Paper / Subject Code: 42171 / MACHINE LEARNING

BE Sem VII (C Scheme, R-19) "Computer" Dec' 2022
8/12/2022
Duration: 3hrs

[Max Marks:80]

N.B.: (1) Question No 1 is Compulsory.
(2) Attempt any **three** questions out of the remaining **five**.
(3) All questions carry equal marks.
(4) Assume suitable data, if required and state it clearly.

Q1. Solve any **four** from following. [20]
a. What are the issues in Machine learning?
b. Explain Regression line, Scatter plot, Error in prediction and Best fitting line.
c. Explain the concept of margin and support vector.
d. Explain the distance metrics used in clustering.
e. Explain Logistic Regression

Q2. a. Explain the steps of developing Machine Learning applications. [10]
b. Explain Linear regression along with an example. [10]

Q3. a. Create a decision tree using Gini Index to classify following dataset. [10]

Sr. No.	Income	Age	Own Car
1	Very High	Young	Yes
2	High	Medium	Yes
3	Low	Young	No
4	High	Medium	Yes
5	Very High	Medium	Yes
6	Medium	Young	Yes
7	High	Old	Yes
8	Medium	Medium	No
9	Low	Medium	No
10	Low	Old	No
11	High	Young	Yes
12	Medium	Old	No

b. Describe Multiclass classification. [10]

Q4. a. Explain the Random Forest algorithm in detail. [10]
b. Explain the different ways to combine the classifiers. [10]

Q5. a. Compute the Linear Discriminant projection for the following two-dimensional dataset. X1= (x1, x2) = {(4,1), (2,4), (2,3), (3,6), (4,4)} and
X2= (x1, x2) = {(9,10), (6,8), (9,5), (8,7), (10,8)} [10]
b. Explain EM algorithm. [10]

Q6. Write detailed note on following. **(Any two)** [20]
a. Performance Metrics for Classification
b. Principal Component Analysis for Dimension Reduction
c. DBSCAN

****