PVP 14

Code: CS7T1

IV B.Tech - I Semester –Regular / Supplementary Examinations
JANUARY - 2022

BIG DATA CONCEPTS
(COMPUTER SCIENCE AND ENGINEERING)

Duration: 3 hours Max. Marks: 70

PART – A

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

11 x 2 = 22 M

1.

a) State the core components of Hadoop?
b) List out the various use cases of Hadoop?
c) What is the use of combiner function in Hadoop?
d) Extrapolate the Master components: Name node, Secondary Node and Job Tracker?
e) Explain about Data Locality in MapReduce?
f) What is GenericOptionsParser class in Hadoop?
g) What is the use of command Line Interface in HDFS?
h) What is the main idea of Fair Scheduler in Hadoop?
i) What is the use of shuffle process in Map Reduce?
j) Define Input Split.
k) Generalise the term Record Reader/Writer.

PART – B

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

3 x 16 = 48 M

2. a) Describe the working of the MapReduce analyzation. 8 M

b) Explain the following
i) Mapper class
ii) Reducer class
iii) Scaling out 8 M

3. a) Discuss Hadoop distributed file system in detail. 8 M

b) Explain in detail the Master and Slave components of Hadoop Cluster? 8 M

4. a) Write in detail the concept of developing the Map Reduce Application. 8 M

b) Describe in brief about API for Map reduce framework. Distinguish between the old and new versions of Hadoop API for Map Reduce frame work. 8 M

5. a) Explain the failures in MapReduce. 8 M

b) Explain the Anatomy of a Map Reduce Job Run? 8 M

6. a) Discuss the various types of map reduce & its output formats. 8 M

b) Implement the Input Format for Compute-Intensive applications? 8 M