**Homework Assignment #1  
Due on Friday, 1/29**

**Part 1**

Obtain a copy of the machine learning toolkit, Weka, freely available for download. Familiarize yourself with the Weka input [ARFF](http://weka.wikispaces.com/ARFF+%28stable+version%29) format. In this format, each instance is described on a single line. The feature values are separated by commas, and the last value on each line is the class label of the instance. Each ARFF file starts with a header section describing the features and the class labels. Lines starting with '%' are comments. See the link above for a brief, but more detailed description of the ARFF format.

**Part 2**

Create a data set of your choice in ARFF format with at least 20 examples (instances) and at least 10 features. Your data set can be much larger if you wish. Experiment with some of the classification algorithms (under "Classify" in the Weka gui) on your data set. At minimum use J48 (the tree learner) and 1-Nearest Neighbor (the instance-based learner). Display the ROC curves (visualize threshold curves in the lower left window after running the classifier) for these two methods when run on your data using 10-fold cross-validation (the default). Turn in a 1-page PDF describing your data set and results.