Total No. of Questions : 8]

SEAT No. :

PE-2598

[Total No. of Pages : 2

[6583]-129

T.E. (IT)

MACHINE LEARNING

(2019 Pattern) (Semester - V)(314443)

Time : 2½ Hours] *[Max. Marks : 70*

Instructions to the candidates:

1) ***Answer Q.1 or Q.2, Q.3 or Q.4, Q.5 or Q.6 and Q.7 or Q.8.***
2) ***Neat diagrams must be drawn wherever necessary.***
3) ***Figures to the right side indicate full marks.***
4) ***Assume suitable data, if necessary.***

Q1) a) What do you understand by overfitting data? Gives any two methods to avoid overfitting. **[6]**

b) Given the set of values X = (3, 9, 11, 5, 2)T and Y=(1, 8, 11, 4, 3)T. Evaluate the regression coefficients. **[5]**

c) Compare the linear regression model and the logistic regression model. **[6]**

OR

Q2) a) What are the metrics used to validate the result of the regression and explain each? **[6]**

b) Distinguish between overfitting and underfitting. How it can affect model generalization. **[5]**

c) Discuss linear regression with examples. Explain the role of the hypothesis function in ML. **[6]**

Q3) a) What is Gini impurity? Briefly explain the properties of Gini Impurity. **[8]**

b) [CO4, L_1] How does the decision tree handle missing attribute Values? **[4]**

c) [CO4, L_2] What do you understand by Information gain? Explain the mathematical formula associated with it. **[6]**

OR

P.T.O.

Q4) a) Write short notes on **[12]**

i) Conditional probability

ii) Tree Pruning

iii) Bayesian Network

b) What do the root node, leaf node, and parent node represent in a decision tree model? Represent them pictorially. **[6]**

Q5) a) What is the difference between the Manhattan distance and Euclidean distance? **[4]**

b) Explain in detail Apriori Algorithm. **[8]**

c) Define the following terms concerning K - Nearest Neighbor Learning **[6]**

i) Regression

ii) Residual

iii) Kernel Function

OR

Q6) a) What is the difference between k-Means and K-Medians and when would you use one over another? **[6]**

b) What is hierarchical clustering? Explain the issue of connectivity constraints. **[8]**

c) What are support and confidence measures in data mining? **[4]**

Q7) a) With suitable equations, explain any two types of activation functions used in a neural network. **[8]**

b) Explain why the initialization process of weight and bias is important for neural networks. **[6]**

c) Write short notes on Mean square Error. **[3]**

OR

Q8) a) Explain with a neat diagram how neural networks are modeled. **[6]**

b) Write short notes on **[6]**

i) Sigmoidal Neuron

ii) McCulloch Pitts Neuron

c) Explain the working of a Perceptron with a neat diagram. **[5]**

♣ ♣ ♣