

**R18**

Code No: 156EN

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech III Year II Semester Examinations, March - 2024

**BIG DATA ANALYTICS**  
(Computer Science and Engineering – Data Science)

Time: 3 Hours

Max. Marks: 75

- Note:** i) Question paper consists of Part A, Part B.  
ii) Part A is compulsory, which carries 25 marks. In Part A, Answer all questions.  
iii) In Part B, Answer any one question from each unit. Each question carries 10 marks and may have a, b as sub questions.

**PART – A**

**(25 Marks)**

- 1.a) Define the term “variety” in the context of Big Data. [2]
- b) What is the difference between Data Warehouse and Big Data? [3]
- c) Define the term “Hadoop”. [2]
- d) What role does scalability play in Big Data systems? [3]
- e) Name few programming languages commonly used with Hadoop. [2]
- f) Does Hadoop enable distributed computing? [3]
- g) What is HIVE in Hadoop ecosystem? [2]
- h) How is Hadoop configured? [3]
- i) Define supervised learning. [2]
- j) List the challenges of social media analytics. [3]

**PART – B**

**(50 Marks)**

- 2.a) Explain the concept of Big Data and its significance in modern businesses.
  - b) Discuss the relationship between machine learning and Big Data analytics. [5+5]
- OR**
3. State and explain the applications of Big Data Analytics. [10]
  4. Discuss the process and challenges associated with integrating Mobile Business Intelligence with Big Data systems. [10]
- OR**
5. Explain the following terms:
    - a) Hadoop’s Parallel World.
    - b) Predictive Analytics. [5+5]
  6. Describe the inputs and outputs of the MapReduce framework. [10]
- OR**
- 7.a) Explain the components of the Hadoop ecosystem and their functions.
  - b) Compare Hadoop with traditional relational database systems in terms of handling large datasets. [5+5]

QA QA QA QA QA QA QA G

- 8.a) Discuss the significance of Hadoop distributors in the ecosystem.
- b) Describe the process of file write and read operations in HDFS.

[5+5]

QA QA QA QA QA QA QA G

- 9. Elaborate on the functions and responsibilities of NameNode, Secondary NameNode, and DataNode in HDFS. [10]

- 10. Explain how BigR contributes to the efficiency and scalability of Big Data Analytics. [10]

**OR**

QA QA QA QA QA QA QA G

- 11. Compare and contrast different types of clustering algorithms used in unsupervised learning. [10]

**---ooOoo---**

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G