



R20 Regulation

Subject code:3P7FA

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

(Autonomous, Accredited by NAAC with 'A+' Grade)

B.Tech VII Semester Supplementary Examinations, April 2024

DATA WAREHOUSING AND DATA MINING (IT)

Maximum Marks: 70

Date:29.04.2024 Duration: 3 hours

- Note:
1. This question paper contains two parts A and B.
 2. Part A is compulsory which carries 20 marks. Answer all questions in Part A.
 3. Part B consists of 5 Units. Answer any one full question from each unit which carries 10M.
 4. Each question carries 10 marks and may have a, b, c, d as sub questions.

Part-A

All the following questions carry equal marks

(10X2M=20 Marks)

- 1 Define Data warehousing.
- 2 Give a brief note on Fact-Less-Facts
- 3 What are the steps involved in KDD process?
- 4 Write the strategies for data reduction.
- 5 Differentiate multiple and multilevel association rule mining
- 6 Mention the importance of Association Rule Mining
- 7 State and explain Bayes theorem.
- 8 Write the need for tree pruning in decision tree induction?
- 9 Mention the Key issues in hierarchical clustering
- 10 Define the centroid of the cluster.

Part-B

Answer All the following questions.

(5X10M=50Marks)

- 11 a) Discuss in detail about the steps of knowledge discovery? [5M]
b) What is noisy data? Explain the binning methods for data smoothening. [5M]
OR
- 12 Explain three types of schemas that are used for modeling data warehouse with examples. State the applications of Data mining. What is its need in Business? [10M]
- 13 Explain data mining as a step-by-step process of knowledge discovery. Mention the Functionalities of Data mining. [10]
OR
- 14 a) Explain in detail about data pre-processing. [5M]
b) What is the need of dimensionality reduction? Explain any two techniques for dimensionality reduction. [5M]

- 15 What is Association rule mining? Briefly describe the criteria for classifying association rules.[10]

OR

- 16 Construct the FP-Tree from the given Transactional Database. Explain the procedure in detail with minimum support = 3. [10]

TID	ITEMS
100	F, A, C, D, G, I, M, P
200	A, B, C, F, L, M, O
300	B, F, H, J, O, W
400	B, C, K, S, P
500	A, F, C, E, L, P, M, N

- 17 State Bayes theorem and discuss how Bayesian classifiers work? [10M]

OR

- 18 What are Bayesian classifiers? With an example, describe how to predict a class label using Naive Bayesian classification. [10M]

- 19 A. Differentiate Agglomerative and divisive Hierarchical Clustering. [5M]
B. Explain Partitioning Clustering-K-Means Algorithm with an example. [5M]

OR

- 20 Explain the partitioning methods? Solve the following problem using Partition methods (K-means, K-medoids) for {2, 4, 10, 12, 8, 20, 30, 11, 25} where $k = 2$. [10M]