Roll No ....

MCADD-705(2)
M.C.A. (Integrated), VII Semester
Examination, May 2022
Data Analytics
Time : Three Hours

Maximum Marks : 70

Note :   i)   Attempt any five questions.

         ii)   All questions carry equal marks.

1. a) What do you mean by Randomization Test? Where is it used? Give a suitable example.

    b) What is ANOVA (Analysis of Variance)? Explain with the help of a suitable example.

2. a) Examine how would you implement regression modeling?

    b) What are the features of support vector machines?

3. a) What is Support vector machine? What is the role of C in EVM?

    b) Give some situations where you will use an SVM over a Random forest machine learning algorithm and vice versa.

4. a) Explain supervised and unsupervised learning with examples.

    b) Describe in detail about hypothesis testing?

5. a) What do you mean by Associative Rule Mining Challenges for big data analysis?

    b) Differentiate clustering and classification of Big data. Give suitable examples.

6. a) Compare and contrast the relationship between centroids and clustering.

    b) What is Big data? Describe the main features of a data analytical system?

7. a) What is the relationship between sampling distribution and re-sampling?

    b) What are the best practices in Big Data Analytics?

8. Write short notes on any two of the following:
    i)   Hypothesis
    ii) Hadoop
    iii) Map Reduce
    iv) Hive

******