Reg. No.																

Question Paper Code	13986

B.E. / B.Tech. - DEGREE EXAMINATIONS, NOV / DEC 2025

Fifth Semester

Information Technology

20ITPC502 - BIG DATA ESSENTIALS

Regulations - 2020

Duration: 3 Hours | Max. Marks: 100

PART - A (MCQ) (10 × 1 = 10 Marks)

Answer ALL Questions

		Marks	K – Level	CO
1.	Which industry is known for using Big Data in algorithmic trading? (a) Agriculture (b) Retail (c) Healthcare (d) Finance	1	K1	CO1
2.	Fraud detection using Big Data is most effective due to: (a) Limited data sources (b) Real-time analysis capabilities (c) Small data volume (d) Predictable patterns	1	K1	CO1
3.	What is a key characteristic of Hadoop Streaming? (a) Real-time processing (b) Use of MapReduce framework (c) Batch processing only (d) Integration with SQL databases	1	K1	CO2
4.	Who is credited with the creation of Hadoop? (a) Jeff Dean and Sanjay Ghemawat (b) Doug Cutting and Mike Cafarella (c) Larry Page and Sergey Brin (d) Brendan Eich and Marc Andreessen	1	K1	CO2
5.	In Hadoop YARN, which component is responsible for allocating resources and scheduling tasks across nodes in the cluster? (a) NameNode (b) JobTracker (c) ResourceManager (d) DataNode	1	K1	CO3
6.	Which of the following types of failures can occur in a MapReduce job? (a) Task failure (b) Node failure (c) Application failure (d) All of the above	1	K1	CO3
7.	How many execution modes does Pig support? (a) One (b) Two (c) Three (d) Four	1	K1	CO4
8.	In which scenario would you choose Pig over traditional databases? (a) When real-time transactions are required (b) When handling unstructured data (c) When ACID properties are important (d) When low-latency access is needed	1	K2	CO4
9.	Which shell is used to run Spark interactively? (a) Python shell (b) Scala shell (c) Spark shell (d) R shell	1	K1	CO5
10.	Which function is used to persist an RDD in memory? (a) cache() (b) persist() (c) store( ) (d) hold()	1	K1	CO5

PART - B (12 × 2 = 24 Marks)

Answer ALL Questions

		Marks	K – Level	CO
11.	Is big data good for marketers? Justify your answer.	2	K2	CO1
12.	State Unstructured data.	2	K1	CO1
13.	Define Hadoop.	2	K1	CO2
14.	What is an Avro? List its advantages.	2	K1	CO2
15.	Give the components of Map-Reduce Architecture.	2	K1	CO3
16.	What are all the benefits of Hadoop MapReduce?	2	K1	CO3
17.	List few HDFS commands that is used in Pig Grunt.	2	K2	CO4
18.	Define Hive QL.	2	K1	CO4
19.	List the features of Spark.	2	K1	CO5
20.	What is the role of driver and executor in Spark architecture?	2	K2	CO5
21.	Differentiate GPU with CPU.	2	K2	CO6

22.		State the uses of cudaMemcpy() function.	*2*	*K1*	*CO6*

PART - C (6 × 11 = 66 Marks)

Answer ALL Questions

23.	a) (i)	Explain the different types of Big Data technologies with examples.	*6*	*K2*	*CO1*
	(ii)	Write the significant impact of adding Big Data in Healthcare.	*5*	*K2*	*CO1*
		OR			
	b)	Discuss in detail about the emerging big data ecosystem and new approach to Analytics.	*11*	*K2*	*CO1*
24.	a)	Outline the steps to be followed by the File Based Data structure in detail.	*11*	*K2*	*CO2*
		OR			
	b)	Explain briefly about Data Ingest with Flume and Scoop with suitable example.	*11*	*K2*	*CO2*
25.	a)	Write short notes on (i) Streaming information access and Low latency information access. (ii) Rest and thrift.	*11*	*K2*	*CO3*
		OR			
	b)	Interpret Job Scheduling and task execution in YARN in detail with suitable example.	*11*	*K2*	*CO3*
26.	a)	Explain in details about Apache PIG architecture and PIG Latin data model with necessary diagrams.	*11*	*K2*	*CO4*
		OR			
	b)	Write short notes on: (i) HIVE Metastore (ii) HIVE QL and HIVE Tables	*11*	*K2*	*CO4*
27.	a)	Discuss about how Spark integrates with Hadoop and runs on YARN with suitable example.	*11*	*K3*	*CO5*
		OR			
	b)	Illustrate the execution process of a Spark job, including job stages and tasks with an example.	*11*	*K3*	*CO5*
28.	a)	Explain in detail about GPU Architecture. Compare and contrast the features GPU and CPU with suitable example.	*11*	*K3*	*CO6*
		OR			
	b)	Illustrate in detail about CUDA Programming Model with suitable example.	*11*	*K3*	*CO6*