

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]

- What are the issues in Machine learning?
- Explain Regression line, Scatter plot, Error in prediction and Best fitting line.
- Explain the concept of margin and support vector.
- Explain the distance metrics used in clustering.
- 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. $X_1 = (x_1, x_2) = \{(4,1), (2,4), (2,3), (3,6), (4,4)\}$ and

[10]

 $X_2 = (x_1, x_2) = \{(9,10), (6,8), (9,5), (8,7), (10,8)\}$

b. Explain EM algorithm.

[10]

Q6. Write detailed note on following. (Any two)

[20]

- Performance Metrics for Classification
- Principal Component Analysis for Dimension Reduction
- DBSCAN
