Reg. No. | | | | | | | | |

G. VENKATASWAMY NAIDU COLLEGE (AUTONOMOUS), KOVILPATTI – 628 502.

UG DEGREE END SEMESTER EXAMINATIONS - APRIL 2025.

(For those admitted in June 2021 and later)

PROGRAMME AND BRANCH: B.COM., BUSINESS ANALYTICS

SEM	CATEGORY	COMPONENT	COURSE CODE	COURSE TITLE
VI	PART - III	CORE	U21BA613	DATA WAREHOUSING AND DATA MINING

Date & Session: 29.04.2025/FN **Time : 3 hours** **Maximum: 75 Marks**

Course Outcome	Bloom's K-level	Q. No.	SECTION – A (10 X 1 = 10 Marks) Answer ALL Questions.
CO1	K1	1	What is the primary purpose of data mining? a) To store data b) To extract patterns and knowledge from large datasets c) To delete unnecessary data d) To create new data
CO1	K2	2.	Which of the following best describes 'data pre-processing'? a) The process of archiving old data b) The process of analyzing data after it has been collected c) The process of preparing raw data for analysis d) The process of visualizing data results
CO2	K1	3.	What is the goal of concept description in data mining? a) To generate detailed and accurate predictions from data b) To cluster data into different categories c) To extract higher-level abstractions from data d) To summarize data without changing its format
CO2	K2	4.	Which of the following is a characteristic of Data Mining Query Language (DMQL) a) It is based on XML. b) It is designed for ad hoc and interactive data mining. c) It cannot interact with databases. d) It is only used for statistical analysis.
CO3	K1	5.	What is the implication expression format for an association rule? a) $X + Y \rightarrow Z$ b) $X \rightarrow Y$ c) $X \cap Y = Z$ d) $X < Y$
CO3	K2	6.	What is a single-dimensional Boolean association rule? a) A rule that involves numerical attributes b) A rule that relates a single attribute to another single attribute c) A rule based on multiple items in the dataset d) A rule that involves boolean values (true/false) for items in the transaction
CO4	K1	7.	What is the main purpose of classification in data mining? a) To categorize data into predefined classes b) To visualize data c) To clean data d) To predict future values
CO4	K2	8.	Which classification method uses Bayes' theorem to predict class membership? a) Decision Trees b) K-Nearest Neighbours c) Bayesian Classification d) Support Vector Machines
CO5	K1	9.	Which of the following is a type of clustering method? a) Linear Regression b) Hierarchical clustering c) Decision Trees d) K-nearest neighbours
CO5	K2	10.	What is the key feature of density-based clustering methods like DBSCAN?

			a) Clusters are created by dividing the data into equal parts b) The number of clusters is fixed in advance c) Clusters are formed by dense regions of data points separated by low-density regions d) Clusters are hierarchical and nested
Course Outcome	**Bloom's K-level**	**Q. No.**	**SECTION – B (5 X 5 = 25 Marks)** **Answer ALL Questions choosing either (a) or (b)**
CO1	K3	11a.	Discuss the various methods used for data integration. **(OR)**
CO1	K3	11b.	Explain the process of normalization in data pre-processing.
CO2	K3	12a.	Explain data summarization in data mining. What are the different techniques used for summarizing data? **(OR)**
CO2	K3	12b.	What are the challenges involved in concept description, characterization, and comparison?
CO3	K4	13a.	Describe the Apriori algorithm for mining association rules. **(OR)**
CO3	K4	13b.	Discuss the process of mining multilevel association rules.
CO4	K4	14a.	Discuss the assumptions made by the Naive Bayes classifier. **(OR)**
CO4	K4	14b.	What are the limitations of using Bayesian classification for prediction tasks?
CO5	K5	15a.	Explain the working of the K-means clustering algorithm. **(OR)**
CO5	K5	15b.	Describe the steps involved in agglomerative hierarchical clustering. How does it work from bottom-up?

Course Outcome	**Bloom's K-level**	**Q. No.**	**SECTION – C (5 X 8 = 40 Marks)** **Answer ALL Questions choosing either (a) or (b)**
CO1	K3	16a.	Explain the concept of data mining and discuss its various functionalities. **(OR)**
CO1	K3	16b.	Explain the concept of data warehousing. What are the key components of a data warehouse?
CO2	K4	17a.	Explain the concept of data mining primitives. Discuss the various types of primitives in data mining. **(OR)**
CO2	K4	17b.	What is a Data Mining Query Language (DMQL)? Discuss its importance in data mining.
CO3	K4	18a.	Explain the challenges faced when mining association rules in dynamic or evolving transactional datasets. **(OR)**
CO3	K4	18b.	Discuss the concept of "negative association rules" and their relevance in data mining.
CO4	K5	19a.	How does the back propagation algorithm differ from decision trees in terms of handling continuous and categorical data for classification? **(OR)**
CO4	K5	19b.	What is over fitting in neural networks? How can the back propagation algorithm be modified to prevent over fitting?
CO5	K5	20a.	Explain the DBSCAN algorithm and its working principle. **(OR)**
CO5	K5	20b.	Discuss the differences between single-linkage, complete-linkage, and average-linkage in hierarchical clustering.