Readme

Most important:
PLEASE USE MY readARFF.py, cause there are some bugs in your version, and I fixed them.
Since my code is very fast, I didn't pickle the trees.

What I implemented:
1) All the things that are required in the assignment description.
2) The 5 points extra credit
3) The 3 points extra credit

How to run:
python decisionTree.py #datasetName (replace the datasetName with dataset file name)
to run my code.

For example:
python decisionTree.py nursery.arff

When you run the code, the result will show you the following contains:
1) Round number
2) Class name, precision, recall and accuracy for each classification by using ZeroR
3) For test set, the class name, precision, recall and accuracy for each classification
4) For trainning set, the class name, precision, recall and accuracy for each classification
5) The average of these five across run. (which is show as "summary")

There are some precision and recall cannot be computed since there are some TP and FP are 0, which is meaningless. For this case, I print "----".

For example, you run:
python decisionTree.py breast-cancer.arff

Here is the first round of the result:

Round 1 :

Using ZeroR:
Class: recurrence-events
Precision: ----
Recall: 0.0
Accuracy: 0.672

Class: no-recurrence-events
Precision: 0.672
Recall: 1.0
Accuracy: 0.672

Test Set:
Class: recurrence-events
Precision: 0.474
Recall: 0.474
Accuracy: 0.655

Class: no-recurrence-events
Precision: 0.744
Recall: 0.744
Accuracy: 0.655

Training Set:
Class: recurrence-events
Precision: 0.957
Recall: 1.0
Accuracy: 0.987

Class: no-recurrence-events
Precision: 1.0
Recall: 0.981
Accuracy: 0.987

Round 2 :
omit...
Round 3 :
omit...
Round 4 :
omit...

Round 5 :
omit...
Summary:
omit...