

**B.Tech Sem – VIII (2007 Course) (Inf. Tech.) : WINTER -
2017**

SUBJECT: ELECTIVE- II: DATA MINING & MACHINE LEARNING

Day: **Wednesday**
Date: **22/11/2017**

W-2017-2685

Time: **02.30 PM TO 05.30 PM**
Max Marks: **80**

N.B:

- 1) **Q. No 1 and 5 are COMPULSORY.** Out of remaining attempt any **TWO** questions from each section.
- 2) Figures to the right indicate **FULL** marks.
- 3) Answers to both the sections should be written in **SEPARATE** answer book.
- 4) Assume suitable data if necessary.

SECTION-I

- Q.1**
- a) Compare OLAP with OLTP. (05)
 - b) Why preprocess the data? (05)
 - c) What is constrain based association mining (04)
- Q.2**
- a) What is data warehouse? Explain three architecture of data warehouse with block diagram. (07)
 - b) Define following terms: (06)
 - i) Dimension table
 - ii) Show flake
 - iii) Fact table
- Q.3**
- a) What are the characteristics of the data present in the data warehouse? Explain. (07)
 - b) Describe the methods for the generation of concept hierarchies for categorical data. (06)
- Q.4**
- a) Explain Apriori algorithm? How the association rules are generated from frequent item sets. (07)
 - b) What is meant by market basket analysis? State and explain with formula the meaning of following terms: (06)
 - i) Support
 - ii) Confidence

SECTION-II

- Q.5**
- a) What is decision tree induction? (05)
 - b) What is linear regression? What are linear model and state its use? (05)
 - c) What is the theory of minimum description length learning? (04)
- Q.6**
- a) Describe Naïve Bayesian method of classification. (07)
 - b) What is meant by classification? How the classifier accuracy determined and also explain its various types. (06)
- Q.7**
- a) Explain K mean clustering algorithm with example. (07)
 - b) State nearest neighbor algorithm. (06)
- Q.8**
- a) What is sparse data problem in statistically based natural language processing? (07)
 - b) Explain machine learning approach to knowledge acquisition. (06)

* * * * *