Reg No.										

MANIPAL INSTITUTE OF TECHNOLOGY

(Constituent Institute of Manipal University)

MANIPAL-576104

VII SEMESTER B.E. (CSE) DEGREE EXAMINATION 2009-10
SUBJECT: DATA WAREHOUSE AND DATA MINING(CSE 407.1)
DATE : 1-1-2010

TIME : 3 HOURS **MAX.MARKS : 50**

Note: (i) Answer **ANY FIVE** full questions.

1.(a) For the tables description below, design a star schema. List the number of dimensions formed out of the design. (4 Marks)

Time Table

Time_key
Day
Day_of_the_Week
Month
Quarter

Sales Table

Time_key
Item_key
Location_key
Dollars_sold
Units_sold

Item Table

Item_key
Item_name
Brand
Type
Supplier_Type

Location Table

Location_key
Street
City
State
Country

(b) Explain Roll-up and Slice operations with an example for each. (4 Marks)

(c) Explain loosely coupled and tightly coupled DBMS to a Data Mining application. (2 Marks)

2. (a) For the transactions given below, if the frequent itemset is {TV, Fridge, Camera} and {TV, Fridge, WM}, find all the association rules, given minimum confidence = 70%.

TID	List of Items
T1	TV, Fridge, WM
T2	Fridge, DVD Player
T3	Fridge, Camera
T4	TV, Fridge, DVD Player
T5	TV, Camera
T6	Fridge, Camera
T7	TV, Camera
T8	TV, Fridge, Camera, WM
T9	TV, Fridge, Camera

(5 Marks)

(b) Explain Border algorithm for Association Rule mining. (5 Marks)

3. (a) Write the FP-Tree algorithm for Association Rule mining. (5 Marks)
(b) Explain Naïve Bayesian Classification method. (5 Marks)

4. (a) Write the basic algorithm for decision tree construction. (4 Marks)
(b) Explain any one Attribute selection measure for Decision Tree construction. (2 Marks)
(c) Write a note on Rough Sets and Genetic Algorithms. (4 Marks)

5. (a) State the ways in which distance can be calculated between two Interval Scaled variables. (4 Marks)
(b) Write a note on Nominal and Ordinal variables. (2 Marks)
(c) For a Data Mining application, a Partitioning Algorithm needs to be used. It is found that the test data is free of outliers. Choose a Partitioning Algorithm and explain it. (4 Marks)

6.(a) Define the terms Precision and Recall with respect to Text/Web Mining. (4 Marks)
(b) Explain the mechanism used for Page Ranking. (2 Marks)
(c) Explain Web Usage mining. (4 Marks)

****