TOTAL PAGES: 2

22STP421BN

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc DEGREE EXAMINATION MAY 2025
(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	_____ allows data to be modeled and viewed in multiple dimensions. (a) Data cube (b) Fact table (c) Dimension table (d) Apex cuboid	K1	CO1
	2	A crucial area of data warehousing is which is a kind of data that describes the data warehouse itself. (a) Data Mining (b) Metadata (c) Data Mart (d) Query	K2	CO1
2	3	Data mining often requires ———————— the merging of data from multiple data stores. (a) Data cleaning (b) Data integration (c) Data Transformation (d) Data Reduction	K1	CO2
	4	Support (X) = ? (a) Number of times X appears (b) Total number of transactions (N) (c) Number of times X appears/N (d) N+X	K2	CO2
3	5	Clustering is known as _____because the class label information is not present. (a) data segmentation (b) automatic (c) outlier (d) unsupervised learning	K1	CO3
	6	A decision tree is a tree in which every node is either a _____ or a decision node. (a) Leaf node (b) Root node (c) Both (a) and (b) (d) Sub node	K2	CO3
4	7	What is a perceptron? (a) a single layer feed-forward neural network with pre-processing (b) an auto-associative neural network (c) a double layer auto-associative neural network (d) a neural network that contains feedback	K1	CO4
	8	What are the types of Neural Networks? (a) Feed-forward (b) Radial Basis Functions (c)Recurrent (d) All of the above	K2	CO4
5	9	The bag-of-words (BoW) model in text mining represents a document as: (a) A collection of images (b) A sequence of sentences (c) A set of words disregarding grammar and word order (d) A structured table with rows and columns	K1	CO5
	10	Which mining tries to discover link structure of the hyper links at the inter document level? (a) Data processing (b) Web structure mining (c) Data usage (d) Multiuser system	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.	Examine the characteristics of a data warehouse.	K4	CO1
		(OR)		
	11.b.	List out the functions performed by data warehouse backend tools.		
2	12.a.	Distinguish between KDD and Data Mining.	K4	CO2
		(OR)		
	12.b.	Illustrate the advantages of the pincer-search algorithm with an example database.		
3	13.a.	Explain the clustering paradigms with an example.	K5	CO3
		(OR)		
	13.b.	Discuss the generic algorithm of decision tree construction.		
4	14.a.	Illustrate the architecture of MLP and RBF.	K3	CO4
		(OR)		
	14.b.	State the applications of rough set framework in data mining.		
5	15.a.	Discuss the significant features that are extracted to convert unstructured text to structured form.	K5	CO5
		(OR)		
	15.b.	Elaborate on the Sequence Mining Problem.		

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 basic OLAP operations for a multidimensional model.	K5	CO1
2	17	Discuss the applications of data mining in the banking industry.	K5	CO2
3	18	Examine the working of the DBSCAN Algorithm.	K4	CO3
4	19	Propose a case study in data mining using NN techniques.	K4	CO4
5	20	Categorize the mining techniques on the web.	K3	CO5

Z-Z-Z END