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 2000

Attempt 1(SVM): Train data size=2500, Accuracy=0.789, fscore=0.789

Attempt 2(SVM): Train data size=4500, Accuracy=0.799, fscore=0.799

Attempt 3(SVM): Train data size=6500, Accuracy=0.801, fscore=0.801

Attempt 4(SVM): Train data size=8500, Accuracy=0.807, fscore=0.807

Attempt 5(SVM): Train data size=10500, Accuracy=0.809, fscore=0.809

Attempt 6(SVM): Train data size=12500, Accuracy=0.811, fscore=0.811

Attempt 7(SVM): Train data size=14500, Accuracy=0.815, fscore=0.814

Attempt 8(SVM): Train data size=16500, Accuracy=0.817, fscore=0.816

Attempt 9(SVM): Train data size=18500, Accuracy=0.819, fscore=0.818

Attempt 10(SVM): Train data size=20084, Accuracy=0.820, fscore=0.820

DT:

Test data size is fixed to 2000

Attempt 1(DT): Train data size=2500, Accuracy=0.664, fscore=0.664

Attempt 2(DT): Train data size=4500, Accuracy=0.673, fscore=0.673

Attempt 3(DT): Train data size=6500, Accuracy=0.682, fscore=0.682

Attempt 4(DT): Train data size=8500, Accuracy=0.667, fscore=0.667

Attempt 5(DT): Train data size=10500, Accuracy=0.676, fscore=0.677

Attempt 6(DT): Train data size=12500, Accuracy=0.685, fscore=0.686

Attempt 7(DT): Train data size=14500, Accuracy=0.674, fscore=0.674

Attempt 8(DT): Train data size=16500, Accuracy=0.677, fscore=0.677

Attempt 9(DT): Train data size=18500, Accuracy=0.693, fscore=0.692

Attempt 10(DT): Train data size=20084, Accuracy=0.693, fscore=0.692

NB:

Test data size is fixed to 2000

Attempt 1(NB): Train data size=2500, Accuracy=0.790, fscore=0.790

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

Attempt 3(NB): Train data size=6500, Accuracy=0.795, fscore=0.795

Attempt 4(NB): Train data size=8500, Accuracy=0.797, fscore=0.797

Attempt 5(NB): Train data size=10500, Accuracy=0.802, fscore=0.802

Attempt 6(NB): Train data size=12500, Accuracy=0.806, fscore=0.806

Attempt 7(NB): Train data size=14500, Accuracy=0.803, fscore=0.803

Attempt 8(NB): Train data size=16500, Accuracy=0.804, fscore=0.804

Attempt 9(NB): Train data size=18500, Accuracy=0.805, fscore=0.805

Attempt 10(NB): Train data size=20084, Accuracy=0.805, fscore=0.805

-------------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 2000

Attempt 1(SVM): Train data size=2500, Accuracy=0.881, fscore=0.880

Attempt 2(SVM): Train data size=4500, Accuracy=0.900, fscore=0.900

Attempt 3(SVM): Train data size=6500, Accuracy=0.911, fscore=0.911

Attempt 4(SVM): Train data size=8500, Accuracy=0.917, fscore=0.917

Attempt 5(SVM): Train data size=10500, Accuracy=0.917, fscore=0.917

Attempt 6(SVM): Train data size=12500, Accuracy=0.923, fscore=0.923

Attempt 7(SVM): Train data size=14500, Accuracy=0.925, fscore=0.925

Attempt 8(SVM): Train data size=16500, Accuracy=0.926, fscore=0.926

Attempt 9(SVM): Train data size=18500, Accuracy=0.929, fscore=0.929

Attempt 10(SVM): Train data size=20084, Accuracy=0.930, fscore=0.930

DT:

Test data size is fixed to 2000

Attempt 1(DT): Train data size=2500, Accuracy=0.858, fscore=0.858

Attempt 2(DT): Train data size=4500, Accuracy=0.868, fscore=0.868

Attempt 3(DT): Train data size=6500, Accuracy=0.859, fscore=0.859

Attempt 4(DT): Train data size=8500, Accuracy=0.869, fscore=0.869

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

Attempt 6(DT): Train data size=12500, Accuracy=0.866, fscore=0.866

Attempt 7(DT): Train data size=14500, Accuracy=0.864, fscore=0.864

Attempt 8(DT): Train data size=16500, Accuracy=0.862, fscore=0.862

Attempt 9(DT): Train data size=18500, Accuracy=0.868, fscore=0.868

Attempt 10(DT): Train data size=20084, Accuracy=0.868, fscore=0.868

NB:

Test data size is fixed to 2000

Attempt 1(NB): Train data size=2500, Accuracy=0.925, fscore=0.925

Attempt 2(NB): Train data size=4500, Accuracy=0.931, fscore=0.931

Attempt 3(NB): Train data size=6500, Accuracy=0.927, fscore=0.927

Attempt 4(NB): Train data size=8500, Accuracy=0.927, fscore=0.927

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

Attempt 6(NB): Train data size=12500, Accuracy=0.929, fscore=0.930

Attempt 7(NB): Train data size=14500, Accuracy=0.930, fscore=0.931

Attempt 8(NB): Train data size=16500, Accuracy=0.930, fscore=0.931

Attempt 9(NB): Train data size=18500, Accuracy=0.931, fscore=0.932

Attempt 10(NB): Train data size=20084, Accuracy=0.931, fscore=0.932

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