**Vidyavahini First Grade College**

**Fifth Semester B.C.A Degree Examination(CBCS Scheme), November 2018**

Computer Science

**DATA MINING-Model Question Paper-2**

Time: 3 Hours] [Max marks: 90

Instructions to Candidates: Answer all the Sections

SECTION-A

Answer any **TEN** questions**: [10 x 1 = 10]**

1.What is knowledge in DM?

2. What is noisy data.?

3.What is data mart?

4. Define logistic regression ?

5.What is supervised learning.

6.Expand DBSCAN?

7.What is Pruning?

8.What is Dendrogram?

9.Mention any 2 clustering softwares?

10.What is HDFS in Big Data?

11. What is Web crawler

12.Give an example of unstructured data?

SECTION-B

Answer any **FIVE** questions**: [5 x 3 = 15]**

13. Explain briefly applications of Data mining in any 3 fields.

14. What is Classification ? mention its importance in DM?

15. What are the requirements of Clustering in DM

16. What is Decision Tree? Explain with an example

17. Differentiate between Clustering and Classification tasks of DM.

18. What is web data mining? mention the types .

19. Briefly explain the importance of big data in an existing world

Answer any **SIX** questions**: [6 x 5 = 30]**

20. Explain 3-tier Architecture of Data ware house with neat diagram .

21. Briefly explain Classification by back propagation.

22. Explain snow flake schema model with an example.

23. Briefly explain rule based method of classification with an example.

24.What is Hierarchical Clustering? Explain AGNES type with an example .

25.Write short notes on Data quality and Validation measures of clustering.

26. Explain the architecture and working of Search engine.

27. Explain briefly 5V characteristics of Big data .

.

Answer any **FIVE** questions**: [5 x 7 = 35]**

28.a Differentiate between OLTP and OLAP

b. Define OLAP cube. Briefly explain some OLAP operations.

29.a Write the Bayesian formula for classification and explain the terms

i. likelihood ii. Prior probability iii. Posterior probability

b. explain the steps involved in construction of decision tree for classification

30. a. What is Prediction ? Mention some applications of it in DM?

b. Briefly explain statistical based classification.

31.a. Explain partition method of clustering with an example .

b. Given a Dataset D={2,3,4,8,10,12,15,20,22,25} make to 2 clusters using K-means

32. a. What is cluster analysis? Mention some applications.

b. Briefly explain density method of clustering with an example

33.a.Write a short notes on web content mining .

b. Briefly explain the ranking of web pages in Search engine.

34. Write short notes on Big Data Tools and Techniques

a. Hadoop

b. NoSQL.

c. Map Reduce