

1. What were the accuracy and confusion matrix for your model?

**The accuracy is 0.9578.**

Confusion Matrix and Statistics

Reference

Prediction 0 1

0 4496 89

1 548 9950

**Accuracy : 0.9578**

95% CI : (0.9544, 0.9609)

No Information Rate : 0.6656

P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.9029

McNemar's Test P-Value : < 2.2e-16

Sensitivity : 0.8914

Specificity : 0.9911

Pos Pred Value : 0.9806

Neg Pred Value : 0.9478

Prevalence : 0.3344

Detection Rate : 0.2981

Detection Prevalence : 0.3040

Balanced Accuracy : 0.9412

'Positive' Class : 0

2. Using the confusion matrix, calculated the percentage of spam messages erroneously classified as ham and ham messages erroneously classified as spam. Are the percentages equal? Often times, one type of error is preferably for a business or user. In this case, it would be worse for an important email to end up in the spam folder than for a spam message to end up in the inbox.

Prediction	Reference	
	0	1
	0	4496 89
1	548	9950

Total ham messages:  $4496 + 548 = 5044$

the Percentage is spam messages erroneously classified as ham:  $(548 / 5044) * 100 = \mathbf{10.86\%}$

total spam messages:  $89 + 9950 = 10039$

the Percentage is ham messages erroneously classified as spam:  $(89 / 10039) * 100 = \mathbf{0.89\%}$