

Name:

Enrolment No:



**UPES**

**End Semester Examination, May 2024**

**Course: Artificial Intelligence and Machine Learning**

**Semester: II**

**Program: M. Tech RE and E-Mobility**

**Time : 03 hrs.**

**Course Code: CSAI7016P**

**Max. Marks: 100**

**Instructions: All questions are compulsory. The question paper consists of 11 questions divided into 3 sections A, B and C. Section A comprises 5 questions of 4 marks each, Section B comprises 4 questions of 10 marks each and Section C comprises 2 questions of 20 marks each.**

**SECTION A  
(5Qx4M=20Marks)**

S. No.		Marks	CO
Q 1	Discuss the similarity & differences between two types of Supervised Machine Learning – namely Regression & Classification.	4	CO1
Q 2	Given two features, $x_1$ and $x_2$ , how will you fit (a) a linear decision boundary and (b) a circle decision boundary (c) a more complex decision boundary using logistic regression?	4	CO1
Q 3	Discuss the major differences between Prim's and Kruskal's Algorithm for MST (Minimum Spanning Tree).	4	CO1
Q 4	For given confusion matrix calculate (A)Precision (b)Recall (C)f1-score [[45,1], [11,33]],	4	CO2
Q 5	A layer 'L' in a neural network has 5 neurons and the previous layer has 4 neurons. How many model parameters are associated with layer L?	4	CO1

**SECTION B  
(4Qx10M= 40 Marks)**

Q 6	Describe Linear Regression Model – Univariate & multivariate. Describe the model, its parameters, and cost function.	10	CO2
Q 7	Discuss the implementation of logistic regression mathematically, including definition of logistic regression, cost function & gradient descent.	10	CO2
Q 8	In what way Decision Tree for Classification is different from Decision Tree for Regression? Highlight the role of maximization of information gain vs maximization of decrease in variance.	10	CO2

Q 9 Find the Minimum Spanning Tree of the following graph using Kruskal's algorithm.

10 CO1

**SECTION-C**  
**(2Qx20M=40 Marks)**

Q 10 Discuss the process for choosing a model from many options through the use of Training data, Cross Validation data & Test data. Clearly highlight the role of cost function (J), and various errors such as  $J_{training}, J_{CV}, J_{test}$

20 CO3

Q 11 We have data from questionnaires survey (to ask people opinion) and objective testing with two attributes (acid durability and strength) to classify whether a special paper tissue is good or not. Here is four training sample.

X1 = Acid durability (seconds)	X2 = Strength (kg/square meter)	Y = Classification
7	7	Bad
7	4	Bad
3	4	Good
1	4	Good

Now, the factory produces a new paper tissue that pass laboratory test with  $X1 = 3$  and  $X2 = 7$ . Without another expansion survey, can we guess what the classification of this new tissue is? (Use KNN Algorithms)

**OR**

Implement gradient descent algorithms for the dataset given below:

Age	Salary
30	800
37	950
25	600
43	1050
50	1200
29	740
46	1100

20 CO3