

**CMR ENGINEERING COLLEGE: : HYDERABAD**  
**UGC AUTONOMOUS**

**III-B.TECH-II-Semester End Examinations (Supply) - December- 2025**  
**Big Data Analytics**  
**(CSD)**

**[Time: 3 Hours]**

**[Max. Marks: 60]**

**Note:** This question paper contains two parts A and B.

Part A is compulsory which carries 10 marks. Answer all questions in Part A.

Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

**PART-A**

**(10 Marks)**

1. a)	What is the difference between data and Big Data?	[1M]
b)	What do you mean by data explosion?	[1M]
c)	Write about the components of the Hadoop ecosystem.	[1M]
d)	List out the tools used for data discovery in Big Data.	[1M]
e)	What is the output format in MapReduce?	[1M]
f)	Differentiate between HDFS and MapReduce.	[1M]
g)	In HDFS, where is metadata stored?	[1M]
h)	What is the function of MapReduce in Hadoop?	[1M]
i)	What is classification algorithm?	[1M]
j)	List out the machine learning algorithms supported by R.	[1M]

**PART-B**

**(50 Marks)**

2.	Explain how Big Data contributes to business intelligence and decision-making.	[10M]
<b>OR</b>		
3.	Discuss the major drivers that have led to the growth of Big Data.	[10M]
4.	What is predictive analytics? Explain its process and use in decision-making.	[10M]
<b>OR</b>		
5.	Write short notes on the following:	[10M]
(i) Hadoop Distributed File System (HDFS)		
(ii) Apache Spark		
(iii) Mobile Business Intelligence		
6.	What is data serialization? Discuss the role of serialization frameworks.	[10M]
<b>OR</b>		
7.	Describe how input and output are handled in Hadoop MapReduce. Also discuss the Input Format and Output Format.	[10M]
8.	What is Apache Pig? Explain Pig Latin and its use in data processing.	[10M]
<b>OR</b>		
9.	What is the role of HBase in Big Data processing? Compare HBase with HDFS.	[10M]
10.	What is BigR? Explain how it helps in Big Data Analytics.	[10M]
<b>OR</b>		
11.	Write an R script to perform classification using decision trees. Use any sample dataset.	[10M]

\*\*\*\*\*