MCA II Semester (NEP) Degree Examination, Nov-2023

MATER OF COMPUTER APPLICATION

Paper: Data Mining

Hours: 03 **Max Marks: 70**

Answer any Five of the following questions with Question No. 1 is Compulsory.

1. a) Explain the architecture of data mining. **7**

   b) Given confusion matrix calculate the following; **7**

	Predicted Class	
	Spam	Not Spam
Actual Class	250	10
Actual Class	15	500

   a. Accuracy b. Precision c. Recall d. F1 Score

2. a) Discuss extract, load and transform process in data transformation. **7**

   b) Define discretization? Discuss the types of discretization with examples. **7**

3. a) Bag1 contains 4 white and 8 black balls and Bag2 contains 5 white and 3 black balls. From one of the bag one ball is drawn at random and the ball which is drawn comes out as black. Find the probability that the ball is drawn from bag1. **7**

   b) State and Prove Bayes theorem. **7**

4. a) Describe the classification of data in cluster analysis. **7**

   b) Given cluster points into three clusters A1(2,10), B1(5,8), C1(1,2) (initial cluster) where the points are A1(2,10), A2(2,5), A3(8,4), B1(5,8), B2(7,5), B3(6,4), C1(1,2), C2(4,9) find the centroids by applying K-mean centroids. **7**

5. a) Discuss the characteristics of OLAP. **7**

   b) Define data modeling. List out and explain types of data modeling. **7**

6. a) List out and explain the task and technique in data cleaning. **7**

   b) Differentiate between Knowledge discovery in database and data mining. **7**

7. a) Illustrate the Properties of Clustering in data mining. **7**

   b) Explain the architecture of ROLAP. **7**

8. Write a short note on the following: **5+5+4**
   a. Data pre-processing.
   b. OLTP and OLAP
   c. Web mining.

- o 0 o -