

Roll No.

Total Pages : 3

46215

BT-6/M-23

BIG DATA ANALYTICS

Paper : PE-IT-S306A

Time : Three Hours]

[Maximum Marks : 75

Note: Attempt *five* questions in all, selecting at least *one* question from each unit. All questions carry equal marks.

UNIT-I

1. (a) Define Big Data. Explain the Evolution of Big Data and their characteristics. (7)
- (b) What is structured, semi-structured, unstructured data? Write examples. (8)
2. (a) Illustrate the Hadoop core components with neat diagram. (7)
- (b) Discuss the Hadoop system and ecosystem components in four layers. (8)

UNIT-II

3. (a) Describe the Map Reduce execution steps with neat diagram. (7)

46215/50/KD/1102

146 [P.T.O.
20/6

- (b) How google file system differs from the hadoop file system and explains the google file system architecture with a neat sketch. (8)
4. (a) What is Map Reduce? Write a map reduce program for weather. Draw the Diagram. (7)
- (b) Write Short notes on the following :
- (i) Driver code.
 - (ii) Mapper code.
 - (iii) Reducer code.
 - (iv) Combiner. (2×4=8)

UNIT-III

5. (a) Define HDFS. Explain in brief about the basic building blocks of Hadoop. (7)
- (b) In Map-Reduce Implementation of Word count, Write Types of keys and Values mapper, reducer, combiner (if any) and which of the types need to Writable? (8)
6. (a) Explain the uses of Name node, Data node and Secondary Name node in Hadoop Distributed File system. Draw the Architecture. (7)
- (b) What are the different modes in which Hadoop can be installed and what is the use of each mode from application and developer point of view? (8)

UNIT-IV

7. (a) Write the differences between HDFS high availability and HDFS federation. (7)
- (b) What is Apache Hadoop YARN? What are the additional benefits YARN brings in to Hadoop 2.0 ? (8)
8. (a) Explain the differences between Hadoop 1.x and Hadoop 2.x. (7)
- (b) Explain the Schedulers in the Hadoop Ecosystem. Justify all Schedulers types with example. (8)