

Roll No.

Total No. of Pages : 02

Total No. of Questions : 08

M.Tech.(CSE Engg.)EI-III (2015 to 2017)/(IT)E3 (2015 & Onwards)(Sem.-3)

BIG DATA ANALYTSIS

Subject Code : MTCS-302

Paper ID : [74152]

Time : 3 Hrs.

Max. Marks : 100

INSTRUCTION TO CANDIDATES :

1. Attempt any FIVE questions out of EIGHT questions.
2. Each question carries TWENTY marks.

Q1. a) What is Hadoop Ecosystem? Explain the role of various components in Hadoop Ecosystem.

b) Explain the role of HBase in big data processing with the help of suitable example.

Q2. a) How do we set a configuration property to be unique/constant across the cluster nodes and no slave nodes should override this? Write pseudo code for this implementation.

b) How MapReduce execution can be optimized with combiner? Explain with the help of suitable example.

Q3. a) How does the “sort” operation help in MapReduce? Discuss with the help of suitable example.

b) Explain the difference between NameNode, Backup Node and Checkpoint name node.

Q4. a) How do you define “Block” in HDFS? What is the block size in Hadoop 1 and in Hadoop 2? Can it be changed? What is its significance?

b) State the reason why we can't perform “aggregation” (addition) in a mapper? Why do we need the “Reducer” for this?

Q5. a) How missing values and impossible values are represented in R language? What will be the output of expression `all(NA==NA)`?

b) Explain with the example how to obtain the transpose of a matrix in R. what is the difference between `subset()` and `sample()` function.

Q6. a) How YARN handles the functionalities of JobTracker and TaskTracker? Can YARN run application that do not follow MapReduce model?

- b) Write the code to initialize and start a YarnClient. After that create an application and get its application id.
- Q7. How Hadoop's CLASSPATH plays vital role in starting or stopping in Hadoop daemons. Is there any impact on MapReduce jobs if there is no mapred-site.xml file created in HADOOP_HOME/conf directory but all the necessary properties are defined in yarn-site.xml?
- Q8. a) Explain the handling of data in R Workspace. What are the different data objects in R? Which data object in R is used to store and process categorical data?
- b) Explain various types of data visualization used in R. How graphical analysis is performed in R?

www.FirstRanker.com