

Total No. of printed pages = 2

MCA 202303

Roll No. of candidate

--	--	--	--	--	--	--	--	--	--

7/21

2021

BINA CHOWDHURY CENTRAL LIBRARY
(MCA & FIPS)
BINA, HATKI, WAPARA,
GOSWAMI - 705-017

M.C.A. 3rd Semester End-Term Examination

INTRODUCTION TO DATA SCIENCE

Full Marks – 70

Time – Three hours

The figures in the margin indicate full marks
for the questions.

Answer question No. 1 and any *four* from the rest.

1. Answer the following (MCQ/ Fill in the blanks) : (10 × 1 = 10)
- (i) What are probability and non-probability sampling techniques?
 - (ii) Define Range and Interquartile Range.
 - (iii) What is high dimensional data?
 - (iv) Name the two types of supervised learning techniques.
 - (v) How rule-based algorithm works?
 - (vi) Why clustering is done in Data Science?
 - (vii) What is block in deep learning?
 - (viii) Write the concept of layers and blocks in Deep Learning.
 - (ix) Why Neural Networks are used in Data Science?
 - (x) How important is Natural Language Processing in Data Science?
2. (a) Write the different techniques used for measuring central tendency and how they help in dealing with data. (3+3 = 6)
- (b) How CLIQUE algorithm works? (5)
- (c) Write the basic functionalities of each block in a Custom Block. (4)

[Turn over

3. (a) What is supervised and unsupervised learning? Write the differences between them. (2+2 = 4)
- (b) Explain the architecture of Artificial Neural Network. Write the advantages of it. (3+2 = 5)
- (c) Why Natural Language Processing (NLP) is used? Write the advantage of NLP. (2+2 = 4)
- (d) Explain Binomial Distribution. (2)
4. (a) What is Sampling? Why it is used in Data Science? Write down the steps involved in Sampling. (1+1+3 = 5)
- (b) Describe and compare Euclidean, Manhattan and Minkowski distance in measuring similarity and dissimilarity. (3+3 = 6)
- (c) Write the differences between Semantic and Lexical Analysis. What is Pragmatic analysis in Natural Language Processing (NLP)? (2+2=4)
5. (a) Give the brief of different levels of Natural Language Processing (NLP). (5)
- (b) What is KNN? Write the intuitive steps for KNN. (1+4 = 5)
- (c) What is clustering? How K Mean clustering algorithm works? (1+4 = 5)
6. (a) Write the importance of descriptive statistics in Data Science? (4)
- (b) List the major tasks in data preprocessing. Explain briefly each of these tasks. (2+4=6)
- (c) How Feed forward Neural Network works? Explain the two phases of this network. (3+2 = 5)
7. (a) Why ensemble methods are required? Explain how Bagging works. (2+3=5)
- (b) What is MADALINE neural network? How MADALINE network is formed? (2+3 =5)
- (c) Write the importance of Correlation and Regression in Data Science. (5)
-