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

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

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

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

SVM:

Test data size is fixed to 1000

Attempt 1(SVM): Train data size=2500, Accuracy=0.782, fscore=0.782

Attempt 2(SVM): Train data size=4500, Accuracy=0.795, fscore=0.795

Attempt 3(SVM): Train data size=6500, Accuracy=0.796, fscore=0.796

Attempt 4(SVM): Train data size=8500, Accuracy=0.801, fscore=0.801

Attempt 5(SVM): Train data size=10500, Accuracy=0.805, fscore=0.805

Attempt 6(SVM): Train data size=12500, Accuracy=0.809, fscore=0.809

Attempt 7(SVM): Train data size=14500, Accuracy=0.810, fscore=0.810

Attempt 8(SVM): Train data size=16500, Accuracy=0.811, fscore=0.811

Attempt 9(SVM): Train data size=18500, Accuracy=0.815, fscore=0.815

Attempt 10(SVM): Train data size=20500, Accuracy=0.817, fscore=0.817

Attempt 11(SVM): Train data size=21084, Accuracy=0.815, fscore=0.815

DT:

Test data size is fixed to 1000

Attempt 1(DT): Train data size=2500, Accuracy=0.665, fscore=0.665

Attempt 2(DT): Train data size=4500, Accuracy=0.676, fscore=0.676

Attempt 3(DT): Train data size=6500, Accuracy=0.687, fscore=0.687

Attempt 4(DT): Train data size=8500, Accuracy=0.652, fscore=0.652

Attempt 5(DT): Train data size=10500, Accuracy=0.659, fscore=0.659

Attempt 6(DT): Train data size=12500, Accuracy=0.676, fscore=0.676

Attempt 7(DT): Train data size=14500, Accuracy=0.670, fscore=0.670

Attempt 8(DT): Train data size=16500, Accuracy=0.660, fscore=0.660

Attempt 9(DT): Train data size=18500, Accuracy=0.686, fscore=0.686

Attempt 10(DT): Train data size=20500, Accuracy=0.688, fscore=0.688

Attempt 11(DT): Train data size=21084, Accuracy=0.688, fscore=0.688

NB:

Test data size is fixed to 1000

Attempt 1(NB): Train data size=2500, Accuracy=0.791, fscore=0.791

Attempt 2(NB): Train data size=4500, Accuracy=0.800, fscore=0.800

Attempt 3(NB): Train data size=6500, Accuracy=0.800, fscore=0.800

Attempt 4(NB): Train data size=8500, Accuracy=0.800, fscore=0.800

Attempt 5(NB): Train data size=10500, Accuracy=0.804, fscore=0.804

Attempt 6(NB): Train data size=12500, Accuracy=0.808, fscore=0.808

Attempt 7(NB): Train data size=14500, Accuracy=0.805, fscore=0.805

Attempt 8(NB): Train data size=16500, Accuracy=0.809, fscore=0.809

Attempt 9(NB): Train data size=18500, Accuracy=0.808, fscore=0.808

Attempt 10(NB): Train data size=20500, Accuracy=0.808, fscore=0.808

Attempt 11(NB): Train data size=21084, Accuracy=0.808, fscore=0.808

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

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--------------------

SVM:

Test data size is fixed to 1000

Attempt 1(SVM): Train data size=2500, Accuracy=0.875, fscore=0.874

Attempt 2(SVM): Train data size=4500, Accuracy=0.897, fscore=0.897

Attempt 3(SVM): Train data size=6500, Accuracy=0.908, fscore=0.908

Attempt 4(SVM): Train data size=8500, Accuracy=0.914, fscore=0.914

Attempt 5(SVM): Train data size=10500, Accuracy=0.915, fscore=0.915

Attempt 6(SVM): Train data size=12500, Accuracy=0.921, fscore=0.921

Attempt 7(SVM): Train data size=14500, Accuracy=0.924, fscore=0.924

Attempt 8(SVM): Train data size=16500, Accuracy=0.925, fscore=0.925

Attempt 9(SVM): Train data size=18500, Accuracy=0.925, fscore=0.925

Attempt 10(SVM): Train data size=20500, Accuracy=0.925, fscore=0.925

Attempt 11(SVM): Train data size=21084, Accuracy=0.925, fscore=0.925

DT:

Test data size is fixed to 1000

Attempt 1(DT): Train data size=2500, Accuracy=0.864, fscore=0.864

Attempt 2(DT): Train data size=4500, Accuracy=0.878, fscore=0.878

Attempt 3(DT): Train data size=6500, Accuracy=0.865, fscore=0.865

Attempt 4(DT): Train data size=8500, Accuracy=0.880, fscore=0.880

Attempt 5(DT): Train data size=10500, Accuracy=0.867, fscore=0.867

Attempt 6(DT): Train data size=12500, Accuracy=0.862, fscore=0.862

Attempt 7(DT): Train data size=14500, Accuracy=0.862, fscore=0.862

Attempt 8(DT): Train data size=16500, Accuracy=0.863, fscore=0.863

Attempt 9(DT): Train data size=18500, Accuracy=0.873, fscore=0.873

Attempt 10(DT): Train data size=20500, Accuracy=0.876, fscore=0.876

Attempt 11(DT): Train data size=21084, Accuracy=0.876, fscore=0.876

NB:

Test data size is fixed to 1000

Attempt 1(NB): Train data size=2500, Accuracy=0.926, fscore=0.926

Attempt 2(NB): Train data size=4500, Accuracy=0.932, fscore=0.932

Attempt 3(NB): Train data size=6500, Accuracy=0.930, fscore=0.930

Attempt 4(NB): Train data size=8500, Accuracy=0.928, fscore=0.928

Attempt 5(NB): Train data size=10500, Accuracy=0.930, fscore=0.930

Attempt 6(NB): Train data size=12500, Accuracy=0.929, fscore=0.929

Attempt 7(NB): Train data size=14500, Accuracy=0.930, fscore=0.930

Attempt 8(NB): Train data size=16500, Accuracy=0.932, fscore=0.932

Attempt 9(NB): Train data size=18500, Accuracy=0.933, fscore=0.933

Attempt 10(NB): Train data size=20500, Accuracy=0.932, fscore=0.932

Attempt 11(NB): Train data size=21084, Accuracy=0.932, fscore=0.932

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