Register Number:……..

KUMARAGURU
college of technology
character is life

B.E DEGREE EXAMINATIONS: APRIL / MAY 2023

(Regulation 2018)

Sixth Semester

COMPUTER SCIENCE AND ENGINEERING

U18CSE0001**:** Big Data Technologies

COURSE OUTCOMES

CO1: Identify the components of Hadoop Distributed File System for big data processing [K4]
CO2: Develop Big Data Solutions using Hadoop Eco System[K3]
CO3: Examine various framework in Big data Processing [K4]
CO4: Illustrate the big data security issues with Hadoop and the need of AWS for Hadoop environment. [K3]

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

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

1.	Define Big data Analytics.	CO1	[K_1]
2.	Compare structured, semi-structured, unstructured data.	CO1	[K_2]
3.	Differentiate between Hadoop and RDBMS.	CO2	[K_2]
4.	Suppose there is a file having size of 514MB is stored in the Hadoop (Hadoop 2.x) by using the default size-configuration of block and by default replication-factor. Then, how many blocks will be created in total and what will be the size of each block?	CO2	[K_2]
5.	What are the components of Apache HBase's Region Server?	CO2	[K_1]
6.	Differentiate SORT BY and ORDER BY in Hive.	CO2	[K_1]
7.	List some advantages of Content-Based Recommendation paradigm over Collaborative-Based Recommendation.	CO3	[K_1]
8.	Distinguish implicit and explicit ratings.	CO3	[K_2]
9.	What are the different file permissions in HDFS for files or directory levels?	CO4	[K_2]
10.	How to achieve security in Hadoop?	CO4	[K_2]

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

11.	a)	Explain the characteristics of big data and formulate how big data analytics helps business people to increase their revenue with any one real time example.	8	CO1	[K_2]
	b)	Explain the steps to be followed to deploy a Big Data solution.	8	CO1	[K_2]

12.	a)	Explain the core components of Hadoop.	8	CO2	[K_2]
	b)	Mention the five basic operations of Map Reduce programming model and explain the map reduce implementation with suitable example.	8	CO2	[K_2]
13.	a)	Explain in detail about the writing process of HDFS Multiple block pipeline using the Rack Awareness Algorithm.	8	CO2	[K_2]
	b)	Explain HBASE architecture and its data model.	8	CO2	[K_2]
14.	a)	Illustrate the architecture of Hive and explain about its features.	8	CO3	[K_2]
	b)	Explain Pig data Model and discuss how it will help for effective data flow.	8	CO3	[K_2]
15.	a)	Identify a type of recommendation system for an online bookstore that has been launched recently. The bookstore has over 1 million book titles, but its rating database has only 10,000 ratings. Which of the following would be a better recommendation system and justify and elaborate the techniques used in the same.? a) User-user collaborative filtering b) Item-item collaborative filtering c) Content-based recommendation.	8	CO3	[K_3]
	b)	Explain in detail about the knowledge-based recommendation systems.	8	CO3	[K_2]
16.	a)	Explain about the following security concepts in terms of Hadoop, • HDFS file ownership and permissions • Enhanced security with Kerberos • Encrypted HDFS data transfers • HDFS data at rest encryption • Encrypted HTTP traffic	8	CO4	[K_2]
	b)	Explain in detail about running Hadoop using AWS components.	8	CO4	[K_2]

****