**CS 459 Introduction to Machine Learning**

**Assignment**

**Part # 2**

Due: 27th April, 2022

In this assignment, you have to apply the machine learning to a problem and data sets of your choice. Part 2 of the assignment is as follows:

1. Build a decision tree using your dataset. If the data is numeric, pre-process it accordingly. You can take help from the relevant chapter in the book “Introduction to Machine Learning with Python” by Muller and Guido.
2. Implement selection of the top node using i) information gain ii) gain ratio using python commands and making use of the pandas library.
3. Develop a decision tree for your dataset using i) information gain ii) gain ratio. Please note that you can use any relevant library (e.g. scikit-learn) to build the decision tree.
4. You should report the accuracy of the decision tree.
5. Use any library to visualize the resulting tree.
6. Carry out logistic regression using your dataset. If the data is not numeric, pre-process it accordingly. You can take help from the relevant chapter in the book “Introduction to Machine Learning with Python” by Muller and Guido. Please note that you can use any relevant library (e.g. scikit-learn)

**Submission:**

**Softcopy**

1. Dataset
2. Python program to implement part 1 and 2