Reg. No.															

Question Paper Code	14068

B.E. / B.Tech. - DEGREE EXAMINATIONS, NOV / DEC 2025

Fifth Semester

Computer and Communication Engineering

20CSEL501 - DATA WAREHOUSING AND DATA MINING

Regulations - 2020

Duration: 3 Hours | Max. Marks: 100

PART - A (MCQ) (10 × 1 = 10 Marks)

Answer ALL Questions

		Marks	K – Level	CO
1.	Which of the following schema contains multiple fact tables? (a) Star Schema (b) Snowflake schema (c) Fact constellation (d) None of the above	1	K2	CO1
2.	A typical OLAP operation includes: (a) Data Modification (b) Data Replication (c) Data Roll-up and Drill-down (d) Data Serialization	1	K1	CO1
3.	The function of regression analysis is _____. (a) To predict numerical values based on historical data (b) To group similar data points (c) To discover associations among sets of items (d) To classify data into predefined categories	1	K1	CO2
4.	What is the primary goal of a Data Mining System? (a) Data Storage (b) Knowledge Discovery (c) Database Management (d) Data Manipulation	1	K2	CO2
5.	The likelihood that an item is purchased when another item is purchased is known as ___. (a) Support (b) Confidence (c) Lift (d) Item set	1	K1	CO3
6.	What is the main objective of association rule mining? (a) Predicting continuous values (b) Finding hidden relationships between items in a dataset (c) Reducing dimensionality (d) Clustering data points	1	K1	CO3
7.	Decision tree induction is a supervised learning technique used for _____prediction and classification tasks. (a) continuous (b) regression (c) discrete (d) clustering	1	K1	CO4
8.	SVMs can be used for---------- (a) Numeric prediction (b) Classification and Regression Trees (c) Numeric prediction and Classification (d) None of the Mentioned	1	K1	CO4
9.	What is the primary goal of clustering? (a) Maximizeinter-clustersimilarity (b) Minimizeintra-clustersimilarity (c)Maximizebothintra-clusterandinter-clustersimilarity (d) Groupsimilardatapoints together	1	K1	CO5
10.	In Iris plants dataset,each column is ---------- (a) a feature (b) an observation (c) unique (d) None of the Mentioned	1	K1	CO6

PART - B (12 × 2 = 24 Marks)

Answer ALL Questions

		Marks	K – Level	CO
11.	What is a Data Warehouse?	2	K1	CO1
12.	Recall Parallel DBMS Vendors.	2	K1	CO1
13.	Define Data Cube.	2	K1	CO2
14.	Compare data similarity and dissimilarity.	2	K2	CO2
15.	List the five categories of pattern mining constraints.	2	K1	CO3

16.	Show the outline of the frequent pattern growth method.	2	K2	CO3
17.	What is called Bayesian classification?	2	K1	CO4
18.	Tell about information gain.	2	K2	CO4
19.	Illustrate Hierarchical Clustering	2	K2	CO5
20.	Why are outliers used? And list out its Types?	2	K1	CO5
21.	Describe the Key Features of WEKA.	2	K2	CO6
22.	Write the GUI interface components of WEKA.	2	K1	CO6

PART - C (6 × 11 = 66 Marks)

Answer ALL Questions

23.	a)	Summarize the multi-dimensional model with an example.	11	K2	CO1
		OR			
	b)	Explain the types of OLAP operations in detail with examples.	11	K2	CO1
24.	a)	Discuss in detail about data integrity preprocessing.	11	K2	CO2
		OR			
	b)	Classify the Data Visualization Techniques in Data Mining.	11	K2	CO2
25.	a)	Build constraint-based association mining with example.	11	K3	CO3
		OR			
	b)	Elaborate Mining Multi dimensional Association Rules.	11	K3	CO3
26.	a)	Describe in detail Rule based Classification.	11	K2	CO4
		OR			
	b)	Briefly discuss the basic decision tree induction algorithm.	11	K2	CO4
27.	a)	Compare Density-Based Clustering Methods and Grid-Based Clustering Methods.	11	K2	CO5
		OR			
	b)	Explain in detail about Partitioning Methods.	11	K2	CO5
28.	a)	Illustrate various Data format sused in WEKA with an example.	11	K2	CO6
		OR			
	b)	Demonstrate the classifiers selected in WEKA.	11	K2	CO6