioncancarry 7 marks

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions Course

Outcome 1 (CO1):

1. DescribetheroleofinferenceengineinRule-basedexpertsystem.
2. Illustratethedifferencebetweenanexpertsystemandaknowledgebasedsystem?

CourseOutcome2(CO2)

1. Explainconflictresolutiontechniqueinrulebasedsystems.
2. CompareHayes-Roth‘sandClancey‘sclassificationofexpertsystem

.

CourseOutcome3(CO3): :COMPUTERSCIENCEANDENGINEERING-CS2

1. Illustratetheadvantagesandlimitationsofa Knowledgebasedsystem.

2. . Describethecomponentsofknowledgebasesystemandillustratetheneedofeach component with an wexample.

CourseOutcome4(CO4):

1. DiscussknowledgeacquisitionstrategiesinMUDsystem

2. Differentiate between procedural knowledge and factual knowledge

Course Outcome 5 (CO5):

1. Whatmakescase-basedreasoningdifferentfromrulebasedreasoning?Whichone will

   be better for knowledge based system design?

2. Whatarethesourcesofuncertainreasoning?

.

ModelQuestionPaper :COMPUTERSCIENCEANDENGINEERING-CS2			
QPCODE:			
Reg No: ___________________			
Name: _________________________ 4 PAGES:			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode: J1PADT253			
CourseName: KNOWLEDGE BASED SYSTEM DESIGN			
Max.Marks:60 Duration:2.5Hours			
AnswerAllQuestions.EachQuestionCarries5Marks			
1	Howdoesanexpertsystemhelpinsolvingcomplexproblems?		
2	Whydon‘trulebasedsystemsbacktrackmuch?		
3	Why don‘t all expert systems use Bayes‘ Rule to compute the degrees of certaintyassociatedwithconclusions?Isthereanyothermethodtosolvesuch problems?		
4	WhyisKnowledgeacquisitionabottleneck?Whatsolutionshavebeen proposed?		
5	Whatisacaselibraryandhowdoesacasedifferfromabook?		(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			
6	(a)	Whatarethecharacteristicsandcomponentsofanexpertsystem?	(3)

	(b)	Differentiate_searchspace‘and_solutionspace‘.Whatwillbethe searchspaceandsolutionspaceforthegameofchess?	(4)
7	(a)	Howcanweselectaknowledge engineeringtool? How hard such tools aretolearnanduse?	(4)
	(b)	Listouttheclassificationofexpertsystemtasks.	(3)
8.		DescribetheKnowledge-BasedSystemarchitectureanditsworking.	(7)
9.	(a)	Identifysomeknowledgeacquisitionstrategiesforheuristicsclassification.	(3)
	(b)	Whatalldifficultiesmayariseintheimplementationphaseofbuildinga Knowledge-Based System?	(4)
10.		Whenthesearchspaceispotentiallylarge,whichtypeoftestingis beneficial? Explain how it works?	(7)
11.	(a)	Describetheworkingofanyindustrialorcommercialexpertsystems.	(4)
	(b)	WhatdoyoumeanbyKnowledgeacquisition?Whatarethestagesof knowledge acquisition?	(3)
12.		WhatarethecommonpitfallsduringKnowledgeBasedSystems implementation and how we can avoid them?	(7)

Syllabus

	Syllabus	
Module	Content	Hours
I	IntroductiontoKnowledgeEngineering:TheHumanExpertandan Artificial Expert –Knowledge based systems architecture-Types of Knowledge-Characteristics of Knowledge-Components of Knowledge-Basic structure of knowledge based system- Knowledge Base -Inference Engine	9

II	Problem Solving Process- Rule Based Systems – Heuristic Classifications – Constructive Problem Solving – Toolsfor Building Expert Systems - Expert System Architectures	6
III	Knowledge Based Systems: Objectives of KBS - Components of KBS - Categories of KBS – Difficulties- KBS Architecture: Basic Structure- Applications	7
IV	DevelopingKnowledge-Based Systems: Effective - Development Model -KnowledgeAcquisition-DevelopingRelationshipswithExperts- Sharing Knowledge - Dealing with Multiple Experts - Knowledge Representation- Factual Knowledge - Procedural Knowledge, Knowledge Management	10
V	CaseBasedReasoning–SemanticsofExpertSystems–Modellingof Uncertain Reasoning. MachineLearning–RuleGenerationandRefinement–Learning Evaluation – Testing and Tuning.	7

CoursePlan

No	Topic	No.ofLectures (39)
1	Module 1(KnowledgeEngineering)	
1.1	IntroductiontoKnowledgeEngineering	1
1.2	TheHumanExpertandanArtificialExpert	1
1.3	Knowledgebasedsystemsarchitecture	1
1.4	TypesofKnowledge	1
1.5	CharacteristicsofKnowledge	1
1.6	ComponentsofKnowledge	1
1.7	BasicStructureofKnowledge-BasedSystems	1
1.8	KnowledgeBase	1
1.9	InferenceEngine	1
2	Module2(BuildingExpertSystems)	
2.1	ProblemSolvingProcess	1
2.2	RuleBasedSystems	1
2.3	HeuristicClassifications	1
2.4	ConstructiveProblemSolving	1
2.5	ToolsforBuildingExpertSystems	1
2.6	ExpertSystemArchitectures	1
3	Module3(KnowledgeBasedSystems)	
3.1	KnowledgeBasedSystems	1
3.2	Objectives	1

3.3	ComponentsofKnowledgebasedsystem	1
3.4	CategoriesofKnowledgebasedsystem	1
3.5	Difficulties	1
3.6	BasicStructureofKBS	1
3.7	ApplicationsofKBS	1
4	Module 4(DevelopmentofKnowledge-BasedSystems)	
4.1	DevelopingKnowledge-BasedSystems:Difficulties	1
4.2	DevelopmentModel	1
4.3	KnowledgeAcquisition	1
4.4	DevelopingRelationshipswithExperts	1
4.5	SharingKnowledge	1
4.6	DealingwithMultipleExperts	1
4.7	KnowledgeRepresentation	1
4.8	FactualKnowledge	1
4.9	RepresentingProceduralKnowledge	1
4.10	Knowledgemanagement	1
5	Module5(CaseBasedReasoning)	
5.1	CaseBasedReasoning	1
5.2	SemanticsofExpertSystems	1
5.3	ModellingofUncertainReasoning.	1
5.4	MachineLearning	1
5.5.	RuleGeneration	1
5.6	RefinementLearningEvaluation	1
5.7	TestingandTuning.	1

ReferenceBooks

1. PeterJackson.IntroductiontoExpertSystems.3rdEdition,PearsonEducation2007.

2. RajendraArvindAkerkar.Knowledge-BasedSystems.JonesandBartletPublishers2010

3. RobertI.Levine,DianeE.Drang,BarryEdelson.AIandExpertSystems:acomprehensive guide, C language, 2nd edition, McGraw-Hill

4. Jean-LouisErmine:ExpertSystems:TheoryandPractice.4thprinting,Prentice-Hall ofIndia, 2001

5. StuartRussell,PeterNorvig:ArtificialIntelligence:AModernApproach.2ndEdition

6. N.P.Padhy:ArtificialIntelligenceandIntelligentSystems.4thimpression,Oxford University Press, 2007.

.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PADT214	SOFT COMPUTING	PROGRAMME ELECTIVE4	3	0	0	3

Preamble: This course introduces the salient features and techniques used in soft computing. Thiscoursecoversvariousneuralnetworksmodels,fuzzysystems,geneticalgorithm concepts and their applications. On completion of the course, students will be able to acquire key skills required for solving real-life problems using soft computing techniques.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Analyse soft computing techniques and their applications (Cognitive knowledge level: Analyze)
CO2	Make use of various neural network models in problem solving (Cognitive knowledge level: Apply)
CO3	Examinevariousfuzzyconcepts(Cognitiveknowledgelevel:Analyse)
CO4	Illustrate use of fuzzy models in various applications (Cognitive knowledge level: Apply)
CO5	Analyse genetic algorithms and their applications (Cognitive knowledge level: Analyze)
CO6	Design,develop,implement andpresent solutionstosimplereal-lifeproblemswith popularopensourcelibraryusingsoftcomputingtechniques(Cognitiveknowledge level:Apply)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1:Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineeringactivitiesbyinteractingwiththeengineeringfraternityandwithsocietyatlarge.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization ofthe program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of-the-arttooltomodel,analyseandsolvepracticalengineeringproblems.

PO6:Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedtothe stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1			✓	✓		✓	
CO2	✓		✓	✓	✓	✓	
CO3			✓	✓		✓	
CO4	✓		✓	✓		✓	
CO5	✓		✓	✓		✓	
CO6	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	60
Analyse	40
Evaluate	Canbeevaluatedusing Assignments/ projects
Create	Canbeevaluatedusing Assignments/ projects

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluation shall only be based on application, analysis or design based questions (for both internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15marks

iii. Testpaper (1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.

Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldanswerany five.Eachquestioncancarry 7 marks Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. Comparelinearandnon-linearactivationfunctionswithanexampleforeach.

2. Showthatasinglelayerperceptroncannotsimulateasimpleexclusive-ORfunctionand give the reason.

3. Implement logical NOR function using neural network model

Course Outcome 2 (CO2)

1. Designasuitableneuralnetworkforhumanfacerecognition.

2. IllustratewithanexamplehowSupervisedLearningtechniquecanbeusedtosolvea multilabel classification problem

3. Comparesigmoidalandstepactivationfunctions.

CourseOutcome3(CO3):

1. TwofuzzysetsAandBdefinedontheEnglishalphabets(F,E,X,Y,I,T)aredefinedas below:

A={(F,0.2), (E,0.9),(X,0.4),(Y,0.3),(I,0.8),(T,0.5)}

B={(F,0.4),(E,0.6), (X,0.7),(Y,0.2),(I,0.8),(T,0.6)}

Findthefollowing.

i)A∩B ii)AUBciii)VerifyDemorgan‘sLaw(AUB)c=Ac∩Bc

2. Consider a set P={P1,P2,P3,P4} of four varieties of paddy plants,set D={D1,D2,D3,D4} of the various diseases affecting the plants and S={S1,S2,S3,S4}bethecommonsymptomsofthediseases.LetRbea relation on P X D and S be a relation on D X S.

Obtaintheassociationoftheplantswiththedifferentsymptomsofthediseasesusingmax–mincomposition.

3.FuzzybinaryrelationRisdefinedonsetA={1,2,3,4,5}andB={90,91,92,93,94,95} and represents the relation ―a is much smaller than b‖ . It is denoted by the membership function

.

R(a,b)= for a≤ b :COMPUTERSCIENCEANDENGINEERING-CS2

$1 - \frac{a}{b}$

0 Otherwise

Where a €Aand b € B. Find the domain and range of R.

Course Outcome 4 (CO4):

1. Design a fuzzy controller to determine the wash time of a washing machine. Assume the inputisdirtandgrease.Usethreedescriptorsforinputvariableandfivedescriptorsforoutput variable. Derive the set of rules for controller actionand defuzzification. Show that if the dirt are more , the wash time will be more and vice versa.

2. Illustratethemultiatributedecisionmakingwithanexample.

3. DemonstrateproblemsolvingusingMamdanifuzzymodel,withanexample.

CourseOutcome5(CO5):

1. IllustrateRouletteWheelSelectiontechniquewithanexample.

2. Differentiatesingle-pointcrossover,multipointcrossoveranduniformcrossover techniques used in Genetic Algorithm.

CourseOutcome6(CO6):

1. Preparearevievarticlebasedonpeerreviewedoriginalpublications(referringminimum 10 publications) about recent neuro-fuzzy systems based applications.

2. Courseprojecttodesignanddevelophandwrittencharacterrecognitionsystemusing Python/ TensorFlow/ Keras based on MNIST handwritten digit dataset.

3. PresentaseminaronanyoneapplicationofGeneticAlgorithm(OtherthanTraveling Salesman Problem)

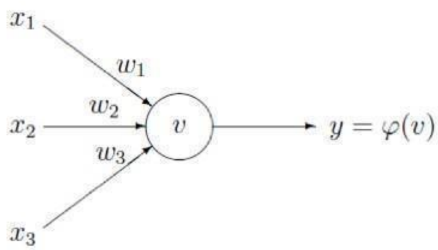
.

ModelQuestionPaper :COMPUTERSCIENCEANDENGINEERING-CS2			
QPCODE:			
Reg No: ___________________			
Name:____PAGES:3			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode: J1PADT214			
CourseName:SoftComputing			
Max.Marks:60 Duration:2.5Hours			
PARTA			
AnswerAllQuestions.EachQuestionCarries5Marks			
1.	Implementthefollowingtruthtableusingsuitableneuralnetworkmodel.		

Input		Output
X1	X2	
0	0	0
0	1	1
1	0	1
1	1	1

2.	DifferentiateSupervisedandUnsupervisedLearningNeuralNetworkswithan example for each.		
3.	DifferentiateGeneralisedBellmembershipfunctionandLeft-RightMembership Function.		
4.	DiscussFuzzyExpertSystemwithanexample		
5.	ExplainthestepsinGeneticAlgorithm,withanexample.		(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			
6.	(a)	Ahandwritten character recognition system for English language has to be designed. Use feedforward neural network system to design the same.	(4)
	(b)	ImplementlogicalX-ORfunctionusingneuralnetworkmodel	(3)
7.	(a)	A company has collected large amount of data in pairs of input output vectors. Design a system using radial basis function network for predicting output for new input.	(4)

(b) Considertheunitshowninthefollowingfigure.



Suppose that the weights corresponding to the three inputs have the following values: w1= 1, w2= - 3, w3= 2 and the activation of the unit is given by thestep-function:

$$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Calculate what will be the output value y of the unit for each of the following input patterns:

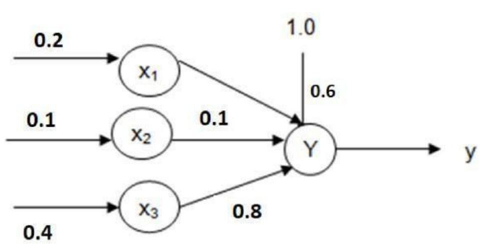
Pattern	P_1	P_2	P_3	P_4
x_1	1	0	1	1
x_2	0	1	0	1
x_3	0	1	1	1

8. (a) Let P and Qbe two fuzzy relations defined on the sets X x Y and Y x Z respectively,whereX={x1,x2},Y={y1,y2,y3,},andZ={z1,z2}.PandQ are given using the following relation matrices. (4)

P= [0.50.20.30.10.70.6] Q= [0.500.80.40.61]

FindtheassociationPoQusingbothmax-mincompositionandmax-product compositions

(b) DifferentiateTriangularandtrapezoidalfuzzymembershipfunctions. (3)

9.	(a)	Consideratwoinput–oneout:pCuOtpMr oPbUl eTmE tRh aSt Ci nIcEluNdCesEfoAlNl oDw iEngNtGhIr Ne eEERIN rules. Rule: 1 - IF x is A3 OR y is B1 THEN z is C1 Rule:2-IFxisA2ANDyisB2THENzisC2 Rule: 3 - IF x is A1 THEN z is C3 AssumeA3=0.2,B1=0.4,A2=0.7,B2=0.9,A1=0.3andtheselectedvaluesof z in c1=(0,11,21), C2=(,40,50,60, 70) & C3=(60, 70,80,90).Find out the value ofzusingMamdanifuzzyinferencesystem.	G-CS 2
	(b)	DerivetheconclusionforMultipleRuleswithmultipleantecedents.	(2)
10.	(a)	Illustratetheproblem-solvingmethodusingSugenofuzzysystem,withan example.	(4)
	(b)	Afuzzyreasoningsystemisprovidedwiththefollowingfactsandrules.)Premise1 (fact):xisA‘andyisB‘)Premise2(Rule1):ifxisAandyisB,thenzisC. DemonstratetheinferenceproceduretofindouttheconclusionzisC‘	(3)
11.	(a)	CalculatetheoutputoftheneuronyforthefollowingnetworkusingBinary Sigmoidal activation function. 	(4)
	(b)	ImplementlogicalANDfunctionusingneuralnetworks.	(3)
12.	(a)	Asalesman has to follow the shortest route to visit N cities exactlyonce andreach the starting city. Apply genetic algorithm to solve this problem.	(4)
	(b)	IllustrateTournamentSelectioninGeneticAlgorithmwithanexample.	(3)

.

Syllabus

:COMPUTERSCIENCEANDENGINEERING-CS2

Module1-Introduction

Soft Computing Constituents, Computational Intelligence and Soft Computing vs Artificial Intelligence and Hard Computing.Introduction to artificial neural networks- biologicalneurons,McCullochandPittsmodelsofneurons,Perceptionnetworks,Multilayer Perceptron. Types of activation function, network architectures, learning process, LearningXOR and other logical Gate functions, Sigmoid Neurons, Gradient Descent Algorithm,Back propagation Neural Network.

Module2–NeuralNetworks

Adaptive Networks, Feed Forward Networks, Supervised Learning Neural Networks, Radial BasisFunctionNetworks,ReinforcementLearning,UnsupervisedLearningNeural Networks, AdaptiveResonance Architectures,BAM,Maxnet,KohonenSelfOrganizing Maps, K-means clustering algorithm, Introduction to Convolutional Neural Networks, Applications of ANNs to solve real life problems. Familiarisation of neural network tools , Building neural networks with Python, TensorFlow, Introduction to Keras.

Module 3–FuzzySets&Logic

Fuzzy versus Crisp, Fuzzy sets, Membership function, linguistic variable, basic operators, properties, Extension principle and Fuzzy relations, Cartesian product, Operations on Fuzzysets , Operations on fuzzy relations, Fuzzy Logic

Module 4–FuzzyReasoning

Fuzzy If-Then Rules, Fuzzy Inference Systems, Fuzzy Expert Systems, Fuzzification and Defuzzification methods, Fuzzy Decision Making , Mamdani Fuzzy Models, Sugeno Fuzzy Models, Applications of Fuzzy logic, Neuro Fuzzy Systems.

Module5–GeneticAlgorithm

Introductiontogeneticalgorithmandhybridsystems-Geneticalgorithms,naturalevolution , properties, classification, GA features, coding, selection, selection methods, reproduction , cross over and mutation operators, basic GA and structure, Solving Travelling Salesman Problem using GA. Applications using GA, Hybrid Systems - GA based BPNN.

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CoursePlan

No	Topic	No. of Lectures(36Hours)
1	Module1(IntroductiontoSoftComputingAndNeuralNetworks)-7Hours	
1.1	Soft Computing Constituents, Computational Intelligence and Soft ComputingvsArtificialIntelligenceandHardComputing.Introduction to artificial neural networks, biological neurons	1
1.2	McCullochandPittsmodelsofneurons,Perceptionnetworks	1
1.3	MultilayerPerceptron,Typesofactivationfunction	1
1.4	Networkarchitectures,Learningprocess	1
1.5	LearningXORandotherlogicalGatefunctions	1
1.6	SigmoidNeurons,GradientDescent Algorithm	1
1.7	BackpropagationNeuralNetwork.	1
2	Module2(NeuralNetworks)8Hours	
2.1	AdaptiveNetworks,FeedForwardNetworks,SupervisedLearning Neural Networks	1
2.2	RadialBasisFunctionNetworks,ReinforcementLearning,	1
2.3	Unsupervised Learning Neural Networks,Adaptive Resonance Architectures,	1
2.4	BAM,Maxnet,KohonenSelfOrganizingMaps,	1
2.5	K-meansclusteringalgorithm,IntroductiontoConvolutionalNeural Networks,	1
2.6	ApplicationsofANNstosolvereallifeproblems.Familiarisationof neural network tools	1
2.7	BuildingneuralnetworkswithPython	1
2.8	TensorFlow, IntroductiontoKeras.	1

.

3	Module3(FuzzySets&Logic)7Hours :COMPUTERSCIENCEANDENGINEERING-CS2	
3.1	FuzzyversusCrisp,Fuzzysets,Membershipfunction,linguisticvariable	1
3.2	Basicoperators,properties	1
3.3	Fuzzyrelations	1
3.4	Cartesianproduct	1
3.5	OperationsonFuzzysets	1
3.6	Operationsonfuzzyrelations	1
3.7	FuzzyLogic	1
4	Module4(FuzzyReasoning)7Hours	
4.1	FuzzyIf-ThenRules	1
4.2	FuzzyInferenceSystems,FuzzyExpertSystems	1
4.3	FuzzificationandDefuzzificationmethods	1
4.4	FuzzyDecisionMaking	1
4.5	MamdaniFuzzyModels	1
4.6	SugenoFuzzyModels	1
4.7	ApplicationsofFuzzylogic,NeuroFuzzySystems.	1
5	Module5(GeneticAlgorithm)7Hours	
5.1	Introductiontogeneticalgorithmandhybridsystems	1
5.2	Geneticalgorithms,Naturalevolution,Properties	1
5.3	Classification,GAfeatures,Coding,Selection,Selectionmethods	1
5.4	Reproduction-CrossoverandMutationoperators	1
5.5	BasicGAandstructure,SolvingTravellingSalesmanProblemusingGA	1
5.6	ApplicationsusingGA	1
5.7	HybridSystems-GAbasedBPNN	1

.

ReferenceBooks :COMPUTERSCIENCEANDENGINEERING-CS2

1. Jyh-ShingRogerJang,Chuen-TsaiSun,EijiMizutani,Neuro-FuzzyandSoftComputing, Prentice-Hall of India, 2003.

2. S.N.Sivanandam,S.N.Deepa,—PrinciplesofSoftComputing‖,2/e,JohnWileyIndia, 2012.

3. JamesA.FreemanandDavidM.Skapura,NeuralNetworksAlgorithms,Applications,and Programming Techniques, Pearson Edn., 2003.

4. SimonHaykin"NeuralNetworksandLearningMachines"PrenticeHall,2008

5. JonKrohn,GrantBeyleveld,AglaeBassens,DeepLearningIllustrated,Pearson,2020

6. TimothyJ.Ross,—FuzzyLogicwithengineeringapplications‖,WileyIndia,2021

7. BartKosko.:NeuralNetworkandFuzzySystems-PrenticeHall,Inc.,EnglewoodCliffs, 1991

8. GoldbergD.E.:GeneticAlgorithmsinSearch,Optimization,andMachineLearning-Addison Wesley, 1989

9. MitchellMelanie,AnIntroductiontoGeneticAlgorithm,PrenticeHall,1998.

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PADT224	HIGH PERFORMANCE COMPUTING	PROGRAMME ELECTIVE4	3	0	0	3

Preamble: This course helps the learners to understand the different architectural features ofhigh-end processors. This course discusses the Basics of high-end processors Architecture, Instruction-Level Parallelism, Data-Level Parallelism, Thread Level Parallelism, and GPU Architectures. This course enables the students to provide solutions to real-world problems making use of the capabilities of HPC systems.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Analysedifferenttypesofmodernprocessingenvironmentsandparallelcomputing hardware(CognitiveKnowledgeLevel:Analyse)
CO2	SummarizetheconceptsofInstructionLevelParallelism(CognitiveKnowledge Level:Analyse)
CO3	AppreciatetheideaofDataLevelParallelism(CognitiveKnowledgeLevel: Apply)
CO4	DemonstratetheconceptofThreadLevelParallelism(CognitiveKnowledgeLevel: Apply)
CO5	ComprehendtheadvancedfeaturesofGPUarchitecture.(CognitiveKnowledge Level:Analyse)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3:Anabilitytodemonstrateadegreeofmasteryovertheareaasperthespecializationof the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedtothe stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO 4	PO5	PO6	PO7
CO1	✔		✔			✔	
CO2	✔		✔			✔	
CO3	✔		✔			✔	
CO4	✔		✔			✔	
CO5	✔		✔			✔	

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	40
Analyse	20
Evaluate	-
Create	-

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesignbasedquestions(forboth internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15marks

iii. Testpaper (1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.

Coursebasedtask/test paperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module,having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity ofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldansweranyfive.Eachquestioncancarry 7 marks
Totaldurationoftheexaminationwillbe150minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

CourseOutcome1(CO1):

1. Differentiatedifferentclassesofcomputer-basedonfeatureslikemicroprocessorcost, system cost, and system design issues.
2. Explainthedifferentmethodsbywhichcomputerhardwareexploitsapplication-level parallelism.
3. Explainindetailtheinstructionsetarchitecture
4. Describe the encoding scheme specified as part of ISA

Course Outcome 2 (CO2):

1. Differentiatedata,name,andcontroldependencieswithsuitableexamples.
2. Explainloopunrollingwithsuitablecodingdemonstration

3. ExplainindetailaboutTournamentPredictors.

4. Describetheuniquefeaturesofverylonginstructionwordprocessors.

CourseOutcome3(CO3):

1. What are the three things conveyed through a data dependence? Explain the Data Dependencies of the following code:

```
Loop:   fld     f0,0(x1)     //f0=array element
        fadd.d  f4,f0,f2     //add scalar in f2
        fsd     f4,0(x1)     //store result
        addi    x1,x1,-8     //decrement pointer 8 bytes
        bne     x1,x2,Loop   //branch x1≠x2
```

2. Assumeasingle-issuepipeline.Unrolltheloopasmanytimesasnecessaryto schedule it without any stalls, collapsing the loop overhead instructions. How many times must the loop be unrolled? Show the instruction schedule. What is the execution time per element of the result?

3. ExplaintheSIMDInstructionSetExtensionsforMultimedia.

CourseOutcome4(CO4):

1. With the help of a neat diagram illustrate a single-chip multicore with a distributed cache.

2. Demonstrate the Implementation of cache coherence in a distributed-memory multiprocessor by adding a directory to each node with a suitable diagram.

3. Consider the following code segments running on two processors P1 and P2. Assume A, and B, are initially 0. Explain how an optimizing compiler might makeitimpossibleforBtobeeversetto2inasequentiallyconsistent execution model.

P1:	**P2:**
`A=1;`	`B=1;`
`A=2;`	`While (A <> 1);`
`While (B == 0);`	`B= 2;`

CourseOutcome5(CO5):

1. ExplainthebenefitsofpotentialGPU.

2. IllustrateGPUsystemasanacceleratedcomputationalplatform.

3. DiscussCPUtoGPUdatatransferoverhead.

ModelQuestionPaper

QPCODE:

Reg No: ___________________

Name:____ PAGES:2

APJABDULKALAMTECHNOLOGICALUNIVERSITY

SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

CourseCode: J1PADT224

CourseName:HighPerformanceComputing

Max.Marks:60 Duration:2.5Hours

PARTA

AnswerAllQuestions.EachQuestionCarries5Marks

<table>
<tr><td>1.</td><td colspan="2">Discussindetailtheimportanceofconsideringprocessorperformanceforthe design of an efficient computer system.</td><td></td></tr>
<tr><td>2.</td><td colspan="2">Illustratehowdata-levelparallelismisachievedinvector,andSIMD architectures.</td><td></td></tr>
<tr><td>3.</td><td colspan="2">Considerthefollowingloop:
<code>for (i=0;i <100;i++) {</code>
<code>A[i] = A[i] + B[i]; /* S1 */</code>
<code>B[i+1] = C[i] + D[i]; /* S2 */</code>
<code>}</code>
Are there existdependencies between S1 and S2?Determine whetherthe aboveloop is parallel? If not, show how to make it parallel.</td><td></td></tr>
<tr><td>4.</td><td colspan="2">Suppose an application running on a 100-processor multiprocessor use 1, 50, or 100processors.Iffor95%ofthetimeall100processorsareused,illustratehow the remaining 5% of the execution time employs 50 processors for a speedup of 80?</td><td></td></tr>
<tr><td>5.</td><td colspan="2">ExplainaboutGPUthreadengine.</td><td>(5x5=25)</td></tr>
<tr><td colspan="4">PartB
(Answeranyfivequestions.Eachquestioncarries7marks)</td></tr>
<tr><td>6.</td><td>(a)</td><td>Illustratehowprocessesareprotectedwiththehelpofvirtualmemory.</td><td>(4)</td></tr>
<tr><td></td><td>(b)</td><td>Describethequantitativeprincipleofcomputerdesignwith Amdahl'slaw.</td><td>(3)</td></tr>
<tr><td>7.</td><td>(a)</td><td>Explainindetaildatadependenceandhazards.</td><td>(7)</td></tr>
</table>

ModelQuestionPaper			
	(b)		
8.	(a)	Describetheuniquefeaturesofverylonginstructionwordprocessors.	(3)
	(b)	Considerathree-waysuperscalarmachinerenamingthesethreeinstructions concurrently: `addi x1, x1, x1` `addi x1, x1, x1` `addi x1, x1, x1` Ifthevalueofx1startsat5,thenwhatwillbeitsvaluewhenafterthis sequence is executed?	4

9.	(a)	The following loop has multiple types of dependences. Find all the true dependences,outputdependencies,andanti-dependencies,andeliminatethe output dependencies and anti-dependencies by renaming.
		`for (i=0; i<100; i=i+1) {`
		`Y[i] = X[i] / c; /* S1 */`
		`X[i] = X[i] + c; /* S2 */`
		`Z[i] = Y[i] + c; /* S3 */`
		`Y[i] = c - Y[i]; /* S4 */`
		`}`
			(7)

10	(a)	Demonstratethedifferenttypesofhardwareapproachesrequiredforthe working of multithreading.	(7)
11	(a)	Consider an 8-processor multicore where each processor has its ownL1 and L2caches.HeresnoopingisperformedonasharedbusamongtheL2caches. AssumethattheaverageL2requestis15cyclesforacoherencemissorother miss and a clock rate of 3.0 GHz, a CPI of 0.7, and a load/store frequency of 40%.If the goal set is that no more than 50% of theL2 bandwidth is consumed by coherence traffic,then what is themaximum coherence miss rate per processor?	(7)
12	(a)	Explainthemulti-GPUplatform.	(7)

Syllabus

Module-1(BasicsofArchitecture)

ClassesofComputers-ClassesofParallelismandParallelArchitectures–Defining Computer Architecture– Dependability – Quantitative Principles of Computer Design –Basics of Memory Hierarchies – Virtual Memory and Virtual Machines – Pipelining

Module-2(Instruction-LevelParallelism)

Instruction-Level Parallelism: Concepts and Challenges – Basic Compiler Techniques for Exposing ILP – Reducing Branch Costs With Advanced Branch Prediction – Hardware-Based Speculation – Multithreading: Exploiting Thread-Level Parallelism to Improve Uniprocessor Throughput

Module-3(Data-LevelParallelism)

VectorArchitecture–SIMDInstructionSetExtensionsforMultimedia–Graphics Processing Units – Detecting and Enhancing Loop-Level Parallelism

Module-4(ThreadLevelParallelism)

MultiprocessorArchitecture:IssuesandApproach–CentralizedShared-Memory Architectures – Performance of Symmetric Shared-Memory Multiprocessors– Distributed Shared-MemoryandDirectory-BasedCoherence–Synchronization:TheBasics– Introduction to Memory Consistency

Module-5(GPUArchitectures)

The CPU-GPU system as an accelerated computational platform – The GPU and the thread engine – Characteristics of GPU memory spaces – The PCI bus: CPU to GPU data transfer overhead – Multi-GPU platforms– Potential benefits of GPU – accelerated platforms

CoursePlan

No	Contents	No. of Lecture Hours (36hrs)
Module1-BasicsofArchitecture(7hours)		
1.1	ClassesofComputers	1hour
1.2	ClassesofParallelismandParallelArchitectures	1hour
1.3	Dependability	1hour
1.4	QuantitativePrinciplesofComputerDesign.	1hour

1.5	BasicsofMemoryHierarchies	1hour
1.6	VirtualMemoryandVirtualMachines	1hour
1.7	Pipelining	1hour
Module-2(IntroductiontoSyntaxAnalysis)(7hours)		
2.1	Instruction-LevelParallelism:ConceptsandChallenges	1hour
2.2	BasicCompilerTechniquesforExposingILP	1hour
2.3	ReducingBranchCostsWithAdvancedBranchPrediction	1hour
2.4	Hardware-BasedSpeculation	1hour
2.5	Multithreading	1hour
2.6	ExploitingThread-LevelParallelismtoImproveUniprocessorThroughput–Lecture1.	1hour
2.7	ExploitingThread-LevelParallelismtoImproveUniprocessorThroughput–Lecture2.	1hour
Module-3-Data-LevelParallelism(7hours)		
3.1	VectorArchitecture-Lecture1	1hour
3.2	VectorArchitecture-Lecture2	1hour
3.3	SIMDInstructionSetExtensionsforMultimedia–Lecture1	1hour
3.4	SIMDInstructionSetExtensionsforMultimedia–Lecture2	1hour
3.5	GraphicsProcessingUnits	1hour
3.6	DetectingandEnhancingLoop-LevelParallelism–Lecture1	1hour
3.7	DetectingandEnhancingLoop-LevelParallelism–Lecture2	1hour
Module4–ThreadLevelParallelism(8hours)		
4.1	MultiprocessorArchitecture:IssuesandApproach	1hour
4.2	CentralizedShared-MemoryArchitectures–Lecture1	1hour
4.3	CentralizedShared-MemoryArchitectures–Lecture2	1hour
4.4	PerformanceofSymmetricShared-MemoryMultiprocessors	1hour
4.5	DistributedShared-Memory	1hour
4.6	Directory-BasedCoherence	1hour
4.7	Synchronization	1hour
4.8	IntroductiontoMemoryConsistency	1hour

Module5–GPUArchitectures(7hours)		
5.1	TheCPU-GPUsystemasanacceleratedcomputationalplatform	1hour
5.2	TheGPUandthethreadengine–Lecture1	1hour
5.3	TheGPUandthethreadengine–Lecture2	1hour
5.4	CharacteristicsofGPUmemoryspaces	1hour
5.5	PCIbus:CPUtoGPUdatatransferoverhead	1hour
5.6	Multi-GPUplatforms	1hour
5.7	PotentialbenefitsofGPU-acceleratedplatforms	1hour

TextBooks

1. JohnL.Hennessy,DavidA.PattersonComputerArchitecture,SixthEditionA Quantitative Approach, Morgan Kaufman, Fifth Edition, 2012.

2. RobertRobey,YulianaZamora,ParallelandHigh-PerformanceComputing, Manning Publications, First Edition, 2021.

ReferenceBooks

1. ThomasSterling,MatthewAnderson,andMaciejBrodowicz,High-PerformanceComputing – Modern Systems and Practices, First Edition, 2017.

2. CharlesSeverance,KevinDowd,High-PerformanceComputing,O'ReillyMedia, Second Edition, 1998.

3. KaiHwang,FayeAlayeBriggs,ComputerArchitectureandParallelProcessing, McGraw-Hill, 1984.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PADT234	CRYPTOGRAPHY AND NETWORK SECURITY	PROGRAMME ELECTIVE4	3	0	0	3

Preamble:Thecourseintroducesthefundamentalconceptsofcryptographyandthesecurity issues. It covers basic symmetric and asymmetric cryptographic algorithms, cryptographic hash functions, message authentication protocols, digital signatures, key management and distribution schemes for symmetric and asymmetric encryption, and case studies involving different types of attacks and conventionalalgorithms. The course enables learners to apply the cryptographicconcepts to real-life problems.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Summarizebasiccryptographicalgorithmsandsecurityissues.(Cognitive KnowledgeLevel: Apply)
CO2	Compareandanalysevarioussymmetricandasymmetrickeycryptographic algorithms. (Cognitive Knowledge Level: Apply)
CO3	Demonstratecryptographichashfunctionsanddigitalsignatureschemes. (CognitiveKnowledge Level: Apply)
CO4	Summarizekeymanagementanddistributionschemesforsymmetricand asymmetricencryption.(CognitiveKnowledgeLevel:Apply)
CO5	Applytheconceptstodesignanefficientcryptographicalgorithm,hashfunction anddigitalsignaturescheme.(CognitiveKnowledgeLevel:Apply)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthecourse.

PO1:Anabilitytoindependentlycarryoutresearch/investigationanddevelopmentworkin engineering and allied streams.

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3:Anabilitytodemonstrateadegreeofmasteryovertheareaasperthespecializationofthe program. The mastery should be at a level higher than the requirements in the appropriate bachelor program.

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworldproblems byfollowing the standards.

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of-the-art tool to model, analyse and solve practical engineering problems.

PO6: An abilitytoengage in life-long learning for the design and development related to the stream related problems takinginto consideration sustainability, societal, ethical and environmentalaspects.

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagementand finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO 4	PO5	PO6	PO7
CO1	✓		✓	✓	✓	✓	
CO2	✓		✓	✓	✓	✓	
CO3	✓		✓	✓	✓	✓	
CO4	✓		✓	✓	✓	✓	
CO5	✓		✓	✓		✓	

AssessmentPattern

Bloom‘sCategory	EndSemester Examination
Apply	60
Analyse	40
Evaluate	-
Create	-

Assignmentsorcourseprojectscanbeusedforhigherlevelassessmentofcourseoutcomes.

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesignbasedquestions(forboth internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation:15marksiii.Test paper

(1 number) : 10 marks

Testpapershallincludeminimum80%ofthesyllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension,application,analysis,synthesis,evaluationandunderstandingofthestudents.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module,having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions(suchquestionsshallbeusefulinthetesting ofoverallachievementandmaturity of the students in a course, through long answer questions relating to theoretical/practicalknowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldansweranyfive.Eachquestioncancarry 7 marks. Total duration of the examination will be 150 minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall benormalized accordingly.
For example if the average end semester mark % for a core course is 40, then the maximum eligiblemark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions Course

Outcome 1 (CO1):

1. UseVigenereCipherwithkeyHEALTHtoencryptthemessage—Cryptographyand networksecurity‖.

2. Inapublickeysystemusing RSA, you intercept the ciphertextC=8 sent to a user whosepublickeyise=13,n=33.WhatistheplaintextM?

3. Alice wants to send a message to Bob, without Eve observing it. Alice and Bob have agreedtouse a symmetric cipher. Key exchange hasalready occurred,andsothey shareakeyK.OutlinethestepsthatAliceandBobmustfollowwhentheyencrypt and decrypt, respectively.

CourseOutcome2(CO2)

1. Withthehelpofanexample,illustratetheconceptbehindElGamalCrypto System.2.

Elaborate on an application which uses Chinese Reminder Theorem.

3. ConsideraDiffieHellmanschemewithacommonprimeq=11andprimitiverootα=2.
   i. Showthat2isaprimitiverootof11.
   ii. IfuserAhaspublickeyY_A=9,whatisA‘sprivatekey?
   iii. If user B has public key Y_B= 3, what is the shared secret key K, shared with

Course Outcome 3 (CO3):

1. DiscussonanattacktowhichMACisvulnerable.HowcanMACbemademoresecure?

2. Imagine that a specific hash function is used for HMAC, and that vulnerabilities havebeenfoundinthehashfunctionssothattheHMACisinsecure.Whichchangesin the HMAC arerequired in order to make it secure again?

3. The Diffie-Hellman key agreement algorithm achieves key agreement by allowing two hoststo create a shared secret.

   a. ClearlyexplaintheoperationoftheDiffie–Hellmankeyexchangeprotocol.

   b. ClearlyexplainwhythebasicDiffie–Hellmanprotocoldoesnotprovideany assurance regarding which other party the protocol is run with.

CourseOutcome4(CO4):

1. ExplaintheauthenticationproceduresdefinedbyX.509certificate.

2. Illustratetheconceptof_certificatechain‘forverificationofdigitalsignatureon X.509certificate.

3. DescribethestepsinfindingthemessagedigestusingSHA-512algorithm.Analyse thesituation where two messages have the same message digest.

CourseOutcome5(CO4):

1. Preparereportonvarioussymmetricandasymmetrickey cryptosystems.Compare featuresof both the systems based on their applications.

2. Presentaseminarondifferentmessageauthenticationprotocolsanddigitalsignature

.

schemes elaborating on their unique features.

3. Courseprojecttodesignaprivateorpubliccryptosystem/messageauthentication protocol ordigital signature scheme.

ModelQuestionPaper		
QPCODE:		
Reg No: ___________________		
Name:___________________ PAGES:4		
APJABDULKALAMTECHNOLOGICALUNIVERSITY		
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR		
CourseCode: J1PADT234		
CourseName: CRYPTOGRAPHY AND NETWORK SECURITY		
Max.Marks:60 Duration:2.5Hours		
PARTA		
AnswerAllQuestions.EachQuestionCarries5Marks		
1.	Encrypt the message —the movie is set to release this Friday‖ using Vigenere cipherwithkey—awards‖.Ignorethespacebetweenwords.Decryptthe messageto get the plain text.	
2.	AliceandBobagreedtouseRSAalgorithmforthesecretcommunication.Alice securely choose two primes, p=5 and q=11 and a secret key d=7. Find the corresponding public key.Bob usesthis public key and sends a ciphertext 18to Alice. Find the plain text.	
3.	Comparedigitalsignatureswithauthenticationprotocols.	

4.		Explainwhymessageauthenticationaloneisinsufficientasproofofmessage originingeneral,andtosettledisputesaboutwhethermessageshavebeensent.	
5.		Quotingsuitableexamples,differentiatebetweenIPspoofingattacksand denialof service attacks.	(5x5=25)
		PartB (Answeranyfivequestions.Eachquestioncarries7marks)	
6.		The encryption key in a transposition cipher is (3,1,4,5,2). Perform encryption and decryption for the message ―meet me after the toga party‖. Addaboguscharacterattheendtomakethelastgroupthesamesizeas theothers.	(7)
7.		Elaborate on the algorithm for generating keys in RSA algorithm. Perform encryption and decryption using RSA for the following. P=7; q=11; e=13; M=8.	(7)

8.		IllustratemaninthemiddleattackonDiffieHellmankeyexchange algorithm.	(7)
9.		Alice wants to send a message M with a digital signature Sig(M) to Bob. AliceandBobhaveanauthenticcopyofeachother‘spublickeys,and haveagreedonusingaspecifichashfunctionh.Outlinethestepsthat Alice must follow when signing M, and the steps that recipient Bob must follow forvalidating the signature Sig(M).	(7)
10.		AlicewantstosendamessagetoBob.AlicewantsBobtobeableto ensure that the message did not change in transit. Briefly outline the cryptographicsteps that Alice and Bob must follow to ensure the integrityof the messageby creating and verifying a MAC.	(7)
11.		ExplainKerberosauthenticationmechanismwithsuitablediagram.	(7)
12.		Withthehelpofanexample,illustrateIPspoofingattacks.	(7)

	Syllabus	
Module	Content	Hours
I	Security Concepts - Introduction, The need for security, Security approaches, Principles of security, Types of Security attacks. Services andMechanism, A model for Network Security. CryptographyConceptsandTechniques:Introduction,plaintext and cipher text, substitution techniques, transposition techniques,encryptionanddecryption,symmetricand asymmetric key cryptography, steganography, key range and key size, possible types of attacks.	9
II	SymmetricKeyCryptography-BlockCipherprinciples,DES, AES,Blowfish, RC5, IDEA, Block cipher operation, Stream ciphers, RC4. Asymmetric Key Cryptography - Principles of public key cryptosystems, RSA algorithm, Elgamal Cryptography, Diffie-Hellman Key Exchange,Knapsack Algorithm.	11
III	CryptographicHashFunctions-MessageAuthentication, SecureHashAlgorithm(SHA-512),Messageauthentication codes: Authentication requirements, HMAC, CMAC, Digital signatures, Elgamal DigitalSignature Scheme.	7
IV	Key Management and Distribution - Symmetric KeyDistribution Using Symmetric& Asymmetric Encryption, Distribution of Public Keys, Kerberos, X.509 Authentication Service, Public – Key Infrastructure.	6
V	Case StudiesofCryptography- Denialofserviceattacks,IP spoofing attacks, Secure inter branch payment transactions, ConventionalEncryptionandMessageConfidentiality, Conventional Encryption Principles, Conventional Encryption Algorithms, Location of Encryption Devices, KeyDistribution.	7

.

CoursePlan :COMPUTERSCIENCEANDENGINEERING-CS2

No	Topic	No. of Lectures
1	SecurityandCryptographyConcepts(9hours)	
1.1	Security Concepts - Introduction, The need for security, Securityapproaches	1
1.2	Principlesofsecurity,TypesofSecurityattacks.	1
1.3	ServicesandMechanism	1
1.4	AmodelforNetworkSecurity.	1
1.5	CryptographyConceptsandTechniques:Introduction,plain text andcipher text, substitution techniques	1
1.6	Transpositiontechniques	1
1.7	Encryptionanddecryption	1
1.8	Symmetricandasymmetrickeycryptography,steganography	1
1.9	Key rangeandkeysize,possibletypesofattacks.	1
2	SymmetricandAsymmetricKeyCryptography(11hours)	
2.1	SymmetricKeyCryptography-BlockCipherprinciples,DES	1
2.2	AES	1
2.3	Blowfish	1
2.4	RC5	1
2.5	IDEA	1
2.6	Blockcipheroperation	1
2.7	Streamciphers,RC4	1
2.8	AsymmetricKeyCryptography-Principlesofpublickey cryptosystems, RSA algorithm	1
2.9	ElgamalCryptography	1
2.10	Diffie-HellmanKeyExchange	1
2.11	KnapsackAlgorithm.	1

3	CryptographicHashFunctions(7hours)	
3.1	CryptographicHashFunctions-MessageAuthentication	1
3.2	SecureHashAlgorithm(SHA-512)	1
3.3	Messageauthenticationcodes:Authenticationrequirements	1
3.4	HMAC	1
3.5	CMAC	1
3.6	Digitalsignatures	1
3.7	ElgamalDigitalSignatureScheme.	1
4	KeyManagementandDistribution(6hours)	
4.1	SymmetricKeyDistributionUsingSymmetric&Asymmetric Encryption	1
4.2	SymmetricKeyDistributionUsingSymmetric&Asymmetric Encryption	1
4.3	DistributionofPublicKeys	1
4.4	Kerberos	1
4.5	X.509AuthenticationService	1
4.6	Public–KeyInfrastructure.	1
5	CaseStudiesofCryptography(7hours)	
5.1	Denialofserviceattacks,IPspoofingattacks	1
5.2	Secureinterbranchpaymenttransactions	1
5.3	ConventionalEncryptionandMessageConfidentiality	1
5.4	ConventionalEncryptionPrinciples	1
5.5	ConventionalEncryptionAlgorithms	1
5.6	ConventionalEncryptionAlgorithms	1
5.7	LocationofEncryptionDevices,KeyDistribution.	1

.

References:

1. WilliamStallings,CryptographyandNetworkSecurity–6thEdition,PearsonEducation, 2013.

2. BehrouzA.Forouzan,DebdeepMukhopadhyay,CryptographyandNetworkSecurity, SpecialIndian Edition, Mc Graw Hill Education, 2007.

3.CryptographyandNetworkSecurity:AtulKahate,McGrawHill,2ndEdition,2008.

4.CryptographyandNetworkSecurity:CKShyamala,NHarini,DrTRPadmanabhan, WileyIndia, 1st Edition, 2011.

5.InformationSecurity,Principles,andPractice:MarkStamp,WileyIndia, 2011. 6. Principles of Computer Security: WM. Arthur Conklin, Greg White, TMH, 2018.

7.IntroductiontoNetworkSecurity:NealKrawetz,CengageLearning,2007.

8.NetworkSecurityandCryptography:BernardMenezes,CengageLearning,2010.

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PADT244	BIG DATA ANALYTICS ON GENOMIC DATA	PROGRAMME ELECTIVE4	3	0	0	3

Preamble:This course helps the learners to provide practical/research solutions to problems inthedomainofBioinformatics.Itenablesthelearnerstounderstand.conceptsof Bioinformatics, Application of AI in Bioinformtics, Big Data Bioinformatics andDataAnalytics with NGS data. This course helps the learners to develop practical solutions to problems in bioinformatics.

CourseOutcomes:TheCOsshownareonlyindicative.Foreachcourse,therecanbe4to6 COs.

Afterthecompletionofthecoursethestudentwillbeableto

CO1	Analyse basic concepts of Biomolecules, Biological databases, Sequence characteristics (Cognitive Knowledge Level :Apply)
CO2	Apply ML/DL Model forRNA /Protein structure Analysis (Cognitive Knowledge Level :Apply)
CO3	ApplyBigdatatechniquesinBioinformatics(CognitiveKnowledgeLevel:Apply)
CO4	Comprehend the Data Analytics pipelines for NGS data (Cognitive Knowledge Level :Apply)
CO5	DesignandDevelopRNASeq/Chip/MetagenomicsseqPipelines(Cognitive Knowledge Level :Apply)
CO6	Develop/suggest a solution for anyresearch problems in the field of Bioinformatics (Cognitive Knowledge Level :Apply)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthe course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization ofthe program.Themasteryshould be at a level higher than the requirements in the appropriate bachelor program

PO4:Anabilitytoapplystreamknowledgetodesignordevelopsolutionsforrealworld problems by following the standards

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedtothe stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO 4	PO5	PO6	PO7
CO1	✓		✓	✓		✓	
CO2	✓		✓	✓		✓	
CO3	✓		✓	✓		✓	
CO4	✓		✓	✓	✓	✓	
CO5	✓			✓	✓	✓	
CO6	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	50-80%
Analyse	20-40%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration

.

100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluation shall only be based on application, analysis or design based questions (for bothinternal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15marks

iii. Testpaper (1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.

Coursebasedtask/test paperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module,having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity ofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldansweranyfive.Eachquestioncancarry 7 marks
Totaldurationoftheexaminationwillbe150minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

.

CourseLevelAssessmentQuestions:COMPUTERSCIENCEANDENGINEERING-CS2

CourseOutcome1(CO1):

1. ComprehendthenatureandscopeofBioinformatics
2. Articulatethedifferentbiomoleculesandthevariousdatabases
3. Comprehend the concept of sequence Alignments

Course Outcome 2 (CO2)

1ComprehendandapplyML/DLmodelsforBioinformaticsdata

2. ApplyLSTM/GANmodels
3. Comprehend Transformer based architecture

Course Outcome 3(CO3):

1. ComprehendtheBigdataconcept,challengesandthevarioustechniquesinit.
2. ApplyandAnalyseMapReduceconceptwithHadoopandSpark
3. Apply Spark SQL for scaling genomic data

Course Outcome 4 (CO4):

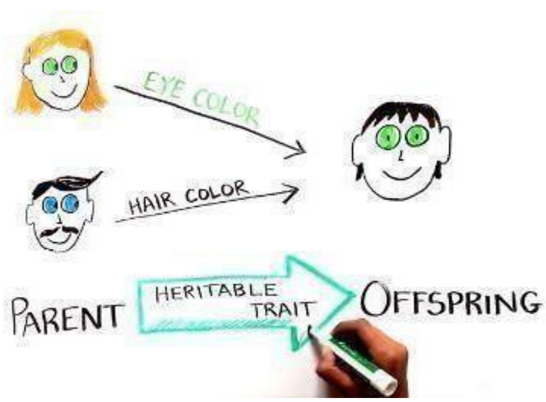
1. ComprehendtheunderstandingofNGSpipelines
2. ArticulateonDenovoassemblies

Course Outcome 5 (CO5):

1. DemonstrateUnderstandingofRNASeq/ChipseqPipelines
2. ApplytheprincipleofQIIME

ModelQuestionPaper
QPCODE:
Reg No: ___________________
Name:____ PAGES:4
APJABDULKALAMTECHNOLOGICALUNIVERSITY
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR
CourseCode: J1PADT244
CourseName:BIGDATAANALYTICSONGENOMICDATA
Max.Marks:60 Duration:2.5Hours
PARTA
AnswerAllQuestions.EachQuestionCarries5Marks

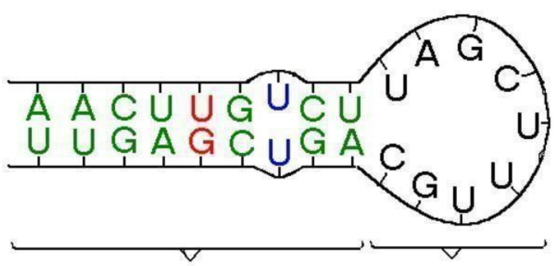
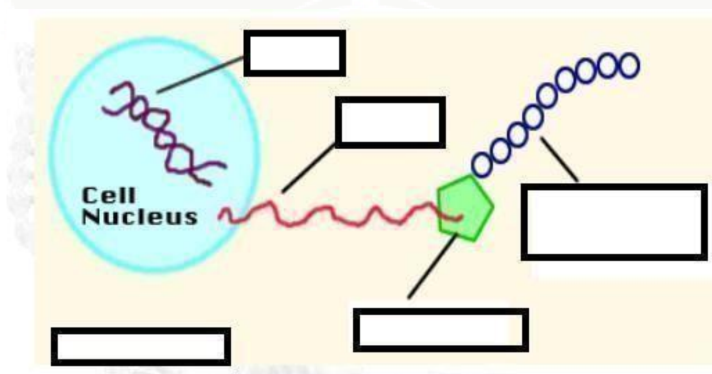
1\. (i) Studythecartoonandcommentontheheritabletraitthattransmittedfrom parent to offspring



(ii) AstructuralgenehastwoDNAstrandsandforthepartialsequenceofthe gene given below, which is the template strand and why?

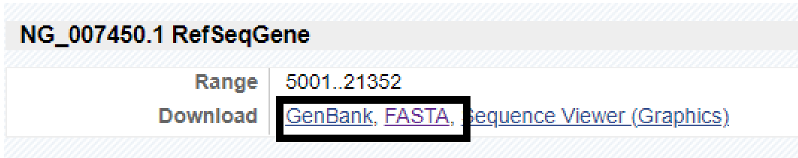
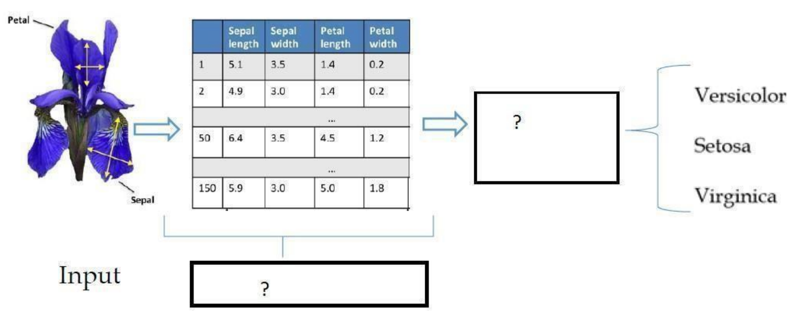
5‘ACTCGATGCTAG3‘
3‘TGAGCTACGATC5‘

.

2.		Bioinformatics is the application of computational technology to handle the rapidlygrowingrepositoryofinformationrelatedtomolecularbiology‖.Givea brief insight based on your knowledge on this definition of bioinformatics	
3.		Write short note on the need and application of Biological databases in sequence analysis :COMPUTERSCIENCEANDENGINEERIN	G-CS2
4.		WriteaPythonprogramusingPysparktoperformthefollowingtask. • Createadataframewiththefollowingattributesstudentid,name,Marks • Findthescholarswhosemark<50%usingsparksqlcomment • Findtheaverageandvarianceofthemarks • Findthemaximumandminimummarks	
5.		DiscussthestepsinvolvedintheMetagenomicsdataanalysispipeline.	(5x5=25)
		PartB (Answeranyfivequestions.Eachquestioncarries7marks)	
6.	(a)	ForthegivenRNAmoleculesecondarystructure,identifythe(i)Stem(ii) Loop and explain the importance of stem-loop structure of RNA. 	3marks
	(b)	Theproteinsynthesisingprocesstheshownbelow.Identifythedifferent stages and explain the role of mRNA?  	4marks

.

7.	(a)	Howdoyousubmitbiologicaldatatoapublicdatabase?Listthemajor submission tools in NCBI.	3marks

	(b)	FromNCBI,whiledownloadingagenesequence,youcandownloadthe sequenceintwoformats-GenBank& FASTA. To know more information about the sequence, which format will be useful? Justify your answer.  	4marks
8.	(a)	AlignAGCTCAGwithAGGTCAbyusingasuitablescoringscheme	5marks
	(b)	DifferentiatebetweenPairwiseandMultiplesequencealignment.	2marks
9.	(a)	Writeasimplepythonprogramforillustratingthemapreducemodelusedin Hadoop?	4marks
		ExplainwithexamplesthePageRankalgorithmusingMapReduce programming concept	3marks
10	(a)	Studythepictorialrepresentationgivenbelow.Commentontheunknown block and its relevance in machine learning  	3marks
	(b)	ExplainarchitectureofCNN.Estimatethenumberofparametersusing2 convolution layer model	4marks
11	(a)	Briefly explainthebigdatachallengesinthefieldofBioinformatics.	3marks
	(b)	DistinguishbetweenBowtieandBWA.	4marks
12	(a)	Whatisacontig?Whatisascaffold?Describedenovoassembly.Namea software used for de novo assembly.	3marks
	(b)	BrieflyexplainthepublicsequencedatabaseswhichsupportNextgeneration sequencing data and the various data formats being supported	4marks

.

:COMPUTERSCIENCEANDENGINEERING-CS2

Syllabus

Module1:IntroductiontoBioinformatics

Informationalviewoflifescience;IntroductiontoBiomoleculesthroughgames-Foldit (Protein overlapping), Eterna Game (RNA Structure), History of Bioinformatics, Biological databases,SequenceSimilarity,identityandhomology,ScoringMatices,Sequence Alignment, Phylogeny

Module2:AIinBioinformatics

AI applications in the field of genomics, the role of deep learning and data mining in computational biology and bioinformatics, ML/DL algorithms: ANN, CNN, LSTM, BERT, GAN, Protein structure prediction (deep neural networks), RNA structure prediction (deep learning models), RNA-protein binding sites prediction with CNN, Deep neural net to predict targetgeneexpression,TranscriptionFactorBindingviaMLP,LSTM,CNN,Protein Contextual Embeddings via BERT

Module3BigDataBioinformatics

OverviewofBigdata,DataStorageandAnalysis,Processing-SCVprinciple,BatchVs Stream processing, Big data Analytics- Typical Analytical Architecture -Types (Descriptive, Inquisitive,Predictive,Perspective);VisualisationandApplications,Computationalfacilities for analysing Big data – Cluster computing vs. Cloud computing, – Challenges in Big Data Analytics,BigdataFrameworks-MapReduce,HadoopandSpark,SparkSQLand dataframes, Spark for Bioinformatics, Big data analytics using Python- PySpark, Big data processing for DNA sequence analysis – PASTASpark

Module4DataAnalyticswithNGSdata

Introduction to next generation sequencing: NGS Platforms, NGS, advantages, limitations and applications, NGS Data sources: NCBI SRA, EBI-ENA, DDBJ-SRA; SRA toolkit; NGS Data analysis: FASTQ files, Quality check, Pre-processing, Mapping - Principles, tools - BWA, Bowtie, SAM tools -output file formats -BAM, SAM; Denovo assemblies - Principles, tools - SOAPdenovo, Velvet; Visualization tools - IGV; Whole Genome/Exome pipeline for Variant calling - VCF files.

Module5AdvancedDataAnalyticswithNGSdata

RNAseq - Gene expression analysis, Differential expression analysis. Alternative splicing - TopHat and Cufflinks for RNAseq, ChIPseq - Introduction and biological theories on ChIPseq analysis.DNAfragmentevaluation.Peakidentification.Twoconditioncomparison.
Saturationanalysis.Motiffinding andrelatedtheories,Metagenomicsanalysisusing QIIME and Picrust

Course Plan (For 3 credit courses, the content can be for 40 hrs and for 2credit courses, the contentcanbefor26hrs.Theauditcourseinthird semestercanhave contentfor30hours).

.

No	:COMTPopUiTcERSCIENCEANDENGINEER	No. of ILNeGc- tures(40hr s)
1	Module1:IntroductiontoBioinformatics	
1.1	Informationalviewoflifescience;Definition;DNA-RNAandProteinas information, Primary and secondary structure of DNA,	1
1.2	Chargaff's Rules, Different forms of DNA, RNA, Introduction to Biomoleculesthroughgames-Foldit(Proteinoverlapping),EternaGame (RNA Structure)	1
1.3	HistoryofBioinformatics,DefinitionofBioinformatics,Bioinformatics versus Computational Biology, Goals of Bioinformatics analysis,	1
1.4	Biologicaldatabases:-Fileformat,conversionoffileformat,Data retrieval system, Genome browsers. Biological data file formats	1
1.5	Basicconceptsofsequencesimilarity,identityandhomology,Scoring matrices- PAM and BLOSUM matrices,	
1.6	Dataretrievalsystem,Sequencedatabases-EMBL,GenBank,DDBJ;	
1.7	Proteindatabases-UniProt,ProteinDataBank.	
1.8	Conceptofsequencealignment-pairwiseandmultiplePairwise-Local and global, ,	1
1.9	Dotplot,BLAST,Multiplesequencealignment(MSA)–CLUSTAL Omega	
1.10	Phylogeny:Basicconceptsofphylogeny,Phylogenetictreeconstruction using MEGA.	1
2	Module2:AIinBioinformatics	
2.1	AIapplicationsinthefieldofgenomics,theroleofdeeplearningand data mining in computational biology and bioinformatics.	1
2.2	ML/DLalgorithms:ANN,CNN,LSTM,	1
2.3	ML/DLalgorithms:BERT,GAN	1
2.4	MLmodelforproteinexpression/Sequenceclassification,	1
2.5	Proteinstructureprediction(deepneuralnetworks),RNAstructureprediction (deep learning models),	1
2.6	RNA-proteinbindingsitespredictionwithCNN,Deepneuralnetto predict target gene expression	1
2.7	TranscriptionFactorBindingviaMLP,LSTM,CNN	1
2.8	ProteinContextualEmbeddingsviaBERT	1

.

3	Module3BigDataBioinformatics	
3.1	OverviewofBigdata–Definition,Characteristics,Sources,Types- Structured,Unstructured&Semi-structured;	1
3.2	DataStorageandAnalysis - NAS,DAS,NoSQLdatabases;Processing- SCVprinciple,BatchVsStreamprocessing	1
3.3	Big data Analytics- Typical Analytical Architecture – Requirement for new analytical architecture -Types (Descriptive, Inquisitive, Predictive, Perspective); Visualisation and Applications,	1
3.4	Computational facilities for analysing Big data – Cluster computing vs. Cloudcomputing,–ChallengesinBigDataAnalytics–Needofbigdata frameworks	1
3.5	BigdataFrameworks-MapReduce,HadoopandSpark,SparkSQLand dataframes,	1
3.6	SparkforBioinformatics,BigdataanalyticsusingPython-PySpark,	1
3.7	BigdataprocessingforDNAsequenceanalysis–PASTASpark	1
4	Module4DataAnalyticswithNGSdata	
4.1	Introductiontonextgenerationsequencing:NGSPlatforms,	1
4.2	NGStechnologies(WGS,ChIP-seq&RNA-seq),advantages,limitations and applications.	1
4.3	NGSDatasources:NCBISRA,EBI-ENA,DDBJ-SRA;	1
4.4	SRAtoolkit;NGSDataanalysis:FASTQfiles,Qualitycheck,Pre- processing,	1
4.4	Mapping-Principles,tools-BWA,Bowtie,SAMtools-outputfile formats - BAM, SAM;	1
4.5	Denovoassemblies-Principles,tools-	1
4.6	SOAPdenovo,Velvet;Visualizationtools-IGV;	1
4.7	WholeGenome/ExomepipelineforVariantcalling-VCFfiles.	1
5	Module5AdvancedDataAnalyticswithNGSdata	
5.1	RNAseq-Geneexpressionanalysis,.	1
5.2	Differentialexpressionanalysis.Alternativesplicing-	1
5.3	TopHatandCufflinksforRNAseq	1
5.4	ChIPseq-IntroductionandbiologicaltheoriesonChIPseqanalysis.	1
5.5	DNAfragmentevaluation.Peakidentification.Twocondition comparison. Saturation analysis.	1

.

5.6	Motiffindingandrelatedtheories	1
5.7	MetagenomicsanalysisusingQIIMEandPicrust	1

.

ReferenceBooks

- Brown,T.A.2002Genome.JohnWileyPress,US.
- Campbell,A.M.&Heyer,L.J.2002DiscoveringGenomics,Proteomicsand Bioinformatics. Benjamin/Cummings
- StuartM.Brown.Next-GenerationDNASequencingInformatics,SecondEdition. New York University School of Medicine (ISBN-13: 978-1621921236)
- XinkunWang.NextGenerationSequencingDataAnalysis,CRCPress.(ISBN13: 9781482217889).
- MuellerJ.P.&MassaronL.(2016),‖Machinelearningfordummies‖,JohnWiley&Sons.
- RussellS.J.&NorvigP.(2003),—ArtificialIntelligenceAModernApproach‖, Pearson Education.
- JordanM.I.&MitchellT.M.(2015),—Machinelearning:Trends,Perspectivesand Prospects.Science‖,349(6245),pp.255-260.(JournalArticle)
- Kamath,U.,Liu,J.,&Whitaker,J.(2019). *DeeplearningforNLPandspeech recognition* (Vol. 84). Cham: Springer.
- Gulli,A.,Kapoor,A.,&Pal,S.(2019).*DeeplearningwithTensorFlow2andKeras: regression, ConvNets,GANs, RNNs, NLP, and more with TensorFlow2 and theKeras API*. Packt Publishing Ltd.

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PADT254	PATTERN RECOGNITION	PROGRAMME ELECTIVE4	3	0	0	3

Preamble: This course aims to impart the fundamentals of statistical pattern recognition and neural network techniques. It introduces to the learner the various pattern recognition algorithms,featureselection,classification,clusteringandtheuseofneuralnetworksin feature extraction. This helps the learner to apply the algorithms in applications that works on pattern recognition and machine intelligence.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Applyprobabilityandnumericalmethodsinstatisticalpatternrecognition. (CognitiveKnowledge Level:Apply)
CO2	Applystatisticalmethodsinfeatureselection.(CognitiveKnowledgeLevel:Apply)
CO3	Applylinearalgebraandstatisticalmethodsinparameterandnon-parameter estimation.(CognitiveKnowledgeLevel:Apply)
CO4	Applythetechniqueofdecisiontreesinpatternrecognition.(CognitiveKnowledge Level:Apply)
CO5	Analyzetheuseofdeeplearningnetworksandartificialneuralnetworksinpattern recognition.(CognitiveKnowledgeLevel:Analyze)
CO6	Design,Develop,ImplementandPresentinnovativeideasinproblemsolvingwith variouspatternrecognitiontechniques.(CognitiveKnowledgeLevel:Create)

ProgramOutcomes(PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1:An ability to independently carry out research/investigation and developmentwork in engineering and allied streams

PO2:An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3:Anability todemonstrateadegreeofmastery overtheareaasperthespecializationof the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4:An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5:Anabilitytoidentify,selectandapplyappropriatetechniques,resourcesandstate-of- the-art tool to model, analyse and solve practical engineering problems.

PO6:An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:Anability todevelopcognitiveloadmanagementskillsrelatedtoprojectmanagement and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO 4	PO5	PO6	PO7
CO1	✓		✓	✓	✓	✓	
CO2	✓		✓	✓	✓	✓	
CO3	✓		✓	✓	✓	✓	
CO4	✓		✓	✓	✓	✓	
CO5	✓		✓	✓	✓	✓	
CO6	✓	✓	✓	✓	✓	✓	

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	70%-80%
Analyze	30%-40%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesign-basedquestions(forboth internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparing a review article based on peer reviewed original publications (minimum 10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15marks

iii. Testpaper (1number) :10marks

Testpapershallincludeminimum80%ofthesyllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module,having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity ofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldansweranyfive.Eachquestioncancarry 7 marks Totaldurationoftheexaminationwillbe150minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example, if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions Course

Outcome 1 (CO1):

1. Illustratethedesigncycleofapatternrecognitionsystemwiththehelpofanexample.
2. Suppose that we have three coloured bottles r (red), b(blue) and g(green).Box r contains 3 apples, 4 oranges and 3 limes. Box B contains 1 apple, 1 orange and 0 limes and box g contains 3 apples, 3 oranges and 4 limes. If a box is chosen at random with probability p(g)=0.2, p(b)=0.2 and p(g)=0.6 and piece of fruit is removed from the box(with equal probability of selecting items from the box), then what is the probability of selecting an apple?Ifweobservethattheselectedfruitisinfactanorange,whatistheprobability that it came from the green box?

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CourseOutcome2(CO2):

1. Illustratefeatureselectionusing t-Testwiththehelpofanexample.

CourseOutcome3(CO3):

1. DerivethefuzzyCsphericalshells(FCSS)algorithmforthecasethatsphericalclusters are to be identified.

CourseOutcome4(CO4):

1. Illustratedecisiontreewiththehelpofanexample.Howdoesitenablepattern classification?
2. Constructadecisiontreeusingthefollowingdata.

Outlook	Temp	Humidity	Windy	PlayGolf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Cloudy	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Cloudy	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Cool	Normal	True	Yes
Cloudy	Mild	High	True	Yes
Cloudy	Mild	Normal	False	Yes
Sunny	Hot	High	True	No

CourseOutcome5(CO5):

1. Howdoartificialneuralnetworksplayasignificantroleinpatternrecognition?Also discuss about its parameter optimisation techniques.

CourseOutcome6(CO6):

1. SupposeanaccidentproneareaisundersurveillanceandrealtimeCCTVvisualsare availabletoyou.Designasolutiontoautomaticallydetectaccidentsontheroadfrom

.

thoserealtimeCCTVvisuals.Explainaboutanyonepatternrecognitionalgorithmyou will make use here and how?

ModelQuestionPaper QP

CODE:

Reg No: ___________________

Name:____PAGES:2

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

Course Code: J1PADT254

CourseName:PatternRecognition

Max.Marks:60 Duration:2.5Hours

PARTA

AnswerAllQuestions.EachQuestionCarries5Marks

1. In a town it was estimated that 3% of people have a particular disease. Adiagnosis testwasconductedforallthepeople,whichyielded8%falsepositiveand92% true positive results. A person is found as positive after the test. What isthe probability that this person is truly having the disease? (5)

2. Howdoesmorphologicaloperationsplayaroleinpatternrecognition? (5)
3. Howcanvisualimagerybeanalysedusingconvolutionalneuralnetworks? (5)
4. Howdoesadecisiontreehandlecontinuousattributes? (5)
5. Definetheterms:weights,bias,activationswithrespecttoneuralnetworks (5)

PartB

(Answeranyfivequestions.Eachquestioncarries7marks)

6. Illustratethedesignprinciplesofpatternrecognitionsystemwithanexample. (7)

7. DerivethefuzzyCsphericalshells(FCSS)algorithmforthecasethatspherical clusters are to be identified. (7)

8. Show that in the case of Gaussian distributions the Chernoff bound becomes (7)

$$\epsilon_{CB} = \exp(-b(s))$$

where

$$b(s) = \frac{s(1-s)}{2}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T [s\Sigma_j + (1-s)\Sigma_i]^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + \frac{1}{2}\ln\frac{|s\Sigma_j + (1-s)\Sigma_i|}{|\Sigma_j|^s |\Sigma_i|^{1-s}}$$

Then take the derivative with respect to s and show that for equal covariance matrices the optimum is achieved for $s = 1/2$. Thus, in this case $b(s)$ equals the Bhattacharyya distance.

9. Let N_1, N_2 be the available values of a feature in two classes, respectively. The feature is assumed to follow a Gaussian distribution with the same variance in each class. Define the test statistic (7)

$$q = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{s_z\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

where

$$s_z^2 = \frac{1}{N_1 + N_2 - 2}\left(\sum_{i=1}^{N_1}(x_i - \bar{x})^2 + \sum_{i=1}^{N_2}(y_i - \bar{y})^2\right)$$

and μ_1, μ_2 are the respective true mean values. Show that q follows the t-distribution with $N_1 + N_2 - 2$ degrees of freedom.

10 Discussthesignificanceofpre-processinginfeatureselection.Illustrateany two methods used for pre-processing. (7)

11 Howcanartificialneuralnetworksbeappliedin Patternrecognition?Also illustrate the features of recurrent neural netwo. (7)

.

12Constructadecisiontreeusingthefollowingdata. (7)

Outlook	Temp	Humidity	Windy	PlayGolf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Cloudy	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Cloudy	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Cool	Normal	True	Yes
Cloudy	Mild	High	True	Yes
Cloudy	Mild	Normal	False	Yes
Sunny	Hot	High	True	No

Syllabus

Module 1: Introduction to Pattern Recognition

Basicsofpatternrecognitionsystems,variousapplications,MachinePerception, classification of pattern recognition systems.Design of Pattern recognition system, Pattern recognitionLifeCycle.StatisticalPatternRecognition:Reviewofprobabilitytheory, Gaussian distribution. Normal density and discriminant functions.

Module2:FeatureSelection

Feature selection – Outlier removal – Data normalization – Missing data, The Peaking phenomenon,Featureselectionusingstatisticalhypothesistesting-Hypothesistestingbasics – Application of t-Test in feature selection. Class separability measures-Divergence-Chernoff bound and Bhattacharya distance-Scatter matrices, Feature subset selection –Scalar feature selection, Feature vector selection.

Module3:ClusteringAlgorithms

Unsupervised learning and clustering - Criterion functions for clustering. Cluster validation. Fuzzyclusteringalgorithms-Pointrepresentatives-quadraticsurfacesandrepresentatives– hyperplanerepresentatives.Binarymorphologyclusteringalgorithms(BMCAs)– Discretization – Morphological operations - Determination of clusters in a discrete binary set-

Assignment of feature vectors to clusters – The algorithmic scheme, Boundary detection algorithms.

Module4:Dimensionalityreduction

Dimensionality reduction: Principal component analysis - its relationship to Eigen analysis. Fisher discriminant analysis - Generalised Eigen analysis. Eigen vectors/Singular vectors as dictionaries. Factor Analysis, Total variability space - a dictionary learning method.Non negative matrix factorisation - a dictionary learning method.
Lineardiscriminantfunctions:Gradientdescentprocedures,Perceptron.

Module5:ArtificialneuralnetworksandPatternClassification

Artificialneuralnetworks:ReviewofArtificialneuralnetworkconcepts,convolutional neural networks, recurrent neural networks.
Non-metric methods for pattern classification: Non-numeric data or nominal data. Decision trees: Classification and Regression Trees (CART).

CoursePlan

No.	Topic	No. of Lectures (40Hours)
1	Module1:IntroductiontoPattern Recognition	7
1.1	Basicsofpatternrecognitionsystems,applications	1
1.2	MachinePerception,Classificationofpatternrecognitionsystems	1
1.3	DesignofPatternrecognitionsystem	1
1.4	PatternrecognitionLifeCycle	1
1.5	StatisticalPatternRecognition	1
1.6	Reviewofprobabilitytheory	1
1.7	Normaldensityanddiscriminantfunctions	1
2	Module2:FeatureSelection	10
2.1	Featureselection–Outlierremoval	1
2.2	Datanormalization–Missingdata	1
2.3	Thepeakingphenomenon	1
2.4	Featureselectionusingstatisticalhypothesistesting	1
2.5	Hypothesistestingbasics–ApplicationoftTestinfeatureselection	1
2.6	Classseparabilitymeasures-Divergence	1
2.7	ChernoffboundandBhattacharyadistance	1
2.8	Scattermatrices	1
2.9	Featuresubsetselection–Scalarfeatureselection	1
2.10	Featurevectorselection	1
3	Module3:ClusteringAlgorithms	9

3.1	Unsupervisedlearningandclustering	1
3.2	Criterionfunctionsforclustering.Clustervalidation.	1
3.3	Fuzzyclusteringalgorithms-Pointrepresentatives	1
3.4	Quadraticsurfacesandrepresentatives–hyperplanerepresentatives.	1
3.5	Binarymorphologyclusteringalgorithms(BMCAs)	1
3.6	Discretization	1
3.7	Morphologicaloperations-Determinationofclustersinadiscrete binaryset	1
3.8	Assignmentoffeaturevectorstoclusters	1
3.9	Thealgorithmicscheme,Boundarydetectionalgorithms.	1
4	Module4:Dimensionalityreduction	8
4.1	Principalcomponentanalysis-itsrelationshiptoEigenanalysis	1
4.2	Fisherdiscriminantanalysis	1
4.3	GeneralisedEigenanalysis	1
4.4	Eigenvectors/Singularvectorsasdictionaries	1
4.5	Totalvariabilityspace-adictionarylearningmethod	1
4.6	Nonnegativematrixfactorisation-adictionarylearningmethod	1
4.7	Lineardiscriminantfunctions:Gradientdescentprocedures	1
4.8	Perceptron	1
5	Module5:ArtificialneuralnetworksandPattern Classification	6
5.1	ReviewofArtificialneuralnetworks,Introductiontodeepneural networks	1
5.2	Convolutionalneuralnetworks	1
5.3	Recurrentneuralnetworks	1
5.4	Non-metricmethodsforpatternclassification:Non-numericdataor nominaldata	1
5.5	Decisiontrees:ClassificationandRegressionTrees(CART)lecture1	1
5.6	Decisiontrees:ClassificationandRegressionTrees(CART)lecture2	1

References

1. S.TheodoridisandK.Koutroumbas,—PatternRecognition‖,4thEd.,AcademicPress, 2009

2. C.M.Bishop,—PatternRecognitionandMachineLearning‖,Springer,2006

3. R.O.Duda,P.E.HartandD.G.Stork,—PatternClassification‖,JohnWiley,2001

4. Hastie,T.,Tibshirani,R.andFriedman,J.—TheElementsofStatisticalLearning‖. Springer.2001.

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PADT264	ADVANCED COMPUTATIONAL LINGUISTICS	PROGRAMME ELECTIVE4	3	0	0	3

Preamble: To familiarize with theconcepts in computational linguistics, modern grammar formalismsandNLgeneration.ItcoversTAG,Statisticalparsing,NLG,Discourse processing,LexicalFunctionalGrammar(LFG),Morphotactics,FiniteStateTransducersand itsuseinmorphology.Studentswillbeabletounderstandthebasicsof TAG,statistical parsing, Logical forms, LFG, well-formedness of formalisms and apply them to Natural Language Generation(NLG) and Parsing

CourseOutcomes:Afterthecompletionofthecourse,thestudentwillbeableto

CO1	UnderstandthebasicconceptsofStatisticalParsingandSemanticKnowledge Representation. (Cognitiveknowledgelevel:Understand)
CO2	MakeuseoftheconceptsofSemanticInterpretationforproblems(Cognitive knowledgelevel:Apply)
CO3	ApplytheconceptsofSpeechUnderstandinginNLproblems.(Cognitiveknowledge level:Apply)
CO4	ApplythefundamentalsofLexicalFunctionalGrammars(Cognitiveknowledgelevel: Apply)
CO5	ApplytheconceptsofMorphologicalParsinginreallifeproblems.(Cognitive knowledgelevel:Apply)

ProgramOutcomes(PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3:An ability todemonstrate a degree ofmastery over the areaas per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7: An ability to develop cognitive load management skills related to project managementand finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1			✓	✓		✓	
CO2			✓	✓		✓	
CO3	✓		✓	✓		✓	
CO4	✓		✓	✓		✓	
CO5	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	40
Analyse	20
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesignbasedquestions(forboth internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15marks

iii. Testpaper(1number) 10 marks

Testpapershallincludeminimum80%ofthesyllabus.
Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

The end semester examination will be conducted by the respective College.
There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module,having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity ofthe students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldansweranyfive.Eachquestioncancarry 7 marks
Totaldurationoftheexaminationwillbe150minutes.
Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Givethedetailsofdependencygrammar.

2. HowdoesSemanticsrelatetoknowledgerepresentation?WhataretheusesofSemantic- Based Knowledge Representation?

CourseOutcome2(CO2)

1. Differentiatebetweensyntacticandsemanticgrammar.

2. Whatiscompositionalsemantics?Explainthesignificanceofdeepstructure.

CourseOutcome3(CO3):

1. How does Natural Language Generation (NLG) works?What are theapplication areas ofNLG?

2. DescribeNaturalLanguageGenerationtechniques.

.

CourseOutcome4(CO4): :COMPUTERSCIENCEANDENGINEERING-CS2

1. DescribetheLexicalFunctionalGrammar.
2. Whatarethewell-formednessconditions?

CourseOutcome5(CO5):

1. Whatismorphotactics?Describetheimportanceofmorphotactics?
2. HowcanweuseLexicon-freeFinite-StateTransducersforstemming?

ModelQuestionPaper			
QPCODE:			
Reg No: ___________________			
Name: ____ PAGES:4			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode: J1PADT264			
CourseName:ADVANCEDCOMPUTATIONALLINGUISTICS			
Max.Marks:60 Duration:2.5 Hours			
PARTA			
AnswerAllQuestions.EachQuestionCarries5Marks			
1.	DifferentiateTreeadjoiningGrammar(TAG)fromContextfreeGrammar(CFG)		
2.	Illustratesemanticknowledgerepresentationwithasuitableexample		
3.	Howeffectively,dependencygrammardoesstatisticalparsing?		
4.	Illustratehow*whmovement*canbehandledinquestions.		
5.	Whatismorphologicalparsing?Explainwithasuitableexample.		
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			
6.	(a)	Howdependencygrammarsareusefulinsyntacticparsing?Explain.	(4)
	(b)	IllustratethederivationusingTAGs	(3)

7.	(a)	Explaincaseframesandcaseframebasedparsing	(4)
	(b)	Explainhowsemanticprocessingcontributestolanguagemodelling	(3)
8.	(a)	DifferentiateNLGfromNLUandthechallengesinthem.	(4)
	(b)	Explaindiscourseprocessingwithanexample	(3)
9.	(a)	Differentiateinflectionalmorphologyfromdativemorphology	(4)
	(b)	ExplainActive,passive,anddativeconstructions.	(3)
1		Whatisafinitestatetransducer?Howisitusefulinmorphological analysis?	(7)
11	(a)	ExplainDiscourseProcessinganditssignificance	(4)
	(b)	ExplainthetechniquesusedinNLG	(3)
1	(a)	ExplainhowmorphologicalparsingisdoneusingFST	(4)
	(b)	Explainthesignificanceoflexiconsandmorphotacticsinmorphologicalparsing	(3)

Syllabus

	Syllabus	
Module	Contents	Hours
I	Tree Adjoining Grammars-Dependency Grammars Statistical Parsing- Introduction to Semantic Processing, Semantic Knowledge Representation	8
II	Deep Structure and Logical Form-Compositional Semantic Interpretation. Semantic Grammars-Case Frames and Case Framebased Parsing	7
III	Natural Language Generation- Problems in NL Generation Basic Generation Techniques-Hard Problems in NLP Speech Understanding and Translation-Discourse Processing	8

IV	Lexical Functional Grammar: Active-Passive and DativeConstructions-Wh-movementinQuestions-Overviewof LFG-LFG Formalism: Well-formedness Conditions Handling Wh-movement in Questions-Computational Aspects	7

V	MorphologyandFiniteStateTransducers: Inflectional Morphology - Derivational Morphology-Finite State Morphological Parsing-The Lexicon and Morphotactics MorphologicalParsingwithFiniteStateTransducers Orthographic Rules and Finite-State Transducers-Combining an FST Lexicon and Rules-Lexicon free FSTs	9

Course Plan (For 3 credit courses, the content can be for 40 hrs and for 2 credit courses, the contentcanbefor26hrs.Theauditcourseinthethirdsemestercanhavecontentfor30 hours).

No	Topic	No.of Lectures(40)
1	Module1(SemanticProcessing)	
1.1	TreeAdjoiningGrammars	1
1.2	DependencyGrammars	1
1.3	CategorialGrammars	1
1.4	StatisticalParsing	1
1.5	ApplicationsofStatisticalParsing	1
1.6	IntroductiontoSemanticProcessing	1
1.7	IntroductiontoSemanticKnowledgeRepresentation	1
1.8	SemanticKnowledgeRepresentation	1
2	Module2(SemanticInterpretation)	
2.1	DeepStructure	1
2.2	LogicalForm	1
2.3	CompositionalSemanticInterpretation	1
2.4	SemanticGrammars	1
2.5	DefiniteClauseGrammar(DCG)	1
2.5	CaseFrames	1
2.6	IntroductiontoFramebasedParsing	1
2.7	FramebasedParsing	1
3	Module3(SpeechUnderstandingandTranslation)	
3.1	NaturalLanguageGeneration	1
3.2	ProblemsinNLGeneration	1
3.3	BasicGenerationTechniques	1
3.4	IntroductiontoHardProblemsinNLP	1

3.5	HardProblemsinNLP	1
3.6	IntroductiontoSpeechUnderstandingandTranslation	1
3.7	SpeechUnderstandingandTranslationExplanation	1
3.8	DiscourseProcessing	1
4	Module4(LexicalFunctionalGrammar)	
4.1	LexicalFunctionalGrammar	1
4.2	Active-PassiveandDativeConstructions	1
4.3	Wh-movementinQuestions	1
4.4	OverviewofLFG-LFGFormalism	1
4.5	Well-formednessConditions	1
4.6	HandlingWh-movementinQuestions	1
4.7	ComputationalAspects	1
5	Module5(MorphologicalParsing)	
5.1	Morphology	1
5.2	FiniteStateTransducers	1
5.3	InflectionalMorphology	1
5.4	DerivationalMorphology	1
5.5	FiniteStateMorphologicalParsing	1
5.6	TheLexiconandMorphotactics	1
5.7	MorphologicalParsingwithFiniteStateTransducers	1
5.8	OrthographicRulesandFinite-StateTransducers	1
5.9	CombininganFSTLexiconandRules-LexiconfreeFSTs	1

ReferenceBooks

1. Jurafsky, D. and J. H. Martin, Speech and language processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Prentice-Hall,2008.

2. Alexander Clark, Chris Fox, and Shalom Lappin (Editors): The Handbook ofComputationalLinguisticsandNaturalLanguageProcessing(Wiley-BlackwellHandbooks in Linguistics), 2010.

3. Akshar Bharathi, Vineet Chaitanya, and Rajeev Sangal: Natural Language Processing:APaninian Perspective. Prentice Hall of India, 1995.

4. JamesAllen:NaturalLanguageUnderstanding.Pearson,2002.

SEMESTERII

INTERDISCIPLINARY ELECTIVE

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PCSO215	INTRODUCTION TO MACHINE LEARNING	INTERDISCIPLINARY ELECTIVE	3	0	0	3

Preamble: This course helps the learners to understand the concepts in Machine Learning. Studentswillbeabletounderstandthebasicsofregression,classificationandclustering. After completing this course students will be able to develop machine learning based solution for real world problems in multidisciplinary environments.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Illustrate the concept, purpose, scope, steps, and applications of ML techniques. (Knowledge level : Apply)
CO2	Understand the concepts of supervised, unsupervised and reinforcement learning to apply in real world problems. (Knowledge level : Apply)
CO3	Illustrate the working of classifiers and clustering techniques for typical machine learning applications. (Knowledge level : Apply)
CO4	Acquire skills to improve the performance of Machine Learning models using ensemble techniques. (Knowledge level : Apply)
CO5	Design and Implement solution for a real world problem using Machine Learning algorithms (Cognitive Knowledge Level: Create)

ProgramOutcomes(PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing the course.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3:An ability todemonstrate a degree ofmastery over the areaas per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5:Anability toidentify,selectandapply appropriate techniques,resourcesand state-of- the-art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:An abilitytodevelopcognitive load managementskills related to project management and finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1							
CO2	✓		✓	✓	✓		
CO3	✓		✓	✓	✓		
CO4	✓		✓	✓	✓		
CO5	✓	✓	✓	✓	✓	✓	✓

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	80%
Analyse	20%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:
ContinuousInternalEvaluation:40marks
Microproject/Coursebasedproject:20marks
Coursebasedtask/Seminar/Quiz:10 marks
Testpaper,1no. :10marks
Theprojectshallbedoneindividually.Groupprojectsnotpermitted. Test paper shall include minimum 80% of the syllabus.
Coursebasedtask/testpaperquestionsshallbeusefulinthetestingofknowledge,skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:
Total :60marks
TheendsemesterexaminationwillbeconductedbytherespectiveCollege. There will be two parts; Part A and Part B.
Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of

the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one questionfromeachmoduleofwhichstudentshouldanswerany five.Eachquestioncancarry 7 marks Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevelAssessmentQuestions

Course Outcome 1 (CO1):

1. Suppose 10000 patients get tested for flu; out of them, 9000 are actually healthy and 1000 areactuallysick.Forthesickpeople,atestwaspositivefor620andnegativefor380.Forthe healthy people, the sametestwas positive for 180and negative for 8820. Constructa confusion matrix for the data and compute the precision and recall for the data.

2. DistinguishbetweensupervisedlearningandReinforcementlearning.Illustratewithan example.

3. Discussanyfourexamplesofmachinelearningapplications.

CourseOutcome2(CO2)

1. StatethemathematicalformulationoftheSVMproblem.Giveanoutlineofthemethodfor solving the problem.

2. Showthefinalresultofhierarchicalclusteringwithcompletelinkbydrawinga dendrogram.

CourseOutcome3(CO3):

1. IdentifythefirstsplittingattributeforthedecisiontreebyusingtheID3algorithmwith the following dataset.

.

2. Considerthe training data in the following table where Play is a class attribute. In the table,theHumidityattributehasvalues—L‖(forlow)or—H‖(forhigh),Sunnyhasvalues —Y‖ (for yes) or —N‖ (for no), Wind has values —S‖ (for strong) or —W‖ (for weak), and Play has values —Yes‖ or —No‖.

Whatistheclasslabelforthefollowingday(Humidity=L,Sunny=N,Wind=W),according to naïve Bayesian classification?

3. ExplainDBSCANalgorithmfordensitybasedclustering.Listoutitsadvantages compared to K-means.

4. ExplainhowSupportVectorMachinecanbeusedforclassificationoflinearlyseparable data.

5. DefineHiddenMarkovModel.Whatismeantbytheevaluationproblemandhowisthis solved?

6. UseK Means clustering to cluster the following data into two groups. Assume cluster centroidarem1=2andm2=4.Thedistancefunctionusedis Euclideandistance.{2,4,10, 12, 3, 20, 30, 11, 25 }

CourseOutcome4(CO4):

1. Explainhowthe*RandomForests*giveoutputfor*Classification*,and*Regression* problems?
2. Is*RandomForest*an*EnsembleAlgorithm*

3. WhyisthetrainingefficiencyofRandomForestbetterthanBagging?

.

:COMPUTERSCIENCEANDENGINEERING-CS2

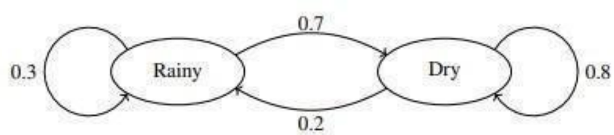
ModelQuestionPaper

Reg No: ___________________			
Name: _____________________ PAGES: 4			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode: J1PCSO215			
CourseName:INTRODUCTIONTOMACHINELEARNING			
Max.Marks:60 Duration:2.5 Hours			
PARTA			
AnswerAllQuestions.EachQuestionCarries5Marks			
1.	Bias-Variancetrade-offisadesignconsiderationwhiletrainingthemachine learningmodel.Justify.		
2.	DerivetheexpressionforsigmoidfunctionassociatedwithLogistic Regression.		
3.	How Optimal Marginal Hyperplane contributes to the accuracy of predictions using SVM. Justify how Kernel functions are used in Linear Inseparable problems		
4.	DiscusshowgoodDBSCANisinclusteringdatapointsavailableindense Euclidean space.		
5.	Suggestanensemblemethodthatgeneratesoneclassifierperround		(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			
6.		ExplainvariousCostFunctionsassociatedwithRegression&Classification.	(7)
7.		WeightupdationcontributestotheperformanceofaNeuralNetwork model. Justify the statement using the Back propagation algorithm.	(7)
8.		Compute the Principal Components for the 2D data: X=(x1,x2)={(1,2),(3,3),(3,5),(5,4),(5,6),(6,5),(8,7),(9,8)}	(7)
9.		Using Naïve Bayes algorithm, predict whether aRed color car whichis imported as a Sports category will be stolen or not.	(7)

Color	Type	Origin	Stolen?
Red	Sports	Domestic	Yes
Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

1		ConstructDendrogramsbasedonCompleteLinkageandAverage Linkage.(7)	(7)
1		Performk-meansalgorithmonthedatagiveninqn.8. (Given no. of clusters =2, iterations=2).	(7)
1	(a)	IfP(Rain)=0.4andP(Dry)=0.6computetheprobabilityforthe sequence—Rain,Rain,Dry,Dry‖.	(3)
	(b)	ElucidatethethreebasicproblemsofHMM	(4)

ID	1	2	3	4	5
1	0	3	18	10	25
2	3	0	21	13	28
3	18	21	0	8	7
4	10	13	8	0	15
5	25	28	7	15	0



	Syllabus	
Module	Contents	Hours
I	Overview of machine learning: supervised, semi-supervised, unsupervisedlearning,reinforcementlearning.TypesofML problems: Classification, Clustering and Regression, Cost functions: DefinitionandTypes,DataPreProcessing,Bias-Variancetradeoff, Cross validation techniques,Classifier performance measures, ROC Curves	6
I	Introduction to neural network : Linear Regression, Least square Gradients, Logistic Regression, Sigmoid function & differentiation,LogisticRegression–Regularization,Neural Networks–ConceptofperceptronandArtificialneuron,Weight initialization techniques,Feed Forward Neural Network, Back Propagation algorithm	8
III	Classification Methods : Support Vector Machine, Optimal Separatinghyperplane,Kerneltrick,Kernelfunctions,Gaussian classconditionaldistribution,BayesRule,NaïveBayesModel, Decision Tree – ID3, Maximum Likelihood estimation techniques	8

.

IV	Clustering Methods: K-means clustering , Hierarchical clustering techniques, Density Based clustering, Feature Selection techniques: Entropy,CorrelationCoefficient,Chi-squareTest,Forward& Backward Selection,Dimensionality Reduction: PCA,LDA,t- SNE	7
V	Basics of graphical models - Bayesian networks, Hidden Markov model, Ensemble methods – Boosting, Bagging, Random forest, XGBoost (Case study)	6

LessonPlan

1	Introductiontomachinelearning(Hours: 6)	
1.1	Overview of machine learning: supervised, semi-supervised, unsupervised learning, reinforcement learning	1
1.2	TypesofMLproblems:Classification,ClusteringandRegression	1
1.3	Costfunctions:DefinitionandTypes	1
1.4	DataPreProcessing,Bias-Variancetradeoff	1
1.5	Crossvalidationtechniques	1
1.6	Classifierperformancemeasures,ROCCurves	1
2	Introductionto neuralnetwork(Hours:8)	
2.1	LinearRegression,LeastsquareGradients	1
2.2	LogisticRegression	1
2.3	Sigmoidfunction&differentiation	1
2.4	LogisticRegression-Regularization	
2.5	NeuralNetworks–ConceptofperceptronandArtificialneuron	1
2.6	Weightinitializationtechniques	1
2.7	FeedForwardNeuralNetwork	1
2.8	BackPropagationalgorithm	1
3	ClassificationMethods(Hours:8)	
3.1	SupportVectorMachine	1
3.2	OptimalSeparatinghyperplane	1
3.3	Kerneltrick,Kernelfunctions	1
3.4	Gaussianclassconditionaldistribution	1
3.5	BayesRule	1
3.6	NaïveBayesModel	1
3.7	DecisionTree–ID3,	1
3.8	MaximumLikelihoodestimationtechniques	1
4	ClusteringMethods(Hours:7)	
4.1	K-meansclustering	1
4.2	Hierarchicalclusteringtechniques	1
4.3	BIRCH	1
4.4	DensityBasedclustering	1
4.5	Feature Selection techniques: Entropy, Correlation Coefficient, Chi-square Test	1
4.6	Forward&BackwardSelection	1
4.5	DimensionalityReduction:PCA	1
4.6	LDA	1

.

4.7	t-SNE	1
5	Basicsofgraphicalmodels(Hours:6)	
5.1	Basicsofgraphicalmodels-Bayesiannetworks	1
5.2	HiddenMarkovmodel	1
5.3	Ensemblemethods-Boosting	1
5.4	Bagging	1
5.5	Randomforest	1
5.6	XGBoost(Casestudy)	1

ReferenceBooks

1. EthemAlpaydın,―IntroductiontoMachineLearning(AdaptiveComputationandMachine Learning)‖,MITPress,2004.

2. KevinMurphy,MachineLearning:AProbabilisticPerspective(MLAPP),MITPress, 2012

3. Han,Jiawei,andMichelineKamber.DataMining:ConceptsandTechniques.San Francisco: Morgan Kaufmann Publishers

4. ChristopherM.Bishop,―PatternRecognitionandMachineLearning‖,Springer,2006

.

CODE	COURSE NAME	CATEGORY		T	P	CREDIT
J1PCS0225	DATA STRUCTURES	INTERDISCIPLINARY ELECTIVE		0	0	3

Preamble: The purpose of the syllabus is to create awarenessabout Data Structures and their applications. After the completion of the course, the learners should be able to either use existing data structures or design their own data structures to solve real world problems.

CourseOutcomes:Afterthecompletionofthecoursethestudentwillbeableto

CO1	Designalgorithmsforataskandcalculatethetimecomplexityofthatalgorithm (Cognitive Knowledge Level: Apply)
CO2	Usearraysandlinkedlistsforproblemsolving(CognitiveKnowledgeLevel:Apply)
CO3	Representdatausingtrees,graphsandmanipulatethemtosolvecomputational problems. (Cognitive Knowledge Level:Apply)
CO4	Makeuseofappropriatesortingalgorithmstoorderdatabasedonthesituation. (CognitiveKnowledgeLevel:Apply)
CO5	DesignandImplementappropriateDataStructuresforsolvingarealworldproblem (Cognitive Knowledge Level: Create)

ProgramOutcomes(PO)

Outcomesaretheattributesthataretobedemonstratedbyagraduateaftercompletingthecourse.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the-art tool to model, analyse and solve practical engineering problems.

PO6:Anabilitytoengageinlife-longlearningforthedesignanddevelopmentrelatedtothe streamrelatedproblemstakingintoconsiderationsustainability,societal,ethicaland environmental aspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagementand finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourse outcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✔		✔	✔		✔	
CO2	✔		✔	✔		✔	
CO3	✔		✔	✔		✔	
CO4	✔		✔	✔		✔	
CO5	✔	✔	✔	✔		✔	✔

AssessmentPattern

Bloom‘sCategory	End SemesterExamination
Apply	70%
Analyse	30%
Evaluate	
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
00	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluationshallonlybebasedonapplication,analysisordesignbasedquestions(forbothinternal and end semester examinations).

.

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15 marks

iii. Testpaper(1number) :10 marks

Testpaper shallincludeminimum80%ofthesyllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:
TheendsemesterexaminationwillbeconductedbytherespectiveCollege.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 question from each module, having 5 marks for each question. Students should answer all questions.Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the students in a course, through long answer questions relating to theoretical/practical knowledge, derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.
For example if the average end semester mark % for a core course is 40, then the maximum eligible mark % for an elective course is 40+20 = 60 %.

SampleCourseLevelAssessmentQuestions

Course Outcome1(CO1):

1. Writeanalgorithmformatrixmultiplicationandcalculateitstimecomplexity.
2. DefineBig-Onotation.DerivetheBig–Onotationfor$5n^3+2n^2+3^n$.
3. Checkwhetherthefollowingistrueornot.
   2n+1 is O(2n). Give reason.

CourseOutcome2(CO2):

1. Howalinkedlistcanbeusedtorepresentthepolynomial

.

$5x^4y^6+24x^3y^4-17x^2y^3+15xy^2+45$.
WriteaproceduretoaddtwoBivariatepolynomialsrepresentedusinglinkedlists.

2. Write an algorithm/pseudocode to convert a given infix expression to postfix expression. Trace the steps involved in converting the given infix expression ((A +B)^C)-((D*C)/F)topostfixexpression.

3. LetL1beasinglylinkedlistinmemory.Writeanalgorithmthat
   i) FindsthenumberofnonzeroelementsinL1
   ii) AddsagivenvalueKtoeachelementinL1

Course Outcome 3(CO3):

1. CreateaBinaryTreewiththefollowingsequence14,15,4,18,9,16,20,17,3,7,5, 2 and perform inorder, preorder and postorder traversals on the above tree and print the output of the traversals.
2. Inacompletebinarytreeofdepthd(completeincludinglastlevel),giveanexpression to find the number of leaf nodes.

CourseOutcome4(CO4):

1. Writeanalgorithm/pseudocodetosortelementsusingHeapsorttechnique.Illustratethe working of Heap sort algorithm on the following input : 35,15,0,1,60

2. With the help of an algorithm/pseudocode and suitable example, explain how you would performbinarysearchonanarrayofnelements.Findthetimecomplexityofbinarysearch algorithm.

3. Supposeanarraycontainselements{10,13,21,32,35,44,55}.Givethestepstofindan element—35‖using i)linearsearch ii)binary search

CourseOutcome5(CO5):

Designareservationsystemforrailwaysthatincludesawaitinglist.Ifthereservationis full, display —reservation full‖ and put them in the waiting list and give a waiting list number. If a passenger wishes to cancel his ticket, he may do it any time. Then the passenger at the front of the waiting list is allotted a berth automatically.

.

:COMPUTERSCIENCEANDENGINEERING-CS2

ModelQuestionPaper

Reg No: _________________			
Name:_______________________________ PAGES:4			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode: J1PCSO225			
CourseName:DataStructures			
Max.Marks:60 Duration:2.5 Hours			
PARTA			
AnswerAllQuestions. EachQuestionCarries5 Marks			
1.	Writeanalgorithmtoaddanewelementinaparticularpositionofanarray.		
2.	CompareCircularQueuewithaNormalQueue		
3.	Howcanadoublylinked listbeusedtofindpalindromes?		
4.	Writeaniterativealgorithmforin-ordertraversal ofaBinaryTree		
5.	TracetheworkingofQuicksortonthefollowinginput38,8,0,28,45,-12,89,66,42.		(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			
6.	(a)	Howistheperformanceofanalgorithmevaluated?	(4)
	(b)	InthefunctionsO(nlogn)andO(logn),whichoneisbetterintermsof computational complexity and why?	(3)
7.	(a)	Writealgorithmstoinsertanddeleteelementsfromadoubleendedqueue. Illustrate with examples	(7)
8.	(a)	Writeanalgorithmtomultiplytwopolynomialsrepresentedusinglinkedlist	(7)
9.	(a)	ListthepropertiesofBinarySearchTree.Writeanalgorithmtosearchan element in a Binary Search Tree	(7)
10		GivealgorithmsforDFSandBFSofagraphand explainwithexamples.	(7)
11	(a)	WriteandillustratealgorithmsforMergeSortandQuickSort.	(7)
12	(a)	Howismemoryde-allocationdoneinmemorymanagement	(7)

.

	Syllabus	
Module	Contents	Hours
I	BasicConceptsofDataStructures SystemLifeCycle,Algorithms,PerformanceAnalysis,Spacecomplexity, TimeComplexity,AsymptoticNotation,ComplexityCalculationof Simple Algorithms	5
II	ArraysandSearching PolynomialrepresentationusingArrays,Sparsematrix,Stacks,Queues, CircularQueues,PriorityQueues,DoubleEndedQueues,Evaluationof Expressions,LinearSearchandBinarySearch	9
III	LinkedListandMemoryManagement Self Referential Structures, Dynamic Memory Allocation, Singly Linked List-OperationsonLinkedList,DoublyLinkedList,CircularLinkedList, Stacks and Queues using Linked List, Polynomial representation using Linked List,Memory allocation and de-allocation,	8
IV	Trees,BinaryTrees-TreeOperations,BinaryTreeRepresentation,Tree Traversals,Graphs- Representation of Graphs, Depth First Search and Breadth First Search on Graphs, Applications of Graphs	7
V	SearchingTechniques–Linearsearch,BinarysearchSortingTechniques – SelectionSort,InsertionSort,QuickSort,MergeSortandHeapSort	5

CoursePlan

Module1:BasicConceptsofDataStructures		(5hours)
1.1	SystemLifeCycle,	1hour
1.2	Algorithms	1hour
1.3	PerformanceAnalysis,SpaceComplexity,TimeComplexity,	1hours
1.4	AsymptoticNotation	1hour
1.5	ComplexityCalculationofSimpleAlgorithms	1hour
Module2:ArraysandSearching		(9hours)
2.1	PolynomialrepresentationusingArrays	1hour
2.2	Sparsematrix	1hours

.

2.3	Stacks	1hour
2.4	Queues,CircularQueues	1hour
2.5	PriorityQueues,	1hour
2.6	DoubleEndedQueue	1hours
2.7	EvaluationofExpressions	1hour
2.8	LinearSearch	1hour
2.9	BinarySearch	1hour
Module3: LinkedListandMemory Management		(8hours)
3.1	SelfReferentialStructures	1hour
3.2	DynamicMemoryAllocation	1hour
3.3	SingleLinkedList-OperationsonLinkedList,	1hour
3.4	DoubleLinkedList	1hour
3.5	CircularLinkedList	1hour
3.6	StacksandQueuesusingLinkedList	1hour
3.7	PolynomialrepresentationusingLinkedList	1hour
3.8	Memoryde-allocation	1hour
Module4:TreesandGraphs		(7hours)
4.1	Trees,BinaryTrees	1hour
4.2	TreeOperations,BinaryTreeRepresentation,	1hour
4.3	TreeTraversals	1hour
4.4	Graphs	1hour
4.5	RepresentationofGraphs	1hour
4.6	DepthFirstSearchandBreadthFirstSearchonGraphs	1hour
4.7	ApplicationsofGraphs	1hour
Module5: SortingandHashing		(5hours)
5.1	SortingTechniques–SelectionSort	1hour
5.2	InsertionSort	1hour
5.3	QuickSort	1hour
5.4	MergeSort	1hour
5.5	HeapSort	1hour

.

TextBook

1. EllisHorowitz,SartajSahniandSusanAnderson-Freed,UniversitiesPress, Fundamentals of Data Structures in C

ReferenceBooks

1. SamantaD.,ClassicDataStructures,PrenticeHallIndia,2/e,2009.
2. RichardF.Gilberg,BehrouzA.Forouzan,DataStructures:APseudocodeApproach with C, 2/e, Cengage Learning, 2005
3. AhoA.V.,J.E.HopcroftandJ.D.Ullman,DataStructuresandAlgorithms, Pearson Publication,1983.
4. TremblayJ.P.andP.G.Sorenson,IntroductiontoDataStructureswithApplications, Tata McGraw Hill, 1995.
5. PeterBrass,AdvancedDataStructures,CambridgeUniversityPress,2008
6. LipschutsS.,TheoryandProblemsofDataStructures,Schaum‘sSeries,1986.
7. WirthN.,Algorithms+DataStructures=Programs,PrenticeHall,2004.
8. HuggesJ.K.andJ.I.Michtm,AStructuredApproachtoProgramming,PHI,1987.
9. MartinBarrett,CliffordWagner,AndUnix:ToolsForSoftwareDesign,JohnWiley, 2008 reprint

.

:COMPUTERSCIENCEANDENGINEERING-CS2

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PCSO235	SOFTWARE PROJECT MANAGEMENT	INTERDISCIPLINARY ELECTIVE	3	0	0	3

Preamble: This course provides fundamental knowledge in the Software Development Process.It covers Software Development, Quality Assurance, Project Management concepts and technology trends. This course enables the learners to apply state of the art industry practices in Software development.

CourseOutcomes:Afterthecompletionofthecoursethe studentwillbeableto

CO1	DemonstrateTraditionalandAgileSoftwareDevelopmentapproaches(Cognitive Knowledge Level: Apply)
CO2	PrepareSoftwareRequirementSpecificationandSoftwareDesignforagivenproblem. (Cognitive Knowledge Level: Apply)
CO3	Justify the significance of design patterns and licensing terms in software development, prepare testing, maintenance and DevOps strategies for a project. (Cognitive Knowledge Level: Apply)
CO4	Make use of software project management concepts while planning, estimation, scheduling, tracking and change management of a project, with a traditional/agile framework. (Cognitive Knowledge Level: Apply)
CO5	Utilize SQA practices, Process Improvement techniques and Technology advancements in cloud based software models and containers & microservices. (Cognitive Knowledge Level: Apply)

ProgramOutcomes(PO)

Outcomes are the attributes that are to be demonstrated by a graduate after completing thecourse.

PO1: An ability to independently carry out research/investigation and development work in engineering and allied streams

PO2: An ability to communicate effectively, write and present technical reports on complex engineering activities by interacting with the engineering fraternity and with society at large.

PO3: An ability to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program

PO4: An ability to apply stream knowledge to design or develop solutions for real world problems by following the standards

PO5: An ability to identify, select and apply appropriate techniques, resources and state-of-the- art tool to model, analyse and solve practical engineering problems.

PO6: An ability to engage in life-long learning for the design and development related to the stream related problems taking into consideration sustainability, societal, ethical and environmental aspects

PO7:Anabilitytodevelopcognitiveloadmanagementskillsrelatedtoprojectmanagementand finance which focus on Entrepreneurship and Industry relevance.

Mappingofcourse outcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	✓		✓	✓		✓	
CO2	✓		✓	✓		✓	
CO3	✓		✓	✓		✓	
CO4	✓		✓	✓		✓	
CO5	✓		✓	✓		✓	

AssessmentPattern

Bloom‘sCategory	EndSemesterExamination
Apply	70%
Analyse	30%
valuate	
Create	

Markdistribution

TotalMarks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Evaluation shall only bebased on application, analysis or design based questions (for both internal and end semester examinations).

ContinuousInternalEvaluation:40marks

i. Preparingareviewarticlebasedonpeerreviewedoriginalpublications(minimum10 publications shall be referred) : 15 marks

ii. Coursebasedtask/Seminar/Datacollectionandinterpretation :15 marks

iii. Testpaper(1number) :10 marks

Testpaper shallincludeminimum 80%ofthesyllabus.

Course based task/test paper questions shall be useful in the testing of knowledge, skills, comprehension, application, analysis, synthesis, evaluation and understanding of the students.

EndSemesterExaminationPattern:

Theendsemesterexaminationwillbeconductedbytherespective College.

There will be two parts; Part A and Part B.

Part A will contain 5 numerical/short answer questions with 1 questionfrom each module, having 5 marks for each question. Students should answer all questions. Part B will contain 7 questions (such questions shall be useful in the testing of overall achievement and maturity of the studentsin a course, through long answer questions relating to theoretical/practical knowledge,derivations, problem solving and quantitative evaluation), with minimum one question from each module of which student should answer any five. Each question can carry 7 marks

Totaldurationoftheexaminationwillbe150minutes.

Note: The marks obtained for the ESE for an elective course shall not exceed 20% over the average ESE mark % for the core courses. ESE marks awarded to a student for each elective course shall be normalized accordingly.

.

Forexampleiftheaverageendsemestermark%foracorecourseis40,thenthemaximum eligible mark % for an elective course is 40+20 = 60 %.

CourseLevel AssessmentQuestions

CourseOutcome1(CO1):

1. Whataretheadvantagesofanincrementaldevelopmentmodelovera waterfall model?
2. Illustrate how the process differs in agile software development and traditionalsoftwaredevelopmentwithasociallyrelevantcasestudy. (Assignment question)

CourseOutcome2(CO2):

1. Howtoprepareasoftwarerequirementspecification?
2. DifferentiatebetweenArchitecturaldesignandComponentleveldesign.
3. Howdoesagileapproacheshelpsoftwaredeveloperstocaptureanddefinethe user requirements effectively?
4. WhatistherelevanceoftheSRSspecificationinsoftwaredevelopment?
5. Prepareausecasediagramforalibrarymanagementsystem.

CourseOutcome3(CO3):

1. Differentiatebetweenthedifferenttypesofsoftwaretestingstrategies.
2. JustifytheneedforDevOpspractices?
3. Howdodesignpatternshelpsoftwarearchitectscommunicatethedesignofa complex system effectively?
4. Whataretheproactiveapproachesonecantaketooptimiseeffortsinthetesting phase?

CourseOutcome4(CO4):

1. Illustratetheactivitiesinvolvedinsoftwareprojectmanagementforasocially relevant problem?
2. HowdoSCRUM,KanbanandLeanmethodologieshelpsoftwareproject management?
3. Isrollinglevelplanninginsoftwareprojectmanagementbeneficial? Justify your answer.
4. Howwouldyouassesstherisksinyoursoftwaredevelopmentproject? Explain how you can manage identified risks?

.

CourseOutcome5(CO5):

1. JustifytheimportanceofSoftwareProcessimprovement?
2. Explainthebenefitsofcloudbasedsoftwaredevelopment,containersand microservices.
3. Givetheroleofretrospectivesinimprovingthesoftwaredevelopmentprocess.
4. Illustratetheuseofprojecthistorydataasapredictiontooltoplanfuture socially relevant projects.

ModelQuestionPaper

Reg No: _________________			
Name:________________________________ PAGES:4			
APJABDULKALAMTECHNOLOGICALUNIVERSITY			
SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR			
CourseCode: J1PCSO235			
CourseName:SOFTWAREPROJECTMANAGEMENT			
Max.Marks:60 Duration:2.5Hours			
PARTA			
AnswerAllQuestions. EachQuestionCarries5 Marks			
1.	ExplainAgileceremoniesandAgilemanifesto.		
2.	CompareSoftwareArchitecturedesignandComponentleveldesign		
3.	Describetheformalandinformalreviewtechniques		
4.	Explainplandrivendevelopmentandprojectscheduling.		
5.	IllustrateSPIprocesswithanexample.		(5x5=25)
PartB (Answeranyfivequestions.Eachquestioncarries7marks)			
6.		Illustratesoftwareprocessactivitieswithanexample.	(7)
7.		What are functional and nonfunctional requirements? Imagine that you are developing a library management software for your college, list eight functional requirements and four nonfunctional requirements.	(7)

.

8.		ExplainContinuousIntegration,Delivery,andDeploymentCI/CD/CD)	(7)
9.		Whatisacriticalpathanddemonstrateitssignificanceinaprojectschedule with the help of a sample project schedule.	(7)
10.		Whatisalgorithmiccostmodeling?Whatproblemsdoesitsufferfromwhen compared with other approaches to cost estimation?	(7)
11.		ExplainelementsofSoftwareQualityAssuranceandSQATasks.	(7)
12.		CompareCMMIandISO9001:2000.	(7)

Syllabus

Module1:IntroductiontoSoftwareEngineering(7hours)

IntroductiontoSoftwareEngineering-Professionalsoftwaredevelopment,
Softwareengineering

ethics. Software process models - The waterfall model, Incremental development. Process activities - Software specification, Software design and implementation, Softwarevalidation, Software evolution. Coping with change - Prototyping, Incremental delivery, Boehm's Spiral Model. Agile software development - Agile methods, agile manifesto - values and principles. Agile development techniques, Agile Project Management. Case studies : An insulin pump control system. Mentcare - a patient information system for mental health care.

Module2:RequirementAnalysisandDesign(8hours)

Functional and non-functional requirements, Requirements engineering processes. Requirements elicitation, Requirements validation, Requirements change, Traceability Matrix.Developingusecases,SoftwareRequirementsSpecificationTemplate,Personas, Scenarios, User stories, Feature identification. Design concepts - Design within the context of software engineering, Design Process, Design concepts, Design Model. Architectural Design - Software Architecture, Architectural Styles, Architectural considerations, Architectural Design Component level design
- What is a component?, Designing Class-Based Components, Conducting Component level design, Component level design for web-apps. Template of a Design Document as per ―IEEE Std 1016-2009 IEEE Standard for Information Technology Systems Design Software Design Descriptions‖. Case study: The Ariane 5 launcher failure.

Module3:ImplementationandTesting(9hours)

Object-oriented design using the UML, Design patterns, Implementation issues, Open-source development - Open-source licensing - GPL, LGPL, BSD. Review Techniques - Cost impact ofSoftware Defects, Codereviewand statistical analysis. Informal Review, Formal Technical Reviews, Post-mortem evaluations. Software testing strategies - Unit Testing, Integration Testing, Validation testing, System testing, Debugging, White box testing, Path testing, Control Structure testing, Black box testing, TestingDocumentation and Help facilities. Test automation, Test-driven development, Security testing. Overview of DevOps and Code Management - Code management, DevOps automation, Continuous Integration, Delivery, and Deployment (CI/CD/CD). Software Evolution - Evolution processes, Software maintenance.

Module4:Software ProjectManagement(6hours)

Software Project Management - Risk management, Managing people, Teamwork. Project Planning, Software pricing, Plan-driven development, Project scheduling, Agile planning. Estimation techniques, COCOMO cost modeling. Configuration management, Version management, System building, Change management, Release management, Agile software management - SCRUM framework. Kanban methodology and lean approaches.

Module5 :SoftwareQuality, ProcessImprovementandTechnologytrends(6 hours)

Software Quality, Software Quality Dilemma, Achieving Software Quality Elements of Software Quality Assurance, SQA Tasks , Software measurement and metrics. Software Process Improvement(SPI), SPI Process CMMI process improvement framework, ISO 9001:2000 for Software. Cloud-based Software - Virtualisation and containers, Everything as a service(IaaS, PaaS), Software as a service. Microservices Architecture - Microservices, Microservices architecture, Microservice deployment.

TextBooks

1. Book1-IanSommerville,SoftwareEngineering,PearsonEducation,Tenthedition, 2015.
2. Book2-RogerS.Pressman,SoftwareEngineering:Apractitioner'sapproach, McGraw Hill publication, Eighth edition, 2014
3. Book 3 - Ian Sommerville, Engineering Software Products: An IntroductiontoModernSoftwareEngineering,PearsonEducation,First Edition, 2020.

References

2. IEEEStd830-1998-IEEERecommendedPracticeforSoftware Requirements SpeciÞcations
3. IEEEStd1016-2009IEEEStandardforInformationTechnology—Systems Design— Software Design Descriptions
4. DavidJ.Anderson,Kanban,BlueHolePress2010

.

5. DavidJ.Anderson,AgileManagementforSoftwareEngineering,Pearson,2003
6. WalkerRoyce,SoftwareProjectManagement:Aunifiedframework,Pearson Education, 1998
7. Steve.Denning,Theageofagile,howsmartcompaniesaretransformingtheway work gets done. New York, Amacom, 2018.
8. SatyaNadella,HitRefresh:TheQuesttoRediscoverMicrosoft‘sSouland Imagine a Better Future for Everyone, Harper Business, 2017
9. HenricoDolfing,ProjectFailureCaseStudies:Lessonslearnedfromother people‘smistakes,Kindleedition
10. MaryPoppendieck,ImplementingLeanSoftwareDevelopment:FromConcept to Cash, Addison-Wesley Signature Series, 2006
11. StarUMLdocumentation-https://docs.staruml.io/
12. OpenProjectdocumentation-https://docs.openproject.org/
13. BugZilladocumentation-https://www.bugzilla.org/docs/
14. GitHubdocumentation-https://guides.github.com/
15. Jiradocumentation-https://www.atlassian.com/software/jira

TeachingPlan

No	Contents	No of Lecture Hrs
Module1:IntroductiontoSoftwareEngineering(7hours)		
1.1	IntroductiontoSoftwareEngineering.[Book1,Chapter1]	1 hour
1.2	Softwareprocessmodels[Book1-Chapter2]	1 hour
1.3	Processactivities[Book1-Chapter2]	1 hour
1.4	Copingwithchange[Book1-Chapter2, Book2-Chapter4]	1 hour
1.5	Casestudies:Aninsulinpumpcontrolsystem.Mentcare-apatient information system for mental health care. [Book 1 - Chapter 1]	1hour
1.6	Agilesoftwaredevelopment[Book1-Chapter3]	1 hour
1.7	Agiledevelopmenttechniques,AgileProjectManagement.[Book1-Chapter 3]	1 hour
Module2:RequirementAnalysisandDesign(8hours)		
2.1	Functionalandnon-functionalrequirements,Requirementsengineering processes [Book 1 - Chapter 4]	1hour

.

2.2	Requirementselicitation,Requirementsvalidation,Requirementschange, Traceability Matrix [Book 1 - Chapter 4]	1hour
2.3	Developingusecases,SoftwareRequirementsSpecificationTemplate[Book 2 - Chapter 8]	1 hour
2.4	Personas,Scenarios,Userstories,Featureidentification[Book3-Chapter3]	1 hour
2.5	Designconcepts[Book2-Chapter12]	1 hour
2.6	ArchitecturalDesign[Book2-Chapter13]	1 hour
2.7	Componentleveldesign[Book2-Chapter14]	1hour
2.8	DesignDocumentTemplate.Casestudy:TheAriane5launcherfailure.[Ref -2,Book2-Chapter16]	1 hour
Module3:ImplementationandTesting(9hours)		
3.1	Object-orienteddesignusingtheUML,Designpatterns[Book1-Chapter 7]	1 hour
3.2	Implementationissues,Open-sourcedevelopment-Open-sourcelicensing- GPL, LGPL, BSD [Book 1 - Chapter 7]	1 hour
3.3	ReviewTechniques-CostimpactofSoftwareDefects,Codereviewand statistical analysis. [Book 2 - Chapter 20]	1 hour
34	InformalReview,FormalTechnicalReviews,Post-mortemevaluations. [Book 2 - Chapter 20]	1 hour
3.5	Softwaretestingstrategies-UnitTesting,IntegrationTesting,Validation testing, System testing and Debugging (basic concepts only). [Book 2 - Chapter 22]	1 hour
3.6	Whiteboxtesting,Pathtesting,ControlStructuretesting,Blackboxtesting. Test documentation [Book 2 - Chapter 23]	1 hour
3.7	Testautomation,Test-drivendevelopment,Securitytesting.[Book3- Chapter 9]	1 hour
3.8	DevOpsandCodeManagement-Codemanagement,DevOpsautomation, CI/CD/CD. [Book 3 - Chapter 10]	1 hour
3.9	SoftwareEvolution-Evolutionprocesses,Softwaremaintenance.[Book1- Chapter 9]	1 hour
Module4:Software ProjectManagement(6hours)		

.

4.1	SoftwareProjectManagement-Riskmanagement,Managingpeople, Teamwork [Book 1 - Chapter 22]	1hour
4.2	ProjectPlanning-Softwarepricing,Plan-drivendevelopment,Project scheduling, Agile planning [Book 1 - Chapter 23]	1hour
4.3	Estimationtechniques[Book1-Chapter23]	1hour
4.4	Configurationmanagement[Book1-Chapter25]	1hour
4.5	Agilesoftwaremanagement-SCRUMframework[Book2-Chapter5]	1hour
4.6	Kanbanmethodologyandleanapproaches.[Ref9-Chapter2]	1hour
Module5:SoftwareQuality,ProcessImprovementandTechnologytrends(6 hours)		
5.1	SoftwareQuality,SoftwareQualityDilemma,AchievingSoftwareQuality. [Book 2 - Chapter 19]	1hour
5.2	ElementsofSoftwareQualityAssurance,SQATasks,Software measurement and metrics. [Book 3 - Chapter 21]	1hour
5.3	SoftwareProcessImprovement(SPI),SPIProcess[Book2-Chapter37]	1hour
5.4	CMMIprocessimprovementframework,ISO9001:2000forSoftware. [Book 2 - Chapter 37]	1hour
5.5	Cloud-basedSoftware-Virtualisationandcontainers,IaaS,PaaS, SaaS.[Book 3 - Chapter 5]	1hour
5.6	MicroservicesArchitecture-Microservices,Microservicesarchitecture, Microservice deployment [Book 3 - Chapter 6]	1hour

INDUSTRYELECTIVE

222EEXXXX	AI DRIVEN THREAT ANALYSISFORCYBER SECURITY	CATEGORY	L	T	P	CREDIT
		Industry Elective	3	0	0	3

Preamble:

The purpose of this course is to provide the learner a comprehensive understanding of how artificial intelligence (AI) is transforming the field of threat analysis in the context of cyber security. Course explores cutting-edge applications of AI in anomaly analysis, detection and providing predictive analytics. Necessary tools to analyze threats, protect valuable data,malwareanalysis,cloudsecurityandincident response will also be covered in the course.

CourseOutcomes:

Afterthecompletionofthecoursestudentswillbeableto

CO1	Develop a deep understanding of the fundamental concepts, principles, and technologies that underpin AI-driven threat analysis in the realm of cyber security.(Cognitive Knowledge Level: Create)
CO2	Gain insights into the evolving cyber threat landscape, including the various types of threats, attack vectors, and vulnerabilities commonly exploited by malicious actors. (Cognitive Knowledge Level: Create)
CO3	Apply AI techniques to detect, analyze, and mitigate cyber threats, enabling proactive defense and timely incident response.(Cognitive Knowledge Level: Apply)
CO4	Developskillsingathering,analyzing,andinterpretingthreatintelligence to proactively identify and track potential threats, as well as engage in threat hunting activities.(Cognitive Knowledge Level: Analyze)

CO5	Enhance the ability to collaborate with multidisciplinary teams, communicate effectively about cyber threats and their implications, and provide actionable recommendations for cyber security improvements along with an awareness of the ethical considerations and challenges related to AI-driven threat analysis in cyber security (Cognitive Knowledge Level: Analyze)

Mappingofcourse outcomeswithprogramoutcomes:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	X		X	X			
CO2	X		X				
CO3	X		X	X	X	X	
CO4	X	X	X				
CO5	X	X	X	X		X	

AssessmentPattern

Bloom’sCategory	EndSemesterExamination
Apply	60-70%
Analysis	40-30%
Evaluate	
Create	

MarkDistribution

TotalMarks	CIE	ESE	ESEDuration
100	40	60	2.5Hours

Theassessmentpatternfor Industrybasedelectivesisasfollows:

ContinuousInternalEvaluation:40marks

Thecontinuousinternalevaluationwillbedonebytheexpertinthe Industryhandlingthecourse.

Micro project/Course based project: 20 marks

Course based task/Seminar/Quiz: 10

marksTestpaper,1no. :10marks

Theprojectshallbedoneindividually.Groupprojectsnotpermitted.Testpapershallincludeminimum 80% of the syllabus.

EndSemesterExamination:60marks

The examination will be conducted by the respective College with the question paper provided by the Industry. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks. The valuation of the answer scripts shall be done by the expert in the Industry handling the course.

Syllabus

Module1

Cybersecurity-Enterpriselevel-IdentityandAccessManagement-Authentication - Authorization -UserManagement-CentralUserRepository -PrivilegedAccessManagement.

MalwareAnalysis-DetectandAnalyzeModernMalware -HandsonAnalysis of a malware.

Module2

Software Development Security - Understand and Integrate Security in the software development Life cycle (SDLC) - Identify and apply securitycontrols in software development ecosystem - Define and apply secure coding guidelines and standards. - OWASP - Broken Access Control - Cryptographic failures - Injection - Insecure Design - Security Misconfiguration - Vulnerable and outdated components - Identification and Authentication failures - Software and Data Integrity failures - Security Logging and Monitoring - Server side failures Request forgery (SSRF).

Module3

Cryptography&Hashing-DifferentCryptographicmethodsusedinIndustry -Hashingmethods.

Advanced Threat Detection using AI - Introduction to advanced threat detection techniques - Exploring anomaly detection, behavior analysis, and machine learning in threat detection - Case studies and real-world examples of advanced threat detection systems.

AI-driven Network and Asset Mapping - Overview of AI-driven network and asset mapping techniques - Understanding the importance of network and asset visibility in cybersecurity - Exploring AI-powered network and asset mapping and visualization platforms.

Module4

Security Assessment and Penetration Testing - Overview of security assessment methodologies and frameworks - Understanding the role of penetration testing in identifying vulnerabilities - Exploring tools and techniques for security assessment and penetration testing.

Cloud Security - Introduction to cloud computing security challenges and considerations -Understandingcloudservicemodelsanddeploymentmodels -Exploringcloudsecuritycontrolsandbestpractices.

Module5

Incident Response and Forensics - Advanced incident response techniques and forensic investigation methodologies -Understanding memory forensics, loganalysis,andmalware analysisinincident response-Hands-onexercise: Conducting an advanced incident response and forensic investigation.

Emerging Trends and Future of Cybersecurity - Discussion on current and emerging trends in cybersecurity - Exploring topics such as AI in cybersecurity, IoT security, and blockchain security.

CoursePlan

No	Topic	No of Lectures (42)
	MODULE1	
1.1	Cybersecurity-Enterpriselevel	1
1.2	IdentityandAccessManagement	2
1.3	Authentication	3
1.4	Authorization	4
1.5	UserManagement	5
1.6	CentralUserRepository	6
1.7	PrivilegedAccessManagement	7
1.8	MalwareAnalysis	8
1.9	DetectandAnalyzeModernMalware	9
1.10	HandsonAnalysisofamalware	10
	MODULE2	
2.1	SoftwareDevelopmentSecurity	11
2.2	UnderstandandIntegrateSecurityinthesoftwaredevelopmentLife cycle (SDLC)	12
2.3	Identifyandapplysecuritycontrolsinsoftwaredevelopmentecosystem	13
2.4	Defineandapplysecurecodingguidelinesandstandards	14
2.5	OWASP-BrokenAccessControl	15
2.6	Cryptographicfailures-Injection	16
2.7	InsecureDesign-SecurityMisconfiguration	17

MODULE5		
5.1	IncidentResponseandForensics-Advancedincident response techniques and forensic investigation methodologies	36
5.2	Understandingmemoryforensics,loganalysis,andmalwareanalysisin incident response	37
5.3	Hands-onexercise:Conductinganadvancedincidentresponseand forensic investigation	38
5.4	EmergingTrendsandFutureofCybersecurity	39
5.5	Discussiononcurrentandemergingtrendsincybersecurity	40
5.6	CaseStudy:AIincybersecurity,IoTsecurity	41
5.7	AIinblockchainsecurity	42

ReferenceBooks

1. “Metasploit: ThePenetrationTester'sGuide”, DavidKennedy,Jim O'Gorman, Devon Kearns, and Mati Aharoni, *No Starch Press*,US, 2011

2. Cloudsecurityandprivacy,TimMather,SubraKumaraswamy,Shahed Latif, *O'Reilly Media, Inc*., 2009

3. Hands-OnArtificialIntelligenceforCybersecurity,AlessandroParisi,*Packt Publishing,* 2019

4. "AI in Cybersecurity", Leslie F. Sikos(Editor), *Springer International Publishing*, 2018

5. “SecuritySoftwareDevelopmentAssessingandManagingSecurityRisks”, DouglasA.Ash*baugh,CRCPress,2008*

Model Question Paper

QP CODE:

Reg No: ___________________

Name: ____Pages:2

APJABDULKALAMTECHNOLOGICALUNIVERSITY

SECONDSEMESTERM.TECHDEGREEEXAMINATION,MONTH&YEAR

CourseCode:.............

CourseName:......................

Max.Marks:60 Duration:2.5Hours

PARTA

Answerany5Questions.EachQuestionCarries12Marks

1. Discuss the importance of identity and access management in today's digital landscape and explain how it helps organizations enhance security and protect sensitive information.

2. What are some common types of malware and their characteristics? Describe the challenges faced by malware analysts in analyzing advanced and evasive malware.

3. Explain the significance of incorporating security measures throughout the software development life cycle (SDLC) and discuss the key practices and methodologies that can enhance the security of software systems.

4. ExplainAI-drivennetworkandassetmappingtechniquesindetail.

5. Discuss the role and significance of penetration testing in ensuring the security of computer systems and networks.

6. Differentiate between various cloud service models. Also, explain cloud deployment model.

7. Explain the advanced incident response techniques used in handling sophisticated cyber incidents, including targeted attacks and advanced persistent threats (APTs). Also, discuss the challenges faced by incident responders and how these techniques help in effectively mitigating and investigating such incidents.

SEMESTERIII

.

SEMESTERIII

Slot	Coursecode	Courses	Marks		L-T-P	Hours	Credit
			CIE	ESE			
TRACK 1							
A*	25PCSXXX	MOOC	Tobe completed successfully		--	--	2
B	25PAGE3NN	AuditCourse	40	60	3-0-0	3	-
C	25PCSI300	Internship	50	50	--	--	3
D	25PCSP300	Dissertation Phase1	100	--	0-0-17	17	11
TRACK 2							
A*	25PCSXXX	MOOC	Tobe completed successfully		--	--	2
B	25PAGE3NN	AuditCourse	40	60	3-0-0	3	-
C	25PCSI300	Internship	50	50	---	--	3
D	25PCSP301	Research ProjectPhase1	100	--	0-0-17	17	11
TOTAL			**190**	**110**		**20**	**16**

TeachingAssistance:6hours

*MOOCCoursetobesuccessfullycompletedbeforethecommencementoffourth semester (starting from semester 1).

.

AUDITCOURSE

AUDITCOURSE						
SLOT	**SL NO**	**COURSE CODE**	**COURSENAME**	**L-T-P**	**HOURS**	**CREDIT**
B	1	J1AGT312	ACADEMICWRITING	3-0-0	3	-
	2	J1AGT322	ADVANCED ENGINEERING MATERIALS	3-0-0	3	-
	3	J1AGT332	DATASCIENCEFOR ENGINEERS	3-0-0	3	-
	4	J1AGT342	DESIGNTHINKING	3-0-0	3	-
	5	J1AGT352	FUNCTIONALPROGRAMMING IN HASKELL	3-0-0	3	-
	6	J1AGT362	PRINCIPLESOFAUTOMATION	3-0-0	3	-
	7	J1AGT372	REUSEANDRECYCLE TECHNOLOGY	3-0-0	3	-
	8	J1AGT382	SYSTEMMODELING	3-0-0	3	-
	9	J1AGT392	EXPERTSYSTEMS	3-0-0	3	-

MOOCCOURSES

The MOOC course shall be considered only if it is conducted by the agencies namely AICTE/NPTEL/SWAYAM or NITTTR. The MOOC course should have a minimum duration of 8 weeks and the content of the syllabus shall be enough for at least40hoursofteaching.Thecourseshouldhaveaproctored/offlineendsemester examination. The students can do the MOOC according to theirconvenience,but shallcompleteitbythirdsemester.ThelistofMOOCcourses willbeprovidedbythe concerned BoS if at least 70% of the course content match with the area/stream of study.Thecourseshallnotbeconsideredifitscontenthasmorethan50%ofoverlap with a core/elective course in the concerned discipline or with an open elective.

MOOCCoursetobesuccessfullycompletedbeforethecommencementoffourth semester (starting from semester 1). A credit of 2 will be awarded to all students whoever successfully completes the MOOC course as per the evaluation pattern of the respective agency conducting the MOOC.

TEMPLATEFORSYLLABUS

CODE J1AGT312	ACADEMIC WRITING	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	NIL

Preamble: Learningacademicwritingsharpensminds,teachesstudents howtocommunicate,and develops their thinking capacities and ability to understand others. Writing is thinking, and every student deserves to be a strong thinker. It can also make them think more carefully about what they write. Showing work to others can help to foster a better culture of learning and sharing among students. It also gives students a sense of how they are contributing to the body of workthat makes up an academic subject.

CourseOutcomes: TheCOsshownareonlyindicative.Foreachcourse,therecan be4to6COs. After the completion of the course the student will be able to

CO1	Understandtheprinciplesofscientific/academicwriting
CO2	Analysethetechniqueofscientificwritingfromthereader'sperspective
CO3	Applytheconceptsofsettingexpectationsandlayingtheprogressiontracks
CO4	Evaluatethemeritsofatitle,abstract,introduction,conclusionandstructuringofa research paper
CO5	Justifytheneedusingaprojectproposaloratechnicalreport
CO6	Preparareviewpaper,anextendedabstractandaprojectproposal

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1		3	1				
CO2		3	1				
CO3		3	1			2	
CO4		3	1				
CO5		3	2	2		2	
CO6	1	3	3	2		2	

AssessmentPattern

Bloom's Category	EndSemester Examination
Apply	40%
Analyse	30%
Evaluate	30%

.

Markdistribution

TotalMarks	CIE	ESE	ESEDuration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:40marks

Coursebasedtask:15marks
Seminar/Quiz : 15 marks
Test paper, 1 no. : 10 marks
Testpapershallincludeminimum80%ofthesyllabus.

EndSemesterExaminationPattern:60marks

The examination will be conducted by the respective College. The examination willbe for 150minutesandwillcontain7questions,withminimumonequestionfromeachmoduleof which student should answer any Vive. Each question can carry 12 marks.

ModelQuestionpaper

		SET1	**TotalPages:**
RegNo.:__________________		Name:_______________________________	
APJABDULKALAMTECHNOLOGICALUNIVERSITY THIRDSEMESTERM.TECHDEGREEEXAMINATION,MARCH2024			
CourseCode: J1AGT312			
CourseName:AcademicWriting			
Max.Marks:60		Duration:2.5Hours	
Answeranyfivefullquestions,eachcarries12marks.			

1a)	Makeclear-cutdistinctionsbetween6factorsthattaketheirtollonreaders' memory.	6
1 b)	Howcanyousustainthe attentionofthereadertoensurecontinuousreading?	6
2a)	Whatarethedifferentmethodsbywhichyoucancreateexpectationsinthe reader?	6
2b)	Giveanaccountofthetopicand non-topicbasedprogressionschemes.	6
3a)	Bringoutthedifferencesbetweenanabstractandtheintroductionofa researchpaper.	8
3 b)	Howarethetitleoftheresearchpaperanditsstructurerelated?	4
4	Whatare7principlesforincludingvisualsinyourresearchpaper.Whatare therecommendedconstituentsofaconclusionsegmentofaresearchpaper?	12

5	Giveadetaileddescriptionoftheprocessandcontentsofaprojectproposal forfunding.	12
6a)	Whatarethecontextsrecommendedforchoosingbetweenactive andpassive voicesintechnicalwriting?	8
6b)	Whatarethedifferentvisualformsthatarerelevantinaresearchpaperand howdoyouchoosethem?	4
7	Givethe designofa researchpaperwiththe purposeseachpartserves.	12
	****	

SyllabusandCoursePlan(For3credit courses,thecontentcanbefor40 hrsandfor2credit courses,thecontentcan befor26hrs.Theauditcoursein thirdsemester canhavecontent for 30 hours).

Syllabus:

CODE J1AG T312	**ACADEMICWRITING**	**Audit**
Module No.	Topicsinamodule	Hours
1	Fundamentals of Academic writing from a reader‘s perspective: acronyms, synonyms, pronouns, disconnected phrases, background ghettos, abusive detailing,crypticcaptions,longsentences:allthattaketheirtollonreaders‘ memory.	6
2	Fluidreading&readingenergyconsumption:settingexpectationsandlaying Progression tracks; Reading energyconsumption	6
3	HowtowritetheTitle,abstract,introduction;Structurethewritingwith headings & subheadings	6
4	Visuals:Resources,Skills,andMethods;Conclusion;References; Bibliography; Grammar in technical writing	6
5	Techniquesofwriting:Anextendedabstract,aprojectproposal,aresearch paper, a technical report.	6

CoursePlan:

No	Topic	No. of Lectures
1	Fundamentals of Academic writing from a reader‘s perspective: acronyms, synonyms, pronouns,disconnectedphrases,backgroundghettos,abusivedetailing,crypticcaptions, long sentences all take their toll on readers‘ memory.	
1.1	TheReadingtool-kittoreducememoryrequired;reducereadingtime	1
1.2	Acronyms,Pronouns,Synonyms;Background,brokencouple,words overflow	1
1.3	Sustainattention:Keepthestorymovingforward;Twists,shouts,Pause to clarify, recreate suspense	2

1.4	Keepthereadermotivated:FuelandmeetExpectations;Bridge knowledgegap: groundlevel;Titlewords;JustInTimetolocal background	2
2	Fluidreading&readingenergyconsumption:settingexpectationsandlayingProgression tracks; Reading energyconsumption	
2.1	Settingexpectationsofthereaderfrom Grammar,fromtheme	1
2.2	Progressiontracksforfluidreading:Topic&stress;topicandnontopic based progression tracks; pause in progression	2
2.3	Detectionofsentencefluidityproblems:Noexpectations/Betrayed expectations	2
2.4	Controllingreadingenergyconsumption:theenergybill;Energy fuelling stations: Pause	1
3	HowtowritetheTitle,abstract,introduction;Structurethewritingwithheadings&subheadings	
3.1	Title:Faceofthepaper:Techniques,Qualities&Purposeoftitle; Metrics	1
3.2	Abstract:Heartofthepaper:4parts;coherence;tenseofverbs, precision; purpose & qualities of the abstract; Metrics	2
3.3	Structure:Headings&sub-headings:Skeletonofthepaper:principles for a good structure; Syntactic rules; Quality & Purpose of structures; Metrics	1
3.4	Introduction:Handsofthepaper:Start,finish;scope,definitions; answers key reader questions; As a personal active story; Traps, qualities; Metrics	2
4	Visuals:Resources,Skills,andMethods;Conclusion;References;Bibliography;Grammar in technical writing	
4.1	Visualsasthevoiceofyourpaper:principles;purpose&qualitiesof visuals; metrics	2
4.2	Conclusion:contents;purpose,quality;metrics;AbstractsVs. Conclusion; examples, counter-examples	1
4.3	References,Bibliography:Styles,punctuationmarks,quotes,citations	1
4.4	GrammarinTechnicalwriting:Articles,Syntax,Mainandsubordinate clauses; Active & passive voices; some commonly made mistakes in technical writing.	2
5	Techniquesofwriting:Anextendedabstract,aprojectproposal,aresearchpaper,a technical report.	
5.1	Extendedabstract:abstractandkeywords,introductionandobjective, method, findings and argument, conclusion and suggestions and references.	1
5.2	Project Proposal:Types, executive summary, background including status, objectives, solution, milestones, deliverables, timelines, resources, budgeting, conclusion	2
5.3	Research paper: writingan overview article: provide a comprehensive foundation on a topic; explain the current state of knowledge; identify gapsinexistingstudiesforpotentialfutureresearch;highlightthemain methodologies and research techniques	2

.

5.4	WritingTechnical Reports:Titlepage;Summary;Tableofcontents; Introduction; Body; Figures, tables, equations and formulae; Conclusion; Recommendations.	1
		30

Reference Books

1. SCIENTIFICWRITING2.0AReaderandWriter’sGuide:Jean-LucLebrun,WorldScientiVic Publishing Co. Pte. Ltd., 2011

2. HowtoWriteandPublishaScientiVicPaper:BarbaraGastelandRobertA.Day,Greenwood publishers, 2016

3. Grammar,Punctuation,andCapitalisation;ahandbookfortechnicalwritersandeditors. www.sti.nasa.gov/publish/sp7084.pdfwww.sti.nasa.gov/sp7084/contents.html

4. EverythingYouWantedtoKnowAboutMakingTablesandFigures.http://abacus.bates.edu/%7Eganderso/biology/resources/writing/ HTWtableVigs.html

.

J1AGT322	ADVANCED ENGINEERING MATERIALS	CATEGORY	L	T		CREDIT
		AUDIT COURSE	3	0		-

Preamble: This course is designed in a way to provide a general view on typically used advanced classes of engineering materials including metals, polymers, ceramics, and composites.

CourseOutcomes: Afterthecompletionofthecoursethestudentwillbeableto

CO1	Analysetherequirement andfindappropriatesolutionforuseofmaterials.
CO2	Differentiatethepropertiesofpolymers,ceramicsandcompositematerials.
CO3	Recognizebasicconceptsand propertiesoffunctionalmaterials.
CO4	Comprehendsmartandshapememorymaterialsforvariousapplications.
CO5	Appraise materialsused for hightemperature,energyproduction and storage applications.

Mappingofcourseoutcomeswithprogram outcomes

	PO1	PO2	PO3	PO4	PO 5	PO6	PO7
CO1	✓				✓	✓	
CO2	✓				✓	✓	
CO3	✓				✓	✓	
CO4	✓				✓	✓	
CO5	✓				✓	✓	

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Understand	60%
Apply	20%
Analyse	20%

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

ContinuousInternalEvaluationPattern:40marks

Coursebasedtask:15marks
Seminar/Quiz : 15 marks
Test paper, 1 no. : 10 marks
Testpapershallincludeminimum80%ofthesyllabus.

EndSemesterExaminationPattern:60marks

The examination will be conducted by the respective College. The examination willbe for 150 minutesandwillcontain7questions,withminimumonequestionfromeachmoduleofwhich student should answer any Vive. Each question can carry 12 marks.

ModelQuestionpaper

AUDITCOURSE

J1AGT322-ADVANCEDENGINEERINGMATERIALS

(Answeranyfivequestions.Eachquestioncarries12 Marks)

1. a)Statetherelationshipbetweenmaterialselectionandprocessing. **5**

b)Writeaboutthecriteriaforselectionofmaterialswithrespecttothecost and service requirements for engineering applications. **7**

2. a)Differentiatethermosettingandthermoplasticswithsuitableexamples. **5**

b)Brieflydiscussaboutthepropertiesandapplicationsofpolymernano compositematerials. **7**

3. a)Writeaboutthepotentialapplicationareasoffunctionallygradedmaterials. **5**

b)Withaneatsketchdescribeanyoneprocessingtechniqueoffunctionally gradedmaterials. **7**

4. a)―Smartmaterialsarefunctional‖?Justifythestatement. **5**

b)Explainthetermselectrostrictionandmagnetostrictionwithitsapplication. **7**

.

5.	a)Whatarethefactorsinfluencing functionallifeofcomponentsatelevated temperature?	**5**
	b)Whataresuperalloysandwhataretheiradvantages?	**7**
6	a)Whatisashapememoryalloy? What metalsexhibit shape memory characteristics?	**4**
	b) Explainaboutthedetectioncapabilitiesandusesofpyroelectric sensors.	**8**
7	a)Differentiatebetweenconventionalbatteriesand fuelcells.	**4**
	b) ExplaintheconstructionandworkingofaLi-ionbattery.	**8**

Syllabus

Module	Content	Hours	Semester Exam Marks (%)
I	Requirements / needs of advanced materials. Classification of materials, Importance of materials selection,Criteriaforselectionofmaterials;motivation for selection, cost basis and service requirements. Relationship between materials selection and processing.	5	20
II	Classificationofnon-metallic materials. Polymer, Ceramics: Properties, processing and applications. NanoComposites - Polymer nanocomposites (PNCs), Processingandcharacterisationtechniques–properties and potential applications.	7	20
III	Functionally graded materials (FGMs),Potential Applications of FGMs, classification of FGMs, processing techniques. limitations of FGMs.	6	20
IV	Smart Materials: Introduction, smart material types - pyroelectric sensors, piezoelectric materials, electrostrictors and magnetostrictors, shape memory alloys–associatedenergystimulusandresponseforms, applications.	5	20
V	High Temperature Materials: super alloys – main classes, high temperature properties of superalloys, applications. EnergyMaterials: materialsforbatteries.	7	20

.

CoursePlan

No	Topic	No.of Lectures
1	**Selectionofmaterialsforengineeringapplications**	
1.1	Benefitsofadvanced materials,classificationofmaterials, importanceofmaterialsselection	2
1.2	Selectionofmaterialsfordifferentproperties,strength, toughness,fatigueandcreep	1
1.3	Selectionforsurfacedurability,corrosionandwearresistance	1
1.4	Relationshipbetweenmaterialsselectionandprocessing	1
2	**Classificationofnon-metallicmaterials&nanocomposites**	
2.1	Rubber:properties,processingandapplications.	1
2.2	Plastics:thermosettingandthermoplastics,applicationsand properties.	2
2.3	Ceramics:propertiesandapplications.	1
2.4	Introductiontonanocomposites,classification	1
2.5	Processingandcharacterisationtechniquesapplicableto polymernanocomposites.	2
3	**Functionallygradedmaterials**	
3.1	Generalconcept,PotentialApplicationsofFGMs	2
3.2	ClassificationofFGMs	1
3.3	FGMsprocessingtechniques:powdermetallurgyroute,melt-processingroute	2
3.4	LimitationsofFGMs	1
4	**Smartmaterials**	
4.1	Introductiontosmartmaterials,types	1
4.2	Pyroelectricsensors-materialclass,stimulus,detection capabilitiesanduses	1
4.3	Piezoelectricmaterials-materialclass,stimulus,sensingand actuatingapplications	1
4.4	Electrostrictorsandmagnetostrictors-materialclass,stimulus, micropositioningcapabilitiesandapplications	1
4.5	Shapememoryalloys(SMAs)-materialclass,stimulus, temperaturesensingandhighstrainresponses,applications.	1
5	**HighTemperatureMaterialsandEnergy Materials**	
5.1	Characteristicsofhigh-temperaturematerials,superalloysas high-temperaturematerials	1
	superalloys-propertiesandapplications	2
5.2	Introductiontolithium-ionbattery(LIBs),operating mechanismsandapplications	2
5.3	IntroductiontoZn-basedbatterysystem,typesandexisting challenges	2

.

ReferenceBooks

1. DeGarmoetal,—MaterialsandProcessesinManufacturing‖,10thEdition,Wiley,2008.
2. R.E.SmallmanandA.H.W.Ngan,PhysicalMetallurgyandAdvancedMaterials, Seventh Edition, Butterworth-Heinemann, 2007
3. VijayamohananK.PillaiandMeeraParthasarathy,—FunctionalMaterials:Achemist's perspective‖, Universities Press Hyderabad (2012).
4. M.V.Gandhi,B.S.Thompson:SmartMaterialsandStructures,Chapman&Hall,1992.
5. G. W. Meetham and M. H. Van de Voorde, Materials for High Temperature EngineeringApplications(EngineeringMaterials)Springer;1edition(May19,2000)
6. InderjitChopra,JayantSirohi,—SmartStructuresTheory‖,CambridgeUniversityPress, 2013

J1AGT332	DATASCIENCEFOR ENGINEERS	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	0

Preamble: This course covers essentials of statistics and Linear Algebra and how to prepare the data before processing in real time applications. The students will be able to handle missing data and detection of any outliers available in the dataset. This course explores data science,Pythonlibrariesand it also coversthe introductionto machine learning forengineers.

CourseOutcomes: After thecompletionofthecoursethestudentwillbeableto

CO1	StudyDataScienceConceptsandstatistics
CO2	DemonstrateUnderstandingofMathematicalFoundationsneededforDataScience
CO3	UnderstandExploratoryanalysisandDataVisualizationandPreprocessingon given dataset
CO4	ImplementModelssuchasNaiveBayes,K-NearestNeighbors,LinearandLogistic Regression
CO5	Buildrealtimedatascienceapplicationsandtestusecases

Mappingofcourseoutcomeswithprogramoutcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO6	PO7
CO1	2		2			2	
CO2	2		2	1		2	
CO3	2		2	2	2	2	
CO4	2		2	2	3	2	
CO5	2		2	3	3	3	2

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Understand	50%
Apply	30%
Analyse	20%

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

.

ContinuousInternalEvaluationPattern:40marks
Coursebasedtask(Project/Assignments/Simulations/Casestudies):15marks
Seminar/Quiz : 15 marks
Testpaper,1no.:10marks
Testpapershallincludeminimum80%ofthesyllabus.

EndSemesterExaminationPattern:60marks
The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 mark.

Syllabus

Module	Content	Hours	Semester Exam Marks (%)
I	**StatisticsforDatascience** Probability: Basic concepts of probability, conditional probability, total probability, independent events, Bayes‘ theorem, random variable, Population, Sample, Population Mean, Sample Mean, Population Distribution, Sample Distribution and sampling Distribution,Mean,Mode,Median,Range,Measureof Dispersion,Variance,StandardDeviation, Gaussian/NormalDistribution,covariance,correlation.	**6**	**20**
II	**LinearAlgebra** Vectors and their properties, Sum and difference of Vectors,distancebetweenVectors,Matrices,Inverseof Matrix, Determinant of Matrix, Trace of a Matrix, Dot Product, Eigen Values, Eigen Vectors, Single Value Decomposition	**6**	**20**
III	**HypothesisTesting** Understanding Hypothesis Testing, Null and Alternate Hypothesis, Non-directional Hypothesis, Directional Hypothesis Critical Value Method, P-Value Method, Types of Errors-Type1 Error,Type2 Error, Types of Hypothesis Test Z Test, Chi-Square	**6**	**20**

.

IV	**ExploratoryDataAnalysis** Data Collection –Public and Private Data, Data Cleaning-Fixing Rows and Columns, Missing Values, Standardizing values, invalid values, filtering data, Data-Integration,Data-Reduction,Data Transformation	**6**	**20**
V	**Machine Learning and Python for Data Science** Python Data structures-List, Tuple, Set, Dictionary, Pandas,Numpy,Scipy,Matplotlib,MachineLearning-SupervisedMachine Learning, Unsupervised Machine Learning,Regression, Classification, Naïve-Bayes	**6**	**20**

CoursePlan

No	Topic	No.of Lectures
1	**StatisticsforDatascience**	
1.1	Probability:Basicconceptsofprobability,conditional probability, total probability	1
1.2	independentevents,Bayes'theorem, randomvariable,Population	1
1.3	Sample,PopulationMean,SampleMean,PopulationDistribution	1
1.4	SampleDistributionandsamplingDistribution,Mean,Mode, Median, Range, Propositional logic andpredicate logic	1
1.5	MeasureofDispersion,Variance,StandardDeviation	1
1.6	Gaussian/NormalDistribution,covariance,correlation.	1
2	**LinearAlgebra**	
2.1	Vectorsandtheirproperties,	1
2.2	SumanddifferenceofVectors,distancebetween Vectors	1
2.3	Matrices,InverseofMatrix,	2
2.4	DeterminantofMatrix,TraceofaMatrix,DotProduct,Eigen Values, Eigen Vectors, Single Value Decomposition	2
3	**HypothesisTesting**	
3.1	UnderstandingHypothesisTesting,NullandAlternate Hypothesis	1
3.2	Non-directionalHypothesis,DirectionalHypothesisCritical ValueMethod,P-Value Method,	2
3.3	TypesofErrors-Type1Error,Type2 Error,	1
3.4	TypesofHypothesisTestZTest,Chi-Square,	2
4	**ExploratoryDataAnalysis**	
4.1	DataCollection–PublicandPrivateData	1
4.2	DataCleaning-FixingRowsandColumns	1
4.3	MissingValues	1
4.4	Standardizingvalues	1
4.5	Invalidvalues,filteringdata	1
4.6	DataIntegration,DataReduction,DataTransformation	1

5	**MachineLearningandPythonforDataScience**	
5.1	PythonDatastructures-List,Tuple,Set,	1
5.2	Dictionary,Pandas,Numpy,Matplotlib	2
5.3	Machine Learning-Supervised Machine Learning, UnsupervisedMachineLearning	1
5.4	Regression,Classification	1
5.5	Naïve-Bayes	1

ReferenceBooks

1. PythonDataScienceHandbook.EssentialToolsforWorkingwithData,Author(s):Jake VanderPlas, Publisher: O'Reilly Media, Year: 2016
2. PracticalStatisticsforDataScientists:50EssentialConcepts,Author(s):PeterBruce, AndrewBruce,Publisher:O‘ReillyMedia,Year:2017
3. PracticalLinearAlgebraforDataScience,byMikeXCohen,ReleasedSeptember 2022, Publisher(s): O'Reilly Media, Inc.
4. DataSciencefromScratch_byJoelGrus,Released,April2015,Publisher(s):O'Reilly Media, Inc.
5. Hands-OnExploratoryDataAnalysiswithPython,bySureshKumarMukhiya, Usman Ahmed, Released March 2020,Publisher(s): Packt Publishing

.

SET1 **TotalPages:**

Reg No.:_ _ Name:___________________________________

APJABDULKALAMTECHNOLOGICALUNIVERSITY

THIRDSEMESTERM.TECHDEGREEEXAMINATION,MARCH2024

CourseCode: J1AGT332

CourseName:DATASCIENCEFORENGINEERS

Max.Marks:60 Duration:2.5Hours

Answeranyfivefullquestions,eachcarries12marks.

1	a) It is observed that 50% of mails are spam. There is software that filters spam mail before reaching the inbox. It accuracy for detecting a spam mail is99%andchancesoftagginganon-spammailas spammailis5%.Ifa certain mail is tagged as spam finds the probability that it is not a spam mail.	5
	b) Depicttherelevanceofmeasuresofcentraltendencyindata wranglingwitha suitableexample	7
2.	a)CalculatetheinverseoftheMatrix $\begin{matrix} 2 & 4 & -6 \\ 7 & 3 & 5 \\ 1 & -2 & 4 \end{matrix}$	4
	b)FindallEigenvaluesandCorrespondingEigenvectorsforthematrixif $\begin{matrix} 2 & -3 & 0 \\ 2 & -5 & 0 \\ 0 & 0 & 3 \end{matrix}$	8
3.	a)Astatisticianwantsto testthehypothesisH0:μ=120usingthe alternative hypothesis Hα: μ > 120 and assuming that α = 0.05. For that, hetookthesamplevaluesasn=40,σ=32.17and$\bar{x}$ =105.37.Determine the conclusion for this hypothesis?	5
	b) Hypothesistestingisanintegralpartofstatisticalinference,listoutthe varioustypesofhypothesistestingandalsomentionstheirsignificancesin data science.	7
4.	a)Briefindetaildirectionalandnon-directionalhypothesis	6
	b) Differentiatenullandalternatehypothesisandalsoelaborateontype1 andtype 2 errors	6
5.	a)Explain theconceptsofTuple,ListandDirectoryinpythonwith example	6
	b) Elucidatereinforcementlearningandapplicationinrealworld.	6

.

6.	a)WhatisFeatureEngineering,demonstratewithanexample	6
	b)Describeindetaildifferentstepsinvolved indatapreprocessing.	6
7.	a)Illustratesupervised learning modelwithlinearregressionmodel	5
	b)Predicttheprobabilityforthegivenfeaturevectorifanaccident will happen or not? Weathercondition:rain,Roadcondition:good,Trafficcondition:normal, Engine problem: no, the task is to predict using Naïve Bayes classification.	7

SNo.	Weather condition	Road condition	Traffic condition	Engine problem	Accident
1	Rain	bad	high	no	yes
2	snow	average	normal	yes	yes
3	clear	bad	light	no	no
4	clear	good	light	yes	yes
5	snow	good	normal	no	no
6	rain	average	light	no	no
7	rain	good	normal	no	no
8	snow	bad	high	no	yes
9	clear	good	high	yes	no
10	clear	bad	high	yes	yes

.

J1AGT342	DESIGN THINKING	CATEGORY	L	T	P	CREDIT
		AUDITCOURSE	3	0	0	-

Preamble:

Thiscourseoffersanintroductoryexplorationoffundamental engineeringconcepts and techniques, the design process, analytical thinking and creativity, as wellasthe fundamentals and development of engineering drawings, along with their application in engineering problems.

CourseOutcomes:

Afterthecompletionofthecoursethestudentwillbeableto

CO1	Identifyandframedesignchallengeseffectively.
CO2	Generatecreativeideasthroughbrainstormingandideation
CO3	Iterateondesignsbasedonuserinsights
CO4	ApplyDesignThinkingtoreal-worldproblemsandprojects.

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1				2		2	2
CO2	2		2	2			2
CO3		2		2		2	2
CO4	2		2	3	2		2

AssessmentPattern

Bloom'sCategory	EndSemester Examination
Apply	40
Analyse	30
Evaluate	30
Create	

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

AUDIT COURSES

.

ContinuousInternalEvaluationPattern:40marks

Coursebasedtask:15marks
Seminar/Quiz : 15 marks
Test paper, 1 no. : 10 marks
Testpapershallincludeminimum80%ofthesyllabus.

EndSemesterExaminationPattern:60marks

The examination will be conducted by the respective College. The examination willbe for 150 minutesandwillcontain7questions,withminimumonequestionfromeachmoduleofwhich student should answer any five. Each question can carry 12 marks.

ModelQuestionpaper

		SET1	**TotalPages:**
RegNo.:__________________			Name:________________________________
APJABDULKALAMTECHNOLOGICAL UNIVERSITY THIRDSEMESTERM.TECHDEGREEEXAMINATION,MARCH2024			
CourseCode: J1AGT342			
CourseName:DESIGNTHINKING			
Max.Marks:60			Duration:2.5Hours
Answeranyfivefull questions,eachcarries12marks.			
1a)	Howcanamultidisciplinaryteamcollaborateeffectivelyto implement design principles?		7
1b)	Whatarethekeydifferencesbetweenhuman-centreddesignand other design methodologies?		5
2a)	Howdoyoumeasurethesuccessofadesignprojectintermsofuser satisfaction and impact?		7
2b)	Howdoestheiterativenatureofthedesignprocesscontributetobetter outcomes		5

3a)	What are the fundamental principles ofeffectivebrainstorming, and how do they differ from traditional problem-solving approaches?	7
3 b)	Whataresomekeyprinciplesofergonomicdesign,andhowdo they contribute to the usability and comfort of products?	5
4a)	Enumeratesomeexamplesofsuccessfulandunsuccessful market testing scenarios, and what lessons can be learned from theseexperiencestoimprovefutureproductorservicelaunches?	7
4b)	What is theprimarypurposeof creatingprototypesin thedesign and development process?	5
5	What strategies and methodologies can designers use to embrace agility and respond quickly to changing user needs and market dynamics?	12
6	Illustrate any four examples of successful bio-mimicry applications in various industries.	12
7	Whatethicalconsiderationsshoulddesignerskeepinmindwhen designingfordiverseusergroups?	12
****		

.

Syllabus:

Module1

Design process: Traditional design, Design Thinking Approach, Introduction to Design Thinking, History and evolution of Design Thinking, Role of design thinking in the human-centred design process. Design space, Design Thinking in a Team Environment, Team formation.

Module2

Design Thinking Stages: Empathize, Define, Ideate, Prototype and Test. The importance of empathy, Building a user-centred mindset. Problem statement formulation, User needs and pain points, establishing target specifications, Setting the final specifications.

Module3

Generating Ideas, Brainstorming techniques, Application of Aesthetics and Ergonomics in Design. Bio-mimicry, Conceptualization, Visual thinking, Drawing/Sketching, Presenting ideas.

Module4

Useof prototyping, Typesof prototypes, Rapidprototyping techniques, Usertesting and feedback collection, Iterative prototyping, testing to gauge risk and market interest

Module5

Entrepreneurship/business ideas, Patents and Intellectual Property, Agility in design, Ethical considerations in design. Overcoming common implementation challenges

Corse Plan SyllabusandCorsePlan (For 3credit courses, thec ontent can be for 40 hrs and for2credit courses, the content can be for 26 hrs. The audit course in third semester can have content for 30hours).

No	Topic	No.oflectures
1	**Designprocess:**	
1.1	Design process: Traditional design, Design Thinking Approach, Introduction to Design Thinking, History and evolution of Design Thinking.	3
1.2	Roleofdesignthinkinginthehuman-centreddesign process.Designspace,	2
1.3	Design Thinking in a Team Environment, Team formation.	2

2	**DesignThinkingStages:**	
2.1	Design Thinking Stages: Empathize, Define, Ideate, Prototype and Test.	2
2.2	Theimportanceofempathy,Buildingauser-centred mindset.	2
2.3	Problemstatementformulation,Userneedsandpain points,establishingtargetspecifications,Settingthefinal specifications.	3
3	**Ideation**	
3.1	GeneratingIdeas,Brainstormingtechniques.	2
3.2	Applicationof AestheticsandErgonomicsinDesign.Bio-mimicry.	3
3.3	Conceptualization,Visualthinking,Drawing/Sketching, Presentingideas.	2
4	**Prototypingandtesting**	
4.1	Use of prototyping, Types of prototypes, Rapid prototyping techniques.	3
4.2	User testing and feedback collection, Iterative prototyping,testingtogaugeriskandmarketinterest	2
5	**IPRindesign**	
5.1	Entrepreneurship/business ideas, Patents and Intellectual Property.	2
5.2	Agilityin design, Ethical considerationsin design. Overcomingcommonimplementationchallenges	2

ReferenceBooks

1. ChristophMeinel,LarryLeiferandHassoPlattner- “DesignThinking: Understand – Improve – Apply”, Springer Berlin, Heidelberg, 2011.
2. ThomasLockwoodandEdgarPapke–“DesignThinking:IntegratingInnovation, Customer Experience, and Brand Value", Allworth Press, 2009.
3. PavanSoni–“DesignYourThinking”,PenguinRandomHouseIndiaPrivate Limited, 2020.
4. AndrewPressman-“DesignThinking:AGuidetoCreativeProblemSolvingfor Everyone”, Taylor & Francis, 2018.
5. NSivaPrasad,“DesignThinkingTechniquesanApproaches”AneBooksPvt. Ltd.,2023

.

SYLLABUS

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1AGT352	FUNCTIONALPROGRAMMINGIN HASKELL	AUDIT COURSE	3	0	0	-

Preamble: This course introduces a functional programming approach in problem solving. Salient features of functional programming like recursion, pattern matching, higher order functions etc. and the implementation in Haskell are discussed.

CourseOutcomes:

After thecompletionofthecoursethestudentwillbeableto

CO1	Understandthefunctionalprogrammingparadigmwhichisbasedonthemathematicsof lambda calculus.
CO2	DevelopHaskellprogramsusingfunctions,guardsandrecursivefunctions
CO3	Applytheconceptoftuples,listsandstringsinHaskellprogramming
CO4	Applytheconceptofalgebraicdatatypes,abstractdatatypes,modules,recursivedatatypes and user defined data types in Haskell programming
CO5	DevelopHaskellprogramswithfilesforreadinginputandstoringoutput

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1					3		
CO2	2			2	3		
CO3	2			2	3		
CO4	2			2	3		
CO5	2			2	3		

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Apply	40%
Analyse	40%
Evaluate	20%
Create	

Markdistribution

Total Marks	CIE	ESE	ESEDuration
100	40	60	2.5 hours

.

ContinuousInternalEvaluation:40marks

Coursebasedtask :15marks

Seminar/Quiz :15marks

Test paper, 1 no. : 10 marks

Testpapershallincludeminimum80%ofthesyllabus.

End Semester Examination: 60 marks

The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum onequestionfromeach moduleof whichstudent should answer any five. Each question can carry 12 marks.

ModelQuestionpaper

			TotalPages:
RegNo.:________________		Name:____________________________	
APJABDULKALAMTECHNOLOGICALUNIVERSITY THIRDSEMESTERM.TECHDEGREEEXAMINATION,DECEMBER2023			
CourseCode: J1AGT352			
CourseName:FunctionalProgramminginHaskell			
Max.Marks:60		Duration:2.5Hours	
Answeranyfivefullquestions,eachcarries12marks.			

1a.	Explainthebasicdifferencesbetweenimperativestyleprogrammingandfunctional style programming.	3
1b.	Analyse each of the following lambda expressions to clarify its structure. If the expressionisafunction,identifytheboundvariableandthebodyexpression,and thenanalysethebodyexpression.Iftheexpressionisanapplication,identifythe function and argument expressions, and then analyse the function and argument expressions: i) λa.(aλb.(ba)) ii) λx.λy.λz.((zx)(zy)) iii) (λf.λg.(λh.(gh)f)λp.λq.p)	9
2a.	Designarecursivefunctiontofind2^{n}wherenis anaturalnumber.	4

2b.	Explainvariousformsoffunctiondefinitionsin Haskellwiththehelpofexamples.	8
3a.	Explainanythreelistoperationsalongwithfunctiondefinitionsandexamples.	6
3b.	Writeaprogramtoduplicateonlyevennumbersamongtheelementsofalistusinga Haskell function by (i) Recursion (ii) List Comprehension and explain. Example:λ>dupli[1,2,3]ANS:[2,2]	6
4	WriteRecursivedefinitionsalongwithanexplanationforthebelowarithmetic operations. Illustrate the recursive flow with the help of a diagram. i. addx y ii. multxy iii. divx y	12
5	Write the Haskellcode to splitalistinto two listssuch thattheelementswithodd index are in one list while the elements with even index are in the other list.	12
6a	Givethetypedefinitionofabinarytreealongwithexplanationoftwofunctionson binary trees.	6
6b	DefineaqueuedatatypeinHaskellalongwithanytwooperationsonitwith examples.	6
7a.	ExplainthebasicstepsofreadingfromfilesandwritingtofilesinHaskell.	4
7b.	WriteaHaskellprogramtoreadfromthefile—input.txt‖,displaythecontentsonthe screen and write the contents to another file —output.txt‖.	8
****		

Syllabus and Corse Plan (For3creditcourses,thecontent canbefor40 hrsandfor2credit courses, the content can be for 26 hrs. The audit course in third semester can have content for 30 hours).

Module1(5Hrs)

Introduction to Functional Programming: Programming language paradigms, imperative style programming, comparison of programming paradigms.
Functional programming, Functions - Mathematical concepts and terminology, Lambda calculus, Function definitions, programs as functions, Functional programming Languages. Haskell basics, GHCi interpreter.

.

Module2(6Hrs)

ProgramminginHaskell: Expressionsandevaluation,Lazyevaluation,letexpressions,scopes.

BasicdatatypesinHaskell,operators,infixoperators,associativityandprecedence,Arithmetic functions.

types,definitions,curryinganduncurrying,typeabstraction.

Functiondefinitions,patternmatching,guards,anonymousfunctions,higherorderfunctions.

Recursion, Programming exercises.

Module3(7Hrs)

Datatypes:tuplesandlists:Tuples,Lists:buildinglists,decomposinglists,functionsonlists,built- in functions on lists, primitive and general recursion over lists, infinite lists.

Strings:functionsonstrings.

Polymorphismandoverloading,conditionalpolymorphism

Module4(6Hrs)

Typeclasses,Algebraicdatatypes,Modules,Recursivedatatypes.

Userdefineddatatypes,Records,Stacks,Queues,Binarytrees,Constructors,Destructors.

Module5(6Hrs)

Functor,Applicativefunctor,Monad

Programmingwithactions: Functionsvsactions,Basicsofinput/output,thedonotation, interacting with the command line and lazy I/O, File I/O.

No	Topic	No.ofLectures
1	**IntroductiontoFunctionalProgramming**	
1.1	Programming language paradigms, imperative style programming, comparison of programming paradigms	1
1.2	Functionalprogramming,Functions-Mathematicalconceptsand terminology	1
1.3	Lambdacalculus	1
1.4	Functiondefinitions,programsasfunctions,Functionalprogramming Languages	1
1.5	Haskellbasics,GHCiinterpreter	1
2	**Haskellbasics**	
2.1	Expressionsandevaluation,Lazyevaluation	1
2.2	letexpressions,scopes,BasicdatatypesinHaskell	1
2.3	operators,infixoperators,associativityandprecedence,Arithmetic	1

.

	functions	
2.4	types,definitions,curryinganduncurrying,typeabstraction.	1
2.5	Functiondefinitions,patternmatching,Guards	1
2.6	anonymousfunctions,higherorderfunctions,Recursion	1
3	**Datatypes:tuplesandlists**	
3.1	Tuples,Lists:buildinglists,decomposinglists	1
3.2	functionsonlists,built-infunctionsonlists	1
3.3	primitiveandgeneralrecursionoverlists	1
3.4	infinite lists	1
3.5	Strings:functionsonstrings	1
3.6	Polymorphismandoverloading	1
3.7	conditionalpolymorphism	1
4	**Userdefineddatatypes**	
4.1	Typeclasses,Algebraicdatatypes,Modules	1
4.2	Recursivedatatypes	1
4.3	Userdefineddatatypes,Records	1
4.4	Stacks,Queues	1
4.5	Binarytrees	1
4.6	Constructors,Destructors	1
5	**Programmingwithactions**	
5.1	Functor,Applicativefunctor,	1
5.2	Monad	1
5.3	Programmingwithactions:Functionsvsactions,Basicsofinput/ output, the do notation	1
5.4	interactingwiththecommandlineandlazyI/O	1
5.5	FileI/O	2

ReferenceBooks

[1] RichardBird,―Introductiontofunctionalprogrammingusing Haskell',secondedition,Prentice hall series in computer science

[2] BryanO'Sullivan,DonStewart,andJohnGoerzen, ―RealWorldHaskell‖

.

[3] RichardBird,—ThinkingFunctionallywithHaskell‖,CambridgeUniversityPress,2014

[4] SimonThompson,—Haskell:TheCraftofFunctionalProgramming‖,Addison-Wesley,3rd Edition, 2011

[5] H.ConradCunningham,—NotesonFunctionalProgrammingwithHaskell‖,2014

[6] GrahamHutton,—ProgramminginHaskell‖,CambridgeUniversityPress,2ndEdition,2016

[7] AlejandroSerrano Mena,—PracticalHaskell:AReal-WorldGuidetoFunctional Programming‖,3rdEdition,Apress,2022

[8] MiranLipovaca,—LearnYouaHaskellforGreatGood!:ABeginner'sGuide‖,NoStarchPress, 2011

J1AGT372	**REUSE AND RECYCLE TECHNOLOGY**	**CATEGORY**	**L**	**T**	**P**	**CREDIT**
		AUDIT COURSE	**3**	**0**	**0**	**-**

Preamble: "Reuse and Recycle Technology" typically focuses on sustainable practices and technologies aimed at reducing waste, conserving resources, and promoting environmental responsibility.

CourseOutcomes: After thecompletionofthecoursethestudentwillbeableto

CO1	Explaintheprinciplesandtechnologiesbehindwastereduction,resource conservation, and sustainable practices
CO2	DescribeandAnalyzewastegenerationandmanagement.
CO3	Applytheknowledgeofvariousreusestrategiesandtheirapplicationindifferent industriesandAnalyzevariousrecycling technologies
CO4	AppraisethemethodsofE-wastemanagementandEco friendlypackaging
CO5	ComprehendEnvironmentalRegulationsandPolicies**,**Understandtheimportance ofenvironmentalregulationsandpoliciesinaddressingenvironmentalchallenges

Mappingofcourseoutcomeswithprogram outcomes

	PO1	**PO2**	**PO3**	**PO4**	**PO 5**	**PO6**
CO1			3			
CO2				3		
CO3				3		
CO4					3	
CO5			3			

AssessmentPattern

Bloom'sCategory	**EndSemesterExamination**
Understand	60%
Apply	20%
Analyse	20%

Markdistribution

Total Marks	**CIE**	**ESE**	**ESE Duration**
100	40	60	2.5 hours

.

ContinuousInternalEvaluationPattern:40marks

Coursebasedtask:15marks
Seminar/Quiz : 15 marks
Test paper, 1 no. : 10 marks
Testpapershallincludeminimum80%ofthesyllabus.

EndSemesterExaminationPattern:60marks

The examination will be conducted by the respective College. The examination willbe for 150 minutesandwillcontain7questions,withminimumonequestionfromeachmoduleofwhich student should answer any five. Each question can carry 12 marks.

ModelQuestionpaper

AUDITCOURSE

J1AGT372 -REUSEANDRECYCLETECHNOLOGY

Answeranyfivefullquestions,eachcarries12marks.

1.	(a) What arethe3pillarsofsustainability? *(b)*Whatissustainablewastemanagement?Whatmakessustainablewaste managementsoimportant?	5 7
2.	(a)Howdothethreecategoriesofmunicipalsolidwastediffer? *(b)*Discussthemunicipalwastecollectionandmanagement?	5 7
3.	(a)ExplainthemajordifferencesbetweenReuseandRecycle? *(b)*Giveanoverviewofrecyclingtechnologiesusedforanytwomaterials. Discuss the Process involved.	5 7
4.	(a)Whatarethe commonsource ofE-waste *(b)*Whatarethechallengesand opportunitiesinE-wastemanagement	5 7
5.	(a)Whatisthecase lawforwasterecyclinginIndia *(b)*Discusssustainablepackaginganditsenvironmentalimpacts	5 7
6.	ExplainthevariousenvironmentalregulationsinIndiaforaddressing Environmentalchallenges	12
7.	a) Giveexamplesofwaterreusetechnologiesincirculareconomy b) Howcanwereducee-wastewithsustainablesolutions	5 7

.

Syllabus

Module	Content	Hours	Semester Exam Marks (%)
I	**Introduction to Sustainability**, Understanding sustainability and its importance, The three pillars of sustainability: Environmental, Social, and Economic. Biodiversity conservation, Climate change and mitigation Sustainable resource management.	6	20
II	**Waste Management,** Definition and classification of waste, Waste Generation and Composition**,** Waste Collection and Transportation, Waste Segregation and Sorting.Waste Disposal Methods Historical perspectives on waste management, The three Rs: Reduce, Reuse, and Recycle.	6	20
III	**Recycling and Reuse:** Importance of reuse, Application of reuse in various industries, Challenges and opportunities in reuse, Overview of recycling technologies,Circulareconomy,Sortingandprocessing of recyclable materials, Advanced recycling methods. Emerging technologies in recycling.	6	20
IV	**E-waste Recycling,** Challenges and environmental impact of electronic waste, E-waste recycling methods and regulations, Sustainable electronics design, **Sustainable Packaging,** Packaging materials and their environmental impact, Eco-friendly packaging alternatives,Packagingdesignforsustainability	6	20
V	**EnvironmentalRegulationsandPolicies,**Understand the importance of environmental regulations and policies in addressing environmental challenges, National and international waste and recycling regulations,Complianceandenforcement,Industry standardsandcertifications	6	20

CoursePlan

.

No	Topic	No.of Lectures
1	**IntroductiontoSustainability(6)**	
1.1	Understandingsustainabilityanditsimportance	1
1.2	Thethreepillarsofsustainability:Environmental,Social, and Economic.	3
1.3	Biodiversityconservation,Climatechangeandmitigation	1
1.4	Sustainableresourcemanagement	1
2	**WasteManagement(6)**	
2.1	Definitionandclassificationofwaste	1
2.2	WasteGenerationandComposition	1
2.3	WasteCollectionandTransportation.	1
2.4	WasteSegregationandSorting.	1
2.5	WasteDisposalMethods	1
2.6	Historicalperspectivesonwastemanagement,ThethreeRs: Reduce,Reuse,and Recycle.	1
3	**RecyclingandReuse(6)**	
3.1	Importanceofreuse,Examplesofreuseinvariousindustries.	1
3.2	Challengesandopportunitiesinreuse	1
3.3	Overviewofrecyclingtechnologies,Sortingandprocessingof recyclablematerials	2
3.4	Advancedrecyclingmethods	1
3.5	Emergingtechnologiesinrecycling.	1
4	**E-wasteRecycling(6)**	
4.1	Challengesandenvironmentalimpact ofelectronicwaste	1
4.2	E-wasterecyclingmethodsandregulations	1
4.3	Sustainableelectronicsdesign	1
4.4	Packagingmaterialsandtheirenvironmentalimpact	1
4.5	Eco-friendlypackagingalternatives	1
4.6	Packagingdesignforsustainability	1
5	**EnvironmentalRegulationsandPolicies(6)**	
5.1	Importanceofenvironmentalregulationsandpolicies in addressingenvironmentalchallenges	2
5.2	Nationaland internationalwasteandrecyclingregulations	2
5.3	Industrystandardsandcertifications,Compliance and enforcement	2

.

ReferenceBooks

1. SustainableEngineering:Concepts,DesignandCaseStudies,DavidT.Allen,Pearson Publication.
2. AComprehensiveBookonSolidWasteManagementwithApplication,Dr.H.S. Bhatia , Misha Books, 2019
3. "CradletoCradle:RemakingtheWayWeMakeThings"byWilliamMcDonough and Michael Braungart.
4. "RecyclingofPlasticMaterials"editedbyVijayKumar Thakur
5. E-waste:Implications,RegulationsandManagementinIndiaandCurrentGlobalBest Practices, Rakesh Johri, TERI
6. "SustainablePackaging‖,SubramanianSenthilkannanMuthu,Springer Nature.
7. **I**ndianEnvironmentalLaw:KeyConceptsandPrinciples"OrientBlackswanPrivate Limited, New Delhi.

J1AGT392	EXPERT SYSTEMS	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	-

Preamble: The course aims to provide an understanding of the basic concepts of Artificial Intelligence(AI)and ExpertSystems. Thecourse also coverstheknowledgerepresentation in expert systems, classes of expert systems, applications of expert systems.

CourseOutcomes: Afterthecompletionofthecoursethestudentwillbeableto:

CO1	ExplaintheconceptsofArtificialIntelligenceanddifferentwaysof knowledgerepresentations.
CO2	Explainthecomponentsofexpertsystems,developmentstagesofexpertsystems and tools available for expert system design.
CO3	Applytheconceptofknowledgerepresentationinexpertsystems
CO4	Differentiatetheclassesofexpertsystemsandexaminepropertiesofexisting systems

Mappingofcourseoutcomeswithprogram outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	1		2	1	2	2	
CO2	1		1	3	2	2	
CO3	1		1	2	2	2	
CO4	2		2	2	3	2	

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Understand	60%
Apply	20%
Analyse	20%

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

ContinuousInternalEvaluationPattern:40 marks

Coursebasedtask(Project/Assignments/Simulations/Casestudies):15marks

Seminar/Quiz : 15 marks

Testpaper,1no.:10 marks

Testpapershallincludeminimum80%ofthesyllabus.

EndSemesterExaminationPattern:**60marks**

The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 mark.

	APJABDULKALAMTECHNOLOGICALUNIVERSITY THIRDSEMESTERM.TECHDEGREEEXAMINATION,MARCH2024	
	CourseCode: J1AGT392	
	CourseName:EXPERTSYSTEMS	
Max.Marks:60		Duration:2.5Hours
Answeranyfivefullquestions,eachcarries12marks.		
1	a)WhatarethetypesofAI?Explainwithexamples.	6
	b)WhatdoyoumeanbyknowledgeinAIandexplainthedifferentways of knowledge representation used in AI?	6
2.	a)Writenoteonsemanticnetwork.	6
	b)WhatarePredicates?Explainitssyntaxandsemantics.	6
3.	a)Writenotesondifferent toolsavailableforexpertsystemdesign.	6
	b).Whatarethedifferentstagesinthedevelopment ofanexpertsystem?	6
4.	a)IllustrateConceptualDependencieswithanexample.	6
	b)IllustratewithanexampletheStructuredKnowledgerepresentationof an Expert System.	6
5.	a)Whatdo youmeanbyFramebasedExpert System?Explain	6
	b)ExplainthearchitectureofMYCIN	6
6.	a)ExplainFuzzybasedexpertsystems	6
	b)Explaintheneuralnetworkbasedexpertsystems	6
7.	a)Explainanytwoapplicationsofexpertsystems?	6
	b)Whatarethelimitationsofexpertsystem?Explain	6

.

Syllabus

Module	Content	Hours	Semester Exam Marks (%)
I	Overview of Artificial Intelligence (AI): Definition & Importance of AI. Knowledge general concepts: Definition and Importance of knowledge, Knowledge-Based Systems, Knowledge organization, Knowledge Manipulationand acquisition. Knowledge Representation: Introduction, Syntax and Semantics- Propositional logic andpredicate logic.	**6**	**20**
II	Basic concepts of expert systems-Introduction toexpert systems, Components of expert systems. Features of Expert System, Stages in the development of expert system, Types of tools available for expert system design	**6**	**20**
III	Knowledge representation in expert systems: Structured Knowledge representation: Graphs, Frames and related structures, Associative networks, Conceptual dependencies, Examples of structured knowledge representation.	**6**	**20**
IV	Classesofexpertsystems:Rule-basedexpertsystems, Example- MYCIN, Frame-based expert system, terminologies, IF-THEN structure. Fuzzyand Neural network based expert systems(basic concepts)	**7**	**20**
V	Currents trends in expert systems, Advantages and limitations of expert systems, Applications of expert systems.	**5**	**20**

.

CoursePlan

No	Topics	No.ofLectures
1	**OverviewofArtificialIntelligence&Knowledgegeneralconcepts**	
1.1	Definition&ImportanceofAI	1
1.2	DefinitionandImportanceofKnowledge,	1
1.3	Knowledge-BasedSystems,KnowledgeOrganization	1
1.4	KnowledgeManipulationandacquisition	1
1.5	KnowledgeRepresentation:Introduction,SyntaxandSemantics	1
1.6	Propositionallogicandpredicatelogic	1
2	**Basicconceptsofexpertsystems**	
2.1	IntroductiontoExpertSystem,Componentsofexpertsystems	2
2.2	FeaturesofExpertSystem,Stagesinthedevelopmentofexpertsystem	2
2.3	Typesoftoolsavailableforexpertsystem design	2
3	**Knowledgerepresentationinexpertsystems**	
3.1	Structured Knowledgerepresentation	1
3.2	Graphs, FramesandRelatedStructures	2
3.3	AssociativeNetworks,ConceptualDependencies	2
3.4	Examples ofstructuredknowledgerepresentation	1
4	**Classesofexpertsystems**	
4.1	Arule-basedexpertsystem-Introduction	1
4.2	MYCIN	1
4.3	IF-THENstructure	1
4.4	Frame-basedexpertsystem	2
4.5	Fuzzybasedexpertsystems	1
4.6	Neuralnetworkbasedexpertsystems	1
5	**Currentstrendsandapplicationsofexpertsystems**	
5.1	Currentstrendsofexpertsystems	2
5.2	Advantagesandlimitations ofexpertsystems	1
5.3	Applicationsofexpertsystems	2

ReferenceBooks

1. E.Rich&K.Knight-ArtificialIntelligence,2/e,TMH,NewDelhi, 2005.
2. P.H.Winston-ArtificialIntelligence,3/e,PearsonEdition,NewDelhi, 2006.
3. D.W.Rolston-PrinciplesofAI&ExpertSystemDevelopment,TMH,NewDelhi
4. KevinNightandElaineRich,NairB.,—ArtificialIntelligence(SIE)—,McGrawHill – 2010
5. DanWPatterson,_IntroductiontoArtificialintelligenceandExpertsystems‘,Prentice Hall of India Pvt. Ltd,2007
6. Russel(Stuart),_ArtificialIntelligence-Modernapproach,PearsonEducationseries inAI‘,3rdEdition, 2009.
7. I.Gupta,G.Nagpal·ArtificialIntelligenceandExpertSystems,MercuryLearning and Information -2020

.

J1AGT382	SYSTEM MODELLING	CATEGORY	L	T	P	CREDIT
		AUDIT COURSE	3	0	0	-

Preamble: Studyofthis courseprovides the learnersa clearunderstandingoffundamental concepts in simulation and modelling. This course covers the different statistical models, importance of data collection and various types of simulations. The course helps the learners to find varied applications in engineering, medicine and bio-technology.

CourseOutcomes: Afterthecompletionofthecoursethestudentwillbeableto

CO1	Analysetherequirementandfindappropriatetoolforsimulation.
CO2	Differentiatethedifferentstatisticalmodels.
CO3	Discussthedifferenttechniquesforgeneratingrandomnumbers.
CO4	Analysethedifferentmethodsforselectingthedifferentinputmodels..
CO5	Discussthedifferentmeasuresofperformanceandtheirestimation

Mappingofcourseoutcomeswithprogramoutcomes

	PO1	PO2	PO3	PO4	PO5	PO6
CO1	2		1	1	2	
CO2	2		1	1	1	
CO3	1					
CO4	1		1	1		
CO5	2		1	1	1	

AssessmentPattern

Bloom'sCategory	EndSemesterExamination
Understand	60%
Apply	20%
Analyse	20%

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5hours

ContinuousInternalEvaluationPattern:

Coursebasedtask(Project/Assignments/Simulations/Casestudies):15marks
Seminar/Quiz: 15 marks
Testpaper,1no.:10 marks
Testpapershallincludeminimum80%ofthesyllabus.

.

EndSemesterExaminationPattern:
The examination will be conducted by the respective College. The examination will be for 150 minutes and will contain 7 questions, with minimum one question from each module of which student should answer any five. Each question can carry 12 marks.

ModelQuestionpaper

AUDITCOURSE

J1AGT382–SYSTEMMODELLING

AnsweranyfivequestionsEachcarries12marks

PARTA

1. a.Discusstheadvantagesanddisadvantagesof simulation. (5marks)

   b.Whatarethe areasofapplicationsof simulation (7marks)

2. a.A bus arrives every 20 minutes at a specified stop beginning at 6:40 A.M. and continuing until 8:40 A.M. A certain passenger does not know the schedule, but arrives randomly (uniformly distributed) between 7:00A.M. and 7:30 A.M. every morning.Whatistheprobabilitythatthepassengerwaitsmorethan5minutesfora bus? (5 marks)

   b. A production process manufactures computer chips on the average at 2% nonconforming.Everyday,arandomsampleofsize50istakenfromtheprocess. Ifthesamplecontainsmorethantwononconforming chips,the processwillbestopped.Computetheprobabilitythattheprocessisstoppedbythe sampling scheme. (7 marks)

3. a.Discussthedifferenttypesoftestsforrandomnumbers. (5marks)

   b.GeneraterandomnumbersusingmultiplicativecongruentialmethodwithX0= 5,a11, and m = 64. (7 marks)

4. a.Whatarethedifferentmethodsofdatacollection. (4marks)

   b.Recordspertainingtothemonthlynumberofjob-relatedinjuriesatan underground coalrnine were beingstudiedbyafederalagency.Thevaluesforthepast100monthswereas follows:

InjuriesperMonth	FrequencyofOccurrence
0	35
1	40
2	13
3	6
4	4
5	1
6	1

.

(a) Applythechi-squaretesttothesedatatotestthehypothesisthattheunderlying distributionisPoisson.Usethe levelofsignificanceα==0.05.

(b) Applythechi-squaretesttothesedatatotestthehypothesisthatthedistribution is Poisson withmean 1.0. Again let α = 0.05. ·

c) What are the differences between parts (a) and (b), and when might each case arise? (8marks)

5. a.Whatisthedifference betweenvalidationandverification.(5marks)
   b.Discussthedifferentmeasuresofperformanceandtheirestimation(7marks)
6. a.Discussthedifferentmethodsofparameterestimation(5marks)
   b.Withanexample,describethe Poissonprocess.(7 marks)
7. a. Distinguish between discrete and continuous systems(5 marks)
   b.Whatarethedifferentcomponentsofasimulationsystem(7marks)

Syllabus

Module	**Content**	**Hours**	**Semester Exam Marks(%)**
I	When simulation is the appropriate tool. Advantages and disadvantages of Simulation; Areas of application, Systems and system environment; Components of a system; Discrete and continuous systems, Model of a system; Types of Models, Discrete-Event System Simulation, Steps of a simulation study.	**6**	**20**
II	Review of terminology and concepts, Useful statistical models, Discrete distributions. Continuous distributions, Poisson process, Empirical distributions. (basic idea only)	**6**	**20**
III	Properties of random numbers; Generation of pseudo-randomnumbers,Techniquesforgeneratingrandom numbers,TestsforRandomNumbers	**6**	**20**
IV	Data Collection; Identifying the distribution with data, Parameter estimation, Goodness of Fit Tests, Fitting a non-stationaryPoissonprocess,Selectinginputmodels without data, Multivariate and Time-Series input models.	**6**	**20**
V	Measures of performance and their estimation, Output analysis for terminating simulations, Output analysis for steady-state simulations, Verification, calibration and validation	**6**	**20**

.

CoursePlan

No	Topic	No.of Lectures
1	**Introduction**	
1.1	Whensimulationistheappropriatetool	1
1.2	Advantagesanddisadvantagesof Simulation;	1
1.3	Areasofapplication,Systemsandsystemenvironment;	1
1.4	Componentsofasystem;Discreteandcontinuoussystems,	1
1.5	Model of asystem;TypesofModels,	1
1.6	Discrete-EventSystemSimulation,Stepsofasimulationstudy	1
2	**StatisticalModelsinSimulation**	
2.1	Reviewofterminologyandconcepts, Empiricaldistributions. (basicideaonly)	1
2.2	Usefulstatisticalmodels,	1
2.3	Discretedistributions.	1
2.4	Continuousdistributions,.	1
2.5	Poissonprocess	1
2.6	Empirical distributions	1
3	**RandomNumberGeneration**	
3.1	Propertiesof randomnumbers;	1
3.2	Generationofpseudo-random numbers,	
3.3	Techniquesforgeneratingrandom numbers	1
3.4	Techniquesforgeneratingrandom numbers(cont)	1
3.5	TestsforRandomNumbers	1
3.6	TestsforRandomNumbers(cont)	1
4	**InputModelling**	
4.1	DataCollection;	1
4.2	Identifyingthedistributionwithdata.	1
4.3	Parameterestimation,GoodnessofFitTests	1
4.4	Fittinganon-stationaryPoissonprocess	1
4.5	Selectinginputmodelswithoutdata,	1
4.6	MultivariateandTime-Seriesinputmodels	1
5	**MeasuresofPerformanceandtheirEstimation**	
5.1	Measuresofperformanceandtheirestimation	1
5.2	Measuresofperformanceandtheirestimation(cont)	1
5.3	Outputanalysisforterminatingsimulations	1
5.4	Outputanalysisforsteady-statesimulations	1
5.5	Verification,calibrationand validation	1
5.6	Verification,calibrationand validation(cont)	1

.

Textbooks:

1. JerryBanks,JohnS.CarsonII,BarryL.Nelson,DavidM.Nicol:Discrete-EventSystem Simulation, 5th Edition, Pearson Education, 2010.

ReferenceBooks:

1. LawrenceM.Leemis,StephenK.Park:Discrete–EventSimulation:AFirstCourse, Pearson Education, 2006.

2. AverillM.Law:SimulationModelingandAnalysis,4thEdition,TataMcGraw-Hill,2007

3. SystemModellingandResponsebyErnestO.Doebelin

4. AverillMLaw,—SimulationModelingandAnalysis‖,McGraw-HillInc,2007Geoffrey Gorden,—SystemSimulation‖,PrenticeHallof India,1992.

.

J1AGT362	Principles of Automation	CATEGORY	L	T	P	CREDIT
		CREDIT COURSE	3	0	0	0

Preamble:

This course deals in detail with the various aspects of automation such as sensors, actuators, controllers, mechanical and electrical elements and their integration for automating new and existing manufacturing and process industries and applications. This course will be beneficial to students in designing automation schemes for industries and to design automated systems

CourseOutcomes: Afterthecompletionofthecoursethestudentwillbeableto

CO1	Explainthefundamentalsofsensorsystemsandtochooseasuitablesensorsystem for the given application based on the evaluation of the constraints.
CO2	Explainthefundamentalsofsignalconditionsandtodesignasuitablesignal conditioning scheme for given application.
CO3	Describethecharacteristicsofvariousactuatorsystemsandtodecidetheright typeofactuator for thegiven application.
CO4	Describetheimportanceofanindustrialrobotandfundamentalsofnumerical control in automation.
CO5	Explainthefundamentalsofcontrollersusedinindustrialautomationandto construct simple automation schemes by ladder logic programs.

Mappingofcourseoutcomeswithprogram outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	2		2	2	2		
CO2	2		2	2	2		
CO3	2		2	2	2		
CO4	2		2	2	2		
CO5	2		2				

AssessmentPattern

Bloom'sCategory	EndSemester Examination
Understand	70 %
Apply	30 %

Markdistribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	2.5 hours

.

ContinuousInternalEvaluationPattern:40 marks
Coursebasedtask(Project/Assignments/Simulations/Casestudies):15marks
Seminar/Quiz: 15 marks
Testpaper,1no.:10 marks
Testpapershallincludeminimum80%ofthesyllabus.

EndSemesterExaminationPattern:**60marks**
TheexaminationwillbeconductedbytherespectiveCollege.Theexaminationwillbefor150 minutesandwillcontain7questions, withminimumonequestionfromeachmoduleofwhich student should answer any five. Each question can carry 12 marks.

ModelQuestionPaper

J1AGT362– Principles of Automation

Time2.5 Hrs Marks60

AnsweranyfivequestionsEachcarries12marks

1. (a) Differentiate the static and dynamic characteristics of a temperature sensor and explainhowitaffectsthe selectionofasuitabletemperature sensor. (6marks)
   (b) Explaintheworking ofastrain-gauge. (6marks)

2. (a)Explain whyanti-aliasingfiltersareusedinanalogtodigitalconverters.(3marks)
   (b) Designafirst orderlowpassfilterwith acutofffrequency of2 kHz. (9marks)

3. (a) What are the factors to consider while deciding choosing between hydraulic, pneumatic or electrical actuation systems for an automation scheme? (4marks)
   (b) Explaintheworkingofathree-waypressurereducingvalve. (4marks)
   (c) Explain the working of solenoids. In what applications would you use a Solenoid valve. (4marks)

4. (a) Explain the principle of the Touch sensor and also mention how they are usedin robots. (5marks)
   (b) Explainthebasicterminologiesinroboticsystemandalsoexplainthecomponents of robotic system. (7marks)

5. (a)Withneat schematicexplainthearchitectureofthePLC. (6marks)
   (b) Explaintheuse ofan up-downcounterin PLCwith asuitableexample.(6marks)

6. (a) Write short note on SCADA. What is difference PLC and SCADA? (3marks)

   (b) Constructaladderlogicforcontrollingaprocesstank asperthelogicgiven below;
   i. ThetankshouldbefilledbyavalveV1whenlowlevelfloatswitchL1is ON and an external input S1 is received.

ii. V1shouldbeclosedwhentheliquidlevelreachesahigh-levelfloatswitchL2.

iii. Anagitatormotorshouldbeturnedonafteradelayof5secafterL2is triggered.

iv. Afteragitatingfor30mins,contentsofthetankshouldbeemptiedbyopening another valve V2.

v. Thetemperatureshouldbemaintainedat70°Cusingathermostat T1andHeaterH (9marks)

7. (a)Explainthelevelsof Automation. (6marks)
(b) ExplaintheworkingofFlowsensor (6marks)

SyllabusandCoursePlan

No	Topics	No.of Lectures
1	**IntroductiontoIndustrialAutomation**	
1.1	BasicElementsofanAutomatedSystem,Levelsof Automation	2
1.2	HardwarecomponentsforAutomation:Sensors,classification,Static and dynamic behaviour of sensors.	2
1.3	Basic working principle of different sensors: Proximity sensors, Temperaturesensors,flowsensors,Pressuresensors,Forcesensors. Positionsensors	4
2	**Signalconditioning**	
2.1	Needforsignalconditioning,Typesofsignalconditioning.	2
2.2	Signalconditioningusingoperationalamplifier-Amplifier(Inverting and Non-inverting) and Filter circuits (Basic concepts). Design of first order low pass filter.	2
2.3	Signalconditioningfordataacquisitionsystems,anti-aliasingfilters, Analog–Digital Conversions, Analog-to-Digital Converters (ADC)- Steps in analog-to-digital conversion, Successive Approximation Method, Digital-to-Analog Converters (DAC)- Steps in digital to analog conversion, Zero-order and first order data hold circuits	4
3	**Actuators**	
3.1	Typesofactuators-mechanical,electrical,pneumaticandhydraulic actuators. (Basic working principle)	2
3.2	Mechanicalsystemsformotionconversion,transmissionsystems	3
3.3	Solenoids,Electricandsteppermotorscontrol.	3
4	**RoboticsandAutomatedManufacturingSystems**	
4.1	RobotAnatomyandRelatedAttributes:JointsandLinks,Common RobotConfigurations,JointDriveSystems,SensorsinRobotics (Basic concepts)	3
4.2	RobotControlSystems,ApplicationsofIndustrialRobots-Material handling	4
4.3	FundamentalsofNumericalcontrol(NC)Technology	1
5	**DiscreteControlandProgrammableLogicControllers**	

.

5.1	DiscreteProcessControl:LogicandSequence control	2
5.2	LadderLogicDiagrams,ProgrammableLogicControllers: ComponentsofthePLC,PLCOperatingCycle,Programmingthe PLC (Basic concepts only)	4
5.3	IntroductiontoDistributedcontrolsystem(DCS)andSupervisory Control and Data Acquisition Systems (SCADA)	2

ReferenceBooks

1. Mikell Groover, Automation, Production Systems, and Computer-Integrated Manufacturing, 5th Edition, Pearson, 2019.
2. YoramKoren,―ComputerControlofManufacturingSystems‖,TataMcGrawHill Edition2005.
3. S.R.Deb;SankhaDeb.RoboticsTechnologyandFlexibleAutomation,SecondEdition McGraw-Hill Education: New York, 2010.
4. W.Bolton,―Mechatronics:ElectronicControlSystemsinMechanicalandElectrical Engineering‖ - PrenticeHall- 2013 - 5th Edition.
5. Doebelin,E.O.andManic,D.N.,―MeasurementSystems:ApplicationsandDesign‖, 7thEdition,McGrawHill,2019.
6. KrishnaKant,ComputerBasedIndustrialControl-,EEE-PHI,2ndedition,2010.
7. NathanIda,Sensors,Actuators,andTheirInterfaces-Amultidisciplinaryintroduction, 2nd Edition, IET Digital Library, 2020.
8. Salivahanan,S.,andVSKanchanaBhaaskaran.Linearintegratedcircuits.McGraw- Hill Education, 2nd edition, 2014.
9. Petruzella,FrankD.Programmablelogiccontrollers.TataMcGraw-HillEducation, 2005
10. ChapmanandHall,―StandardHandbookofIndustrialAutomation‖,OnsidineDMC& OnsidineGDC‖,NJ,1986

.

J1INT303:INTERNSHIP

A student shall opt for carrying out the Internship at an Industry/Research Organization or at another institute of higher learning and repute (Academia). The organizationforInternshipshallbeselected/decidedbythestudentsontheirownwith prior approval from the faculty advisor/respective PG Programme Coordinator/Guide/Supervisor. Every student shall be assigned an internship Supervisor/Guide at the beginning of the Internship. The training shall be related to their specialisation after the second semester for a minimum duration of six to eight weeks. On completion of the course, the student is expected to be able to develop skills in facing and solving the problems experiencing in the related field.

Objectives

- Exposuretotheindustrialenvironment,whichcannotbesimulatedinthe classroom and hence creating competent professionals for the industry.
- Providepossibleopportunitiestolearn understand andsharpen therealtime technical / managerial skills required at the job.
- Exposuretothecurrenttechnologicaldevelopmentsrelevanttothesubject area of training.
- Createconduciveconditionswithquestforknowledgeanditsapplicability on the job.
- Understandthesocial,environmental,economicandadministrative considerations that influence the working environment.
- Exposestudentstotheengineer'sresponsibilitiesandethics.

BenefitsofInternship Benefits

to Students

- AnopportunitytogethiredbytheIndustry/organization.
- Practicalexperienceinanorganizationalsetting&Industryenvironment.
- Excellentopportunitytoseehowthetheoreticalaspectslearnedinclasses are integrated into the practical world. On-floor experience provides much more professional experience which is often worth more than classroom

.

teaching.

- Helpsthemdecideiftheindustryandtheprofessionisthebestcareer option topursue.
- Opportunitytolearnnewskillsandsupplementknowledge.
- Opportunitytopracticecommunicationandteamworkskills.
- Opportunitytolearnstrategiesliketimemanagement,multi-taskingetcin an industrial setup.
- Makesavaluableadditiontotheirresume.
- Enhancestheircandidacyforhighereducation/placement.
- Creatingnetworkandsocialcircleanddevelopingrelationshipswith industry people.
- Providesopportunitytoevaluatetheorganizationbeforecommittingtoa full time position.

BenefitstotheInstitute

- Buildindustryacademiarelations.
- Makestheplacementprocesseasier.
- Improveinstitutionalcredibility&branding.
- Helpsinretentionofthestudents.
- CurriculumrevisioncanbemadebasedonfeedbackfromIndustry/ students.
- Improvementinteachinglearningprocess.

BenefitstotheIndustry

- Availabilityofreadytocontributecandidatesforemployment.
- Yearroundsourceofhighlymotivatedpre-professionals.
- Studentsbringnewperspectivestoproblemsolving.
- Visibilityoftheorganizationisincreasedoncampus.

- Qualitycandidate’savailabilityfortemporaryorseasonalpositionsand projects.
- Freedomforindustrialstafftopursuemorecreativeprojects.
- Availabilityofflexible,cost-effectiveworkforcenotrequiringalong-term employer commitment.
- Proven,cost-effectivewaytorecruitandevaluatepotentialemployees.
- Enhancementofemployer’simageinthecommunitybycontributingtothe educational enterprise.

TypesofInternships

- IndustryInternshipwith/withoutStipend
- Govt/PSUInternship(BARC/Railway/ISROetc)
- Internshipwithprominenteducation/researchInstitutes
- InternshipwithIncubationcentres/Start-ups

.

Guidelines

- All the students need to go for internship for minimum duration of 6 to 8 weeks.
- Studentscantakeminiprojects,assignments,casestudiesbydiscussingit withconcernedauthorityfromindustry andcanworkonitduring internship.
- All students should compulsorily follow the rules and regulations as laid by industry.
- Every student should take prior permissions from concerned industrial authority if they want to use any drawings, photographs or any other document from industry.
- StudentshouldfollowallethicalpracticesandSOPofindustry.
- Studentshavetotakenecessaryhealthandsafetyprecautionsaslaidby theindustry.
- Studentshouldcontacthis/herGuide/Supervisorfromcollegeonweekly basis to communicate the progress.
- Eachstudenthastomaintainadiary/log book
- Aftercompletionofinternship,studentsarerequiredtosubmit
    - Reportofworkdone
    - Internshipcertificatecopy
    - Feedbackfromemployer/internshipmentor
    - Stipendproof(incaseofpaidinternship).

Total Marks 100: ThemarksawardedfortheInternshipwillbeonthebasisof(i) Evaluation done by the Industry (ii) Students diary (iii) Internship Report and (iv) Comprehensive Viva Voce.

Continuous InternalEvaluation:50marks

Student's diary - 25 Marks

Evaluation done by the industry- 25 Marks

Student'sDiary/DailyLog:Themainpurposeofwritingdailydiaryistocultivatethe habit of documenting and to encourage the students to search for details. Itdevelops thestudents'thoughtprocessand reasoningabilities. Thestudentsshould record in the daily training diary the day to day account of the observations,

.

impressions,informationgatheredandsuggestionsgiven,ifany.Itshouldcontainthe sketches & drawings related to the observations made by the students. The daily training diary should be signed after every day by the supervisor/ in charge of the section where the student has been working. The diary should also be shown to the FacultyMentorvisitingtheindustryfromtimetotimeandgotratifiedonthedayofhis visit. Student's diary will be evaluated on the basis of the following criteria:

- Regularityinmaintenanceofthediary
- Adequacy&qualityofinformationrecorded
- Drawings,design,sketchesanddatarecorded
- Thoughtprocessandrecordingtechniquesused
- Organizationoftheinformation.

Theformatofstudent'sdiary

NameoftheOrganization/Section:

NameandAddressoftheSection Head:

NameandAddressoftheSupervisor:

Nameandaddressofthestudent:

InternshipDuration:From...........................To...

Briefdescriptionaboutthenatureofinternship:

Day	Brief write up about the Activities carried out: Such as design, sketches,resultobserved,issuesidentified,datarecorded,etc.
1	
2	
3	

Signature of Industry Supervisor

SignatureofSectionHead/HRManagerOf

fice Seal

.

AttendanceSheet

NameoftheOrganization/Section:

NameandAddressoftheSection Head:

NameandAddressoftheSupervisor:

Nameandaddressofthestudent:

InternshipDuration:From...........................To...

Month &Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	...
Month &Year																					
Month &Year																					

Signature of Industry Supervisor *SignatureofSectionHead/HRManager*

Office Seal

Note:

- Student's Diary shall be submitted by the students along with attendance record and an evaluation sheet duly signed and stamped by the industry to the Institute immediately after the completion of the training.
- Attendance Sheet should remain affixed in daily training diary. Do not remove or tear it off.
- Studentshallsignintheattendancecolumn.Donotmark'P'.
- Holidaysshouldbemarkedinredinkintheattendancecolumn.Absent shouldbemarkedas 'A'inredink.

.

EvaluationdonebytheIndustry(Marks25)

FormatforSupervisorEvaluationofIntern

Student Name :________________________________Date:________________________________

Supervisor Name :_____________________________Designation:__________________________

Company/Organization:____

Internship Address: ______________________________

____Dates ofInternship:From

____To_______________________________

Please evaluateintern by indicating the frequency with which you observed the following parameters:

Parameters **Marks**	Needs improvement (0 - 0.25 mark)	Satisfactory (0.25-0.50 mark)	Good (0.75 mark)	Excellent (1mark)
Behavior				
PerformsinadependableManner				
Cooperateswithcoworkersandsupervisor				
Showsinterestinwork				
Learnsquickly				
Showsinitiative				
Produces high quality work				
Acceptsresponsibility				
Acceptscriticism				
Demonstratesorganizationalskills				
Usestechnicalknowledgeandexpertise				
Showsgoodjudgment				
Demonstratescreativity/originality				
Analyzesproblemseffectively				
Isself-reliant				
Communicates well				
Writeseffectively				
Hasaprofessionalattitude				
Givesaprofessionalappearance				
Ispunctual				
Usestimeeffectively				

Overallperformanceofstudent

Intern (Tick one) : Needsimprovement(0-0.50mark)/Satisfactory(0.50-1.0mark)/ Good (1.5 mark) / Excellent (2.0 mark)

Additionalcomments,ifany(2marks):

Signature of Industry Supervisor *SignatureofSectionHead/HRManager*

Office Seal

.

EndSemesterEvaluation(ExternalEvaluation):50Marks

InternshipReport	-	25Marks
VivaVoce	-	25Marks

InternshipReport:Aftercompletionoftheinternship,thestudentshouldprepare a comprehensive report to indicate what he has observed and learnt in the training period and should be submitted to the faculty Supervisor. The student may contact Industrial Supervisor/ Faculty Mentor for assigning special topics and problems and should prepare the final report on the assigned topics. Daily diary will also help to a greatextentin writing theindustrial report sincemuchoftheinformation has already been incorporated by the student into the daily diary. The training report should be signed by the Internship Supervisor, Programme Coordinator and Faculty Mentor.

TheInternshipreport(25Marks)willbeevaluatedonthebasisoffollowingcriteria:

- Originality
- Adequacy andpurposefulwrite-up
- Organization,format,drawings,sketches,style,languageetc.
- Varietyandrelevanceoflearningexperience
- Practicalapplications,relationshipswithbasictheoryandconceptstaught in the course

VivaVoce(25Marks)willbedonebyacommitteecomprisingFacultySupervisor, PG Programme Coordinator and an external expert (from Industry or research/academic Institute). This committee will be evaluating theinternship report also.

.

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT
J1PADP304	RESEARCH PROJECT PHASE I	PROJECT WORK	0	0	17	11

ResearchProject: Studentschoosingtrack2shallcarryouttheresearchproject in their parent Institution only under the guidance of a supervisor assigned by the DLAC.

Dissertation: Allcategoriesofstudentsintrack1aretocarryoutthedissertation in the Institute they are studying or can work either in any CSIR/Industrial R&D organization/any other reputed Institute which have facilities for dissertation work in the area proposed.

Mark Distribution:

Phase1:Totalmarks:100,only CIA

SEMESTER-IV

.

CODE	COURSENAME	CATEGORY	L	T	P	CREDIT
J1PADD401	**DISSERTATIONPHASE II**	**Project Work**	**0**	**0**	**24**	**16**

All categories of students in track 1 are to carry out the DISSERTATION PHASE II in the institute they are studying or in any Industrial/ R&D organization/any other reputed institute which have facilities for dissertation work in the area proposed. DISSERTATION PHASE II shall not compulsorily continuation of DISSERTATIONPHASEI. The student has to publisha research article in a conference or a reputed journal before appearing for the end-semester examination. The eligibility criteria for registering to the end semester examination are attendance in the course and no pending disciplinary action. The minimum attendance for appearing for the end semester examination is 75%. Students who do not meet these eligibility criteria are ineligible (identified by FE grade) to appear for the ESE. Students, who have completed a course but could not appear for the end semester examination, shall be awarded _AB' Grade, provided they meet other eligibility criteria The pass minimum for the course is 45% for ESE and 50% for (CIA and ESE) put together.

ContinuousInternalAssessment(CIA)TotalMarks:100

Theevaluationcommitteecomprises

1- ProjectCoordinator(s)
2- ASenior facultymember
3- Supervisorofthestudent

Pattern(CIA)

Zerothevaluationbythe EvaluationCommittee	-
InterimevaluationbytheEvaluationCommittee	30 marks
FinalevaluationbytheEvaluation Committee	50 marks
Projectprogressevaluationbysupervisor	20 marks

Evaluationby thesupervisor

The guide/supervisor shall monitor the progress being carried out by the student on a regular basis. In case it is found that progress is unsatisfactory it shall be reported to the Department Evaluation Committee for necessary action.

Student's Diary/ Log book: The main purpose of writing diary/log book is to cultivate the habit of documenting and to encourage the students to search for details. The activity diary shall be signed afterevery week by the supervisor.

EndSemesterEvaluation(ESE)TotalMarks:100

Theevaluationcommitteecomprises

1- ProjectCoordinator(s)
2- Anexternalexpert(fromIndustryorresearch/academic institute)
3- Supervisorofthestudent

.

Pattern(ESE)

1. **InnovationandOriginality(10marks):**
Assessmentoftheuniquenessandinnovationdemonstratedintheprojectwork.
Original contributions, if any, to the field or problem area.
2. **ImplementationandExecution(20marks):**
Evaluationoftheactualimplementationorexecutionoftheproject,including: Quality
of work done
Demonstrated skills and techniques applied
Adherencetoprojecttimelinesandmilestones
3. **ProjectDocumentation(25marks):**
Comprehensive project report evaluation including:
Introduction and problem statement
Literature review
Methodologyandapproach
Results and analysis
Conclusionandrecommendations References
and citations
Detailsofthepublications
Plagiarism certificate

ThePlagiarismlevelintheprojectreportshallbelessthan25%.

4. **PresentationandDefence(40marks):**
Oralpresentationoftheprojecttoapanelofexaminers,including: Clarity
and effectiveness of the presentation
Ability to explain the project objectives, methodologies, and findings
Handlingquestionsandprovidingsatisfactoryanswersduringthedefence

5. **Publicationoftheworkeitherinaconferenceorinajournal(5marks)**

SYLLABUS:

DETAILS	HOURS
` 1. Literaturestudy/surveyofpublishedliteratureontheassigned topic 2. TopicSelectionandProposal 3. Formulationofobjectives 4. ResearchandPlanning 5. Formulationofworkplanandtaskallocation. 6. Execution 7. DocumentationandReporting 8. ProjectShowcase reflectingon the projectexperience andlessons learned	200

.

Dissertation outside the Institute: For doing dissertation outside the Institution, the following conditions are to be met:

i. They have completed successfully the course work prescribed in the approved curriculum up to the second semester.
ii. The student has to get prior approval from the DLAC and CLAC.
iii. Facilities required for doing the dissertation shall be available in the Organization/Industry (A certificate stating the facilities available in the proposed organization and the time period for which the facilities shall be made available to the student, issued by a competent authority from the Organization/Industry shall be submitted by the student along with the application).
iv. They should have an external as well as an internal supervisor. The internal supervisor should belong to the parent institution and the external supervisorshould be Scientists or Engineers from the Institution/Industry/ R&D organization with which the student is associated for doing the dissertation work. The external supervisor shall be with a minimum post graduate degree in the related area.
v. The student has to furnish his /her monthly progress as well as attendance report signed by the external supervisor and submit the same to the concerned internal supervisor.
vi. The external supervisor is to be preferably present during all the stages of evaluation of the dissertation.