

300818(22)

BE 8th Semester (New Scheme)

Examination April-May 2020

Branch: Computer Science & Engineering

Big Data & Hadoop

Time Allowed : Three Hours

Maximum Marks : 80

Passing Marks : 28

Note : Answer all questions. Part (a) of 2 marks is compulsory. Attempt any two parts from (b), (c) and (d) which are of 7 marks each.

UNIT-1

- (a) What is Big Data? Discuss it in terms of 4Vs.
- (b) Explain the different categories of Big Data Analytics? How Conventional Data Analytics Systems differ from Big Data Systems.
- (c) Differentiate between Reporting and Analysis.
- (d) What are the different features provided by big data platforms? List four big data platforms.

UNIT-2

- (a) What do you mean by filtering?
- (b) What is the role of Stream Computing in Big Data Analytics? Explain Data Stream Management System in detail.
- (c) What is the of DGIM algorithm? Explain the algorithm with suitable example.
- (d) What is the role of Decaying Windows in Stream Data Analysis? Explain Sampling in Data Streams along with its types.

UNIT-3

- (a) What is the default block size in Hadoop 1 and in Hadoop 2 ?
- (b) Using suitable block diagram explain the anatomy of Map Reduce Job Run.
- (c) Describe the following in terms of HDFS:
 - i. Name Node and Data Node
 - ii. Blocks and Input Splits
 - iii. Hadoop Streaming
 - iv. Hadoop Pipes
 - v. Data Locality Optimization
- (d) What do you mean by HDFS federation, failover and fencing?

UNIT-4

- (a) Define Hadoop Cluster.
- (b) Write detail to setup and configure a Hadoop Cluster.
- (c) For enabling security in Hadoop, elaborate:
 - i. Kerberos
 - ii. Delegation Token access
- (d) What are the functions required for administering Hadoop? Name four standard Hadoop Benchmarks along with its use.

UNIT-5

- (a) What is HBASE?
 - (b) Explain Data Processing Operations in PIG.
 - (c) What is Zookeeper? List the services provided by Zookeeper. Explain the architecture of Zookeeper.
 - (d) Write down the process involved in interactive visualization strategies? Briefly explain Data Visualization techniques.
-