Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

M.Tech Degree S1 (S,FE) May 2024 (2022 scheme) / S1 (WP) (R) December 2023 Examination

Discipline: COMPUTER SCIENCE AND ENGINEERING

Course Code & Name: 221TCS100 ADVANCED MACHINE LEARNING

Max. Marks: 60 Duration: 2.5 Hours

PART A

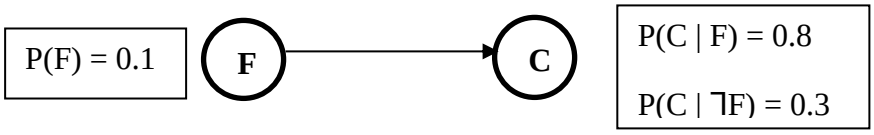
Answer all questions. Each question carries 5 marks Marks

1 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? (5)

2 With a proper diagram, describe the basic components of a perceptron. Give any two activation functions used in a perceptron. (5)

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

4 Consider the following Bayesian network, where *F = 'having the flu'* and *C =* '*coughing'*. (5)



Write down the joint probability distribution for the variables, F and C.

5 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 (5)

a) Increase precision.

b) Increase recall. Justify your answer.

PART B

Answer any 5 questions. Each question carries 7 marks

6 a) Differentiate between supervised and unsupervised learning. Give 2 examples each for supervised and unsupervised learning. (4)

b) How is Maximum Likelihood Estimate different from Maximum A Posteriori estimation? If the observed data size is small, which method is better for parameter estimation? (3)

7 Suppose a linear regression model is represented by $y = \alpha + \beta x$, where *y* is the dependent variable and *x* is the independent variable. If $\beta = \frac{Cov(x,y)}{Var(x)}$ and $\alpha = \bar{y} - \beta \bar{x}$, calculate α and β for the following data: (7)

x	1	2	3	4	5
y	1	2	1.3	3.75	2.25

8 Given the following frequency table of the data on a certain set of patients seen by a doctor, can the doctor conclude that a person having chills, fever, mild headache and without running nose has the flu? Substantiate using the Naïve Bayes' algorithm. (7)

	Chills		Running Nose		Headache			Fever		Total
Classes	Yes	No	Yes	No	No	Mild	Strong	Yes	No	
Flu -Yes	3	2	4	1	1	2	2	4	1	5
Flu - No	1	2	1	2	1	1	1	1	2	3
Total	4	4	5	3	2	3	3	5	5	8

9 Given the dataset {a, b, c, d, e} and the distance matrix given in the table below, construct a dendrogram by single-linkage hierarchical clustering using the agglomerative method. (7)

	{a}	{b}	{c}	{d}	{e}
{a}	0	9	3	6	11
{b}	9	0	7	5	10
{c}	3	7	0	9	2
{d}	6	5	9	0	8
{e}	11	10	2	8	0

10 a) Define overfitting. Illustrate any one mechanism that can be adopted to prevent overfitting. (4)

b) List any three characteristics of Gaussian Mixture Models. (3)

11 Consider a support vector machine whose input space is 2-D, and the inner products are computed by means of the kernel $K(x, y) = (x.y + 1)^2 - 1$, (7)

where x.y denotes the ordinary inner product. Show that the mapping to feature space that is implicitly defined by this kernel is the mapping to 5-D given by

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \varphi(x) = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2}x_1x_2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \end{bmatrix}$$

12 How do you define the accuracy of a machine learning model? Illustrate the concept behind any two mechanisms used to improve the accuracy of machine learning models. (7)