Roll No. | | | | | | | | | | | | Total No. of Pages : 02

Total No. of Questions : 09

B.Tech. (Computer Science Engg. / ECE) (Sem.–7,8)

INTRODUCTION TO BIG DATA

Subject Code : BTEC-909A-18

M.Code : 90683

Date of Examination : 14-07-22

Time : 3 Hrs. Max. Marks : 60

INSTRUCTIONS TO CANDIDATES :

1. **SECTION-A is COMPULSORY consisting of TEN questions carrying TWO marks each.**
2. **SECTION-B contains FIVE questions carrying FIVE marks each and students have to attempt any FOUR questions.**
3. **SECTION-C contains THREE questions carrying TEN marks each and students have to attempt any TWO questions.**

SECTION-A

1. Write briefly :

a. Define Big Data.

b. Define HDFS.

c. What do you mean by distributed file system?

d. Define clustering.

e. What is the importance of decision tree?

f. What do you mean by real time database?

g. Define scalability.

h. Define GPFS.

i. What do you mean by data analytics?

j. Define Indexing.

SECTION-B

2. Explain characteristics of big data in detail.

3. Explain architecture of Hadoop.

4. Explain Decision tree algorithm in detail with suitable example.

5. What do you mean by stream computing? Explain it in detail with examples.

6. Explain RTAP in detail.

SECTION-C

7. Explain the K-Means clustering algorithm with an example.

8. Explain how the data flow takes places in Map Reduce framework.

9. Explain challenges of conventional system and solve a real time analytics problem using conventional system.

NOTE : Disclosure of Identity by writing Mobile No. or Making of passing request on any page of Answer Sheet will lead to UMC against the Student.