PRN:		Course Code:	AMT 1602

Third Year BTECH (AIML) End Semester Examination Summer 2023

Big Data Analytics

Day and Date: Thursday, 15/06/23 Max Marks: **100**

Time: 10:30 - 1:30 pm Duration: **03** hrs

Instructions: All Questions are compulsory

			Marks	Bloom's Level	CO
Q.1		**Attempt the following.**			
	a)	Define Big Data. Explain various characteristics of Big Data	08	L1	CO1
	b)	With a neat diagram, Explain the Hadoop Ecosystem in detail.	08	L2	CO1
		OR			
	b)	Explain Hadoop Architecture in detail.	08	L2	CO1
Q.2		**Attempt the following.**			
	a)	Illustrate data serialization in Hadoop.	08	L2	CO2
	b)	What is the role of Data node and Name node in HDFS. Give commands with appropriate arguments to perform data transfer between local file system and HDFS.	08	L1	CO1
		OR			
	b)	How data ingest with Apache Sqoop in Hadoop.?	08	L1	CO1
Q.3		**Attempt any two from the following.**			
	a)	Build word count Map Reduce program for counting occurrences of specific words in the input text file(s). Also write the commands to compile and run the code.	08	L3	CO2
	b)	Classify and explain the different types of schedulers in MapReduce.	08	L4	CO2
	c)	Explain Map-reduce framework in brief.	08	L2	CO2
Q.4		**Attempt any two from the following.**			
	a)	Summarize the need of Apache Pig and explain its features.	08	L2	CO3
	b)	Illustrate the architecture of Pig with neat diagram.	08	L2	CO3
	c)	List and explain different Pig commands.	08	L4	CO3
Q.5		**Attempt any Two of the following**			
	a)	Explain various datatypes and schemas supported by HiveQL	09	L2	CO3
	b)	How the Hive Data Manipulation Language works?	09	L1	CO3
	c)	Compare Pig with Hive.	09	L5	CO3
Q.6		**Attempt any Two of the following**			
	a)	Compare supervised learning with unsupervised learning	09	L5	CO4
	b)	Build and explain Bloom filter for filter on stream data.	09	L3	CO4
	c)	Build and explain Flajolet-Martin algorithm for counting distinct elements in stream data.	09	L3	CO4

*****

ESE

SGU Library Copy