Roll No ....

MCADD-702

M.C.A. (Integrated), VII Semester

Examination, May 2022

Data Sciences in Mining

Time : Three Hours

Maximum Marks : 70

Note: i) Attempt any five questions.

ii) All questions carry equal marks.

1. a) What is Data Sciences? Explain role of Data Science in Mining. 6

   b) Discuss Exploratory Data Analysis (EDA). List and discuss basic tools of EDA. 8

2. a) What is Machine Learning? Discuss Supervised and Unsupervised Machine Learning. 6

   b) Explain k-NN algorithm in detail. How do we choose the factor 'k' in k-NN algorithm? 8

3. a) How the Principle Component Analysis (PCA) is used for Data Compression? Discuss. 7

   b) What is Feature Selection? How Decision trees are used in extracting meaning from data? 7

4. What is Data Visualization? Briefly discuss the usage of all visualization tools in Data Science? 14

5. Briefly explain Benefits, Ethical Issues and Challenges in Data Sciences? 14

6. Write a short note on following:
   a) Big data and its tools   5
   b) 'R' Programming basics   5
   c) Data Mining   4

7. a) Discuss k-means algorithm in detail.   6
   b) Discuss why Naïve Bayes algorithm is suitable for Filtering Span.   8

8. a) Discuss Random Forest Algorithm. Explain how it useful for classification problem?   8
   b) Discuss the methods for mining social network graphs.   6

******

MCADD-702