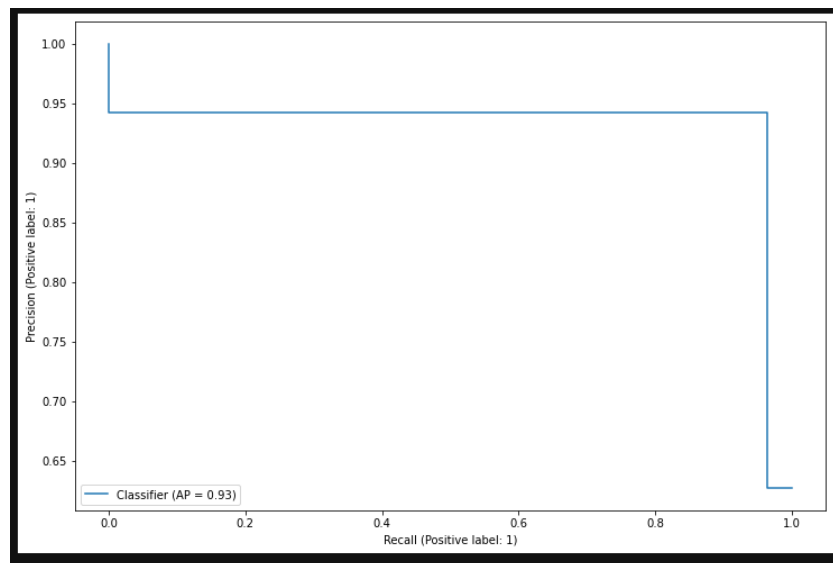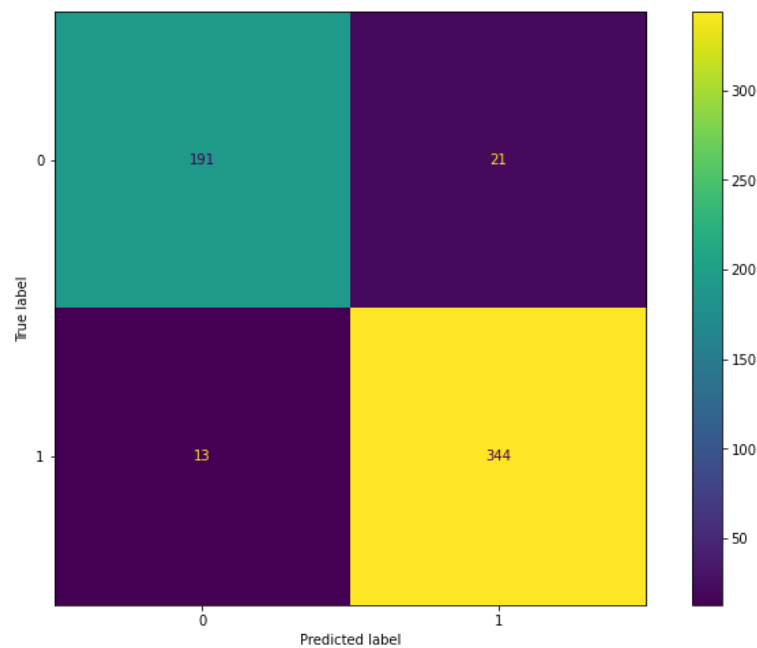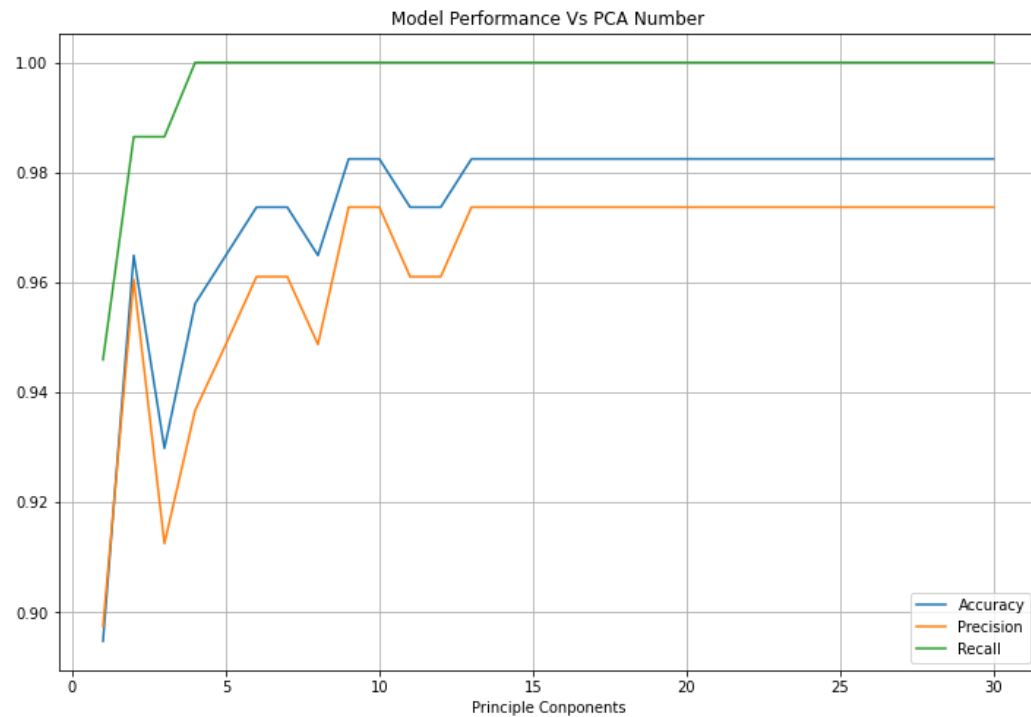# Homework 3

GitHub: https://github.com/pballou/ECGR_4105/tree/master/Homework/homework_3

1. Naïve Bayes classifier on cancer dataset. Pretty accurate overall, especially for a model only based on probabilities.

```
              precision    recall  f1-score   support

         0.0       0.94      0.90      0.92       212
         1.0       0.94      0.96      0.95       357

    accuracy                           0.94       569
   macro avg       0.94      0.93      0.94       569
weighted avg       0.94      0.94      0.94       569
```
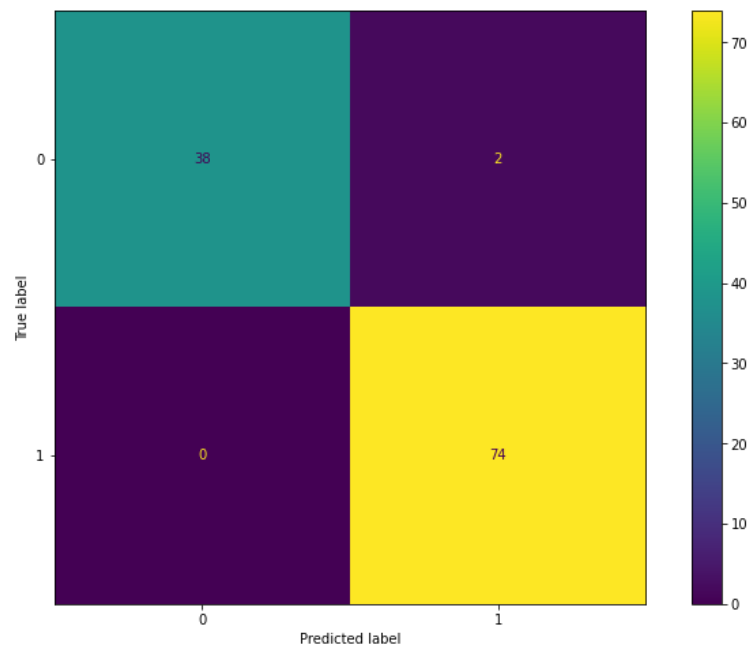
2. Logistic regression with PCA (k=1:30) on cancer dataset. The optimal k is 9 in this situation.
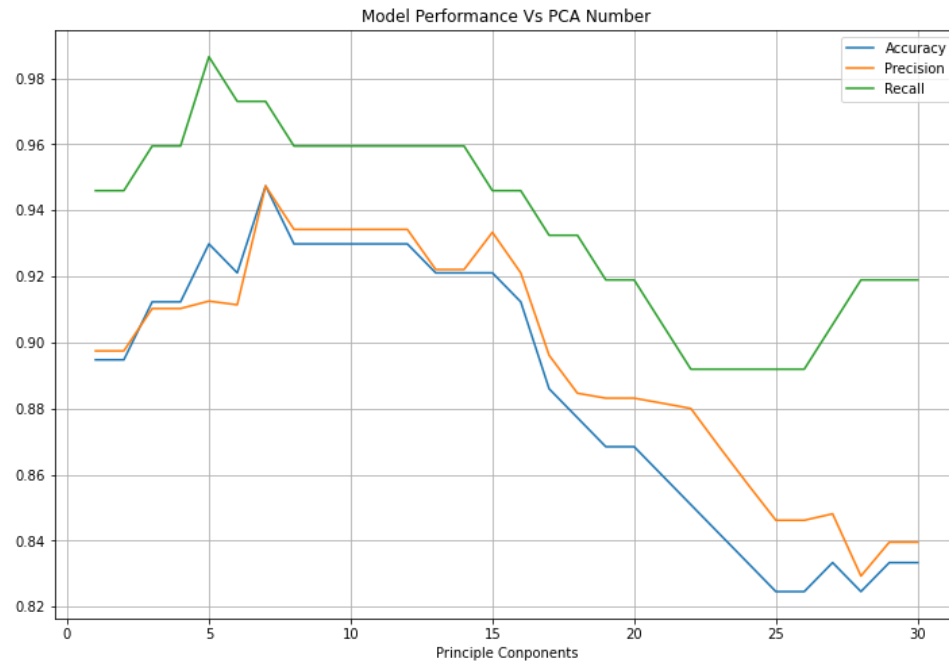

Results with k=9:

	precision	recall	f1-score	support
0	1.00	0.95	0.97	40
1	0.97	1.00	0.99	74
accuracy			0.98	114
macro avg	0.99	0.97	0.98	114
weighted avg	0.98	0.98	0.98	114

3. Naïve Bayes with PCA (k=1:30) on cancer dataset. The optimal k is 7 in this situation.

Model Performance Vs PCA Number

Results with k=7:

	precision	recall	f1-score	support
0	0.95	0.90	0.92	40
1	0.95	0.97	0.96	74
accuracy			0.95	114
macro avg	0.95	0.94	0.94	114
weighted avg	0.95	0.95	0.95	114