



Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
B.Tech Degree S4 (R) Examinations April 2026 (2024 Scheme)

Course Code: PEECT413
Course Name: MACHINE LEARNING

Max. Marks: 60

Duration: 2 hours 30 minutes

PART A

(Answer all questions. Each question carries 3 marks)

		CO	Marks
1	Describe model-based learning with an example.	1	(3)
2	When is underfitting said to happen in machine learning?	1	(3)
3	Differentiate between linear and multivariate regression.	2	(3)
4	What do you mean by "Support Vectors" in the context of Support Vector Machines (SVM) classifier?	2	(3)
5	How can the number of clusters be determined in k-means clustering?	3	(3)
6	What is exploration vs exploitation in Reinforcement Learning?	4	(3)
7	Draw the structure of a biological neuron.	2	(3)
8	Describe the RELU activation function and its properties.	2	(3)

PART B

(Answer any one full question from each module, each question carries 9 marks)

Module -1

- 9 a) An AI system monitors credit card transactions for potential fraud. In a single day, the system flagged 500 transactions as "Fraudulent." Internal audits confirmed that 480 of those flagged transactions were indeed fraudulent. However, the system completely missed 20 fraudulent transactions, labeling them as "Legitimate." Calculate:
- i) True Positives ii) False Positives iii) Precision iv) Recall

- b) What is an ROC curve? Explain what the AUC (Area Under the Curve) value indicates about a classifier's performance. 1 (3)
- 10 a) With necessary plots explain bias-variance trade off. 1 (5)
- b) Distinguish between batch learning and online learning. 1 (4)

Module -2

- 11 a) You are developing a linear regression model to predict housing prices. For a small validation set of 4 properties, the actual values (y) and the predicted values (y') are:
 Actual (y): [50000,60000,70000,80000]
 Predicted (y'): [48000,65000,68000,85000]
- i) Calculate the Mean Squared Error (MSE) for this model.
- ii) The model has a learned weight $w=4$. If a regularization parameter $\lambda=0.5$ is applied, determine which regularization term—L1 or L2—would result in a higher penalty value added to the loss function.
- b) A logistic regression model is used to predict if a transaction is fraudulent ($y=1$). For a specific transaction that was actually legitimate ($y=0$), the model predicted a probability of fraud of 0.60. Calculate the Binary Cross-Entropy loss for this specific sample. (use the natural logarithm) 2 (3)
- 12 a) Random Forest is just Bagging applied to Decision Trees — then why is Random Forest better than simple Bagging of Decision Trees? 2 (6)
- b) Define the Naive Bayes Classifier. What fundamental assumption does it make about the input features? 2 (3)

Module -3

- 13 a) Given two feature vectors representing two different documents, $A=(1,2)$ and $B=(4,6)$, calculate the following between A and B:
 i) The Minkowski distance of order $p=2$.
 ii) The Cosine similarity distance.
- b) A marketing firm uses two features—Annual Spend and Loyalty Score—to segment customers. A new customer has a feature vector $X=(8,14)$. The 4 (4)

existing cluster centroids are: Budget Shoppers: (5,5), Premium Shoppers: (10,15), Occasional Shoppers: (15,2). Using the Manhattan distance, determine which segment the new customer X belongs to.

- 14 a) Differentiate between agglomerative and divisive approaches in clustering with suitable diagrams. 4 (6)
- b) What is a Markov Decision Process? 4 (3)

Module -4

- 15 a) Describe the structure of a perceptron and how does it help in making decisions? 2 (4)
- b) Explain Gradient Descent algorithm. Narrate the role of the learning rate in the algorithm and state what happens when it is set too high or too low with necessary plots. 2 (5)
- 16 a) A single artificial neuron is used to predict whether a student will pass an exam based on three features — hours studied (x_1), attendance percentage (x_2), and previous grade (x_3). 2 (5)

The input values, weights, and bias are given as:

$$x = \begin{bmatrix} 6 \\ 80 \\ 65 \end{bmatrix}, w = \begin{bmatrix} 0.5 \\ 0.2 \\ 0.3 \end{bmatrix}, b = 0.5$$

Apply the sigmoid activation function to obtain the neuron output. Also if the threshold for passing is $T \geq 0.5$, will the student be predicted to pass or fail? Obtain your answer showing the neuron model.

- b) How can the perceptron be used to represent XOR function? 2 (4)
