

**B.E. / B.Tech. - DEGREE EXAMINATIONS, NOV/DEC 2022**

Fifth Semester

**Information Technology**

**20ITPC502 - BIG DATA ESSENTIALS**

(Regulations 2020)

Duration: 3 Hours

Max. Marks: 100

**PART - A (10 × 2 = 20 Marks)**

Answer ALL Questions

	<i>Marks, K-Level, CO</i>
1. List out the best practices of Big Data Analytics.	2,K1,CO1
2. What is V3?	2,K1,CO1
3. What is Hadoop?	2,K1,CO2
4. What is HDFS?	2,K1,CO2
5. Define Map Reduce Framework.	2,K1,CO3
6. Write down the two major functionalities of Hadoop YARN.	2,K1,CO3
7. List out two execution environments to run a Pig Latin Programs.	2,K1,CO4
8. What is Hive?	2,K1,CO4
9. Define GPU Computing.	2,K1,CO5
10. What is Spark Shell?	2,K1,CO5

**PART - B (5 × 13 = 65 Marks)**

Answer ALL Questions

11. a) (i) Write the characteristics and Structure of Big Data. 6, K1,CO1  
(ii) Describe evolution of Big Data. 7,K1,CO1
- OR**
- b) Discuss the Case Study: Big Data and Healthcare, Big Data and Medicine. 13,K1,CO1
12. a) What is Hadoop? Explain its architecture with components. 13,K2,CO2
- OR**
- b) Explain in detail about HDFS. 13,K2,CO2
13. a) What is MapReduce Concept? Consider you have following input data for your MapReduce in Big data Program. 13,K2,CO3  
**Welcome to the Hadoop world**  
**Hadoop is Good**  
**Hadoop is Bad**  
Generate the Step by Step Task of MapReduce. With neat diagram show the output.



**OR**

b) How YARN runs an application. Explain in detail. 13,K2,CO3

14. a) Write short notes on:

(i) User Defined Functions 7,K2,CO4

(ii) Data Processing Operators in Pig Latin. 6,K2,CO4

**OR**

b) (i) Draw the Architecture of Hive. 7,K2,CO4

(ii) What is HBase? Discuss HBase Basics & Concept Model. 6,K2,CO4

15. a) Elaborate CUDA Programming model. Explain CUDA Memory model. 13,K2,CO6

**OR**

b) Explain in detail about the steps of converting the Vector addition program to CUDA. 13,K2,CO6

**PART - C (1 × 15 = 15 Marks)**

16. a) Discuss the following

(i) Data Analysis with Spark Shell. 8,K2,CO5

(ii) Writing Spark Application. 7,K2,CO5

**OR**

b) (i) Write a Spark API to count the number of words in the given text file. 7,K2,CO5

(ii) Write a Spark API Program to display the error message from the log file. 8,K2,CO5