

Register No:

Name:

SAINTGITS COLLEGE OF ENGINEERING (AUTONOMOUS)

(AFFILIATED TO APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY, THIRUVANANTHAPURAM)

**SIXTH SEMESTER INTEGRATED MCA DEGREE EXAMINATION(R,S), MAY 2024
(2020 SCHEME)****Course Code : 20IMCAT304****Course Name : Fundamentals of Data Science****Max. Marks : 60****Duration:3 Hours****PART A***(Answer all questions. Each question carries 3 marks)*

1. What are some of the practical applications of data mining?
2. Write a short remark about the evolution of big data.
3. What factors led to the crucial role of storage in the context of Big Data?
4. Explain Brewer's theorem.
5. What are the different types of semantic analysis utilized for interpreting textual data, and how do they work?
6. What are some of the positive aspects and benefits of employing Big Data analytics in various industries and domains?
7. Identify and explain the core components of Hive.
8. Illustrate the Hadoop Heartbeat Message system with the help of a diagram.
9. What are the key characteristics of the storage layer in a Big Data architecture?
10. What security considerations should be taken into account when designing a Big Data stack?

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

11. Describe any four data mining techniques commonly used in Big Data analytics with examples. 6

OR

12. Describe the various types of business analytics in Big Data. 6

MODULE II

13. How does each component of the BASE model influence the design and functioning of distributed systems? 6

OR

14. What are the primary categories of NoSQL storage devices, and how do they differ in their approach to data storage and retrieval? 6

MODULE III

15. How can classification, regression, clustering, and outlier detection approaches distinguish different aspects of machine learning, and what are their primary goals and real-world applications? 6

OR

16. What are the various processes involved in the discovery phase of a data analytics project, and how do they contribute to uncovering insights and understanding the data? 6

MODULE IV

17. Describe the Hadoop ecosystem and illustrate it with a clear diagram. 6

OR

18. Give an overview of the HDFS component in the Hadoop architecture. 6

MODULE V

19. Could you elaborate on the various real-world applications of Big Data? 6

OR

20. Briefly explain the following layers in Big Data architecture 6
i)Physical infrastructure layer
ii)Platform management layer
