Parsed data exists

Reading file(311 lines) ../parsedData/functionWords.txt

Reading file(22084 lines) ../parsedData/alldata45.txt

Starting feature vector classification

Reading parsed data

There are 11040 Native samples and 11044 Non-Native

Train set size 17667 - native=8842, non-native=8825

Test set size 4417 - native=2198, non-native=2219

Train data 'all feature vector' size is 311x17667

Train data size is 17667

Test data 'all feature vector' size is 311x4417

Test data size is 4417

-----------FUNC\_WORDS\_START---------------------------

Running SVM...

non-native native

precision[0.83302326 0.81120423] - tp/(tp+fp)

recall [0.80712032 0.8366697 ] - tp/(tp+fn)

fscore [0.81986725 0.8237402 ] - harmonicAvg(prec + recall)

fscore weighted =0.8217945173308785 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =82.1824767942042%

Running Decision Tree...

non-native native

precision[0.70096022 0.69237668] - tp/(tp+fp)

recall [0.69085174 0.70245678] - tp/(tp+fn)

fscore [0.69586927 0.69738031] - harmonicAvg(prec + recall)

fscore weighted =0.6966211961495747 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =69.66266696853067%

Running NB...

non-native native

precision[0.79536335 0.799908 ] - tp/(tp+fp)

recall [0.80396575 0.79117379] - tp/(tp+fn)

fscore [0.79964142 0.79551693] - harmonicAvg(prec + recall)

fscore weighted =0.7975889758011868 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =79.76001811184061%

-------------FUNC\_WORDS\_END-------------------------

Starting top x words classification

top\_words\_list size is 343

Reading parsed data

There are 11040 Native samples and 11044 Non-Native

Train set size 17667 - native=8842, non-native=8825

Test set size 4417 - native=2198, non-native=2219

Train data 'all feature vector' size is 343x17667

Train data size is 17667

Test data 'all feature vector' size is 343x4417

Test data size is 4417

------------------TOP\_WORDS\_START--------------------

Running NB...

non-native native

precision[0.93351303 0.93564582] - tp/(tp+fp)

recall [0.93645786 0.93266606] - tp/(tp+fn)

fscore [0.93498313 0.93415357] - harmonicAvg(prec + recall)

fscore weighted =0.9345703184367629 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =93.45709757754132%

Running Decision Tree...

non-native native

precision[0.87232143 0.87827285] - tp/(tp+fp)

recall [0.88057684 0.86988171] - tp/(tp+fn)

fscore [0.87642969 0.87405714] - harmonicAvg(prec + recall)

fscore weighted =0.8752490577826715 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =87.52546977586597%

Running NB...

non-native native

precision[0.93351303 0.93564582] - tp/(tp+fp)

recall [0.93645786 0.93266606] - tp/(tp+fn)

fscore [0.93498313 0.93415357] - harmonicAvg(prec + recall)

fscore weighted =0.9345703184367629 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =93.45709757754132%

---------------TOP\_WORDS\_END-----------------------

duration(formatted HH:MM:SS): 00:03:31