U.S.N | | | | | | | | | | |

P.E.S. College of Engineering, Mandya - 571 401

(An Autonomous Institution affiliated to VTU, Belgaum)

Eighth Semester, B.E. - Information Science and Engineering

Semester End Examination; June/July - 2015

Data Mining

Time: 3 hrs *Max. Marks: 100*

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

PART - A

1. a. Discuss different types of attributes and give one example for each. 10

   b. With a neat block diagram, explain the process of knowledge discovery in databases. 10

2. a. Explain different approaches for feature subset selection. 6

   b. What is entropy? How it is used for supervised Discretization? 10

   c. Let X and Y be vectors

   X = (3, 2, 0, 5, 0, 0, 0, 2, 0, 0)

   Y = (1, 0, 0, 0, 0, 0, 0, 1, 0, 2) 4

   Calculate cosine similarity.

3. a. How do you measure for selecting the best split in classification? 10

   b. Discuss the characteristics of Nearest – Neighbor classifiers. 10

4. a. Write and explain frequent itemset generation with Apriori algorithm. 10

   b. What is maximal frequent itemset? Give one example. What are its disadvantages? 10

PART – B

5. a. Explain the procedure that is used in FP-tree construction. 10

   b. How do you handle concept hierarchy? Explain. 10

6. a. Distinguish between Hierarchical and Partitional clustering. 4

   b. Explain the basic K-mean algorithm used in clustering. 10

   c. How DBSCAN works with clusters of varying density. 6

7. a. Discuss basic measures used for text retrieval. 10

   b. What are the challenges observed in mining the world wide web? 10

8. a. Explain any two applications on data mining. 10

   b. What are the factors considered in choosing a data mining system? Explain. 10

* * * * *