Total No. of Questions : 8]

SEAT No. :

P3195

[Total No. of Pages : 2

[5670]-297

B.E. (Computer)

DATA MINING TECHNIQUES AND APPLICATIONS

(2012 Pattern) (Semester - I) (Elective - I)

Time : 2½ Hours] *[Max. Marks :70*

Instructions to the candidates:

1) *Answer questions 1 or Q2, Q.3 or Q.4, Q.5 or Q6, Q.7 or Q.8.*
2) *Neat diagrams must be drawn wherever necessary.*
3) *Figures to the right indicate full marks.*
4) *Assume suitable data, if necessary.*

Q1) a) Explain data integration, data transformation and data reduction in short. **[6]**

b) Write and explain FP Growth Algorithm for calculating frequent item sets. How it is better than Apriori algorithm. **[8]**

c) Describe performance metrics for evaluating classifiers with equations. **[6]**

OR

Q2) a) Explain following procedures for attribute subset selection: **[8]**
   i) Stepwise forward selection.
   ii) Stepwise backward elimination.

b) Explain the following in short: **[8]**
   i) Regression Classification Algorithm.
   ii) ID3 Decision trees algorithm.

c) Write a short note on constraint based association rule mining. **[4]**

Q3) a) Explain cluster analysis. Write and explain K-Means Clustering algorithm. **[6]**

b) What is Hierarchical Clustering? Explain Agglomerative and Divisive Hierarchical Clustering with neat diagrams. **[5]**

c) Explain different distance measures in brief. **[6]**

OR

P.T.O.

Q4) a) Write typical requirements of clustering in data mining. Describe any five. **[5]**

b) Write and explain PAM Clustering algorithm. **[8]**

c) Using the two given objects represented by the tuples (22, 1 , 42, 10) and (20, 0, 36, 8), compute the following: **[4]**

i) Euclidean distance

ii) Manhattan distance

Q5) a) Explain the following basic measures for text retrieval: **[6]**

i) Precision

ii) Recall

iii) F-score

b) Write and explain Hyperlink-Induced Topic (HITS) algorithm. **[6]**

c) Which methods are used for Dimensionality Reduction of Text in text mining? Describe any one in brief. **[5]**

OR

Q6) a) Describe the following terms in brief: Feature vector, Term frequency, Bag of words and Inverse document frequency. **[8]**

b) Write a short note on: **[6]**

i) Web content mining

ii) Web usage mining.

c) What is document ranking? **[3]**

Q7) a) Write a short note on: **[10]**

i) Reinforcement learning

ii) Wholistic Learning

b) Describe multi-perspective learning along with diagram. **[6]**

OR

Q8) a) Explain the following with respect to reinforcement learning: Intelligent agent and environment, Learning agents, Rewards and Adaptive learning. **[8]**

b) Describe the advanced techniques for big data mining. **[8]**