Code: 15A05506 **R15**

B.Tech III Year I Semester (R15) Regular Examinations November/December 2017

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) What is a Daemon thread?
(b) What are the benefits of cloud computing?
(c) List the three modes in which Hadoop can run.
(d) What is HDFS?
(e) What is shuffling in map reduce?
(f) What is “mapper” and “reducer” in Hadoop?
(g) Enumerate the objectives of fair scheduler.
(h) Which interface needs to be implemented to create mapper and reducer for Hadoop?
(i) List machine learning algorithms exposed by Mahout for clustering and classification.
(j) List out Hadoop’s configuration files.

PART – B

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

UNIT – I

2 What is multithreading? Explain different ways of implementing threads in java.

OR

3 Discuss the techniques for developing distributed applications in java.

UNIT – II

4 Explain components in Hadoop stack in detail.

OR

5 Explain the goals of HDFS and discuss the functionalities of Name node and Data node.

UNIT – III

6 (a) Explain the need for MapReduce programming model in detail.
(b) Discuss the functionalities of YARN.

OR

7 Discuss the process of map and reduce part of MapReduce framework with an example.

UNIT – IV

8 Explain the following operations performed by Hadoop MapReduce framework:

(a) Partition.
(b) Shuffle.
(c) Sort.

OR

9 Discuss different types of input and output formats of MapReduce with an example.

UNIT – V

10 Explain k-means clustering algorithm with an example.

OR

11 Discuss how Mahout libraries can be used for big data analytics with example.

*****