

Roll No. 40

Exam Code : D-24

Subject Code—62440

B. Sc. EXAMINATION

(Main/Reappear) (Batch 2020 Onwards)

(Fifth Semester)

COMPUTER SCIENCE

CCsL-504

Paper—II

Data Analytics

Time : 3 Hours

Maximum Marks : 80

Note : Attempt *Five* questions in all, selecting *one* question from each Unit. Q. No. 1 is compulsory. All questions carry equal marks.

1. (i) What is the purpose of data warehouse ?
- (ii) Describe the Business Intelligence.
- (iii) What is a splitting variable in decision tree ?

P.T.O.

(5-27-1-1224) J-62440

- (iv) What is a decision tree ?
- (v) What are frequent itemsets ?
- (vi) What are association rules ?
- (vii) What are support vectors ?
- (viii) Explain the kernel method. 8×2=16

Unit I

- 2. What is Data Mining ? What are supervised and unsupervised learning techniques ? Explain in detail. 16
- 3. Explain BI (Business Intelligence) skills and BI applications in detail. 16

Unit II

- 4. What is logistic regression ? What are advantages and disadvantages of regression algorithm ? 16
- 5. Write down the steps of developing an ANN and Advantages and disadvantages of using ANN. 16

Unit III

6. Write down K-Means algorithm for Clustering with its advantages and disadvantages. 16
7. Explain in detail with example web content, web structure and web usage mining. 16

Unit IV

8. What is Naïve-Bayes Model ? Write advantages and disadvantages of Naïve-Bayes. 16
9. Write short notes on the following : 16
 - (a) Big data and Big data landscape
 - (b) Business implications and technology implications of big data.