

data shape:
(1150, 20)

data head

accuracy score
accuracy of different models
0.4956521739130435
0.5652173913043478
0.5347826086956522
0.5043478260869565

cross validation
0.6243478260869566

best parameters
{'max_depth': 20, 'max_features': 10, 'min_samples_leaf': 20}
0.6521739130434783

accuracy of the model
0.6034782608695652

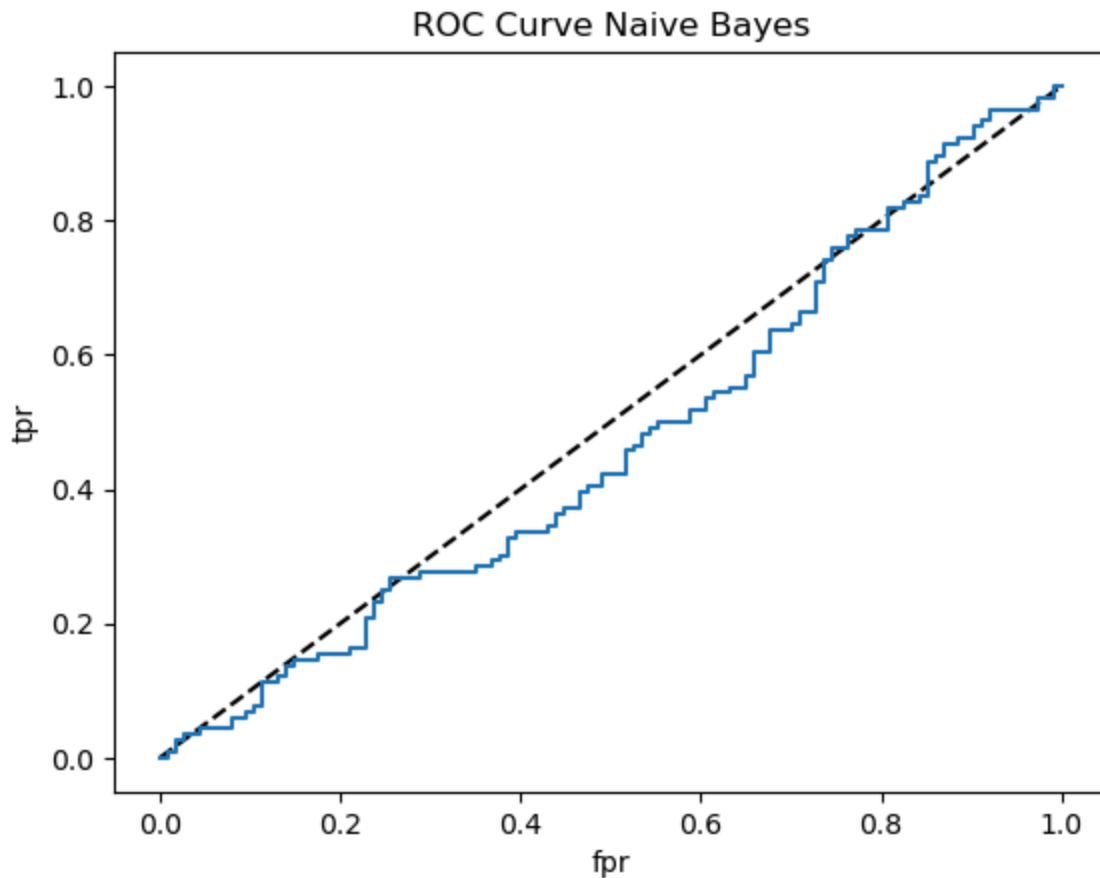
accuracy of gaussian cross validation
0.5947826086956522

```
[[500 39]
 [422 189]]
      precision  recall f1-score  support

0      0.54      0.93      0.68      539
1      0.83      0.31      0.45      611

      accuracy                0.60      1150
macro avg      0.69      0.62      0.57      1150
weighted avg    0.69      0.60      0.56      1150
```

area under the curve score, \ (how reliable our model is\)
0.46748336358136716



accuracy of cross val of pipeline
0.6182608695652174

optimal parameters for accuracy with pipeline model
{'kn__n_neighbors': 23, 'pca__n_components': 14}
0.6617391304347826

accuracy of cross validation on gSearch2
0.6417391304347827

accuracy of cross validation on gSearch3
0.717391304347826

accuracy of cross validation on gSearch4
0.7278260869565217

classification report

	precision	recall	f1-score	support
0	0.54	0.93	0.68	539
1	0.83	0.31	0.45	611

accuracy			0.60	1150
macro avg	0.69	0.62	0.57	1150
weighted avg	0.69	0.60	0.56	1150

classification report with updated parameters

	precision	recall	f1-score	support
--	-----------	--------	----------	---------

0	0.66	0.71	0.68	539
1	0.72	0.67	0.70	611

accuracy			0.69	1150
macro avg	0.69	0.69	0.69	1150
weighted avg	0.69	0.69	0.69	1150

Positive for disease