Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

M.Tech Degree S1 (R,S) (FT/WP/PhD) Examination November/December 2025 (2022 scheme)

Course Code & Name:
221TCS100 - ADVANCED MACHINE LEARNING

Max. Marks: 60 Duration: 2.5 Hours

PART A
Answer all questions. Each question carries 5 marks Marks

1 Describe the mean square error loss function and R^2 metric of a linear regression model. (5)

2 Explain the backpropagation algorithm. (5)

3 Give two differences between the KMeans clustering and hierarchical clustering. List any two methods to choose the cluster centres in the KMeans clustering algorithm. (5)

4 List two applications of SVM. Explain the "kernel trick" with a suitable example. (5)

5 What is the difference between Maximum Likelihood estimation (MLE) and Maximum a Posteriori (MAP) estimation? Suppose I flip a coin 5 times and get the following data: H, H, H, T, H. What is the probability of obtaining a head estimated with MLE and MAP, if I know that the coin is fair? (5)

PART B
Answer any 5 questions. Each question carries 7 marks

6 a) Differentiate between semi-supervised and reinforcement learning. Give one example each for semi-supervised and reinforcement learning. (4)

b) How does the learning rate affect the convergence time of the gradient descent algorithm? (3)

7 a) Compare Batch Gradient Descent algorithm with Stochastic Gradient Descent algorithm. (4)

b) Suppose that you have a linear support vector machine (SVM) binary classifier. Consider a point that is currently classified correctly, and is far (3)

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.

8 Consider the given data set. Apply the Naïve Bayes' algorithm and predict that if the fruit has the following properties, then which type of fruit it is. Fruit = {Yellow, Sweet, Long} (7)

Frequency Table:

Fruit	Yellow	Sweet	Long	Total
Mango	350	450	0	650
Banana	400	300	350	400
Others	50	100	50	150
Total	800	850	400	1200

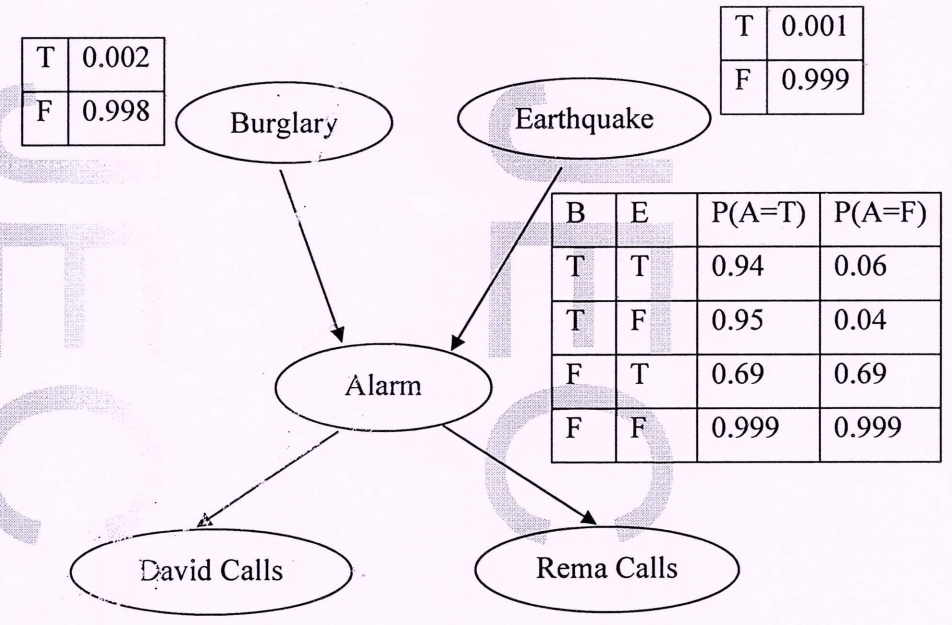
9 a) Compare Gaussian Mixture models with K-Means clustering algorithm. (4)

b) How does Principal Component Analysis achieve dimensionality reduction? (3)

10 a) Calculate the output of a 3 input neuron with input, X = [2, 2, 1], weights, W = [0.5, 0.3, 0.2] and b = -0.04. Use the sigmoid activation function. (4)

b) If you have a higher dimensional data set, what mechanism can be adopted to reduce the complexity of the model? If the dataset contains features that have different ranges, units of measurement, or orders of magnitude, what pre-processing step should be done? (3)

11 Harry installed a new burglar alarm at his home to detect burglary. The alarm reliably responds at detecting a burglary but also responds for minor earthquakes. Harry has two neighbours David and Rema, who have taken a responsibility to inform Harry at work when they hear the alarm. (7)



T	0.002
F	0.998

T	0.001
F	0.999

B	E	P(A=T)	P(A=F)
T	T	0.94	0.06
T	F	0.95	0.04
F	T	0.69	0.69
F	F	0.999	0.999

A	P(D=T)	P(D=F)
T	0.91	0.09
F	0.05	0.95

A	P(S=T)	P(S=F)
T	0.75	0.25
F	0.02	0.98

What is the probability that Rema calls, David calls, the alarm rings, but there was no burglary or earthquake?

12 A database contains 80 records on a particular topic of which 55 are relevant to a certain investigation. A search was conducted on that topic and 50 records were retrieved. Of the 50 records retrieved, 40 were relevant. Construct the confusion matrix for the search and calculate the precision and recall scores for the search. (7)

***