Registration No: | | | | | | | | | | |

Total Number of Pages: 02 **Course: B.Tech**
Sub_Code: RIT7D002

7th Semester Regular/Back Examination: 2024-25
SUBJECT: Big Data Analytics
BRANCH(S): CSEAI, CSE, EEE, ELECTRICAL, ECE, ETC
Time: 3 Hours
Max Marks: 100
Q.Code: R357

Answer Question No.1 (Part-1) which is compulsory, any eight from Part-II and any two from Part-III.
The figures in the right-hand margin indicate marks.

Part-I

Q1 **Answer the following questions:** **(2 x 10)**

a) What is Big Data?
b) What are the characteristics of big data?
c) What is HDFS?
d) What is YARN?
e) Define the various file formats supported by HIVE.
f) Mention different ways of data discovery in data analytics.
g) What is mapreduce?
h) What is HIVE?
i) Which are the application areas of Big Data analytics.
j) State the common input formats in Hadoop.

Part-II

Q2 **Only Focused-Short Answer Type Questions- (Answer Any Eight out of Twelve)** **(6 × 8)**

a) Give a detail note on HBASE.
b) What are the three V's of Big Data system with examples?
c) What is benchmarking how it works in Hadoop?
d) Explain how map reduce jobs run on YARN.
e) Discuss Data Analytics Lifecycle with examples.
f) How does security is done in Hadoop? Justify.
g) Explain about Big Data Computational Limitations.
h) Discuss Hadoop 2 architecture.
i) Explain the map reduce data flow with single reduce and multiple reduce.
j) Discuss the NoSQL Data Architectural Patterns in Big-Data System.

k) What are the various operational modes of Hadoop cluster configuration? Explain in detail about configuring/installing Hadoop in fully distributed mode.

l) What is the difference between the 'set' & a 'map' data structure and explain about various map implementations in java with suitable examples?

Part-III

Only Long Answer Type Questions (Answer Any Two out of Four)

Q3 Explain the Hadoop distributed file system architecture with a neat sketch. How google file system differs from the Hadoop file system? Explains the google file system architecture with a neat sketch. **(16)**

Q4 Explain about k-means Clustering in detail. **(16)**

Q5 Briefly discuss about MapReduce and YARN. **(16)**

Q6 How to Build a Real-Time Twitter Analysis Application Using Big Data Tools? **(16)**