Male/Female train/test

=== Summary ===

Correctly Classified Instances 33 66 %

Incorrectly Classified Instances 17 34 %

Kappa statistic 0.32

Mean absolute error 0.3393

Root mean squared error 0.5745

Relative absolute error 67.8624 %

Root relative squared error 114.9041 %

Total Number of Instances 50

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.72 0.4 0.643 0.72 0.679 0.705 F

0.6 0.28 0.682 0.6 0.638 0.709 M

Weighted Avg. 0.66 0.34 0.662 0.66 0.659 0.707

=== Confusion Matrix ===

a b <-- classified as

18 7 | a = F

10 15 | b = M

Accuracy = 0.66

Precision = 0.662

Recall = 0.66

F-measure = 0.659

Spam/ham 10-fold Cross-validation

=== Summary ===

Correctly Classified Instances 2826 97.6841 %

Incorrectly Classified Instances 67 2.3159 %

Kappa statistic 0.9182

Mean absolute error 0.0231

Root mean squared error 0.1513

Relative absolute error 8.3139 %

Root relative squared error 40.6258 %

Total Number of Instances 2893

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.981 0.044 0.991 0.981 0.986 0.981 ham

0.956 0.019 0.909 0.956 0.932 0.989 spam

Weighted Avg. 0.977 0.04 0.978 0.977 0.977 0.982

=== Confusion Matrix ===

a b <-- classified as

2366 46 | a = ham

21 460 | b = spam

Accuracy = 0.977

Precision = 0.978

Recall = 0.977

F-measure = 0.977

J48 Spam/ham 10 cross fold

=== Summary ===

Correctly Classified Instances 2775 95.9212 %

Incorrectly Classified Instances 118 4.0788 %

Kappa statistic 0.8516

Mean absolute error 0.0478

Root mean squared error 0.1962

Relative absolute error 17.2267 %

Root relative squared error 52.6839 %

Total Number of Instances 2893

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.978 0.133 0.974 0.978 0.976 0.938 ham

0.867 0.022 0.885 0.867 0.876 0.938 spam

Weighted Avg. 0.959 0.115 0.959 0.959 0.959 0.938

=== Confusion Matrix ===

a b <-- classified as

2358 54 | a = ham

64 417 | b = spam

Accuracy = 0.959

Precision = 0.959

recall = 0.959

F-Measure = 0.959

BayesianLogisticRegression Spam/ham 10fold- cross validation

=== Summary ===

Correctly Classified Instances 2877 99.4469 %

Incorrectly Classified Instances 16 0.5531 %

Kappa statistic 0.98

Mean absolute error 0.0055

Root mean squared error 0.0744

Relative absolute error 1.9936 %

Root relative squared error 19.9743 %

Total Number of Instances 2893

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.998 0.023 0.995 0.998 0.997 0.988 ham

0.977 0.002 0.989 0.977 0.983 0.988 spam

Weighted Avg. 0.994 0.019 0.994 0.994 0.994 0.988

=== Confusion Matrix ===

a b <-- classified as

2407 5 | a = ham

11 470 | b = spam

Accuracy = 0.994

Precision = 0.994

recall = 0.994

F-Measure = 0.994

BayesianLogisticRegression Male/Female train/test

=== Summary ===

Correctly Classified Instances 32 64 %

Incorrectly Classified Instances 18 36 %

Kappa statistic 0.28

Mean absolute error 0.36

Root mean squared error 0.6

Relative absolute error 72 %

Root relative squared error 120 %

Total Number of Instances 50

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.72 0.44 0.621 0.72 0.667 0.64 F

0.56 0.28 0.667 0.56 0.609 0.64 M

Weighted Avg. 0.64 0.36 0.644 0.64 0.638 0.64

=== Confusion Matrix ===

a b <-- classified as

18 7 | a = F

11 14 | b = M

Accuracy = 0.64

Precision = 0.644

recall = 0.64

F-Measure = 0.638

J48 Male/Female train/test

=== Summary ===

Correctly Classified Instances 32 64 %

Incorrectly Classified Instances 18 36 %

Kappa statistic 0.28

Mean absolute error 0.3646

Root mean squared error 0.5829

Relative absolute error 72.9209 %

Root relative squared error 116.5794 %

Total Number of Instances 50

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.6 0.32 0.652 0.6 0.625 0.647 F

0.68 0.4 0.63 0.68 0.654 0.647 M

Weighted Avg. 0.64 0.36 0.641 0.64 0.639 0.647

=== Confusion Matrix ===

a b <-- classified as

15 10 | a = F

8 17 | b = M

Accuracy = 0.64

Precision = 0.641

recall = 0.64

F-Measure = 0.639

NBC:

Male/Female train/test: Spam/Ham 10- folds cross-validation

a b <-- classified as a b <-- classified as

18 7 | a = F 2366 46 | a = ham

10 15 | b = M 21 460 | b = spam

Accuracy = 0.66 Accuracy = 0.977

Precision = 0.662 Precision = 0.978

Recall = 0.66 Recall = 0.977

F-measure = 0.659 F-measure = 0.977

J48:

Male/Female train/test: Spam/Ham 10- folds cross-validation

a b <-- classified as a b <-- classified as

15 10 | a = F 2358 54 | a = ham

8 17 | b = M 64 417 | b = spam

Accuracy = 0.64 Accuracy = 0.959

Precision = 0.641 Precision = 0.959

recall = 0.64 recall = 0.959

F-Measure = 0.639 F-Measure = 0.959

BayesianLogisticRegression:

Male/Female train/test: Spam/Ham 10- folds cross-validation

a b <-- classified as a b <-- classified as

18 7 | a = F 2407 5 | a = ham

11 14 | b = M 11 470 | b = spam

Accuracy = 0.64 Accuracy = 0.994

Precision = 0.644 Precision = 0.994

recall = 0.64 recall = 0.994

F-Measure = 0.638 F-Measure = 0.994

As you can see for the Male/Female data the NBC is the best classifier although it is close, this is because it is very hard to say if an email is male or female this gives a lot of variation and logistic regression is not very good when that happens. Decision trees are in general not ideal for text which you can see in the results as well since it is the worst in all cases. This is why for determining if a certain text is written by a Male or Female NBC is the best solution out of these three. Spam and Ham is a lot easier to separate, the line between them is a lot clearer which makes it so that Logistic regression becomes really good and you can see that in the result as it is nearly perfect, while the others still have errors >2% of the time.