

BTS-VI-04.19-0618

Reg. No.

--	--	--	--	--	--	--	--

A

***B.Tech. Degree VI Semester Examination April 2019***

**CS 15-1604 DATA MINING**  
(2015 Scheme)



Time: 3 Hours

Maximum Marks: 60

**PART A**

(Answer *ALL* questions)

(10 × 2 = 20)

- I. (a) How is a data warehouse different from a database? How are they similar?
- (b) Briefly compare snowflake schema, fact constellation and star query model.
- (c) Elucidate two phases involved in decision trees.
- (d) List and describe two interesting measures for association rules.
- (e) The naïve Bayes classifier makes what assumption that motivates its name?
- (f) Compare clustering and classification.
- (g) What is meant by hierarchical clustering?
- (h) How does Hadoop Map reduce work?
- (i) What are the components of Apache spark ecosystem?
- (j) What are the basic differences between relational database and HDFS?

**PART B**

(4 × 10 = 40)

- II. Explain the data mining functionalities, and the kinds of patterns they can discover.

**OR**

- III. Why is data preprocessing required? Also describe the real-world data that are to be analyzed by data mining techniques.

- IV. Discuss the single-dimensional Boolean association rule mining for transactional databases.

**OR**

- V. Explain support vector machines in details.

- VI. Explain one of the centroid-based technique used in clustering with proper algorithm.

**OR**

- VII. Write notes on time series mining and graph mining.

- VIII. Give an introduction to cloud computing and its services.

**OR**

- IX. Explain Hadoop architecture and ecosystem.

\*\*\*