

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GITAM SCHOOL OF TECHNOLOGY

GITAM (Deemed to be University,VISAKHAPATNAM-530045)

| | | | | | |
|------------------------------|-----------------------|---------------------------|---------------------------|--------------------|------------------|
| Program | M. Tech | Category of Course | PE | Semester | II |
| Branch | CSE/DS | Course Code | 19ECS767 | Term | 2023-24 |
| | | Course Title | BIG DATA ANALYTICS | | |
| Continuous Evaluation | ASSIGNMENT -II | Max Marks | 6 M | Start Date | 18-3-2024 |
| | | | | Cut of Date | 29-3-2024 |

UNITS III, IV, and V

| | Answer all questions. | | |
|----------|---|------------|-----------|
| 1 | a) Write about the design of the Hadoop distributed file system (HDFS). Describe different HDFS concepts in detail. b) Write a note on Hadoop File-Based Data Structures | CO3 | L1 |
| 2 | What is the concept of Unit test? Explain with an example. Discuss the case study on MRUnit. | CO4 | L2 |
| 3 | What is YARN? Analyze the YARN architecture in detail with a suitable block diagram. | CO4 | L3 |
| 4 | Analyze the Anatomy of a YARN Application Run and how YARN runs an application. | CO4 | L4 |
| 5 | How does the YARN job scheduling work, Explain different job scheduling approaches in YARN. | CO4 | L3 |
| 6 | Discuss core HBase concepts. Write a note on HBase clients with HBase examples | CO5 | L4 |

Note: Please write the assignment on A4 sheets only