

Register No.: ..... Name: .....

**SAINTGITS COLLEGE OF ENGINEERING (AUTONOMOUS)**

(AFFILIATED TO APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY, THIRUVANANTHAPURAM)

**THIRD SEMESTER M.TECH DEGREE EXAMINATION (Regular), FEBRUARY 2022****COMPUTER SCIENCE & SYSTEMS ENGINEERING****(2020 Scheme)****Course Code : 20CSSET223****Course Name: Data Mining****Max. Marks : 60****Duration: 3 Hours****PART A***(Answer all questions. Each question carries 3 marks)*

1. How is a data warehouse different from a database?
2. List any three applications of Naive Bayes Classification.
3. Differentiate linear regression and non-linear regression.
4. Define a clustering problem. Given clusters  $K_i$  and  $K_j$ , suggest some alternatives to calculate the distance between clusters.
5. Define Association Rule Mining with an example.
6. Define large itemset, support and confidence.
7. What is the significance of personalization in web content mining?
8. Differentiate Markov model and Hidden Markov model.

**PART B***(Answer one full question from each module, each question carries 6 marks)***MODULE I**

9. List and explain the steps involved in data mining as a process of knowledge discovery (6) with a neat diagram.

**OR**

10. Illustrate the basic data mining tasks with examples. (6)

**MODULE II**

11. Explain point estimation with a statistical perspective on data mining with examples. (6)

**OR**

12. Illustrate classification using Bayes theorem with an example. (6)

**MODULE III**

13. "The KNN technique is extremely sensitive to the value of K". Justify the statement (6) with an example.

**OR**

14. Explain rule extraction from a decision tree. How does it differ from generating rules from a neural network? (6)

**MODULE IV**

15. How does DBSCAN find clusters? What are the advantages of DBSCAN clustering compared with other clustering methods? (6)

**OR**

16. Compare various agglomerative algorithms with examples. (6)

**MODULE V**

17. Explain Apriori-gen algorithm with example transactions. (6)

**OR**

18. Illustrate the use of partition algorithm using market basket data example. (6)

**MODULE VI**

19. What are the advantages of Web content mining? Describe the working of any one of the web content mining methods. (6)

**OR**

20. Explain the spatial data structures involved in spatial mining. (6)

\*\*\*\*\*