Code: IT8T2D

IV B.Tech - II Semester – Regular / Supplementary Examinations MAY-2022

BIG DATA ANALYTICS
(INFORMATION TECHNOLOGY)

Duration: 3 hours   Max. Marks: 70

PART – A

Answer *all* the questions.   All questions carry equal marks

11x 2 = 22 M

1.

a) What is "Big Data Analytics"?

b) What do you mean by Apache Hadoop?

c) Classify the Basic file system operations in HDFS.

d) What do you mean by Interface in Hadoop?

e) What is the use of shuffle process in MapReduce?

f) Write the general form of map function in Hadoop MapReduce.

g) What do you mean by Job Completion in MapReduce?

h) Classify the Output formats in MapReduce.

i) Write the significance of Hive Ql Operators.

j) What is meta store in Hive Shell?

k) What is Hive Ql ?

PART – B

Answer any *THREE* questions. All questions carry equal marks.

$$3 \times 16 = 48 \text{ M}$$

2. a) Explain the MapReduce model with an example. 10 M

   b) Demonstrate the scope of VMWare in Hadoop. 6 M

3. Demonstrate the interaction between the Client and HDFS.

   16 M

4. a) Illustrate how do you Analyze data with Hadoop
      MapReduce? 8 M

   b) Explain about Running a distributed MapReduce Job.

   8 M

5. a) Explain about Progress and status updates in MapReduce.

   8 M

   b) Demonstrate the Job submission and Job Initialization in
      MapReduce. 8 M

6. a) Explain whirlwind tour of the data Model. 10 M

   b) Demonstrate about Hive services. 6 M