

Code: CS6T4

III B.Tech - II Semester – Regular Examinations – May 2017

**DATA WAREHOUSING AND DATA MINING
(COMPUTER SCIENCE & ENGINEERING)**

Duration: 3 hours

Max. Marks: 70

PART – A

Answer *all* the questions. All questions carry equal marks

11x 2 = 22 M

1.

- a) What is a data ware house?
- b) Define fact table.
- c) What is snowflake schema?
- d) What are the various operations on Data cube?
- e) Can you list few statistical techniques?
- f) Define the term pattern.
- g) How can you predict association from large databases?
- h) What is attribute selection measure?
- i) What are the fields in which clustering techniques are used?
- j) Write about the feature selection in outlier detection.
- k) What is the use of Clustering?

PART – B

Answer any *THREE* questions. All questions carry equal marks.

3 x 16 = 48 M

2. Can you list the characteristic differences between OLAP and OLTP? 16 M

3. a) Explain various issues in data mining. 8 M

b) Explain various tasks in data mining. 8 M

4. a) Given the following transactional database 8 M

TID	Items
1	C, B, H
2	B, F, S
3	A, F, G
4	C, B, H
5	B, F, G
6	B, E, O

Find all frequent item sets using Apriori algorithm. With minimum support of 60% and confidence as 80%

b) What examples can you find for bayes theorem? Write them with explanation. 8 M

5. Explain the hierarchical and density based methods of clustering. 16 M

6. Explain about the Time series and streaming outlier detection. 16 M