CSE313

Enroll. No. -

[ET]

END SEMESTER EXAMINATION : NOVEMBER - 2022

FUNDAMENTALS OF MACHINE LEARNING

Time :03 Hrs. Maximum Marks :60

Note: Attempt questions from all sections as directed. Use of Scientific calculator is allowed.

Section - A : Attempt any Four questions out of Five . Each question carries 06 marks. **[24 Marks]**

Q1. What are a training set and test set in Machine Learning, and why are they important?

Q2. How could stochastic gradient descent save time compared to standard gradient descent? Explain the Gradient Descent algorithm.

Q3. Explain hypothesis space searching in decision tree learning.

Q4. Explain the Difference Between Classification and Regression?

Q5. Explain the procedure to construct decision trees

Section – B : Attempt any two questions out of three. Each question carries 10marks. **[20 Marks]**

Q6. The values of independent variable x and dependent value y are given below:

X	Y
0	2
1	3
2	5
3	4
4	6

Find the least square regression line y=ax+b. Estimate the value of y when x is 10.

Q7. Explain the principle of the gradient descent algorithm. Accompany your explanation with a diagram. Explain the use of all the terms and constants that you introduce and comment on the range of values that they can take.

Q8. (a) Illustrate the idea of PCA for a two dimensional data using suitable diagrams. (5)

(b) Why do you prefer Euclidean distance over Manhattan distance in the K means Algorithm? (5)

[16 Marks]

Section - C : Compulsory question

Q9. (a) State the mathematical formulation of the SVM problem. Give an outline of the method for solving the problem. (5)

(b) Show the final result of hierarchical clustering with complete link by drawing a dendrogram. (5)

	A	B	C	D	E	F
A	0					
B	0.12	0				
C	0.51	0.25	0			
D	0.84	0.16	0.14	0		
E	0.28	0.77	0.70	0.45	0	
F	0.34	0.61	0.93	0.20	0.67	0

(c) Compare feature extraction and feature selection techniques. Explain how dimensionality can be reduced using subset selection procedures. (6)

****