**Data set**

* Random 10% sample of “spambase” email dataset (57 attributes).
* 480 emails in sample

**Table & chart of results**

|  |  |
| --- | --- |
| **Stacking Ensemble** | 0.972019248 |
| **Random Forest** | 0.971147944 |
| **Gen. Boosting** | 0.970256837 |
| **W Avg Ensemble** | 0.968712252 |
| **Avg Ensemble** | 0.967543912 |
| **Sparse Log Reg** | 0.961137844 |
| **SVM** | 0.917641934 |
| **Rule-Based Method** | 0.880175845 |
| **Decision Tree** | 0.879680786 |
| **Neural Net** | 0.847066279 |
| **Gaussian Pr** | 0.753816907 |
| **k-NN** | 0.718875621 |

**Sparse Logistic Regression (**cv.glmnet()**)**

(ROC)

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

AUC = 0.96 (0.94 - 0.98)

sensitivity = 0.91 (0.86 - 0.95)

specificity = 0.91 (0.87 - 0.94)

[1] 0.9604249589



**Random Forest (**randomForest()**)**

(ROC)

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

AUC = 0.97 (0.95 - 0.99)

sensitivity = 0.93 (0.88 - 0.96)

specificity = 0.93 (0.89 - 0.96)

[1] 0.9709301174



**Generalised Boosting Method (**gbm()**)**

(ROC)

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

AUC = 0.97 (0.95 - 0.99)

sensitivity = 0.92 (0.87 - 0.95)

specificity = 0.92 (0.88 - 0.95)

[1] 0.9699399988

**Gaussian Process (**gausspr()**)**

(ROC)

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

AUC = 0.75 (0.71 - 0.8)

sensitivity = 0.68 (0.61 - 0.75)

specificity = 0.68 (0.62 - 0.74)

[1] 0.7536089824



**Support Vector Machine (**ksvm()**)**

(ROC)

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

AUC = 0.92 (0.89 - 0.95)

sensitivity = 0.86 (0.8 - 0.91)

specificity = 0.86 (0.81 - 0.9)

[1] 0.9195231589



**Neural Network (**neuralnet()**)**

(ROC)

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

AUC = 0.82 (0.78 - 0.86)

sensitivity = 0.77 (0.7 - 0.83)

specificity = 0.7 (0.64 - 0.75)

[1] 0.8232836294



**Decision Tree (**C5.0()**)**

(ROC)

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

AUC = 0.88 (0.84 - 0.92)

sensitivity = 0.85 (0.78 - 0.89)

specificity = 0.87 (0.82 - 0.9)

[1] 0.8796807858



**Rule-Based Method (**C5.0()**)**

(ROC)

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

AUC = 0.88 (0.85 - 0.91)

sensitivity = 0.84 (0.78 - 0.89)

specificity = 0.84 (0.79 - 0.88)

[1] 0.8801758451



**k-Nearest Neighbour (**knn3()**)**

(ROC)

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

AUC = 0.72 (0.67 - 0.77)

sensitivity = 0.66 (0.58 - 0.73)

specificity = 0.64 (0.58 - 0.7)

[1] 0.7188756213



**Average Ensemble (**self**)**

(ROC)

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

AUC = 0.97 (0.95 - 0.99)

sensitivity = 0.93 (0.88 - 0.96)

specificity = 0.93 (0.89 - 0.96)

[1] 0.9682964019



**Weighted Average Ensemble (**self**)**

(ROC)

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

AUC = 0.97 (0.95 - 0.99)

sensitivity = 0.93 (0.88 - 0.96)

specificity = 0.93 (0.89 - 0.96)

[1] 0.9686528446



**Stacking (**self**)**

(ROC)

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

AUC = 0.97 (0.96 - 0.99)

sensitivity = 0.93 (0.89 - 0.97)

specificity = 0.94 (0.9 - 0.96)

[1] 0.9720192479

