



Roll No.																			
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

PRESIDENCY UNIVERSITY

BENGALURU

End - Term Examinations – December 2025

Date: 08-12- 2025

Time: 01:00pm – 04:00pm

School: SOCSE	Program: B.Tech (CSE)		
Course Code: CSE3509	Course Name: Emerging Technologies in Big Data		
Semester: V	Max Marks: 100	Weightage: 50%	

CO - Levels	C01	C02	C03
Marks	46	86	48

Instructions:

- (i) Read all questions carefully and answer accordingly.
- (ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks.

10Q x 2M=20M

1.	List characteristics of Big Data.	2 Marks	L1	C01
2.	Outline the evolution of Big Data.	2 Marks	L1	C01
3.	List the features of NoSQL.	2 Marks	L1	C01
4.	List the steps of export process in Sqoop.	2 Marks	L2	C02
5.	List DML commands of Hive.	2 Marks	L1	C02
6.	Define Machine Learning.	2 Marks	L1	C02
7.	List key features of Apache Spark.	2 Marks	L1	C03
8.	List two types of operations in RDD programming with an example.	2 Marks	L2	C03
9.	Describe Scala.	2 Marks	L1	C03
10.	List storage layers of Spark.	2 Marks	L2	C03

Part B

Answer the Questions.

Total Marks 80M

11.	a.	Explain Anatomy of File Write in HDFS.	10 Marks	L3	C01
Or					
12.	a.	Explain steps for installation of Hadoop single node cluster configuration.	10 Marks	L2	C01
13.	a.	Explain design of HDFS.	10 Marks	L2	C01
Or					
14.	a.	<p>Consider Tamil Nadu Crop production dataset. The overall area of crop production for years 2008 and 2009 are given below:</p> <p>(2008, 8608), (2008, 3), (2008, 73), (2008, 236), (2008,6), (2008, 17294), (2009, 1244), (2009, 10), (2009, 13), (2009,91), (2009, 344), (2009, 943)</p> <p>Apply mapper to find maximum area of crop production.</p> <p>.</p>	10 Marks	L3	C01
15.	a.	Explain Apache Mahout and its features.	10 Marks	L2	C02
Or					
16.	a.	Explain architecture of HBase.	10 Marks	L2	C02
17.	a.	Explain operations in Mahout.	10 Marks	L2	C02
Or					
18.	a.	Explain the architecture of HBase.	10 Marks	L2	C02
19.	a.	<p>Reliance wants to save and analyse data using Hbase as it provides the feature of record look up.</p> <p>a. Create a hbase table named Reliance-rollno with column families Customer-data(CustomerId,Age,Profession) and Product Detail(ProductName, Quantity, Amount)</p> <p>b. Insert 3 rows into the table and display the table</p> <p>c. Delete the Age column from row 2 and display the table.</p> <p>d. Delete the complete row 2 and display the table.</p>	10 Marks	L3	C02

		e. Disable the table and check whether it is disabled or not			
Or					
20.	a.	Explian types of Joins in Hive.	10 Marks	L3	C02
21.	a.	<p>SBI bank wants to analyse their customer data stored as csv file using Hive.</p> <p>a. Create a data set with 10 records with fields CustomerId,Name,City(Bangalore, Mysore), Profession</p> <p>b. Create an internal table, load the dataset into this table and display the contents of the table</p> <p>c. Bank wants to analyse the data based on the city so perform partition on dataset based on city and display the partitioned data.</p> <p>d. Write a hive query to display all the rows with city Bangalore.</p>	10 Marks	L3	C02
Or					
22.	a.	Explain bucketing and partitioning in Hive.	10 Marks	L2	C02
23.	a.	Explain architecture of Spark.	10 Marks	L2	C03
Or					
24.	a.	Explain passing functions to Spark.	10 Marks	L2	C03
25.	a.	Explain control structures in Scala with examples.	10 Marks	L3	C03
Or					
26.	a.	Explain Maps and Tuples in Scala.	10 Marks	L3	C03