Code No.: AI732PE

R20 | H.T.No. | | | 8 | R | | | | |

CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS
IV–B.TECH–I–Semester End Examinations (Regular) - November- 2023
BIG DATA ANALYTICS
(CSM)

[Time: 3 Hours]   [Max. Marks: 70]

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 20 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   (20 Marks)

1. a)	Define data preprocessing.	[2M]
b)	What is Tableau data management?	[2M]
c)	What is the full form of Spark?	[2M]
d)	Give a brief note on Hadoop file system.	[2M]
e)	What is the importance of running descriptive statistics?	[2M]
f)	What makes Big Data Analytics different from analyzing a big database?	[2M]
g)	Define Machine Learning.	[2M]
h)	What is the map hypothesis in machine learning?	[2M]
i)	How do you Analyze data Visualization?	[2M]
j)	What is the concept of data visualization?	[2M]

PART-B   (50 Marks)

2. Describe the workplace safety.   [10M]

OR

3. Elaborate the AWS/Rack space in detail.   [10M]

4. List and explain the various operational modes of Hadoop cluster configuration. [10M]

OR

5. Describe the data ETL process.   [10M]

6. Explain the outlier detection and elimination?   [10M]

OR

7. How does Big Data change our analytics organization and architecture? Discuss. [10M]

8. List and explain the perspectives and issues in machine learning. [10M]

OR

9. Describe the Hypotheses Testing and explain various analytical mythologies in [10M]
machine learning.

10. Is QlikView better than Tableau? Justify the statement.   [10M]

OR

11. Describe the QlickView and D3.   [10M]

****