USN | | | | | | | | | | |

06IS74

Seventh Semester B.E. Degree Examination, June/July 2013
Data Mining

Time: 3 hrs. Max. Marks:100

Note: ***Answer FIVE full questions, selecting at least TWO questions from each part.***

PART – A

1 a. What is data mining? Explain the various data mining task with example. **(08 Marks)**
b. List and explain general characteristics of data sets. **(06 Marks)**
c. Explain measurement errors and data collection errors with example. **(06 Marks)**

2 a. Explain different sampling approaches with example. **(08 Marks)**
b. List the benefits of dimensionality reduction. **(04 Marks)**
c. Explain feature subset selection process with a flow chart. **(08 Marks)**

3 a. Write the Hunt's algorithm used to build decision tree. **(06 Marks)**
b. Explain characteristics of rule-based classifiers. **(06 Marks)**
c. Explain nearest neighbor classifier with algorithm. **(08 Marks)**

4 a. What is apriori principle? Explain. **(06 Marks)**
b. State and explain the requirements for candidate generation procedure. **(06 Marks)**
c. Explain different factors affecting the complexity of Apriori algorithm. **(08 Marks)**

PART – B

5 a. Explain how the FP tree is represented. **(08 Marks)**
b. Describe the effect of skewed support distribution. **(06 Marks)**
c. Explain Simpson's paradox. **(06 Marks)**

6 a. Write and explain basic-Kmeans algorithm. **(08 Marks)**
b. Explain different applications of cluster analysis. **(06 Marks)**
c. Write a note on DBSCAN. **(06 Marks)**

7 Explain the following in brief,
a. Spatial data mining.
b. Multimedia data mining.
c. Text mining.
d. Web mining. **(20 Marks)**

8 a. Explain the social impact of data mining. **(10 Marks)**
b. Describe the trends in data mining. **(10 Marks)**

* * * * *