(b) What do you understand by the following metrics. Explain in context of Machine Learning.

(i) Mean Squared Error, Mean Absolute Error and Root Mean Squared Error.

(ii) Accuracy Score.

(iii) Confusion Matrix, Accuracy, Precision, Recall. (6,9)

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 5097 **K**

Unique Paper Code : 2922104701

Name of the Paper : Data Science

Name of the Course : **B.A. (H) Business Economics**

Semester : VII

Duration : 2 Hours Maximum Marks : 60

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt any **four** questions, all questions carry equal marks.
3. Use of simple calculator is allowed.
4. If any assumption is made while solving a problem, the same must be stated clearly.

1. (a) What do you understand by Machine Learning and how does it differ from Deep Learning.

(b) State different Machine Learning Algorithm which are commonly used in business and economics. Which of these algos stated are parametric.

(c) Explain step by step the Process of Data Science. (4,4,7)

2. (a) Discuss different libraries commonly used for Machine Learning. Which one of these Libraries is used for Data Visualization.

(b) What are toy data sets. Can you state examples of five different toy data sets commonly used in Machine Learning.

(c) Distinguish between supervised and unsupervised learning, State two algorithm falling in each of these categories. (5,5,5)

3. (a) What do you understand by Data Cleaning. What can be different types of data problems which can be addressed using Machine Learning. State ML-Python Codes employed for data cleaning.

(b) Explain the need for carrying out a Principal Components Analysis . What codes are required to import a toy data set from SKlearn and perform a PCA on this data set. (8,7)

4. (a) What is Normalization. Explain LI and L2 Normalization and what import codes from SKlearn are needed to achieve the data normalization.

(b) Explain Classification as a ML Algorithm. Clearly explain the role of decision tree in classification.

(c) Explain the concepts of (i) Entropy (ii) Gini Coefficient in Decision Tree Classification. (5,5,5)

5. (a) Explain why Random Forest is considered as an extension of decision tree.

(b) What is K-Means clustering. What is the principle behind the clustering process.

(c) Write down ML-Python codes which generates random data of 100 observations and forms four clusters showing centroid, each with clear division. You can choose your own colour combinations and labels, but codes for all must be supported by codes. (5,5,5)

6. (a) Explain clearly the following terms with respect to support vector machines (i) Margin (ii) Hyper plane (iii) Support Vectors (iv) Kernel and Radical Based Function as used for Classification.