

**GUJARAT TECHNOLOGICAL UNIVERSITY****BE - SEMESTER-VI(NEW) EXAMINATION – WINTER 2022****Subject Code:3163209****Date:19-12-2022****Subject Name:Data Mining and Business Intelligence****Time:02:30 PM TO 05:00 PM****Total Marks:70****Instructions:**

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Simple and non-programmable scientific calculators are allowed.

- Q.1** (a) Define: Information Gain, Gain Ratio & Gini Index **03**  
 (b) Explain Rule-based Classification in brief. **04**  
 (c) Differentiate OLTP and OLAP. **07**

- Q.2** (a) What is meta data repository? Explain. **03**  
 (b) A data warehouse is a subject-oriented, integrated, time-variant, and nonvolatile collection of data – Justify. **04**  
 (c) What is Cuboid? Explain various OLAP operations on data cube with example. **07**

**OR**

- (c) With the help of a neat diagram explain the 3-tier architecture of a data warehouse. **07**

- Q.3** (a) Discuss click-stream analysis using data mining. **03**  
 (b) Explain Snowflake schema with an example. **04**  
 (c) List and Explain Data mining task primitives. **07**

**OR**

- Q.3** (a) Discuss fraud detection analysis using data mining. **03**  
 (b) Explain Fact Constellation Schema for Multidimensional Database with an example. **04**  
 (c) Define the term “Data Mining”. With the help of a suitable diagram explain the process of knowledge discovery from databases. **07**

- Q.4** (a) What is MapReduce? Explain in brief. **03**  
 (b) Explain text mining using example. **04**  
 (c) What is noise? Describe the possible reasons for noisy data. Explain the different techniques to remove the noise from data. **07**

**OR**

- Q.4** (a) What is Big Data? What is big data analytic? Explain in brief. **03**  
 (b) Discuss the main features of Hadoop Distributed File System. **04**  
 (c) What is data transformation in data mining? Explain. **07**

- Q.5** (a) Describe web mining using example. **03**  
 (b) What is classification? Explain classification as a two step process with diagram. **04**  
 (c) Explain Apriori algorithm for discovering frequent item sets for mining Boolean Association Rules. **07**

**OR**

- Q.5** (a) Define: Name Node, Data Node, Spatial Mining **03**  
 (b) What is clustering? List out the General applications of Clustering. **04**  
 (c) What is Market Basket Analysis? Explain Association Rules with Confidence & Support. **07**

\*\*\*\*\*