

Evaluating the performance of the 3 classification models

I evaluate the performance of the models using the `confusionMatrix()` function from the `caret` package. I use this function because it gives you a lot of feedback, including a confusion matrix and 13 other results based off the confusion matrix. The table it returns is easy to understand, and the 13 other results are also quick and easy to understand (assuming you are familiar with the terms).

Confusion Matrix

Ridge Regression

```
Confusion Matrix and Statistics

      Reference
Prediction Benign Malignant
Benign      70         0
Malignant   4         50

      Accuracy : 0.9677
      95% CI : (0.9195, 0.9911)
      No Information Rate : 0.5968
      P-Value [Acc > NIR] : <2e-16

      Kappa : 0.9338

      Mcnemar's Test P-value : 0.1336

      Sensitivity : 1.0000
      Specificity : 0.9459
      Pos Pred value : 0.9259
      Neg Pred value : 1.0000
      Prevalence : 0.4032
      Detection Rate : 0.4032
      Detection Prevalence : 0.4355
      Balanced Accuracy : 0.9730

      'Positive' Class : Malignant
```

LASSO Regression

```
Confusion Matrix and Statistics

      Reference
Prediction Benign Malignant
Benign      69         0
Malignant   5         50

      Accuracy : 0.9597
      95% CI : (0.9084, 0.9868)
      No Information Rate : 0.5968
      P-Value [Acc > NIR] : < 2e-16

      Kappa : 0.9176

      Mcnemar's Test P-value : 0.07364

      Sensitivity : 1.0000
      Specificity : 0.9324
      Pos Pred value : 0.9091
      Neg Pred value : 1.0000
      Prevalence : 0.4032
      Detection Rate : 0.4032
      Detection Prevalence : 0.4435
      Balanced Accuracy : 0.9662

      'Positive' Class : Malignant
```

Elastic Net

```
Confusion Matrix and Statistics

      Reference
Prediction Benign Malignant
Benign      71         0
Malignant   3         50

      Accuracy : 0.9758
      95% CI : (0.9309, 0.995)
      No Information Rate : 0.5968
      P-Value [Acc > NIR] : <2e-16

      Kappa : 0.9502

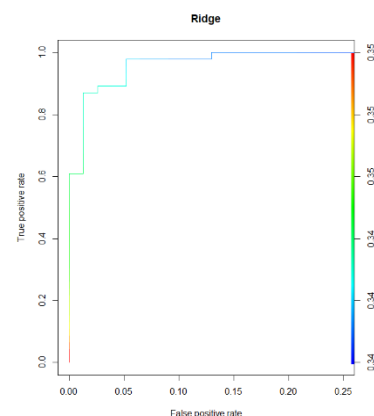
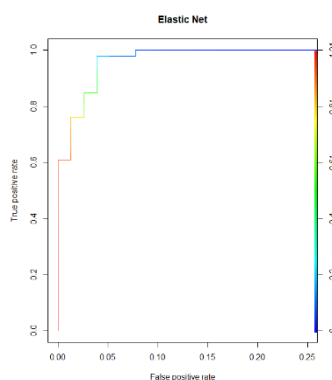
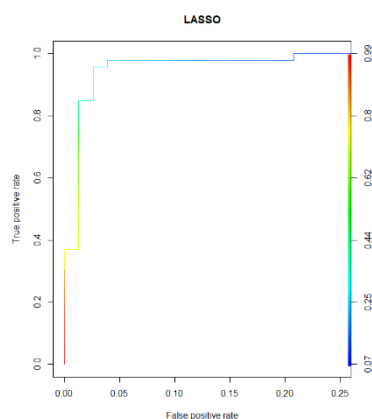
      Mcnemar's Test P-value : 0.2482

      Sensitivity : 1.0000
      Specificity : 0.9595
      Pos Pred value : 0.9434
      Neg Pred value : 1.0000
      Prevalence : 0.4032
      Detection Rate : 0.4032
      Detection Prevalence : 0.4274
      Balanced Accuracy : 0.9797

      'Positive' Class : Malignant
```

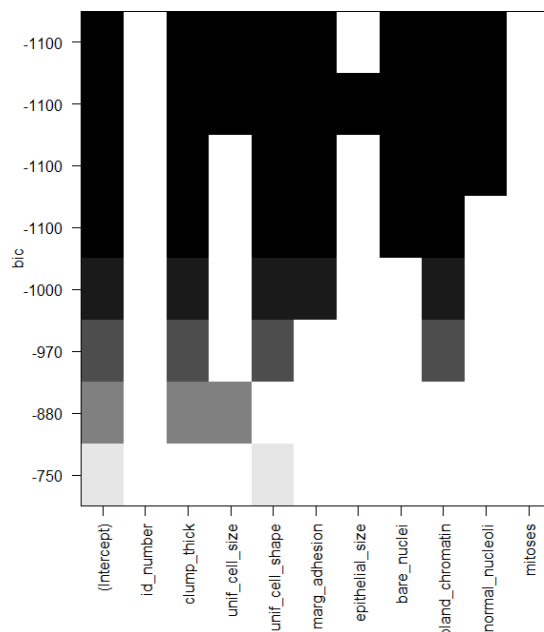
The 3 models were roughly as accurate as one another, with accuracy of between 96-98%. The difference between the models is statistically insignificant.

ROC plot



We can see from the ROC plots that the Elastic Net model is the best at minimizing the false positive rate while maintaining a high true positive rate. This is consistent with what we see in our confusion matrix, which showed the Elastic Net model to be most accurate. I must stress that this is an insignificant difference and well within the margin of error.

BIC plot & Coefficients



The BIC plot shows what variables influenced the classification the most. As we can see, the intercept, clump_thick, unif_cell_shape, and bland_chromatin are the most strongly correlated to the classification. By comparing the BIC to the coefficient matrix, we see that the Elastic Net model was the closest to the BIC plot. This would suggest that it is more accurate than the other 2 models, more so than its slightly better accuracy in the Confusion Matrix would suggest.

```
> Betas
10 x 3 sparse Matrix of class "dgCMatrix"
              Ridge      Lasso Elastic Net
(Intercept)  -0.6675306569 -3.36465628 -9.4291188
clump_thick   0.0004747274  0.12604058  0.4320427
unif_cell_size 0.0005088763  0.13508914  0.2418572
unif_cell_shape 0.0005298392  0.21843027  0.3359744
marg_adhesion  0.0004913624  0.07131694  0.3065716
epithelial_size 0.0005827808  .          0.1181045
bare_nuclei    0.0004785993  0.02618180  0.3307079
bland_chromatin 0.0005977216  0.14292094  0.3809395
normal_nucleoli 0.0004542458  0.01772895  0.1151563
mitoses        0.0004971262  .          0.2538592
```

Summary

There were no false negatives, which is very important because we are dealing with potentially life-or-death here. It's good that the models err on the side of caution and have more false positives. However, the models should be limited to assistive use only. They are not accurate enough to replace the diagnoses of a trained doctor. I recommend the Elastic Net model because it consistently outperformed the other 2 models. The takeaway message from this is that these models are very accurate, but not quite accurate enough for medical usage.