(1)

510110A - Business Intelligence and Data Mining
M.E. (Computer Engineering)

(Elective - II) (2013 Course)

Time : 3 Hours Max. Marks: 50

Instructions to the candidates:

May 2016

1) Draw labeled diagrams if necessary.

2) Assume suitable data if necessary.

Q 1. a) What is Business Intelligence? Explain Business Intelligent System components 4

b)What are the characteristics and components of Decision Support System 4

OR

Q 1. a)What is a need of Data warehouse. Explain ETL process in detail. 4

b)Explain role of BI in Banking sector 4

Q 2. a) What is the need of Data cleaning process. Explain its steps. 3

b) With the neat diagram, explain three tiers Data Warehousing Architecture. 3

OR

Q 2. Explain OLAP and its types. Also explain the advantages of OLAP 6

Q 3. For the Olympics Game System ,draw a multidimensional cube. What specific OLAP operations should you perform for the following operations? 10

a. List the total games and its type(group and individual)
b. List the total participants for each game
c. List the Country who scored highest medals in each game.
d. List top 3 countries who won the medals
e. List the average of each country

OR

Q 3. Compare Star, Snowflake, and Fact Constellations schemas. Consider the 10

example of Sales department of an electronic goods.

Q 4. How are association rules mined from large databases? 10

OR

Q 4. Describe following terms in detail. 10

1. Frequent Item set Generation
2. Support and Confidence
3. Infrequent Pattern
4. Negative pattern
5. Concept Hierarchy

Q 5. a) Explain different attribute selection methods 6
b) What is tree pruning and its necessacity. 2

OR

Q 5. a) Solve the following using Naive Bayes classification 8
Attributes are Color, Type, Origin, and the subject, stolen can be either yes or no.

Data Set:

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Find output for the tuple : **Red Domestic SUV**

Q 6. a) Consider following data of variable A and B . Apply K means algorithm 8
A: 1.0, 1.5, 3.0, 5.0, 3.5, 4.5, 3.5
B: 1.0, 2.0, 4.0, 7.0, 5.0, 5.0, 4.5
Assume K=2

OR

Q 6. a) What are outliers? Explain types of outliers 4
b) What are the challenges in clustering high dimensional and big data . 4