

GUJARAT TECHNOLOGICAL UNIVERSITY

BE - SEMESTER-VII (OLD) EXAMINATION – WINTER 2018

Subject Code: 171601

Date: 03/12/2018

Subject Name: Data Warehousing And Data Mining

Time: 10:30 AM TO 01:00 PM

Total Marks: 70

Instructions:

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.

Q.1 (a) What is Data Mining? Explain Data mining as one step of Knowledge Discovery Process. **07**

(b) Differentiate between Operational Database System and Data Warehouse. **07**

Q.2 (a) Explain Data Cleaning process for missing values and noisy data. **07**

(b) Explain three-tier data warehouse architecture. **07**

OR

(b) Explain mining in following Databases with example. **07**

1. Temporal Databases

2. Sequence Databases

3. Spatial Databases and Saptiotemporal Databases.

Q.3 (a) What is Cuboid? Explain various OLAP Operations on Data Cube with example. **07**

(b) Explain Fp -Growth Algorithm for finding Frequent Item-sets. **07**

OR

Q.3 (a) State the Apriori Property. Generate large itemsets and association rules using Apriori algorithm on the following data set with minimum support value and minimum confidence value set as 50% and 75% respectively. **07**

TID Items Purchased

T101 Cheese, Milk, Cookies

T102 Butter, Milk, Bread

T103 Cheese, Butter, Milk, Bread

T104 Butter, Bread

(b) What is Concept Hierarchy? List and explain types of Concept Hierarchy. **07**

Q.4 (a) Explain k-means and k-medoids algorithm of clustering. **07**

(b) What is "Information Gain"? Explain the steps required to generate a Decision Tree from a training data set. **07**

OR

Q.4 (a) Explain Baye's Theorm and Naïve Bayesian Classification. **07**

(b) Explain linear regression? What are the reasons for not using the linear regression model to estimate the output data? **07**

Q.5 (a) Discuss the application of data warehousing and data mining in government sector. **07**

(b) What are neural networks? Describe the various factors which make them useful for classification and prediction in data mining. Explain how the topology of neural network is designed. **07**

OR

Q.5 (a) Explain different types of Web Mining with example. **07**

(b) What are the challenges for effective resource and knowledge discovery in mining the world wide web? **07**
