PCS111225AB-1

Reg No.: _____ Name: _____

Jyothi Engineering College(Autonomous)

M.Tech Degree S1 (R) Examination, December 2025 (2025 Scheme)

25PCST100- ADVANCED MACHINE LEARNING

Total Mark: 60 Total Time: 2hr 30min

CO MARK

PART A

Answer All Questions

1. Suppose that the lifetime of an electrical device is modelled by an exponential distribution with (unknown) parameter λ. We test 5 devices and find they have lifetimes of 10, 8, 7, 9, and 4 years, respectively. What is the MLE for λ? The probability mass function of exponential distribution f(x) = λe- λx. CO1 (5)

2. For models like linear regression, we can use Gradient Descent and Normal Equation to estimate the parameter. How do these techniques differ from each other? CO2 (5)

3. What is dimensionality reduction in Machine Learning? How does Principal Component Analysis achieve this? CO2 (5)

4. Define Precision and Recall. Suppose we are using Logistic regression with a threshold of 0.5 to predict the probability of having disease. How will we change the threshold to
a) Increase precision.
b) Increase recall. Justify your answer. CO4 (5)

5. Find the Jaccard coefficient of the following data set. CO4 (5)

Object	Spherical	Sweet	Sour	Crunchy
Apple	Yes	Yes	Yes	Yes
Banana	No	Yes	No	No

PART B

Answer Any Five Question(s)

6. Compare and contrast supervised learning, unsupervised learning, and reinforcement learning in terms of their applications, and describe the major challenges in reinforcement learning and how they are addressed in real-world scenarios. CO2 (7)

7. Given the two-dimensional patterns (2, 1), (3, 5), (4, 3), (5, 6), (6, 7), and (7, 8), compute the principal component using the PCA algorithm. CO3 (7)

8. How do LASSO and RIDGE regularization techniques differ in terms of their impact on the model's complexity and generalization? Explain with examples. CO2 (7)

9. The following table consists of training data from an employee database. Let "status" be the class label attribute. For a given row entry, "count" represents the number of data tuples having the values for department, status, age range, and salary range given in that row. Given a data tuple having the values "sales", "31…35" and "46…50k" for the attributes department, age range, and salary range respectively. Predict the status for the tuple using Naïve Bayes classifier. CO3 (7)

Department	Status	Age Range	Salary Range	Count
Purchase	Senior	46…..50	36k….40k	4
Purchase	Junior	26….30	26k….30k	10
Systems	Junior	21…..25	46k….50k	20
Systems	Senior	31….35	66k….70k	5
Systems	Junior	26….30	46k….50k	3
Systems	Senior	41…..45	66k….70k	3
Marketing	Senior	36…..40	46k….50k	10
Marketing	Junior	31….35	41k….45k	4
Sales	Senior	31….35	46k….50k	30
Sales	Junior	26….30	26k….30k	40
Sales	Junior	31….35	31k…..35k	40

10. Consider the following data points: A1(3, 9), A2(3, 4), A3(9, 4), A4(6, 8), A5(8, 5), A6(7, 4), A7(3, 2). Identify the cluster centres by applying the k-means algorithm, with k = 2. Initially assign A1 and A4 as the clusters centre respectively. The distance function between two points a = (x1, y1) and b = (x2, y2) is defined as D (a, b) = |x2 – x1| + |y2 – y1| CO5 (7)

11. Given the following data, construct the Receiver Operator Characteristic (ROC) curve of the data. Compute the AUC. CO4 (7)

Threshold	TP	TN	FP	FN
1	0	5	0	5
2	1	5	0	4
3	1	4	1	4
4	3	4	1	2
5	3	3	2	2
6	4	3	2	1
7	4	2	3	1
8	4	1	4	1
9	5	0	5	0

12. Find the root splitting attribute of the given dataset using any method you have learned. CO4 (7)

Outlook	Temperature	Humidity	Wind	Played football(yes/no)
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes

****