**Introduction to Machine Learning Program assignment #1**

Problem:

The classifier module, with inputs for the training data. The output should indicate the class label assigned to each instance in the testing data. For this assignment, you need to implement decision tree module and using pruning to avoid overfitting.

For the validation step, output the classification performance. Use confusion matrix to summarize the classification result for the testing data set. You should implement both resubstitution validation (i.e., identical training and testing data sets) and K-fold cross validation, K being a value that you can select (K>1).

Option:

You can use more than one decision tree module to do this assignment.

Model ensembles: boosting and bagging.

Data:

<https://archive.ics.uci.edu/ml/datasets/Iris>

including 150 number of instances with 4 attributes.

Attribute Information:

1. sepal length in cm

2. sepal width in cm

3. petal length in cm

4. petal width in cm

5. class:

-- Iris Setosa

-- Iris Versicolour

-- Iris Virginica

Submit your source code and report.

The report should include the results, environment, using library and language, explain of your code and how to use it.