



# 23CIPC323

## MANGALORE INSTITUTE OF TECHNOLOGY & ENGINEERING

(A Unit of Rajalaxmi Education Trust®, Mangalore)  
Autonomous Institute affiliated to VTU, Belagavi, Approved by AICTE, New Delhi  
Accredited by NAAC with A+ Grade & ISO 9001:2015 Certified Institution

### Model Question Paper

### Sixth Semester BE Degree Examination

### Big Data Analytics

**Time: 3 Hours (180 Minutes)**

**Max. Marks: 100**

*Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.  
2. M: Marks, L: RBT (Revised Bloom's Taxonomy) level, C: Course outcomes.*

Module -1			M	L	C
Q1	a.	Define Big Data. Explain its 3V's characteristics	06	L2	CO1
	b.	Explain the applications of Big Data Analytics in following domain 1. Web Domain. 2. Healthcare	08	L2	CO1
	c.	List & explain the various classification of Big Data Types	06	L2	CO1
<b>OR</b>					
Q2	a.	How do you classify data as structured, semi-structured, unstructured and multi-structured? And give example for each.	06	L2	CO1
	b.	With a neat diagram, explain the client interaction with namenode/datanode to perform read/write operation in HDFS	08	L2	CO1
	c.	Explain HDFS Block Replication	06	L2	CO1
<b>Module- 2</b>					
Q3	a.	With a neat diagram, illustrate the working of Apache Hadoop parallel MapReduce data flow for wordcount example.	08	L3	CO3
	b.	Explain the working of combiner to the map task output in MapReduce	06	L2	CO2
	c.	Discuss fault tolerance and speculative execution in MapReduce	06	L2	CO2
<b>OR</b>					
Q4	a.	With an example, illustrate the working of YARN Hadoop V2 Architecture and compare its advantages over Hadoop Version 1 architecture.	08	L3	CO3
	b.	List & explain MapReduce Advantages	06	L2	CO2
	c.	With a neat diagram, explain Sqoop Import operation	06	L2	CO2
<b>Module - 3</b>					
Q5	a.	What is NoSQL data store ? explain CAP Theorem in Big Data Solutions.	06	L2	CO2
	b.	Explain characteristics of Schema-less Models	06	L2	CO2
	c.	By considering example student database, Illustrate how NoSQL databases provide increased flexibility compared to traditional relational databases	08	L3	CO4
<b>OR</b>					
Q6	a.	Explain the any 6 advantages of Key-Value Pair NoSQL Data Store.	06	L2	CO2
	b.	Explain MongoDB architecture and its advantages.	06	L2	CO2
	c.	With an example illustrate HBase operations: create, insert, scan, get.	08	L3	CO4
<b>Module – 4</b>					
Q7	a.	Explain HIVE architecture and its components.	06	L2	CO2
	b.	Compare HIVE with traditional RDBMS.	06	L3	CO3
	c.	Write a HiveQL statement to Create a table named as student with four columns	08	L3	CO4

		as USN, Name, Dept, Phone and load the data by considering the following text file named as “ <b>ss.txt</b> ” stored in directory /home/user1 <div style="border: 1px solid black; padding: 5px; margin: 5px 0;"> 4MT15CS001, SS1, CSE, 8978987894  4MT15CS002, SS2, CSE, 8978987895  4MT15CS003, SS3, CSE, 8978987896 </div> Write a query to alter the table name to “ <b>CSEStudent</b> ” and third column name to “ <b>Department</b> ”.			
<b>OR</b>					
Q8	a.	With a neat diagram, explain Apache Pig Architecture.	06	L2	CO2
	b.	Compare Apache Pig versus MapReduce.	06	L3	CO3
	c.	Considering the following student.csv file <div style="border: 1px solid black; padding: 5px; margin: 5px 0;"> USN            NAME    DEPT    Phone  4MT15CS001,   SS1,    CSE,    8978987894  4MT15CS002,   SS2,    CSE,    8978987895  4MT15CS003,   SS3,    CSE,    8978987896 </div> a) Write & illustrate the working of the Pig Latin Script Statement to load a student.csv file into relation “A” b) Write & illustrate the working of the Pig Latin “foreach” statement to dump values of first two columns of student.csv file	06	L3	CO4
<b>Module - 5</b>					
Q9	a.	Explain Apache Spark architecture and features.	06	L2	CO3
	b.	With a neat diagram, explain Spark Streaming.	06	L2	CO3
	c.	Write a Spark program for word count using PySpark.	07	L3	CO4
<b>OR</b>					
Q10	a.	With a neat diagram, explain components of Apache Spark	07	L2	CO3
	b.	Explain any two Spark RDD transformation operations.	06	L2	CO3
	c.	Demonstrate analysis of weather data using Spark SQL	07	L3	CO4

\_\*\*\*\*\*\_