

PARUL UNIVERSITY
FACULTY OF ENGINEERING & TECHNOLOGY
B.Tech. Winter 2022-23 Examination

Semester: 7
Subject Code: 03105430
Subject Name: Data Mining and Warehousing

Date: 06/10/2022
Time: 10:30am to 1:00pm
Total Marks: 60

Instructions:

1. All questions are compulsory.
2. Figures to the right indicate full marks.
3. Make suitable assumptions wherever necessary.
4. Start new question on new page.

Q.1 Objective Type Questions -(All are compulsory) (Each of one mark) **(15)**

1. Which is the correct sequence
 - a. Data- Information – Knowledge
 - b. Knowledge- Data- Information
 - c. Data- Knowledge- Information
 - d. Knowledge- Information- Data
2. Which of the following is not an application of Data Mining?
 - a. Market Analysis and Management
 - b. Corporate Analysis & Risk Management
 - c. Fraud Detection
 - d. To store data in Database
3. Which of the following is not a derived model from Classification?
 - a. Mathematical Formulae
 - b. Neural Networks
 - c. Decision Trees
 - d. Graph
4. Which of the following issues is considered before investing in Data Mining?
 - a. Functionality
 - b. Vendor consideration
 - c. Compatibility
 - d. All of the above
5. Identify the options below that a data warehouse can include.
 - a. Database table
 - b. Online data
 - c. Flat files
 - d. All of the above
6. What is significance of Business Intelligence?
7. DSS in data warehouse stands for?
8. What is KDD in Data Mining?
9. What is Clustering?
10. What is Data Warehouse?
11. _____ is the set of processes used to get data from the data source into the data warehouse.
12. _____ is a data mining technique used to determine the groups of data that may exist in a set of data based on the characteristics of the data.
13. In order to perform text mining, analysts must first apply _____ to unstructured data
14. Classification is the process of finding a model that describes _____
15. Data mining is the crucial step in which clever techniques are applied to _____.

Q.2 Answer the following questions. (Attempt any three) (15)

- A) Describe the steps involved in data mining when viewed as a process of knowledge discovery.
- B) Explain Mean, Median, Mode, Variance, Standard Deviation & five number summary with suitable database example.
- C) Draw and Explain Snowflakes and Fact constellations Schema
- D) In data pre-processing why we need data smoothing? Discuss data smoothing by Binning.

Q.3 A) Generate frequent itemsets and generate association rules based on it using apriori algorithm. (07)
Minimum support is 50% and minimum confidence is 70%

TID	ITEMSET
100	1, 3, 4
200	2, 3, 5
300	1,2,3,5
400	2,5

B) Explain Data warehouse architecture. (08)

OR

B) Explain various OLAP operations. (08)

Q.4 A) Why naïve Bayesian classification is called “naïve”? Describe naïve Bayesian classification with example. (07)

OR

A) Explain basic concepts of text mining and web mining. (07)

B) List out and explain Requirements of Clustering in Data Mining. and applications of it. (08)