Parsed data exists

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

Reading file(22084 lines) ../parsedData/alldata10.txt

Starting feature vector classification

Reading parsed data

There are 11040 Native samples and 11044 Non-Native

Train set size 17667 - native=8832, non-native=8835

Test set size 4417 - native=2208, non-native=2209

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[0.74347158 0.69277597] - tp/(tp+fp)

recall [0.657311 0.77309783] - tp/(tp+fn)

fscore [0.69774147 0.7307363 ] - harmonicAvg(prec + recall)

fscore weighted =0.7142351509263892 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =71.51913063165044%

Running Decision Tree...

non-native native

precision[0.62146385 0.62328767] - tp/(tp+fp)

recall [0.62652784 0.61820652] - tp/(tp+fn)

fscore [0.62398557 0.6207367 ] - harmonicAvg(prec + recall)

fscore weighted =0.6223615033129279 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =62.23681231605161%

Running NB...

non-native native

precision[0.71083297 0.73238095] - tp/(tp+fp)

recall [0.74558624 0.69655797] - tp/(tp+fn)

fscore [0.72779496 0.71402043] - harmonicAvg(prec + recall)

fscore weighted =0.7209092540394721 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =72.10776545166404%

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

Starting top x words classification

top\_words\_list size is 356

Reading parsed data

There are 11040 Native samples and 11044 Non-Native

Train set size 17667 - native=8832, non-native=8835

Test set size 4417 - native=2208, non-native=2209

Train data 'all feature vector' size is 356x17667

Train data size is 17667

Test data 'all feature vector' size is 356x4417

Test data size is 4417

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

Running SVM...

non-native native

precision[0.93852459 0.77630122] - tp/(tp+fp)

recall [0.72566772 0.95244565] - tp/(tp+fn)

fscore [0.81848353 0.85539963] - harmonicAvg(prec + recall)

fscore weighted =0.836937404753904 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =83.90310165270546%

Running Decision Tree...

non-native native

precision[0.80695497 0.81646734] - tp/(tp+fp)

recall [0.81937528 0.80389493] - tp/(tp+fn)

fscore [0.8131177 0.81013236] - harmonicAvg(prec + recall)

fscore weighted =0.8116253677191829 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =81.16368575956531%

Running NB...

non-native native

precision[0.84284497 0.88883519] - tp/(tp+fp)

recall [0.89588049 0.83288043] - tp/(tp+fn)

fscore [0.86855387 0.85994856] - harmonicAvg(prec + recall)

fscore weighted =0.8642521917318375 - (fscore1\*realPortion+fscore2\*realPortion)

Total accuracy =86.43875933891782%

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

duration(formatted HH:MM:SS): 00:05:36