

Course: Machine Learning  
Course no. CAP6610  
Instructor: LiMin Fu

### **Project: The Decision Tree Approach**

The ID3 algorithm developed by Ross Quinlan is an inductive classification algorithm. The knowledge acquired is represented in the form of decision trees. It uses a set of training examples to build a decision tree representing the concepts learned. The goodness of the decision tree is tested on a different set of test examples.

In this exercise, you will learn how to use the ID3 tool for machine learning. The program, the training and test data files can be downloaded from the course web page ([www.cise.ufl.edu/~fu/learn.html](http://www.cise.ufl.edu/~fu/learn.html)). Read the document posted on the web page first to get yourself familiar with the use of the program. Both the program and the data were obtained from the public domain for instructional use only.

To invoke the program, type

```
>gcl  
>(load "id3")  
>(run)
```

Now, the program starts running. Then, load the data files and follow the prompts.

Please hand in

- Two decision trees built from the training data, using respectively 95% and 65% statistical thresholds (the concept strength) as the stopping condition. The trees must be put in a graphic format.
- The test accuracy for the two decision trees.
- Give discussion on the program performance and results.