A

221TCS100022301

Pages: 4

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

First Semester M.Tech Degree Examination December 2022 (2022 Scheme)

Discipline: COMPUTER SCIENCE AND ENGINEERING

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 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 $f(x) = \lambda e^{-\lambda x}$. (5)

2 Given training set includes four features, Age, Income, Student, Credit_rate and the class label is Buys_computer. Apply Naive Bayes algorithm and estimate the value for Buys_Computer for the test dataset (Age: youth, Income: high, Student: yes, Credit_rate: excellent) (5)

Age	Income	Student	Credit_rate	Buys_computer
medium	high	no	fair	yes
medium	Low	Yes	Excellent	Yes
senior	medium	no	Excellent	no
medium	high	Yes	fair	yes
medium	medium	no	Excellent	yes
senior	medium	Yes	fair	yes
senior	low	Yes	Excellent	no
youth	medium	no	fair	no

3 Apply the DBSCAN algorithm on (5)

Point	X	Y
P1	3	7
P2	4	6
P3	5	5
P4	6	4
P5	7	3
P6	6	2

where minimum point = 3 and epsilon = 1.5.

4 What is meant by kernel trick in the context of support vector machines? Show that the kernel function $K(\vec{x}, \vec{y}) = (\vec{x}.\vec{y})^2$is equivalent to $\phi(\vec{x})^T.\phi(\vec{y})$, where $\phi(\vec{z}) = z_1^2 + z_2^2 + \sqrt{2}z_1z_2$, where each vector is represented in 2 dimensions. (5)

5 Find the Jaccard coefficient of the following data set. *Hint:* $JC = \frac{|A \wedge B|}{|A \vee B|}$ (5)

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

PART B

Answer any 5 questions. Each question carries 7 marks

6 The details about technicians' experience in a company (in several years) and their performance rating are in the table below. Using least square linear regression , estimate the performance rating for a technician with 30 years of experience. (7)

Experience of Technician (in Years)	Performance Rating
15	87
10	88
18	89
4	64
3	75
12	89
5	63

7 Implement AND function with bipolar input and target using perceptron. (7)

8 Cluster the following eight points (with (x, y) representing locations) into three clusters using K means algorithm : A1(5, 8), A2(2, 10), A3(6, 4), A4(7, 8), A5(5, 7), A6(8, 3), A7(1, 2), A8(4, 9). Calculate using Manhattan distance. Show only the first 2 iterations. (7)

9 Suppose we are given the following positively labelled data point and negatively labelled data points. Show that this is not linearly separable. (7)

$$\left\{\begin{pmatrix}2\\2\end{pmatrix},\begin{pmatrix}2\\-2\end{pmatrix},\begin{pmatrix}-2\\-2\end{pmatrix},\begin{pmatrix}-2\\2\end{pmatrix}\right\}$$

$$\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix},\begin{pmatrix}-1\\-1\end{pmatrix},\begin{pmatrix}-1\\1\end{pmatrix}\right\}$$

Φ is a nonlinear mapping from input space into some feature space.

$$\Phi_1\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{cases}\begin{pmatrix}4-x_2+|x_1-x_2|\\4-x_1+|x_1-x_2|\end{pmatrix} & \text{if } \sqrt{x_1^2+x_2^2}>2\\ \begin{pmatrix}x_1\\x_2\end{pmatrix} & \text{otherwise}\end{cases}$$

Find the support vectors.

10 Given a confusion matrix (7)

		TRUE LABEL	
		POSITIVE	NEGATIVE
PREDICTED LABEL	POSITIVE	55	5
	NEGATIVE	10	30

Find a)Accuracy

b)Precision

c)Recall

d)F1 score

e)False positive rate

11 How does the splitting attribute affect the outcome of a decision tree? Explain how the splitting attribute of the given dataset can be determined using any one of the methods you have learned. (7)

Instance no.	Class label	x_1	x_2
1	1	T	T
2	1	T	T
3	0	T	F
4	1	F	F
5	0	F	T
6	0	F	T

12 Assuming that there is three clusters and clusters have the following data points.: (7)

Cluster 1 ={{1,0},{1,1}}

Cluster 2 ={{1,2},{2,3},{2,2},{1,2}},

Cluster 3 ={{3,1},{3,3},{2,1}}

Find the Silhouette Coefficient of the first data point in Cluster 1. What can you comment about the cohesion and separation of this data point as evident from the Silhouette Coefficient? *Hint: The Silhouette Coefficient for a sample is (b-a)/max(a,b), where a is the average intra-cluster distance and b is the distance between a sample and the nearest cluster.*

*****