

(AUTONOMOUS)
MSc DEGREE EXAMINATION MAY 2026
(Fourth Semester)

Branch - **STATISTICS**

MAJOR ELECTIVE COURSE – II DATA MINING AND WAREHOUSING

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** marks

(10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	OLAP stands for (a) Online Analytical Processing (b) Online Analysis Processing (c) Online Transaction Processing (d) Online Aggregation Processing	K1	CO1
	2	Metadata is a _____ (a) Data about mining (b) Data about query (c) Data about data (d) Data about no data	K2	CO1
2	3	The synonym for data mining is (a) data warehouse (b) knowledge discovery in database (c) ETL (d) Business intelligence	K1	CO2
	4	An itemset whose support is greater than or equal to a minimum support threshold is _____ (a) Itemset (b) Frequent Itemset (c) Infrequent items (d) Threshold values	K2	CO2
3	5	Which of the following is a randomized algorithm? (a) CLARANS (b) BIRCH (c) AGES (d) CLARA	K1	CO3
	6	Decision tree is a _____ model. (a) predictive (b) descriptive (c) learning (d) both (a) and (b)	K2	CO3
4	7	What is the basic building block of a neural network? (a) Layer (b) Weight (c) Node (d) Neuron	K1	CO4
	8	In which type of neural network, the data is grouped based on its distance from a center point? (a) Convolution Neural Network (b) Recurrent Neural Network (c) Modular Neural Network (d) Radial Basis Functions Neural Network	K2	CO4
5	9	A _____ technique is used to crawl through various web sources to collect required information, which enables a individual or company to promote business. (a) Text mining (b) Web mining (c) Spatial mining (d) Time series mining	K1	CO5
	10	Text mining is the process of: (a) Extracting information from images (b) Analyzing structured data (c) Analyzing unstructured text data (d) Identifying relationships between variables	K2	CO5

Cont...

SECTION - B (35 Marks)
 Answer ALL questions
 ALL questions carry EQUAL Marks

(5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Identify the problems in designing data warehouse schemas.	K3	CO1
	(OR)			
	11.b.	Identify the significant features of a data warehouse.		
2	12.a.	Classify Data Mining techniques.	K4	CO2
	(OR)			
	12.b.	Examine the application of data mining in Science data.		
3	13.a.	Compare CLARANS with CLARA.	K4	CO3
	(OR)			
	13.b.	Analyze the difficulties which arise when one uses a decision tree in a real-life situation.		
4	14.a.	Explain the principles and applications of rough set theory in data mining.	K4	CO4
	(OR)			
	14.b.	Discuss the learning techniques of MLP and RBF.		
5	15.a.	Explain Web Structure Mining.	K3	CO5
	(OR)			
	15.b.	Explain temporal data mining.		

SECTION - C (30 Marks)
 Answer ANY THREE questions
 ALL questions carry EQUAL Marks

(3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO
1	16	Explain the different models of data warehouse servers.	K4	CO1
2	17	Discover Association Rules from market-basket data.	K4	CO2
3	18	Construct a decision tree using classical algorithms.	K4	CO3
4	19	Outline a case study to illustrate the potential application of NN for Data Mining.C	K4	CO4
5	20	List out the significant features that are extracted to convert unstructured text to structured form.	K4	CO5

Z-Z-Z

END