**Prudential US Mortality rate case study**

Objective –

The classification goal is to predict if the client will subscribe a term deposit (variable y)

Dataset –

<http://archive.ics.uci.edu/ml/datasets/Bank+Marketing>

(bank-additional-full.csv)

Description –

The dataset contains 41,188 rows and 21 columns.

<https://www2.1010data.com/documentationcenter/prod/Tutorials/MachineLearningExamples/BankMarketingDataSet.html>

Tool used – Anaconda for Python 2.7, spyder2

Language – Python 2.7

**Tool setup**

1. Download and install Anaconda for python 2.7 and

<https://www.anaconda.com/download/#windows>

1. Open Anaconda command prompt and run below command

spyder

1. Once spyder started please import below .py file



**Diatatic(Approach) –**

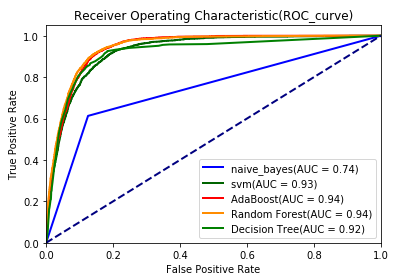
1. data\_preprocessing – created a LabelEncoder to deal with categorical variables.
2. Model\_fitting – applied random\_forest,adaboost,decision\_tree,naïve\_bayes,svm
3. Evaluation – uses a ROC to evaluate the model.

**Results –**

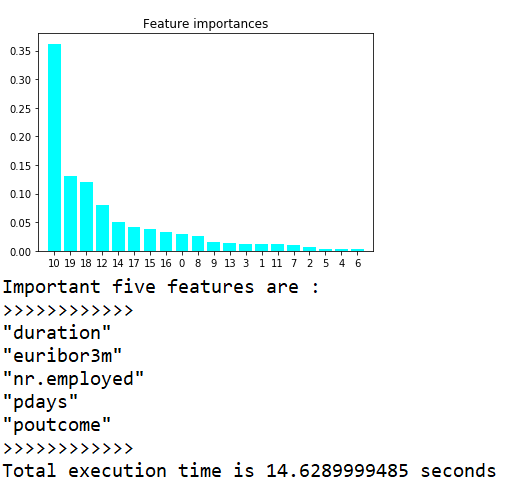
After running all above model below are the top model that are giving highest accuracy

Top models are

1. AdaBoost
2. RandomForest



Top features are



**How banks can use above analysis**

1. Banks should try to connect more on calls with customer asking relative question that are beneficial for banks and equally important for customers.
2. Banks should check for the customer who are financially weak as those customers more likely to open a term deposit.
3. Banks should check for occupation of employee as there may be a chances that customer will go for long term deposit.
4. Age can also be a factor to open long term deposit like pension plans, education loans.
5. Bank should try to increase number of calls to customer as it help to understand customer more.
6. Bank should keep track of people whenever they are opening any policy or any other funds to get a rough idea of the season so that they can give offer during that time.