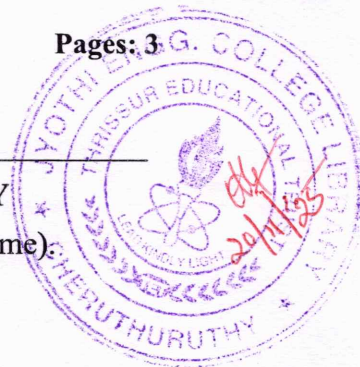


Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
B.Tech Degree S5 (R,S) Examination November 2025 (2019 Scheme)



Course Code: AMT 305

Course Name: INTRODUCTION TO MACHINE LEARNING

Max. Marks: 100

Duration: 3 Hours

PART A

(Answer all questions; each question carries 3 marks)

		Marks
1	Define machine learning with suitable examples. What makes machine learning different from traditional programming.	3
2	Write short notes on classification and regression with example.	3
3	Is principal component analysis a supervised learning problem? Justify your answer.	3
4	How can Naive Bayes classifier using Bayes' Theorem can be derived and proved using the concepts of conditional probability and the law of total probability?	3
5	Discuss the significance of validation data during training phase.	3
6	A neuron with 4 inputs has the weights 0.3, 0.8, 0.2, 0.5 and bias 1. The activation function is linear, $f(x) = 2x$. If the inputs are 4, 3, 5, 2, compute the output. Draw a diagram representing the neuron.	3
7	Differentiate between Maximum Likelihood Estimation (MLE) and Maximum A Posteriori (MAP).	3
8	Show that the function $K(\vec{x} \circ \vec{y}) = (\vec{x} \circ \vec{y})^2$ is a kernel function.	3
9	Explain voting and the advantages of voting in ensemble method.	3
10	Explain any three different ways to measure the similarity while using k-means clustering algorithm.	3

PART B

(Answer one full question from each module, each question carries 14 marks)

Module -1

- | | | |
|----|--|---|
| 11 | a) Write a short note on hypothesis space, version space, most specific and most general hypothesis. | 8 |
| | b) Consider the following scenarios and determine which machine learning | 6 |

paradigm (Supervised, Unsupervised, or Reinforcement learning) would be most appropriate. Justify your answer for each case.

Scenario 1: An e-commerce platform wants to group its products into categories based on customer browsing and purchase patterns.

Scenario 2: A self-driving car needs to learn how to navigate through city traffic while obeying traffic rules.

Scenario 3: A healthcare company wants to predict the likelihood of a patient developing a specific disease based on their medical history and genetic information.

- 12 a) Define VC dimension. Show that an axis aligned rectangle can shatter 4 points 2 dimensions. 8
- b) Explain the methods used to learn multiple classes for a K class Classification Problem. 6

Module -2

- 13 a) Construct a linear regression model to predict the annual home insurance premium based on the given data. What would be the predicted premium for a home that is 2000 square feet? 8

Home size (in square feet)	1200	1500	1800	2200	2500	3000
Annual home insurance premium (\$)	900	1100	1250	1450	1600	1850

- b) Explain various dimensionality reduction techniques. Why dimensionality reduction technique is important? 6
- 14 a) Identify the first splitting attribute for decision tree by using ID3 algorithm with the following dataset. 8

A1	A2	A3	Class
True	Hot	High	No
True	Hot	High	No
False	Hot	High	Yes
False	Cool	Normal	Yes
False	Cool	Normal	Yes
True	Cool	High	No
True	Hot	High	No
True	Hot	Normal	Yes
False	Cool	Normal	Yes
False	Cool	High	Yes

- b) Discuss logistic Regression. Write the pros and cons of logistic regression 6

Module -3

- 15 a) Describe the basic outline of the backpropagation algorithm. 8
- b) Suppose there are three classifiers X, Y, and Z. The (FPR, TPR) measures of the three classifiers are as follows: Classifier X: (0, 0.9), Classifier Y: (0.5, 0.75), Classifier Z: (0, 1). Which of these classifiers can be considered the best, and why? 6
- 16 a) Explain cross-validation in machine learning. Explain the different types of cross-validations. 8
- b) Describe the working of a multilayer feed forward network 6

Module -4

- 17 a) Describe soft-margin SVM with neat diagram. Differentiate between hard-margin SVM and soft-margin SVM. 8
- b) Derive the equation for margin between support vectors in SVM? 6
- 18 a) What is a kernel in SVM? Why do we use kernels in SVM? 8
- b) What is "Kernel trick"? Explain different types of kernel functions. 6

Module -5

- 19 a) Explain any two ensemble methods used for learning. 4
- b) For the given data, compute two clusters using K-means algorithm for clustering where initial cluster centers are (1, 1) and (5, 7). Execute for two iterations. 10

x	1	1.5	3	5	3.5	4.5	3.5
y	1	2	4	7	5	5	4.5

- 20 a) Illustrate Single linkage and Complete linkage in hierarchical clustering. 4
- b) Given the dataset {a, b, c, d, e} and the following distance matrix, construct a dendrogram by complete-linkage hierarchical clustering using the agglomerative method. 10

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
<i>a</i>	0	9	3	6	11
<i>b</i>	9	0	7	5	10
<i>c</i>	3	7	0	9	2
<i>d</i>	6	5	9	0	8
<i>e</i>	11	10	2	8	0
