B.Tech (CSE-AI & ML) 5th Semester (G-Scheme)
Examination, November-2023
BIG DATA & ANALYTICS
Paper–PCC-DS-306G

Time allowed : 3 hours]　　　　　[Maximum marks : 75

Note : Question No. 1 is compulsory. Attempt any **five** questions selecting one question from each Unit. All questions carry equal marks.

1. (a) Explain 3 Vs of Big Data

    (b) Discuss the Non-Definitional traits of Big Data

    (c) Compare Hadoop and SQL

    (d) Explain CAP Theorem

    (e) Explain HQL in Hadoop Ecosystems

    (f) Discuss the alter commands in Cassandra

Unit-I

2. What is the Data Science Process? Write and explain Data science process steps.

3. What is Big Data? What are the Applications and Challenges with Big Data? Discuss Big Data Workflow Management.

Unit-II

4. What is Hadoop? Draw its architecture with its components. Why we prefer Hadoop over SQL. Write the Difference Between RDBMS and Hadoop.

5. What is NewSQL? Differentiate SQL, NoSQL, and NewSQL.

Unit-III

6. Discuss Mapper, Reducer, Combiner, and Partitioner in the context of MapReduce.

7. Explain the MongoDB. Is MongoDB better than SQL? Write its Features and Data types.

Unit-IV

8. Write and explain the Data Types in Cassandra.

9. What is Cassandra? Write its Architecture with its components. How Cassandra stores its data using its data model? What is the role of CQLSH? Discuss the CRUD operations.