# Yelp Dataset

## Random Classifier

**Training F1-Measure**

F1 Score: 0.180419612228

Confusion Matrix:

[[104 102 114 97 105]

[115 134 136 124 132]

[225 197 183 195 197]

[490 496 481 487 514]

[459 486 461 479 487]]

**Validation F1-Measure**

F1 Score: 0.169737750897

Confusion Matrix:

[[15 16 15 22 16]

[12 21 22 21 20]

[47 33 24 28 32]

[71 74 78 67 66]

[69 72 47 55 57]]

**Test F1-Measure**

F1 Score: 0.186163692511

Confusion Matrix:

[[ 38 31 28 25 21]

[ 41 36 34 34 45]

[ 60 48 58 76 58]

[148 157 134 124 139]

[119 149 124 125 148]]

## Most Frequent Classifier Score

**Training F1-Measure**

F1 Score: 0.104267004647

Confusion Matrix:

[[ 0 0 0 522 0]

[ 0 0 0 641 0]

[ 0 0 0 997 0]

[ 0 0 0 2468 0]

[ 0 0 0 2372 0]]

**Validation F1-Measure**

F1 Score: 0.105014749263

Confusion Matrix:

[[ 0 0 0 84 0]

[ 0 0 0 96 0]

[ 0 0 0 164 0]

[ 0 0 0 356 0]

[ 0 0 0 300 0]]

**Test F1-Measure**

F1 Score: 0.103923019985

Confusion Matrix:

[[ 0 0 0 143 0]

[ 0 0 0 190 0]

[ 0 0 0 300 0]

[ 0 0 0 702 0]

[ 0 0 0 665 0]]

## Bernoulli Naive Bayes

**Training F1-Measure**

F1 Score: 0.577269725617

Confusion Matrix:

[[ 275 8 15 22 202]

[ 10 323 23 85 200]

[ 9 25 471 153 339]

[ 36 155 86 1211 980]

[ 53 124 73 248 1874]]

**Validation F1-Measure**

F1 Score: 0.330091223067

Confusion Matrix:

[[ 29 9 5 7 34]

[ 9 16 12 26 33]

[ 2 23 28 54 57]

[ 8 33 27 123 165]

[ 13 25 9 69 184]]

**Test F1-Measure**

F1 Score: 0.341137874213

Confusion Matrix:

[[ 42 24 5 18 54]

[ 22 46 32 31 59]

[ 9 36 44 99 112]

[ 10 58 61 246 327]

[ 10 37 27 150 441]]

## Decision Tree

**Training F1-Measure**

F1 Score: 1.0

Confusion Matrix:

[[ 522 0 0 0 0]

[ 0 641 0 0 0]

[ 0 0 997 0 0]

[ 0 0 0 2468 0]

[ 0 0 0 0 2372]]

**Validation F1-Measure**

F1 Score: 0.261495450896

Confusion Matrix:

[[ 16 15 15 14 24]

[ 14 10 24 22 26]

[ 12 18 40 57 37]

[ 19 35 56 135 111]

[ 20 15 33 115 117]]

**Test F1-Measure**

F1 Score: 0.279169541182

Confusion Matrix:

[[ 29 21 31 34 28]

[ 28 23 45 59 35]

[ 17 30 59 127 67]

[ 35 39 97 298 233]

[ 22 38 60 251 294]]

## Linear SVC

**Training F1-Measure**

F1 Score: 0.997870801179

Confusion Matrix:

[[ 520 0 0 0 2]

[ 0 640 0 0 1]

[ 0 0 996 1 0]

[ 0 0 0 2454 14]

[ 0 0 0 2 2370]]

**Validation F1-Measure**

F1 Score: 0.413976371409

Confusion Matrix:

[[ 36 23 8 9 8]

[ 14 25 16 27 14]

[ 7 23 50 65 19]

[ 10 11 31 171 133]

[ 4 9 18 102 167]]

**Test F1-Measure**

F1 Score: 0.400687403274

Confusion Matrix:

[[ 55 33 15 20 20]

[ 33 58 41 36 22]

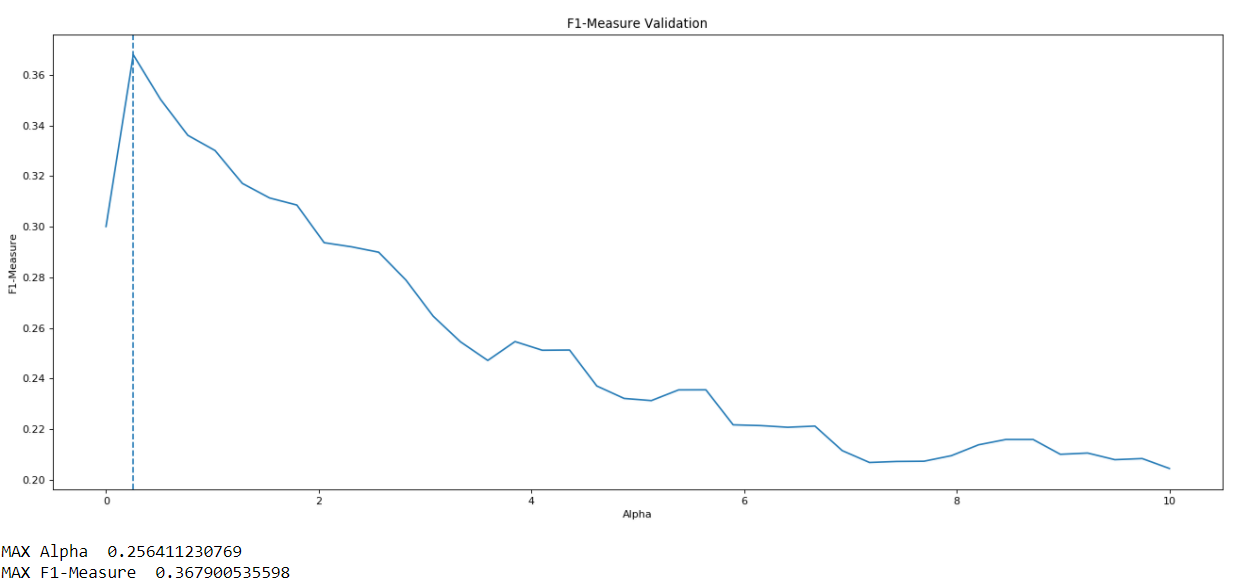
[ 12 42 77 118 51]

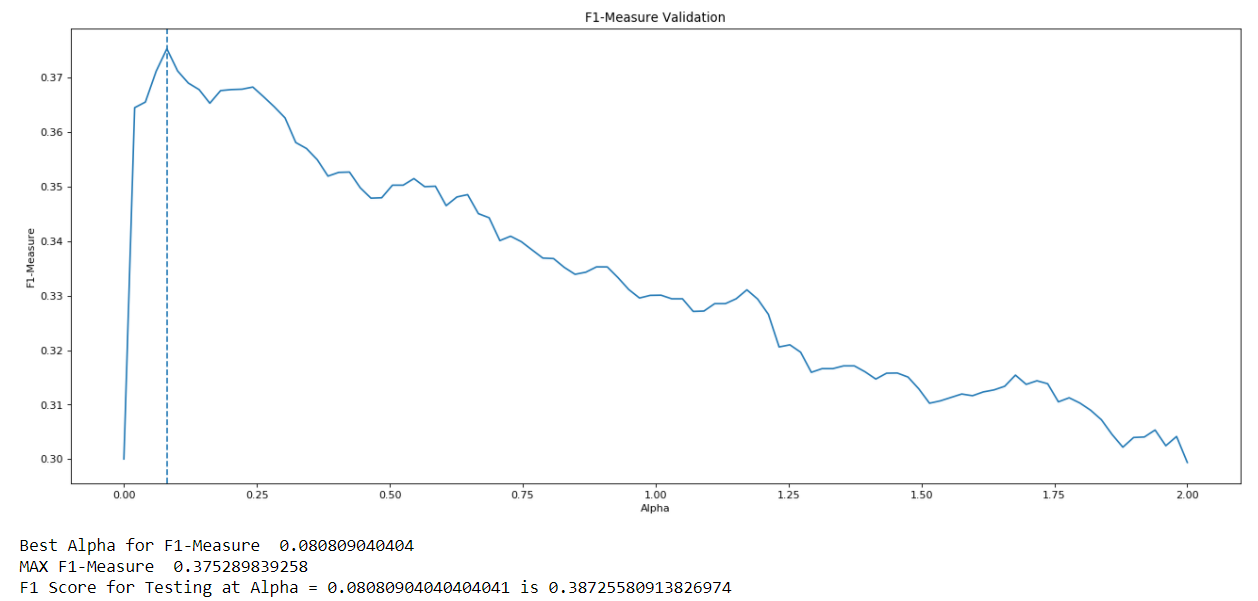
[ 12 31 83 341 235]

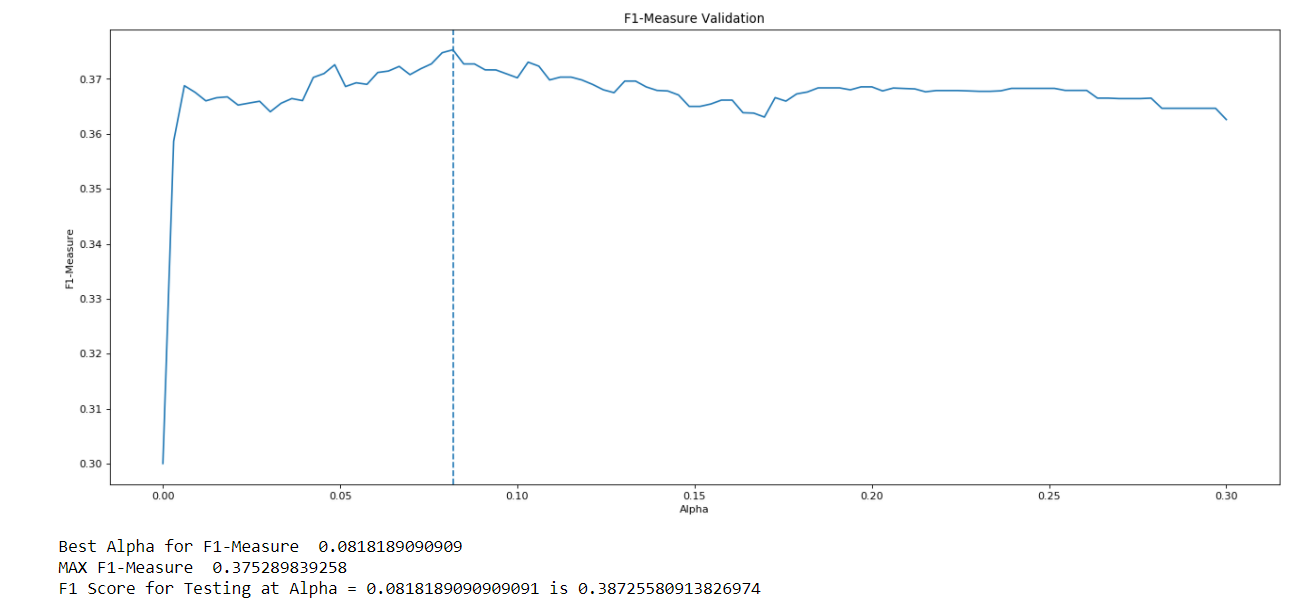
[ 15 12 41 236 361]]

# Yelp

## Bernoulli Naive Bayes Hyper-Parameter Tuning







## Linear SVC

Validation F1-Measure when Squared-Hinge Loss: 0.414261158207

Validation F1-Measure when Hinge Loss: 0.411135458873

Linear SVC

Iteration 1

Model with rank: 1

Mean validation score: 0.486 (std: 0.003)

Parameters: {'C': 0.1, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 2466, 'tol': 0.001}

Model with rank: 2

Mean validation score: 0.485 (std: 0.006)

Parameters: {'C': 0.1, 'dual': True, 'loss': 'hinge', 'max\_iter': 5709, 'tol': 0.001}

Model with rank: 3

Mean validation score: 0.456 (std: 0.002)

Parameters: {'C': 1.0, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 9370, 'tol': 0.01}

Model with rank: 3

Mean validation score: 0.456 (std: 0.002)

Parameters: {'C': 1.0, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 2459, 'tol': 0.01}

Iteration 2

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 4710, 'tol': 0.001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 4066, 'tol': 0.0001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 9689, 'tol': 0.001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 776, 'tol': 0.0001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 993, 'tol': 0.0001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 6630, 'tol': 0.01}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 5624, 'tol': 0.001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 931, 'tol': 0.0001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 6287, 'tol': 0.001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'dual': True, 'loss': 'squared\_hinge', 'max\_iter': 9933, 'tol': 0.0001}

Iteration 3

yelp

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'max\_iter': 7449, 'tol': 1e-06}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'max\_iter': 6358, 'tol': 0.0001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'max\_iter': 3775, 'tol': 0.0001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'max\_iter': 9580, 'tol': 0.001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'max\_iter': 5457, 'tol': 0.0001}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'max\_iter': 3337, 'tol': 1e-06}

Model with rank: 1

Mean validation score: 0.514 (std: 0.005)

Parameters: {'C': 0.01, 'max\_iter': 4428, 'tol': 0.001}