**Vidyavahini First Grade College**

**Fifth Semester B.C.A Degree Preparatory Examination (CBCS Scheme), Oct/Nov 2019**

Computer Science

**DATA MINING QPCODE 68506**

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 Data Mining?

2.What is pattern matching?

3.What is Associative Classification in DM?

4.What is ETL in Data warehousing?

5.Mention any two Data mining tools or S/W.

6.What is Clustering?

7.What is Pruning?

8.What is prediction in DM?

9.what is Sigmoid function ?

10.What is an Outlier?

11.Define Search Engine

12.What is Big Data?

SECTION-B

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

13.Differentiate B/W operational Databases and Data warehouses.

14.What are Data mining Tasks? Mention the types.

15.Briefly explain steps involved in Data preprocessing.

16. Explain 2 steps of Classification process

17. Explain some OLAP operations

18. What is cluster analysis? Mention its requirements.

19. Explain briefly different data structure formats of big data.

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

20. Explain briefly application s of Data mining in different fields.

21. Briefly explain 3-tier Architecture of Data ware house

22. What is Star Schema? Explain with an example.

23. Given a Dataset D={2,5,6,8,10,12,15,25,30} make to 2 clusters using K-means.

24.Explain the steps involved in Agglomerative Hierarchical Clustering with a neat diagram.

25.What is Web Data Mining? Explain the types.

26. Briefly explain rule based classification with an example

27. Explain briefly 5V characteristics of Big data.

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

28.Explain the Steps involved in construction of Decision Tree.

Construct a Decision tree for a given training dataset

29.a. Explain web usage mining

b. Explain briefly the architecture of Search Engine

30. What is Naïve Bayesian classification. From the table given beside evaluate the statement. Male with salary > 18000 buys widget. By finding the posterior probabilty.

31.a.write short notes on i. Hadoop ii. NoSQL

b. Mention some tools used in Big data

32. a. briefly explain back propagation method of classification

b. Briefly explain density method of clustering with an example

33. What is KDD? Explain the process with a neat Diagram.

34. a. Explain the metrics used for Data Mining

b. What is machine Learning? Explain the types.