Code No: **R194205E** **R19** **Set No. 1**

IV B.Tech II Semester Supplementary Examinations, May/June – 2024

BIG DATA ANALYTICS

(Common to Computer Science &Engineering and Information Technology)

Time: 3 hours **Max. Marks: 75**

Answer any FIVE Questions
ONE Question from Each unit
All Questions Carry Equal Marks

*****

UNIT I

1	a)	Define big data? Outline various characteristics of a big data platform.	[7]
	b)	What is analytic process and tools? Explain with list of data analytics tools used in the industry.	[8]
		(OR)	
2	a)	Compare five key differences between Analysis and Reporting.	[7]
	b)	Summarize Traditional / RDBMS failed in Big data Analytics.	[8]

UNIT II

3	a)	Outline Stream data model and its Architecture with a neat diagram.	[7]
	b)	Explain the Decaying Window Algorithm with an example.	[8]
		(OR)	
4	a)	Outline Real Time Analytics platform application for Stock Market Predictions.	[7]
	b)	Name different applications of data streams. Explain in detail.	[8]

UNIT III

5	a)	Define Hadoop Distributed File System (HDFS). Compare HDFS and GFS.	[7]
	b)	Explainthe various types of map reduce & its formats.	[8]
		(OR)	
6	a)	Summarize the Java interfaces to HDFS Basics.	[7]
	b)	List the various typesof failures in running a Map Reduce job. Demonstrate the anatomy of Map Reduce program.	[8]

UNIT IV

7	a)	Define HiveQL. Explain all its features.	[7]
	b)	Define Zookeeper and explain all its features with applications.	[8]
		(OR)	
8	a)	Illustrate distributed modes of running PIG Scripts.	[7]
	b)	Summarize the funList and HIVE services. How will you query the data in HIVE discuss with suitable example?	[8]

UNIT V

9	a)	Explain the importance of Regression in Data Science and Data Analytics.	[7]
	b)	Explain about Predictive Analysis.	[8]
		(OR)	
10	a)	Explain Multiple Linear Regressions.	[7]
	b)	Explain Visual data analysis techniques.	[8]