Register Number:……..

B.E DEGREE EXAMINATIONS: NOV /DEC 2024

(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
CO2: Develop Big Data Solutions using Hadoop Eco System
CO3: Examine various framework in Big data Processing
CO4: Illustrate the big data security issues with Hadoop and the need of AWS for Hadoop

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

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

1.	Summarize how digital data is categorized based on its types.	CO1	[K_2]
2.	Identify advantages of Hadoop compared to traditional platforms.	CO1	[K_3]
3.	List any two Hadoop shell commands and describe their functionalities.	CO2	[K_2]
4.	Outline any two tools available for cluster monitoring.	CO2	[K_2]
5.	What are the different execution modes of PIG?	CO2	[K_2]
6.	List the advantages and limitations of using Hive for querying.	CO2	[K_1]
7.	What is the purpose of the Hive Meta-store?	CO3	[K_2]
8.	Recall the basic principle behind collaborative filtering.	CO3	[K_2]
9.	What is authorization, and how does it differ from authentication?	CO4	[K_2]
10.	Differentiate between security vulnerabilities and security threats, providing examples of each.	CO4	[K_2]

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

11.	a)	Analyze the difference between parallel and distributed computing architectures in the context of Big Data processing.	8	CO1	[K_3]
	b)	Describe the impact of the five V's on big data analytics.	8	CO1	[K_3]
12.	a)	Explain how MapReduce works to distribute and process data across a Hadoop.	8	CO2	[K_2]

		cluster.			
	b)	Define the roles of Name Node, Secondary Name Node and Data Node in the Hadoop architecture.	8	CO2	[K_2]
13	a)	Discuss the key components of the Hadoop ecosystem and their roles in distributed computing.	8	CO2	[K_2]
	b)	Explain the importance of SSH in Hadoop cluster configuration.		CO2	[K_2]
14.	a)	Explain the role of the Hive Shell in interacting with Hive services.	8	CO3	[K_2]
	b)	Design a workflow using Oozie to orchestrate Pig and Hive jobs for a data processing pipeline.	8	CO3	[K_3]
15.	a)	Design a content-based recommendation system for recommending news articles based on their textual content.	8	CO3	[K_3]
	b)	How does knowledge-based recommendation differ from collaborative filtering and content-based approaches?	8	CO3	[K_2]
16.	a)	Explain different types of authentication methods used in IT environments.	8	CO4	[K_2]
	b)	Describe the AWS services typically employed for operating Hadoop clusters.	8	CO4	[K_2]

****