CBCS SCHEME

USN | | | | | | | | | | |

22SCS21

Second Semester M.Tech. Degree Examination, June/July 2023
Big Data Analytics

Time: 3 hrs. Max. Marks: 100

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

		Module – 1	M	L	C
Q.1	a.	Explain MapReduce data flow with multiple tasks with a diagram.	10	L1	CO1
	b.	Briefly explain the following HDFS concepts: (i) Blocks (ii) Name Node and Data Nodes (iii) HDFS High Availability	10	L1	CO1
		OR			
Q.2	a.	Explain the Hadoop File system and Basic File system operations commands.	10	L1	CO1
	b.	How a client writing data to HDFS with a diagram.	10	L1	CO1
		Module – 2			
Q.3	a.	Explain Anatomy of a YARN Application Run with a neat diagram.	10	L1	CO2
	b.	Explain scheduling in YARN with a neat diagram.	10	L1	CO2
		OR			
Q.4	a.	Explain Data Integrity and Compression in HDFS.	10	L1	CO2
	b.	With a neat diagram, explain sequence file format with record compression and block compression.	10	L1	CO2
		Module – 3			
Q.5	a.	Explain (i) Types of Hadoop Logs (ii) Tuning a Job checklist	10	L2	CO3
	b.	Briefly discuss about Apache Oozie and program for Oozie work flow definition to run the maximum temperature MapReduce Job.	10	L2	CO3
		OR			
Q.6	a.	Explain the Anatomy of a MapReduce Job RUN with a neat diagram.	10	L2	CO3
	b.	How status updates are propagated through the MapReduce system with a neat diagram.	10	L2	CO3
		Module – 4			
Q.7	a.	Explain the Input Format class hierarchy with a neat diagram.	10	L3	CO4
	b.	Explain the Output Format class hierarchy with a neat diagram.	10	L3	CO4
		OR			
Q.8	a.	With a neat diagram, explain how FLUME Agent with a spooling directory source and a logger sink connected by a file channel.	10	L3	CO4
	b.	Describe how Load balancing between two Agents in FLUME with a diagram.	10	L3	CO4
		Module – 5			
Q.9	a.	Discuss about PIG. Explain the two execution modes of PIG and three ways of executing PIG program.	10	L3	CO5
	b.	Explain Loading, Storing, Grouping, Joining and Splitting Data in PIG.	10	L3	CO5
		OR			
Q.10	a.	Discuss about SPARK. How SPARK Runs a Job with a diagram.	10	L3	CO5
	b.	Explain SPARK on YARN client mode with a diagram.	10	L3	CO5

* * * * *