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

SAINTGITS COLLEGE OF ENGINEERING (AUTONOMOUS)

(AFFILIATED TO APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY, THIRUVANANTHAPURAM)

SECOND SEMESTER M.TECH DEGREE EXAMINATION (Regular), JULY 2022

COMPUTER SCIENCE AND SYSTEMS ENGINEERING

(2021 Scheme)

Course Code: **21SE204-B**

Course Name: **Data Mining**

Max. Marks: **60** **Duration: 3 Hours**

PART A

(Answer all questions. Each question carries 3 marks)

1. Give an insight about preprocessing in KDD process.
2. What is bias in Point Estimation?
3. List out any 3 issues related to clustering.
4. What are outliers? How is outlier detection done?
5. Define confidence of an association rule.
6. 'Large number of database scans is a weakness of the Apriori approach'. Justify with necessary comments.
7. What are crawlers? How is a periodic crawler different from incremental crawler?
8. What is temporal database? How is it different from a traditional database?

PART B

(Answer one full question from each module, each question carries 6 marks)

MODULE I

9. Explain any 6 data mining issues. (6)

OR

10. Interpolate data mining from a database perspective. (6)

MODULE II

11. Given the training data in the table below (Buy Computer data, predict the class of the following new example using Naïve Bayes classification: age<=30, income=medium, student=yes, credit-rating=fair. (6)

RID	age	income	student	credit_rating	Class: buys_computer
1	<=30	high	no	fair	no
2	<=30	high	no	excellent	no
3	31 . . . 40	high	no	fair	yes
4	>40	medium	no	fair	yes
5	>40	low	yes	fair	yes
6	>40	low	yes	excellent	no
7	31 . . . 40	low	yes	excellent	yes
8	<=30	medium	no	fair	no
9	<=30	low	yes	fair	yes
10	>40	medium	yes	fair	yes
11	<=30	medium	yes	excellent	yes
12	31 . . . 40	medium	no	excellent	yes
13	31 . . . 40	high	yes	fair	yes
14	>40	medium	no	excellent	no

OR

12. List out the advantages and disadvantages of Neural Networks algorithm for classification. (6)

MODULE III

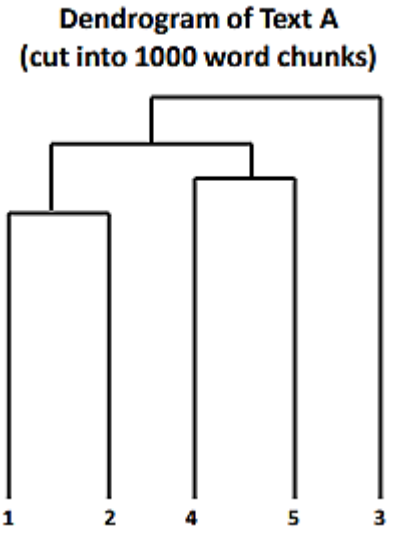
13. How will you generate rules from a Neural Net? Give the algorithm for the same. (6)

OR

14. With an algorithm, demonstrate how classification is done with simple distance metric. (6)

MODULE IV

15. Interpret the below dendrogram and explain detail.



(6)

OR

16. Explain Minimum Spanning Tree algorithm in hierarchical clustering with an example. (6)

MODULE V

17. How will you measure the quality of an association rule? Explain. (6)

OR

18. Explain incremental rules. (6)

MODULE VI

19. Explain the process of Page Ranking with an example. (6)

OR

20. With the help of an algorithm, explain STING. (6)

****