Reg. No. : | | | | | | | | | | |

Question Paper Code: 95R04

Ph.D. COURSE WORK EXAMINATION, NOV 2019

Elective

Computer Science and Engineering

19PCS504 - BIG DATA ANALYTICS

(Regulation 2019)

Duration: Three hours | Maximum: 100 Marks

Answer ALL Questions

PART - A (10 x 2 = 20 Marks)

1.	Define Fraud.	CO1- U	(2)
2.	What are the industry applications of big data?	CO1- U	(2)
3.	Define key-value storage.	CO2- U	(2)
4.	What is Schema-less database?	CO2- U	(2)
5.	How mapReduce phase works in hadoop data analysis?	CO3- U	(2)
6.	What are the benefits of block?	CO3- U	(2)
7.	What are the steps in testing the mappers?	CO4- U	(2)
8.	List the types of job scheduler.	CO4- U	(2)
9.	Write a note on the use of Zookeeper?	CO5- U	(2)
10.	List any four complex data types of Pig.	CO5- U	(2)

PART - B (5 x 16 = 80 Marks)

11.	(a)	Discuss Industry Examples of Big data in detail	CO1- U	(16)
		Or		
	(b)	Explain Near Real time event processing frame work for fraud detection with the help of neat diagram?	CO1- U	(16)
12.	(a)	Discuss briefly about key-value and document data model.	CO2- U	(16)
		Or		
	(b)	Discuss Master slave and peer-peer replication in in detail.	CO2- U	(16)

13.	(a)	Discuss serialization in detail.	CO3- U	(16)
		Or		
	(b)	Discuss Avro in detail.	CO3- U	(16)
14.	(a)	Discuss Map reduce job scheduling in detail with neat diagram.	CO4- U	(16)
		Or		
	(b)	Discuss YARN (Mapreduce 2) in detail with failures in classic Map-reduce.	CO4- U	(16)
15.	(a)	Explain Pig data Model in detail and Discuss how it will help for effective data flow.	CO5- U	(16)
		Or		
	(b)	Explain Cassandra data model in detail and discuss how it will help for effective data flow.	CO5- U	(16)