Seat No.:	Enrolment No
-----------	--------------

## GUJARAT TECHNOLOGICAL UNIVERSITY MCA - SEMESTER- IV • EXAMINATION - WINTER - 2016

Date:26/10/2016 Subject Code: 640005 Subject Name: Data Warehousing and Data Mining Time: 10.30 AM TO 01.00 PM **Total Marks: 70 Instructions:** 1. Attempt all questions. Make suitable assumptions wherever necessary. 3. Figures to the right indicate full marks. 0.1 i) Give any four differences between OLAP and OLTP. 02 (a) ii) Explain meaning of 4 keywords used in the definition of Data Warehouse. 02 iii) Define Concept hierarchy with example. 03 **(b)** i) Explain in brief Knowledge Discovery from Database with diagram. 05 ii) Explain Star schema and Snow flake schema with suitable example. 02 **Q.2** (a) Explain outlier analysis with its importance. 07 Why real world data is dirty? How to clean the Noisy data? Elaborate it with **(b)** 07 appropriate example OR What do you mean by data reduction? Explain in brief any two data reduction 07 **(b)** strategies. **Q.3** What is SelfExp, PathExp and InExp? How these measures are used in discovery driven (a) 05 exploration of data cubes? Discuss Decision Tree Induction algorithm. **(b)** 05 Define Terms: Data Mart, Loose coupling, Tight coupling, Prediction (c) 04 OR 0.3 Explain Bagging and Boot strap methods. 05 (a) Discuss Bayes' Theorem. 05 **(b)** Define Terms: Data Cube, Semi-tight coupling, no-coupling, Classification (c) 04 **Q.4** What is Apriori property? How the Apriori property is used in finding frequent itemset. 07 (a) Explain Join and Prune step Justify the Statement with proper example: "Strong rules are not necessarily **(b)** 07 interesting". OR **Q.4** Discuss Major issues in Data Mining. 07 (a) Justify the statement: "All patterns derived by Data Mining algorithm are Interesting". **(b)** 07 Write down the k-medoid algorithm for clustering and explain its working. How it is **Q.5** 07 better than k-mean clustering algorithm? Write Short Note: 07 **(b)** i) Data Mining in Retail Industry ii) Types of OLAP Servers OR Explain K-mean clustering method with its advantages and disadvantages. Q.5 07 (a) Write Short Note: **(b)** 07 iii) Role of Data Mining in Financial Data Analysis iv) OLAP Operations \*\*\*\*\*\*

1