

Total No. of Questions : 8]

SEAT No. :

**P3226**

[5461]-267

[Total No. of Pages : 2

**B.E. (Computer)**

**DATA MINING TECHNIQUES AND APPLICATIONS  
(2012 Pattern) (Semester - I) (End Sem.) (Elective - I) (410444D)**

*Time : 2½ Hours]*

*[Max. Marks : 70*

*Instructions to the candidates:*

- 1) Answer Q.1 or Q.2, Q.3 or Q.4, Q.5 or Q.6, Q.7 or Q.8.
- 2) Neat diagram must be drawn wherever necessary.
- 3) Figures to the right side indicate full marks.
- 4) Assume suitable data if necessary.

**Q1) a)** Use the two methods below to normalize the following group of data :

200, 300, 400, 600, 1000 [6]

- i) min-max normalization by setting min = 0 and max = 1
- ii) z-score normalization

**b)** A database has five transactions. Let min sup = 60% and min conf = 80%. [6]

TID	Items
T100	{M, O, N, K, E, Y}
T200	{D, O, N, K, E, Y}
T300	{M, A, K, E}
T400	{M, U, C, K, Y}
T500	{C, O, O, K, I, E}

Calculate all frequent item sets using Apriori algorithm.

- c)** Write short note on : [8]
- i) Extracting Rules from decision trees.
  - ii) KNN approach

OR

**Q2) a)** Explain data preprocessing steps in short. [8]

**b)** Explain performance metrics : Accuracy, Precision, Recall and F-Measure with required equations. [8]

**c)** What do you mean by frequent patterns? Explain constraint based association rule based mining frequent item sets. [4]

*P.T.O.*

- Q3)** a) Write equations defining Manhattan, Minkowski and Euclidean distance measures. [6]  
b) Explain AGNES and DIANA (Agglomerative and Divisive Hierarchical Clustering). [5]  
c) Write and explain K-means clustering algorithm. [6]

OR

- Q4)** a) Explain k-Medoid Clustering algorithm. [8]  
b) Explain how k-Medoid is extended to CLARANS for handling large data sets. [4]  
c) What are typical requirements of clustering in data mining (explain any five)? [5]

- Q5)** a) Explain basic measures for text retrieval like Precision, Recall etc. in the terms of retrieved and relevant documents. [8]  
b) Compare: Web content mining and Web usage mining. [6]  
c) What are methods for Dimensionality Reduction of Text in text mining? List the methods. [3]

OR

- Q6)** a) Compare different text mining approaches. [5]  
b) Write and explain Hyperlink-Induced Topic (HITS) algorithm. [6]  
c) Explain these terms in short: Feature vector, Bag of words, Tf-Idf. [6]

- Q7)** a) Write notes on [10]  
i) Reinforcement learning,  
ii) Systematic Learning  
b) Write a note on multi-perspective learning along with diagram for multi-perspective learning. [6]

OR

- Q8)** a) What is Big Data? How it is characterized? What are the challenges in Big data Analysis? [8]  
b) What are techniques for big data mining? Explain in short. [8]

