



VIT

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science Engineering and Information Systems
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025

REG.NO.:

SLOT: F1+TF1

Programme Name & Branch : B.Tech. Information Technology
Course Code and Course Name : BITE411L & BIG DATA ANALYTICS
Faculty Name(s) : Prof. Daphne Lopez, Prof. Pounambal M, Prof. Vivekananda GN
Class Number(s) : VL2024250502645, VL2024250502669, VL2024250502676
Date of Examination : 01-02-2025
Exam Duration : 90 minutes
Maximum Marks: 50

General instruction(s):

- Answer All Questions
- M - Max mark; CO - Course Outcome; BL - Blooms Taxonomy Level (1 - Remember, 2 - Understand, 3 - Apply, 4 - Analyse, 5 - Evaluate, 6 - Create)

Q. No	Question	M	CO	BL
1.	A retail company wants to analyze customer purchasing patterns. They gather data from various sources, such as online transactions, in-store purchases, social media reviews, and IoT-enabled smart shelves in stores. Identify the types of data and categorize them based on their sources. Explain how the characteristics of Big Data (volume, velocity, variety, veracity) apply to this scenario.	10	1	3
2.	How can Big Data address challenges and drive applications in the following industries? • Communication Media and Entertainment • Education • Healthcare • Government • Transportation	10	1	4
3.	A logistics company plans to use HDFS to store and process data related to package tracking, delivery routes, and customer feedback. • Explain how the architecture of HDFS (NameNode, DataNodes, and replication) can benefit the company in ensuring data availability and fault tolerance. • How can you read, write, upload, and retrieve data in HADOOP?	5	2	4
4.	Illustrate a dashboard which uses artificial intelligence methods to predict e-commerce transaction by applying various phases of data analytics lifecycle to ensure corrective results for its intended customers.	5	1	3
5.	Perform the MapReduce technique for the following input. Coke → Pepsi, Sprite, Fanta Pepsi → Coke, Sprite, Fanta, 7Up Sprite → Coke, Pepsi, Fanta, 7Up Fanta → Coke, Pepsi, Sprite, 7Up 7Up → Pepsi, Sprite, Fanta • Explain how the Map phase will extract each name from the input lines. • Describe how the Shuffle and Sort phase groups and eliminates duplicates. • Demonstrate how the Reduce phase combines the names into a single sorted string of unique names.	10	2	6
