APPLIED MACHINE LEARNING

Assignment 1

(Weightage – 12 Marks)

Last Date for submission: 17th October 2023

INSTRUCTIONS:

1) **Problem statement 1 is compulsory for all and you can choose between Problem statement 2 or 3(3 + 9 = 12 Marks)**

2) **Need to implement and upload both in one python notebook only together with data sets in a zipped folder**

3) **Each submission should contain the following**

   Your name, BITS ID, Course name and number, Selection of problem statements info

   Need to go through all the pre-processing steps needed

4) **Need to maintain integrity in completing this task also. In case of any similarity between submissions will be considered as plagiarism**

Problem statement 1: 3 Marks - Compulsory

Key words: Assignment- Python - Data Exploration - Iris Dataset

1. Download Iris dataset from UCI website.

2. For all numeric attributes, compute Mean and Standard Deviation

3. Create following scatter plots

> **a) Sepal Length vs Petal Length**

> **b) Petal width vs Sepal width**

> **c) Petal length vs Sepal length**

4. Create 5-point summary for all numeric attributes and visualize it.

5. Create histogram (with 5 bins) from Sepal Length and Petal Width

Problem statement: 2 – 9 Marks - Classification

Predict if a customer is satisfied or dissatisfied with their banking experience.

Tags: **E-commerce, banking, binary classification**

Description: **You are given an anonymized dataset with a large number of numeric Features.**

The output variable is a customer satisfied or unsatisfied.

1. https://www.kaggle.com/competitions/santander-customer-satisfaction/data

2. Identify right Machine Learning model to implement

3. Implement either using Jupiter Notebook or in colab.

Using the hierarchical sales data from Walmart across various geo-locations, forecast daily sales for the next 15 days.

Tags: retail, Regression

Description: To forecast daily sales for the next 28 days given the dataset, which consists of item level, product categories, store details and explanatory variables such as price, promotions, day of the week, and special events.

1. https://www.kaggle.com/competitions/walmart-recruiting-store-sales-forecasting/data

2. Identify right Machine Learning model to implement

3. Implement either using Jupiter Notebook or in colab.