Assignment 3 Random Forest
Hailun Zhu    ID[hailunz]

1.  Parameters setting:
    Number of the trees : N = 5;
    Number of the features: featureNum = 3, total 6 features. I choose ceiling $\sqrt{6}$ = 3.
2.
   The best accuracy that I got is:
     Accuracy:
     0.5756385068762279

3.    Another experiment
Accuracy:
0.5399476096922069
feature: [0,3,5] : 0.5756385068762279
feature: [2,4,5]: 0.5140798952193845
feature: [0,1,4]:0.518664047151277
feature:[1,2,4]: 0.5225933202357563
feature:[3,4,5] :0.5756385068762279

As can be seen from this experiment, the tree that uses feature 3 and feature 5 got higher accuracy, this means that these two features are better. And the feature 4 is in four different trees, and the accuracies are different, which may imply feature 4 is not so important.

4. For the training data, each time I randomly choose 2/3 of the training set. Total rows of training are 12212, so randomly choosing the dataset the total number of rows are approximately 8141.