Register Number:……..

KUMARAGURU
college of technology
character is life

B.TECH DEGREE EXAMINATIONS: APRIL / MAY 2023

(Regulation 2018)

Sixth Semester

INFORMATION TECHNOLOGY

U18ITI6304: Big Data Analytics

COURSE OUTCOMES

CO1: Outline the big data technologies used for storage, analysis and manipulation of data

CO2: Explain Big Data eco system and its components

CO3: Analyze the Big Data stored in HDFS using Hadoop Map Reduce framework

CO4: Understand the Pig scripting and HBase architecture

CO5: Apply the Hive concepts, Hive Data types, loading and querying for Big Data

CO6: Explain the MongoDB architecture and its operations

Time: Three Hours **Maximum Marks: 100**

Answer all the Questions: -
PART A (10 x 2 = 20 Marks)
(Answer not more than 40 words)

1.	What are the benefits offered by big data to an organization in increasing its value?	CO1	[K_2]
2.	Mention the contents of the metadata maintained by the name node.	CO1	[K_2]
3.	Suppose there is a file of size 516 MB stored in HDFS (Hadoop 2. x) using default block size configuration and default replication factor. How many blocks will be created in total and what will be the size of each block?	CO2	[K_2]
4.	What is the role of the job tracker?	CO2	[K_2]
5.	Difference between Name node and Data node.	CO3	[K_2]
6.	How is the limitation of MapReduce overcome in a future version of Hadoop?	CO3	[K_2]
7.	List out the five basic operations of the Map reduce programming Model.	CO3	[K_1]
8.	Does Pig differ from Hbase? Justify it.	CO4	[K_2]
9.	Mention the data types in Hive.	CO5	[K_1]
10.	What is the use of object identifiers in MongoDB?	CO6	[K_2]

Answer any FIVE Questions:-
PART B (5 x 16 = 80 Marks)
(Answer not more than 400 words)

11.	a)	Discuss the use of Big Data Analytics in Business with suitable real-world example.	8	CO1	
	b)	Write a short note on the Hadoop ecosystem.	8	CO1	[K_2]
12.		Explain briefly about Map Reduce data flow with map reduce tasks with neat diagram. We have a text file with data Orange Apple, grapes, Orange, Orange, Pineapple, Pears, Jackfruit and Pear. Perform the operation of a word count on the sample.txt using Map Reduce.	16	CO2	[K_3]
13.	a)	Describe about how-to setup single node Hadoop cluster and installation and discuss about Hadoop configuration files.	8	CO3	[K_2]
	b)	Briefly discuss the need of HDFS federation in block management.	8	CO3	[K_2]
14.	a)	Discuss in detail about the HBase architecture and component with neat diagram.	8	CO4	[K_2]
	b)	i) Elaborate the two modes used for running the Pig scripts?	4	CO4	[K_3]
		ii) Illustrate the use of the FILTER and DISTINCT operator in Pig Latin with relevant example.	4		
15.	a)	i) Write the Hive command to create a table with columns: Employee id, Name, Salary, Designation.	2	CO5	[K_3]
		ii) Assume employee table as given below, with the fields named Id, Name, Salary, Designation, and Dept. -Generate a query to retrieve the employee details in order by using Department name. -Generate a query to retrieve the number of employees in each department. -A query is used to retrieve employee details whose Department is TP and Salary is more than Rs 40000.	6		
	b)	Outline Hive architecture and its major components with neat sketch.	8	CO5	[K_2]
16.	a)	Elaborate about the sharding operation in MongoDB with relevant diagrams and examples.	10	CO6	[K_2]
	b)	Explain the basic operations of MongoDB with relevant examples.	6	CO6	[K_2]

****