**Applied Machine Learning**

**Spring 2025**

**Assignment # 1**

**Max Marks: 50 Due Date: February 14, 2025**



**Q1.** Understand, Download and Run the code for the dataset YBI-Foundation. <https://www.kaggle.com/code/ybifoundation/tutorial-decision-tree-regression-scikit-learn> (10)

**Q2.** Imagine that you are a medical researcher compiling data for a study. You have collected data about a set of patients, all of whom suffered from the same illness. During their course of treatment, each patient responded to one of 5 medications, Drug A, Drug B, Drug c, Drug x and Drug y. Part of your job is to build a model to find out which drug might be appropriate for a future patient with the same illness. The features of this dataset are Age, Gender, Blood Pressure, and the Cholesterol of the patients, and the target is the drug that each patient responded to.

<https://www.kaggle.com/datasets/pablomgomez21/drugs-a-b-c-x-y-for-decision-trees/data?select=drug200.csv>

It is a sample of multiclass classifier, and you can use the training part of the dataset to build a decision tree, and then use it to predict the class of a unknown patient, or to prescribe a drug to a new patient. (20)

**Q3.** From the given code sample and datasets, run the following sections on relevant dataset: Submit the running file for Jupyter Notebook along with the used datasets. You can add or modify the code where required. (20)

<https://www.kaggle.com/code/antoreepjana/statistics-for-ml-data-analysis>

<https://www.kaggle.com/code/smsmibrahim/decision-tree-id3-implementation-using-play-tennis>

<https://www.kaggle.com/code/antoreepjana/statistics-for-ml-data-analysis#1.-CART-Algorithms>

<https://www.kaggle.com/discussions/getting-started/279349>

<https://www.kaggle.com/code/mcmuralishclint/decision-tree-regression>

<https://www.kaggle.com/code/sndorburian/decision-tree-for-trading-using-python>

<https://www.kaggle.com/code/hamelg/python-for-data-29-decision-trees>

<https://www.geeksforgeeks.org/decision-tree-implementation-python/>

<https://www.kaggle.com/code/sahilwagh/drug-dt>

<https://www.w3schools.com/python/python_ml_decision_tree.asp>

<https://www.javatpoint.com/decision-tree-in-python-sklearn>

a. Run the code for ID3.

b. Run the code for C4.5.

Compare and analyze the results of the above two algorithms.

**GOOD LUCK**☺