Code No: 157BC

R18

Time: 3 Hours   **Max. Marks: 75**

Note: i) Question paper consists of Part A, Part B.
ii) Part A is compulsory, which carries 25 marks. In Part A, answer all questions.
iii) In Part B, Answer any one question from each unit. Each question carries 10 marks and may have a, b as sub questions.

PART – A

(25 Marks)

1.a)	What is data warehouse?	[2]
b)	List out the applications of data mining.	[3]
c)	What is meant by association rule mining?	[2]
d)	Write a short note on SPM algorithm?	[3]
e)	Why are decision trees useful?	[2]
f)	List the advantages of using decision trees.	[3]
g)	Discuss the two approaches to improve quality of hierarchical clustering.	[2]
h)	List the applications of cluster analysis.	[3]
i)	Define data stream mining.	[2]
j)	Give the taxonomy of web mining.	[3]

PART – B

(50 Marks)

2.a) Explain how to integrate data mining system with a data warehouse.
b) "Data preprocessing is necessary before data mining process". Justify your answer. [5+5]

OR

3.a) Differentiate between data mining and data warehouse.
b) Discuss the major issues in data mining. [5+5]

4.a) Write a short notes on constraint based association mining.
b) Describe various types of association rules. [5+5]

OR

5. Explain in detail about frequent pattern mining in data mining. [10]

6. Describe Bayesian Belief Network with an example. [10]

OR

7.a) Briefly explain classification problems and general approaches to solve them.
b) Explain the merits and de-merits of the lazy learning method. [5+5]

8. Explain the following.
   a) Cluster analysis.
   b) Grid–based methods. [5+5]

OR

9.a) How density based method is used for clustering?
  b) Illustrate K-mean algorithm with an example. [4+6]

10. Explain the following.
    a) Spatial data mining.
    b) Text mining. [5+5]

OR

11. Discuss various kinds of patterns to be mined from web/server logs in web usage mining.
    [10]

---ooOoo---