# How to make it to the top 8% in the leaderboard on Kaggle Leaf Classification problem

December 26, 2016

```
In [2]: #How I achieved top 8% in the world on Kaggle Leaderboard on the Leaf Class
        #Play and work with the tuning parameters to attain higher accuracy
        #The code I implemented is as follows:

In [3]: #Import the necessary modules for doing the analysis


        #import pandas as pd
        #from sklearn.preprocessing import LabelEncoder
        #from sklearn.neural_network import MLPClassifier

In [4]: #The Label Encoder module was brought up to convert the species type into c

        #We read the file from the source and then transform the species output int

In [6]: #train=pd.read_csv("enter your file path here")
        #train.drop(("id"),inplace=True,axis=1)
        #encoded=LabelEncoder().fit(train.species)
        #encodedtrans=encoded.transform(train.species)

In [7]: #Now drop the species column from train as we have encoded the labels and r

In [8]: #train.drop(("species"),axis=1,inplace=True)
        #testid=test["id"]
        #test.drop(("id"),axis=1,inplace=True)

In [10]: #Here I used deep learning artifical neural network algorithm called "Mult
         #Then this algorithm was fit on the training data with the following speci

In [11]: #clf = MLPClassifier(solver='lbfgs',hidden_layer_sizes=(6000,),activation=

In [12]: #The learning rate was minimized to make the algorithm stable

In [13]: #clf.fit(train,encodedtrans)

In [14]: #Then I did the prediction on the test data to find the probability of cla

In [15]: #pred=clf.predict_proba(test)
```

```
In [16]: #Then the results are stored as a pandas dataframe

In [17]: #submission = pd.DataFrame(pre, index=testid, columns=encoded.classes_)

In [18]: #Then finally this is the csv file that will take you to the top.

In [ ]: #submission.to_csv("leaf.csv")
```