

# Assignment 1

## **Objective**

In this assignment, you are required to implement a predictive modeling approach based on the decision tree.

## **Detailed Requirement**

We have introduced a predictive modeling approach based on the decision tree in the class. In this assignment, you will implement and evaluate this approach on two data sets. You do not need to create your own data set, since there are many such data sets available in the UCI Machine Learning Repository:

<http://archive.ics.uci.edu/ml>

In this repository, there are data sets associated with different problem domains, including credit card application approval, medical diagnosis and document analysis.

## **Assignment Submission**

### **Program (40%)**

You should submit your program and a readme file with instructions on how to run the program.

### **Report (60%)**

You should summarize your work in the form of a report which should include descriptions of:

1. The data set you have chosen and its main characteristics.
2. The data preprocessing operations you have applied to improve data quality.
3. The structure of the decision tree you have implemented, and the reason of choosing this particular structure.
4. The design of the different modules of your program. In describing each module, you should relate your design to the theories you have learnt in the class.
5. The prediction results and your interpretation of the results.
6. Limitations and possible improvements of the program.