R15

B.Tech III Year I Semester (R15) Supplementary Examinations June 2018
INTRODUCTION TO BIG DATA
(Common to CSE & IT)

Time: 3 hours   Max. Marks: 70

PART – A
(Compulsory Question)
*****

1   Answer the following: (10 X 02 = 20 Marks)
   (a)   Define cloud computing.
   (b)   What is socket? Explain.
   (c)   Write note about limitations of HAR files.
   (d)   What are namenodes?
   (e)   What is the role of combiner and partitioner in map reduce application?
   (f)   Write note mapper.
   (g)   What is the function of logs in map reduce jobs?
   (h)   What is a Daemon thread?
   (i)   List some big data analytics libraries in Apache Mahout.
   (j)   Define clustering.

PART – B
(Answer all five units, 5 X 10 = 50 Marks)

UNIT – I

2   Explain in detail about thread and its role in cloud computing.
OR
3   Describe the concept and architecture of cloud computing.

UNIT – II

4   Elucidate the internals of Hadoop distributed files system (HDFS).
OR
5   Discuss the goals of HDFS and explain the concept of Hadoop architecture.

UNIT – III

6   (a)   Elaborate the design issues of setting up cluster in HDFC.
    (b)   Discuss the functionalities of YARN.
OR
7   I Discuss the process of map and reduce part of MapReduce framework with an example.

UNIT – IV

8   List and explain briefly the different logs produced by map reduce.
OR
9   Discuss different types of input and output formats of MapReduce with an example.

UNIT – V

10   Discuss how Mahout libraries can be used for big data analytics with example.
OR
11   Elucidate the concept of K-means clustering with an example.

*****