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

Total No. of Questions : 09

B.Tech. (CSE/ECE)   (Sem.–7,8)
INTRODUCTION TO BIG DATA
Subject Code : BTEC-909A-18
M.Code : 90683
Date of Examination : 23-01-2023

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.   Why do you need to tame big data?

    b.   What is the need of Map Reduce model?

    c.   Define YARN.

    d.   Define Hadoop MR Engine.

    e.   What do you mean by Segmentation?

    f.   Define Linear Regression.

    g.   Define RTSA.

    h.   Why is HDFS preferred to RDBMS?

    i.   Define thread pooling?

    j.   Explain in brief perception of value.

SECTION-B

2. Explain distributed file system in detail.

3. List the main characteristics of Big Data.

4. Explain linear regression with suitable example.

5. Explain benefits of stream computing in big data world.

6. Explain HADOOP Implementation process in detail.

SECTION-C

7. Why does one choose analytical system over conventional system and explain limitations of conventional system in detail.

8. Explain Indexing Techniques in detail and create inverted index using JAQL.

9. Explain Hadoop Architecture and Deployment in detail.

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.