Velagapudi Ramakrishna Siddhartha Engineering College::Vijayawada
(Autonomous)

VR20

III /IV B Tech Degree Examinations
Sixth Semester (CBCS)
Department of Information Technology
20IT6404B BIG DATA

Time:3Hrs	MODEL QUESTION PAPER	Max Marks:70

Part – A is Compulsory
Answer one (01) question from each unit of Part – B
Answers to any single question or its part shall be written at one place only

Cognitive Levels(K): K1-Remember;K2-Understand; K3-Apply; K4-Analyze; K5-Evaluate; K6-Create

Q. No		Question	Marks	Course Outcome	Cog. Level
Part – A			10X1=10M		
1	a	List out different types of Data in Big data.	1	CO1	K2
	b	RDBMS versus Hadoop	1	CO1	K2
	c	List the types of NoSQL databases.	1	CO2	K2
	d	Suppose there is file of size 514 MB stored in HDFS (Hadoop 2.x) using default block size configuration and default replication factor. Then, how many blocks will be created in total and what will be the size of each block	1	CO2	K4
	e	Write the data types in MongoDB	1	CO2	K2
	f	Describe Jobtracker and Tasktracker	1	CO3	K2
	g	What is FUSE	1	CO3	K2
	h	Write the advantages of Pig over MapReduce	1	CO3	K4
		Define RDD.		CO4	K2
	I	What is HIVE.	1	CO4	K2
	j	Write the features of Spark	1	CO4	K4
Part – B			4X15 =60M		
UNIT – I					
2	a	Discuss about the three characteristics defined in Big Data in detail	8	CO1	K2
	b	Contrast between traditional fraud detection patterns and modern-day fraud detection with a neat sketch in fraud detection patterns for big data deployment	7	CO1	K4
(OR)					
3	a	Discuss the history of Hadoop and its ecosystem. List out Hadoop latest release.	8	CO1	K2
	b	Explain the solution taken to overcome the problem for data storage and analysis in Hadoop.	7	CO1	K3
UNIT – II					
4	a	What is HDFS. List out two types of nodes operating in HDS cluster.	7	CO2	K1

	b	Discuss the term commodity hardware in design of HDFS.	8	CO2	K3
		(OR)			
5	a	a. Create the mongoDB collection by the name food and then insert documents into food collection. Each document should have a fruits array. (i) Find those documents from the food collection which have the fruits array having grape in the first index position (ii) Find those documents from the food collection where the size of the array is three. (iii) Find those documents from the food collection which have the element orange in the 0 index position in the array fruits.	8	CO2	K5
	b	Describe the advantages of NoSQL and Compare SQL versus NoSQL 7M	7	CO2	K4
		UNIT – III			
6	a	Explain about Hadoop streaming and pipes.	7	CO3	K2
	b	Explain MapReduce dataflow with single, multiple and no resuce tasks with a neat sketch.	8	CO3	K3
		(OR)			
7	a	Demonstrate pig latin structure and schema with an example. 7M	7	CO3	K4
	b	List out the functions of pig and give description of each function in a neat table.	8	CO3	K2
		UNIT – IV			
8	a	List out different types of Joins in Hive and explain with an example.	7M	CO4	K3
	b	Elaborate the procedure to create and manage tables in Hive with example .	8M	CO4	K3
		(OR)			
9	a	Describe the operations of RDDs and Implement word count program using Spark RDDs.	8M	CO4	K4

	b	Design the scheme and analysis using Spark SQL for any usecase.	7M	CO4	K3

Designation	Name in Capitals	Signature with Date
Course Coordinators	Dr. Shaik Fathimabi	
Module Coordinator	Y. Sandeep	
Program Coordinator	Dr.G.Kalyani	
Head of the Department	Dr.M.Suneetha	