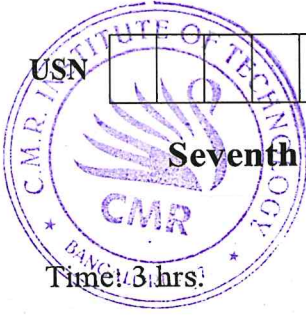USN | | | | | | | | | |

18CS72

## Seventh Semester B.E. Degree Examination, Feb./Mar.2022
## Big Data Analytics

Time: 3 hrs.

Max. Marks: 100

**Note: Answer any FIVE full questions, choosing ONE full question from each module.**

### Module-1

1  a. Discuss the Evolution of Big Data. (06 Marks)
   b. Explain the characteristics of Big Data. (04 Marks)
   c. With a neat block diagram, explain Data Architecture Design. (10 Marks)

**OR**

2  a. Write notes on Analytics Scalability to Big Data and Massive Parallel Processing Platforms. (12 Marks)
   b. Highlight Big Data Analytics applications with one case study. (08 Marks)

### Module-2

3  a. What are the core components of Hadoop? Explain in brief its each of its components. (10 Marks)
   b. Explain Hadoop Distributed File System. (10 Marks)

**OR**

4  a. Define MapReduce Frame work and its functions. (06 Marks)
   b. Write down the steps on the request to MapReduce and the types of process in MapReduce. (10 Marks)
   c. Write short notes on Flume Hadoop Tool. (04 Marks)

### Module-3

5  a. Discuss the characteristics of NoSQL data store along with the features in NoSQL transactions. (08 Marks)
   b. With neat diagrams, explain the following for shared-Nothing Architecture for Big Data Tasks,
      (i) Single Server model
      (ii) Sharding very large databases
      (iii) Master Slave distribution model.
      (iv) Peer-to-Peer distribution model. (12 Marks)

**OR**

6  a. Define key-value store with example. What are the advantages of key-value store? (10 Marks)
   b. Write down the steps to provide client to read and write values using key-value store. What are the typical uses of key value store? (10 Marks)

### Module-4

7  a. With a neat diagram, explain the process in MapReduce when client submitting a Job. (10 Marks)
   b. Explain Hive Integration and work flow steps involved with a diagram. (10 Marks)

**OR**

8  a.  Using HiveQL for the following:
      (i)      Create a table with partition.
      (ii)     Add, rename and drop a partition to a table.     **(10 Marks)**
   b.  What is PIG in Big Data? Explain the features of PIG.     **(10 Marks)**

## Module-5

9  a.  In Machine Learning explain linear and non-linear relationship with essential graphs.
                                        **(10 Marks)**
   b.  Write the block diagram of text mining process and explain its phases.     **(10 Marks)**

**OR**

10  a.  Define multiple regressions. Write down the examples involved in forecasting and optimization in regression.     **(10 Marks)**
    b.  Explain the parameters in social graph network topological analysis using centralities and PageRank.     **(10 Marks)**

* * * * *