

MSc Sem.-2 Examination

409

DS-MLA

May-2025

Time : 2-30 Hours]

[Max. Marks : 70

Instructions: All questions are compulsory. Use of non-programmable scientific calculator is allowed.

- Q.1** (a) What is Machine Learning? Explain its types. (07)
 (b) Give the difference between Supervised and Unsupervised Machine Learning. (07)

OR

- (a) What is Supervised Machine Learning? Explain with Example. (07)
 (b) What is Over fitting and Under fitting ? Explain with appropriate example. (07)

- Q.2** (a) What is Naïve Bayes Classification? Explain in detail. (07)
 (b) Explain: Support Vector Machine (SVM) Algorithm (07)

OR

- (a) What is Regression? Explain in detail. (07)
 (b) What is Decision Tree? Explain detail. (07)

- Q.3** (a) What is distance Measure Methods. Explain All. (07)
 (b) Given the following 2D dataset of points: (07)

$$\text{Points}=\{(1,2),(2,3),(3,4),(6,7),(7,8),(8,9)\}$$

cluster these points into 2 clusters using the K-Means algorithm.

OR

- (a) What is Apriori algorithm? Explain in detail. (07)
 (b) Using the distance matrix, perform single-link hierarchical clustering. Draw a dendrogram. (07)

	P1	P2	P3	P4	P5
P1	0	0.10	0.41	0.55	0.35
P2	0.10	0.00	0.64	0.47	0.98
P3	0.41	0.64	0.00	0.44	0.85
P4	0.55	0.47	0.44	0.00	0.76
P5	0.35	0.98	0.85	0.76	0.00

- Q.4** (a) Explain the concept of Singular Value Decomposition (SVD). (07)
(b) What is Random Forest? Explain detail. (07)

OR

- (a) What is Logistic Regression? Explain in detail. (07)
(b) What is Bagging and Boosting? Explain with an example. (07)

Q.5 Attempt any **SEVEN** out of **TWELVE**: (14)

- (1) What is Pseudo labelling?
- (2) What are the steps of Machine Learning?
- (3) What is Confusion Matrix?
- (4) Explain over-fitting and under-fitting.
- (5) What is R^2 or R squared?
- (6) What is Ensemble Learning?
- (7) What is linkage in hierarchical clustering?
- (8) What is Random Forest?
- (9) Define: Bagging and Boosting
- (10) Define: Kernel trick
- (11) What is Data Preprocessing?
- (12) Explain: Features & Target
