

Roll No.

Total Pages : 3

46292

BT-6/M-23

BIG DATA ANALYTICS

Paper : PC-CS-AIDS-306A

Time : Three Hours]

[Maximum Marks : 75

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

UNIT-I

1. (a) What do you mean by big data? How is it different from traditional data? (7)
(b) Why do we need a big data analytics in the business? What is the role of drivers in big data? Justify with example. (8)
2. (a) What do you understand by .Algorithm using Map Reduce? How pig is used for this purpose? Explain. (7)
(b) Write short note on 4 Vs in Big Data and explain the Matrix Vector Multiplication by Map Reduce. (8)

UNIT-II

3. (a) What do you mean by Apache Hadoop Ecosystem? Draw the Architecture of Hadoop. (7)

46292/100/KD/1098

46 [P.T.O.
20/6

- (b) What are the major configuration parameters required in a Map Reduce program? Justify. (8)
4. (a) Who takes care of replication consistency in a Hadoop cluster and what do under/over replicated blocks mean? Draw the Architecture of Hadoop Cluster. (7)
- (b) If you have an input file of 350 MB, how many input splits would HDFS create and what would be the size of each input split? Explain with Diagram. (8)

UNIT-III

5. (a) What is Map reduce programming Model ? Where is it used and why? (7)
- (b) Explain Map Reduce Word Count Process. Suppose HDFS Input file is Panipat Institute is affiliated by Kurkshetra University. Draw the diagram (8)
6. (a) Explain the architecture of HDFS. Justify all Daemons of HDFS. (7)
- (b) Write short notes :
- (i) Mapper.
 - (ii) Reducer.
 - (iii) Secondary Name Node.
 - (iv) Metadata. (2*4=8)

UNIT-IV

7. (a) What is Pig ? Explain architecture of Apache Pig. (7)
- (b) What do you mean by the Schedulers in YARN? Draw the Architecture of YARN. (8)

8. (a) Write Five Differences between Hadoop 1.0 and Hadoop 2.0. Explain MRV1 and MRV2 in terms of Hadoop Version. (7)
- (b) What is the Flume? Explain the Architecture of Flume. Explain the role of Data Channel in the Flume. Justify. (8)
-