

Format:****Name of Classifier**

Notes:

- notes

No NAN

Accuracy Score: [paste here, decimal form]

[Insert Confusion Matrix]

Filled Data

Accuracy Score: [paste here, decimal form]

[Insert Confusion Matrix]

Gaussian Process Classifier

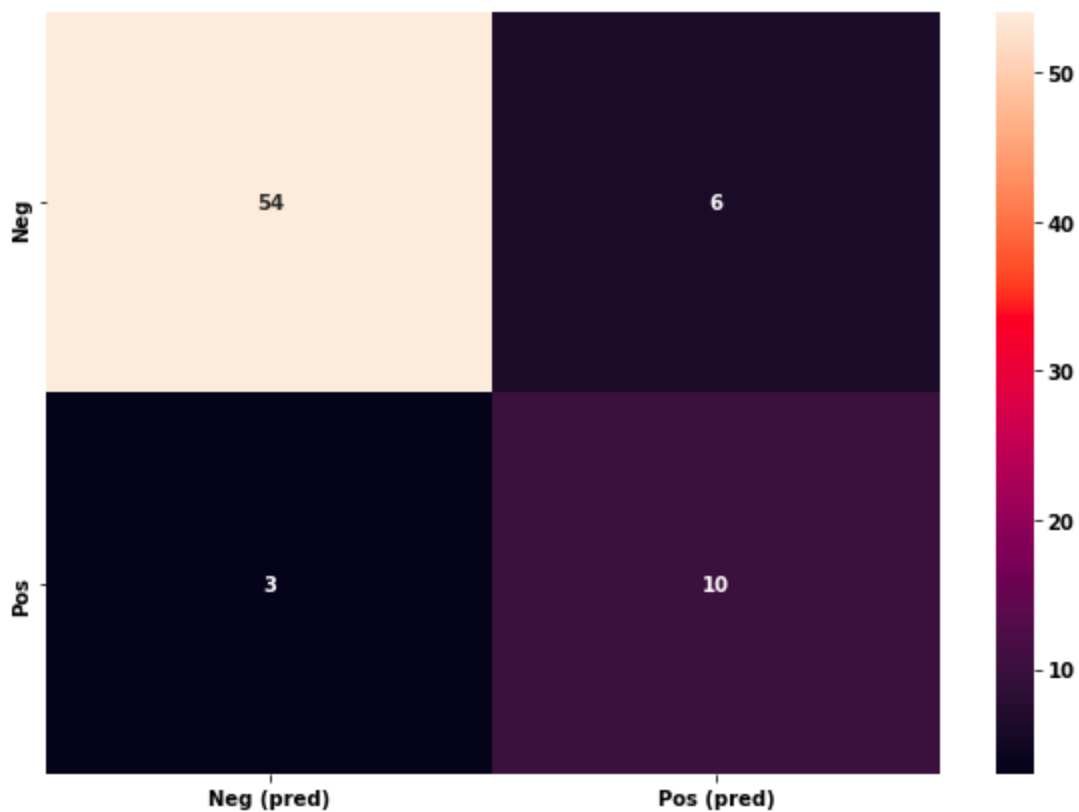
Notes:

- special parameter: $\text{kernel} = 1.0 * \text{RBF}(1.0) \rightarrow$ "The kernel used for prediction. In case of binary classification, the structure of the kernel is the same as the one passed as parameter but with optimized hyperparameters. In case of multi-class classification, a CompoundKernel is returned which consists of the different kernels used in the one-versus-rest classifiers."

(https://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessClassifier.html#sklearn.gaussian_process.GaussianProcessClassifier)

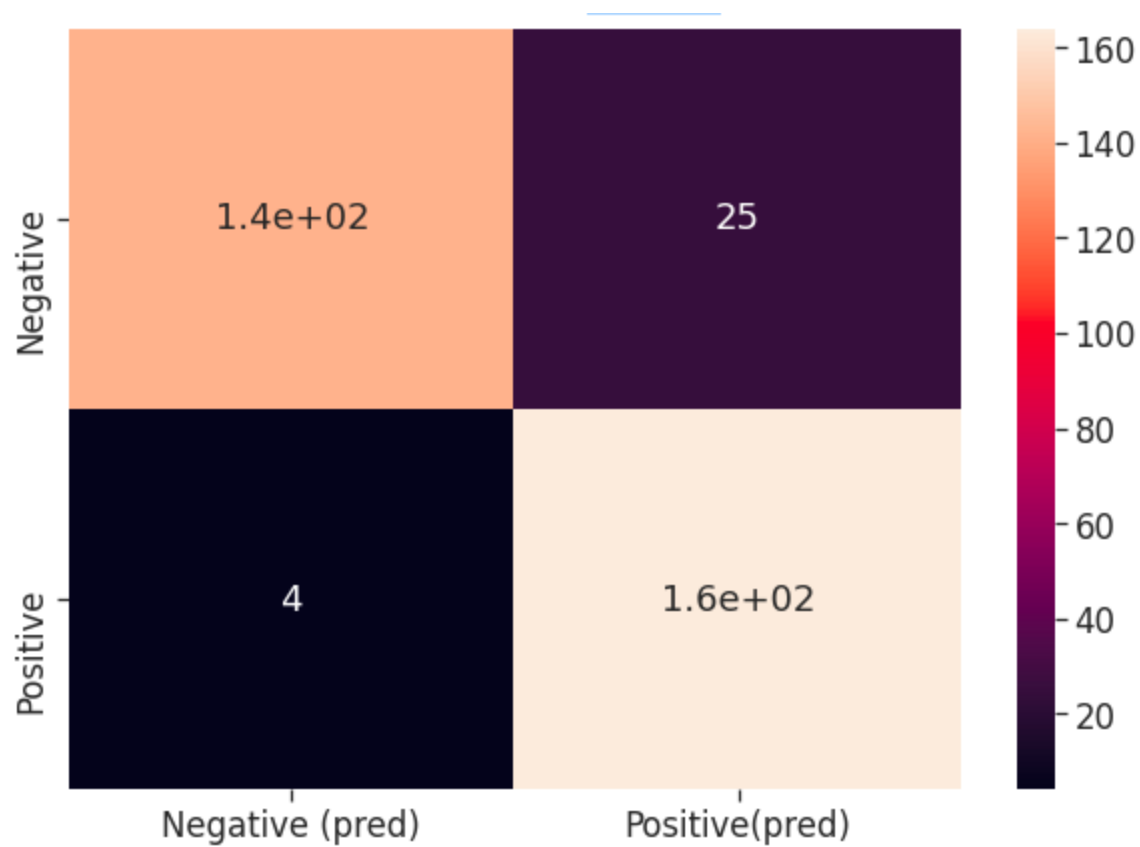
NO NAN

Accuracy Score: 0.9214876033057852



Gaussian Filled Data

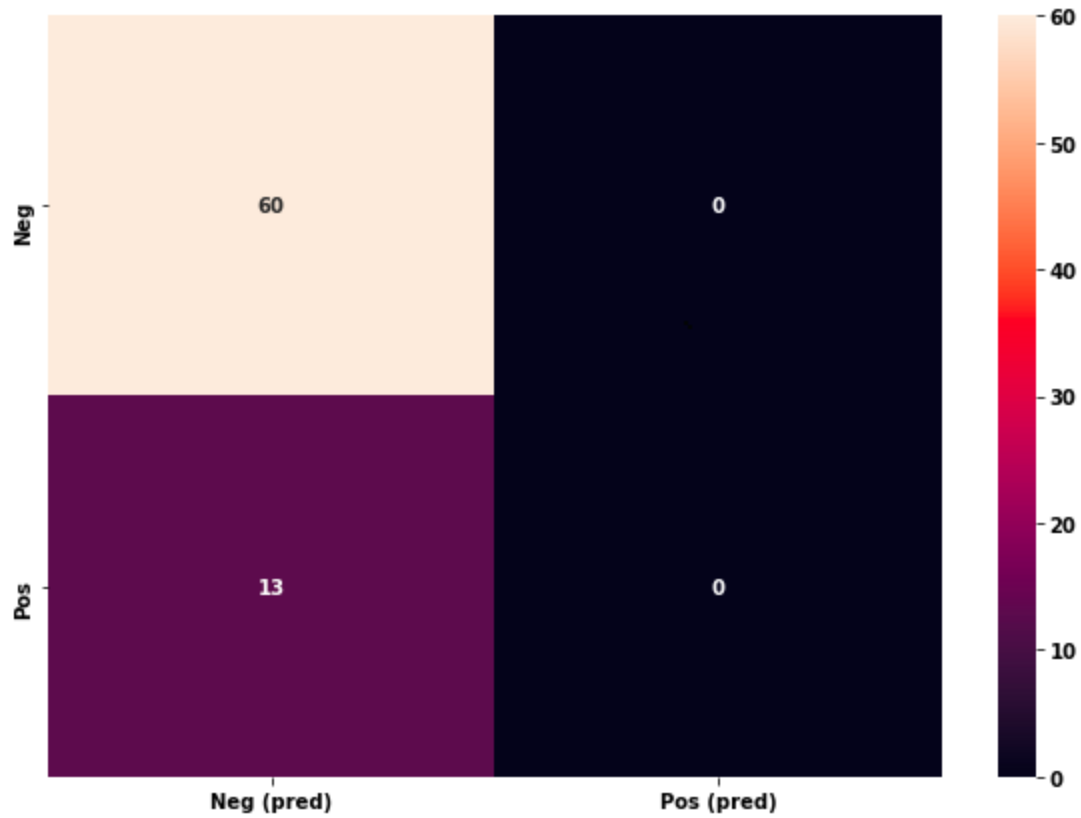
Accuracy Score: 0.913433



Support Vector Machines

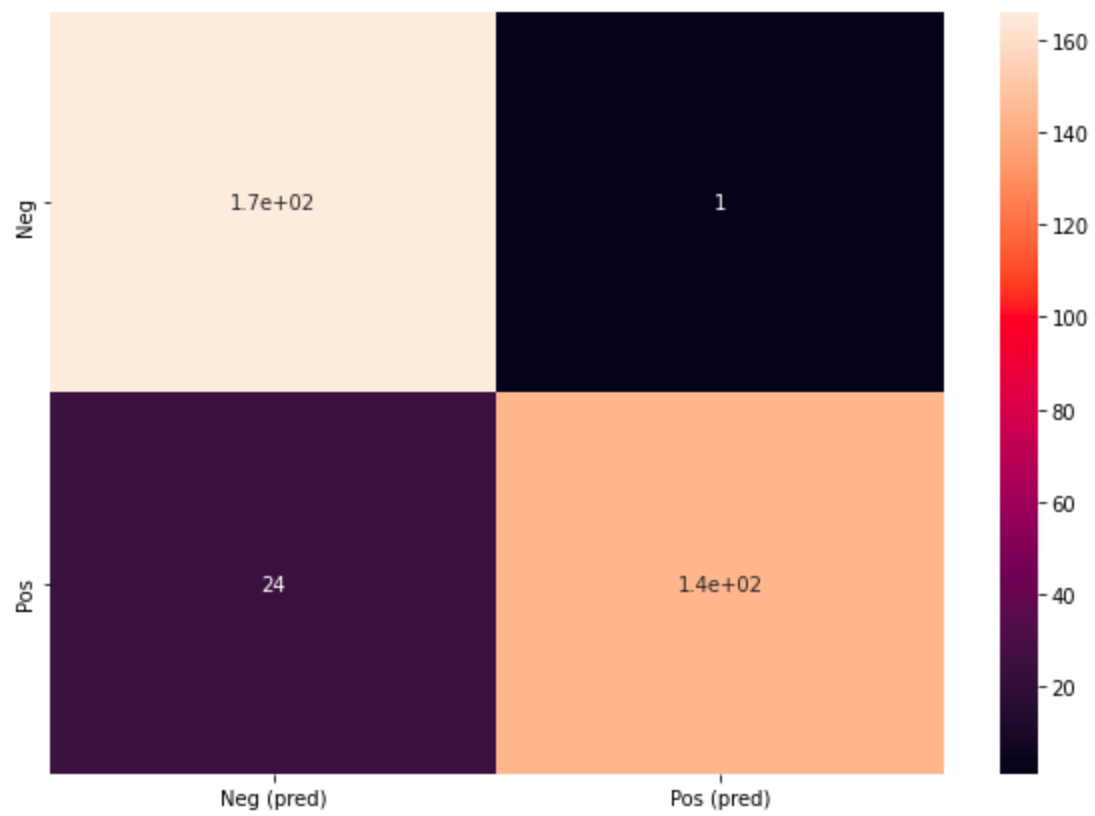
No NAN

Accuracy Score: 0.821917808219178



SVM Filled Data

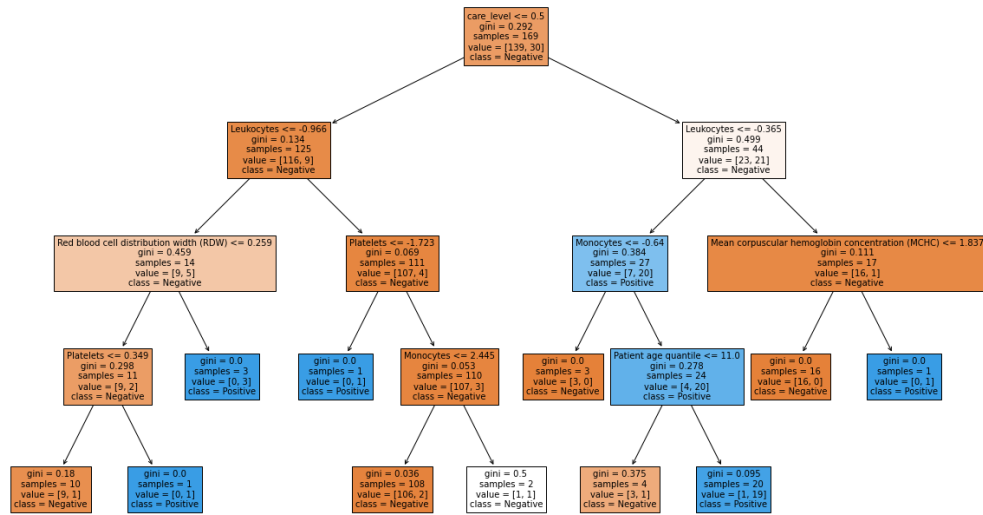
Accuracy Score: 0.9253731343283582



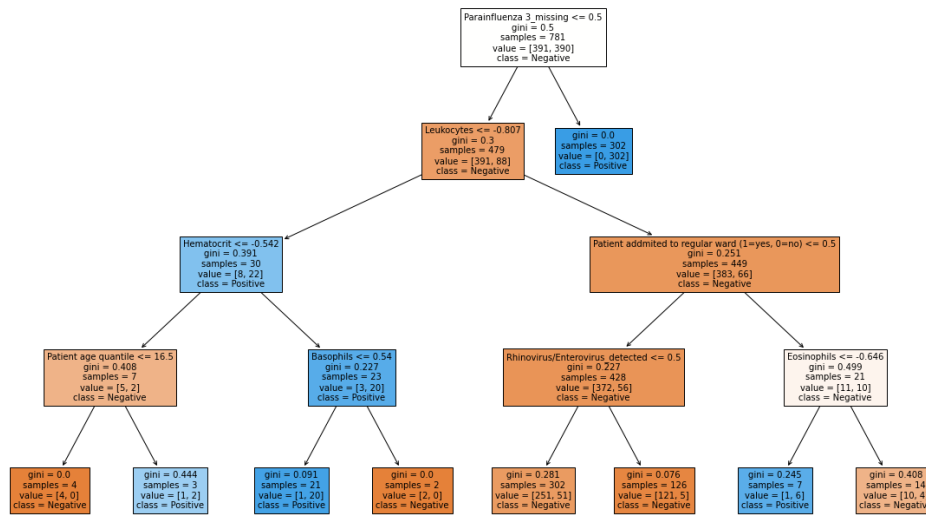
Decision Tree Classifier

Notes:

- max depth = 4 got the highest accuracy score
- decision tree for NAN removed:

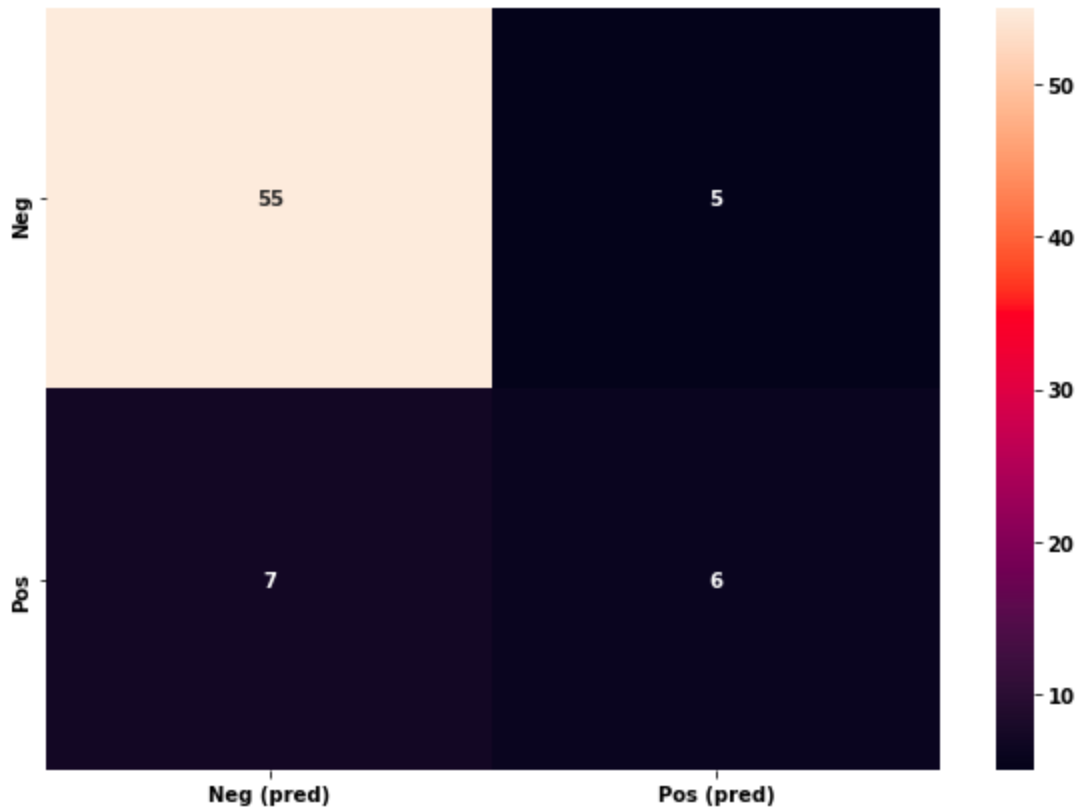


- decision tree for filled data:



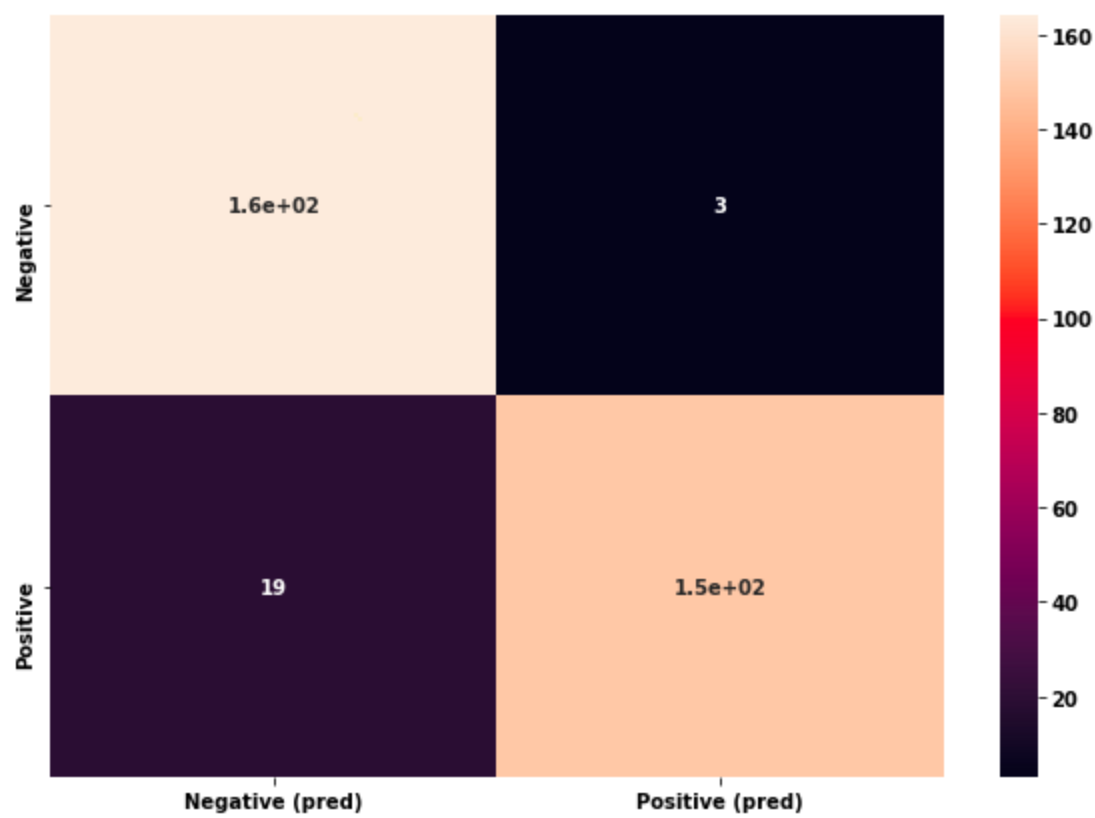
No NAN

Accuracy Score: 0.8356164383561644



DTC Filled Data

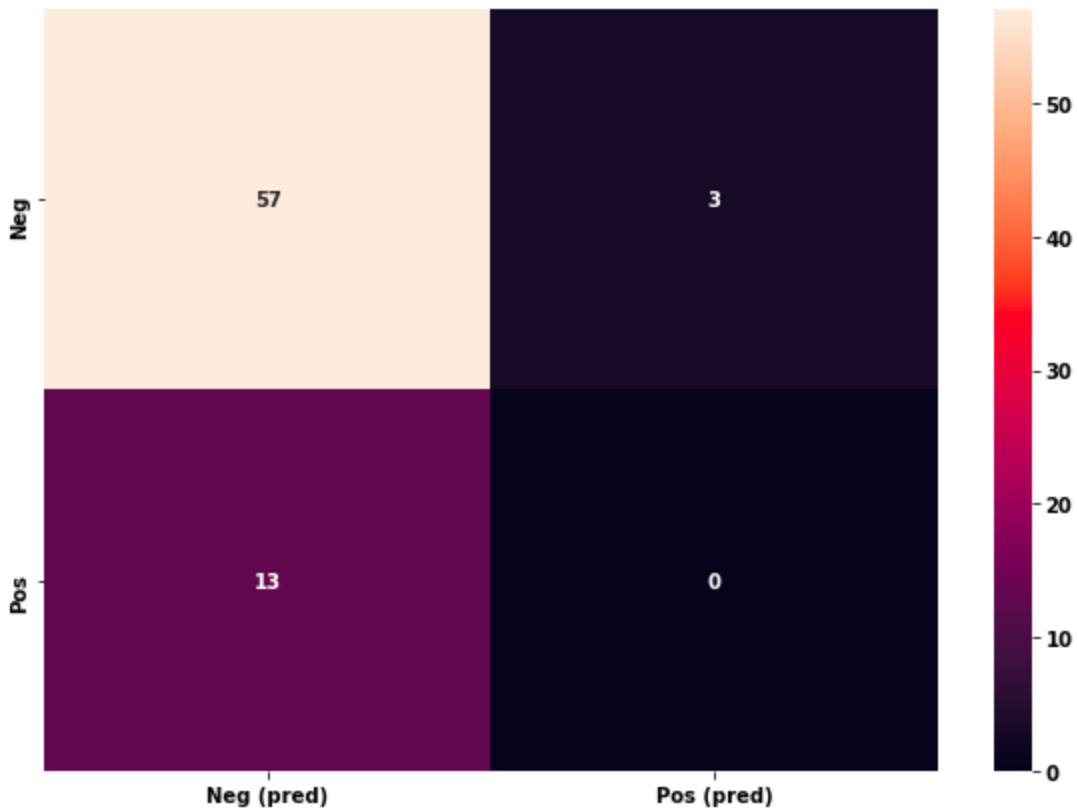
Accuracy Score: 0.9343283582089552



Quadratic Discriminant Analysis

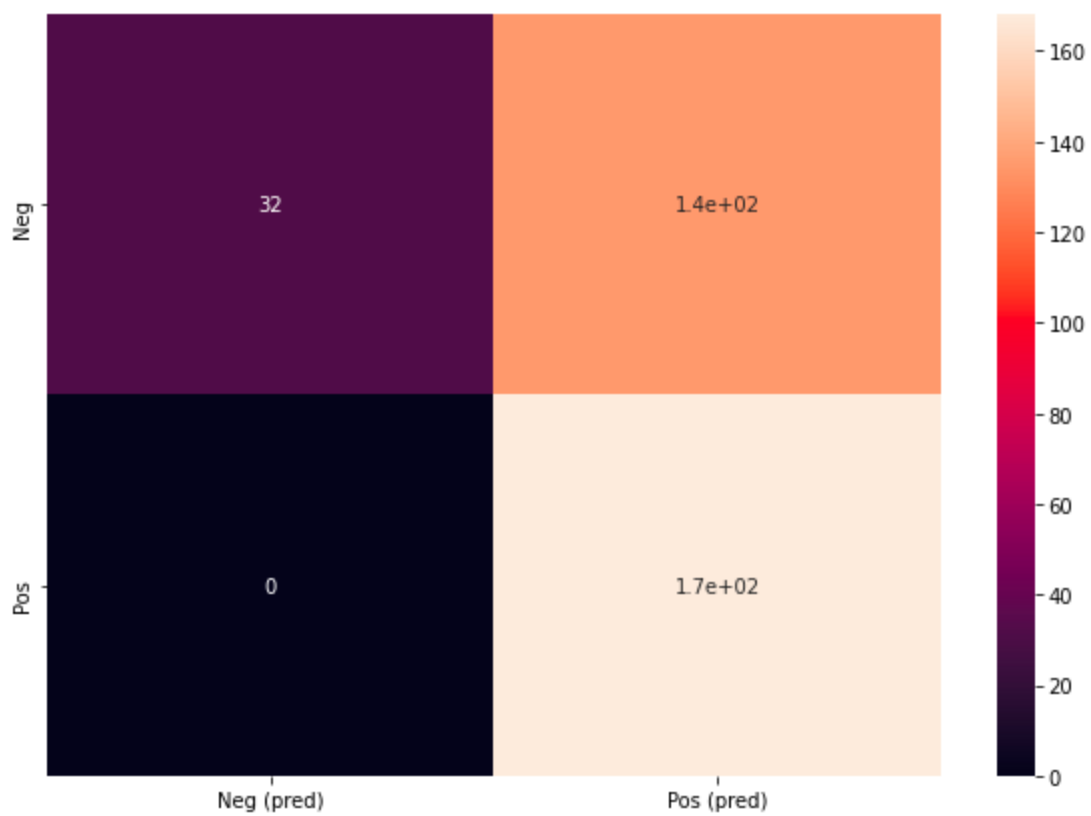
No NAN

Accuracy Score: 0.7808219178082192



QDA Filled Data

Accuracy Score: 0.5970149253731343



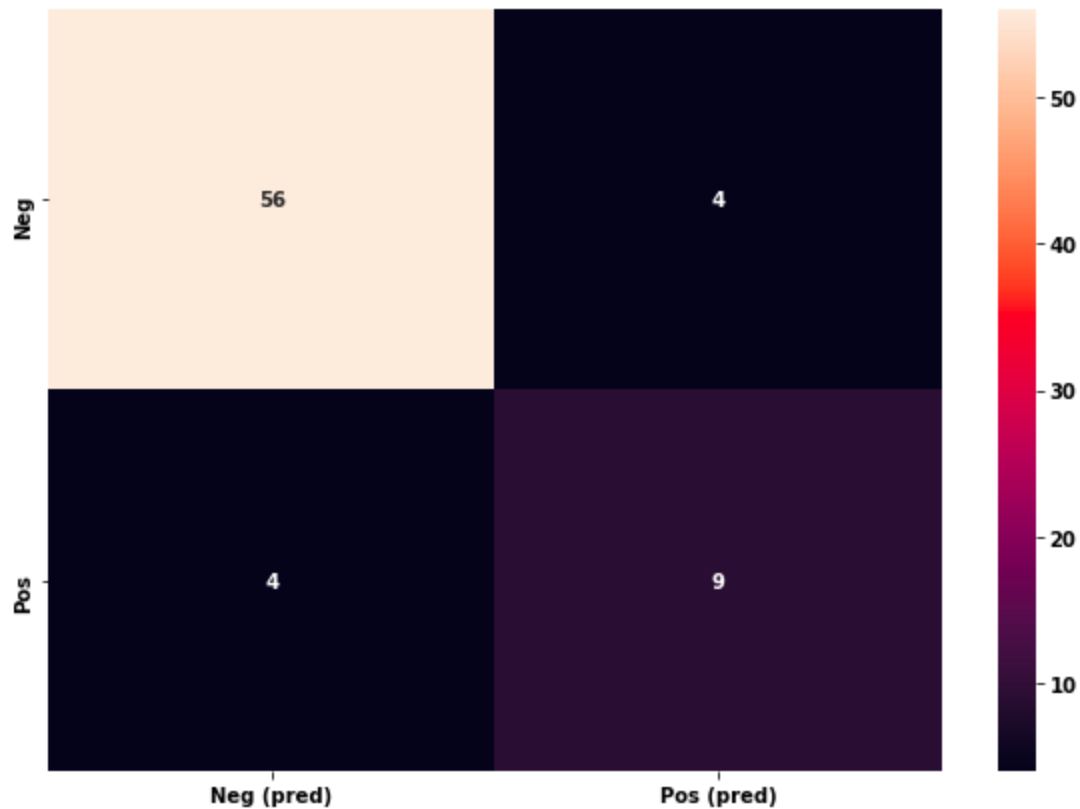
Multi-layer Perceptron Classifier

Notes:

- `max_iter = 100` → `max_iter` denotes the number of *epochs* → an epoch is one cycle through the full training dataset
(<https://analyticsindiamag.com/a-beginners-guide-to-scikit-learns-mlpclassifier/>)
- 100 epochs got the highest accuracy score after some messing around w the parameter; 50, 200+ were in the lower 80% accuracy range

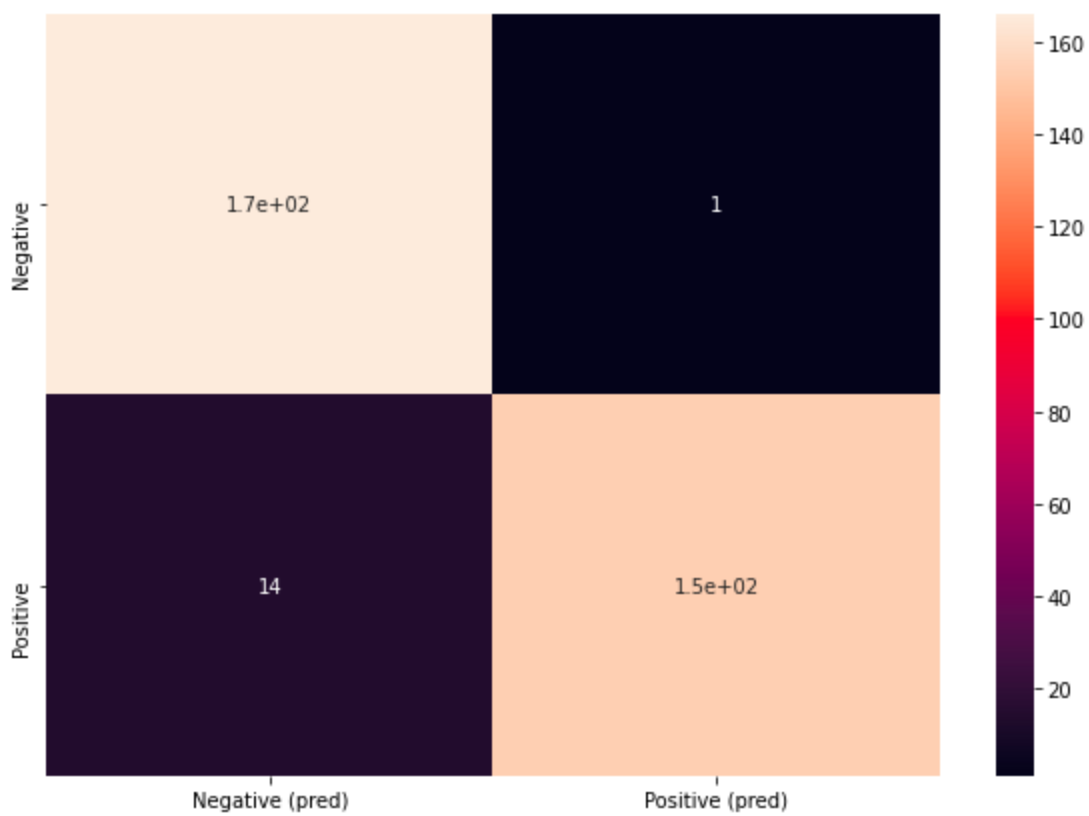
No NAN

Accuracy Score: 0.8904109589041096



MLP Filled Data

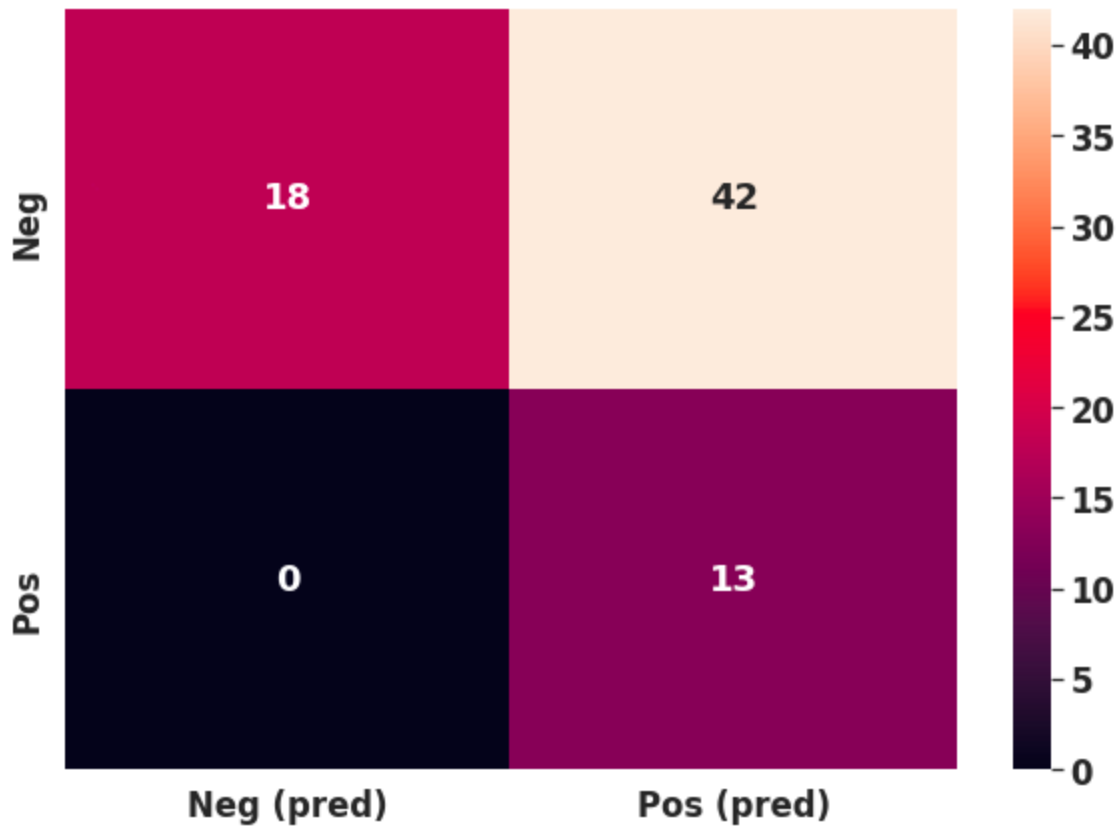
Accuracy Score: 0.955224



Naive Bayes

NB No NAN

Accuracy Score: 0.4246575342465753



NB Filled Data

Accuracy Score: 0.9253731343283582



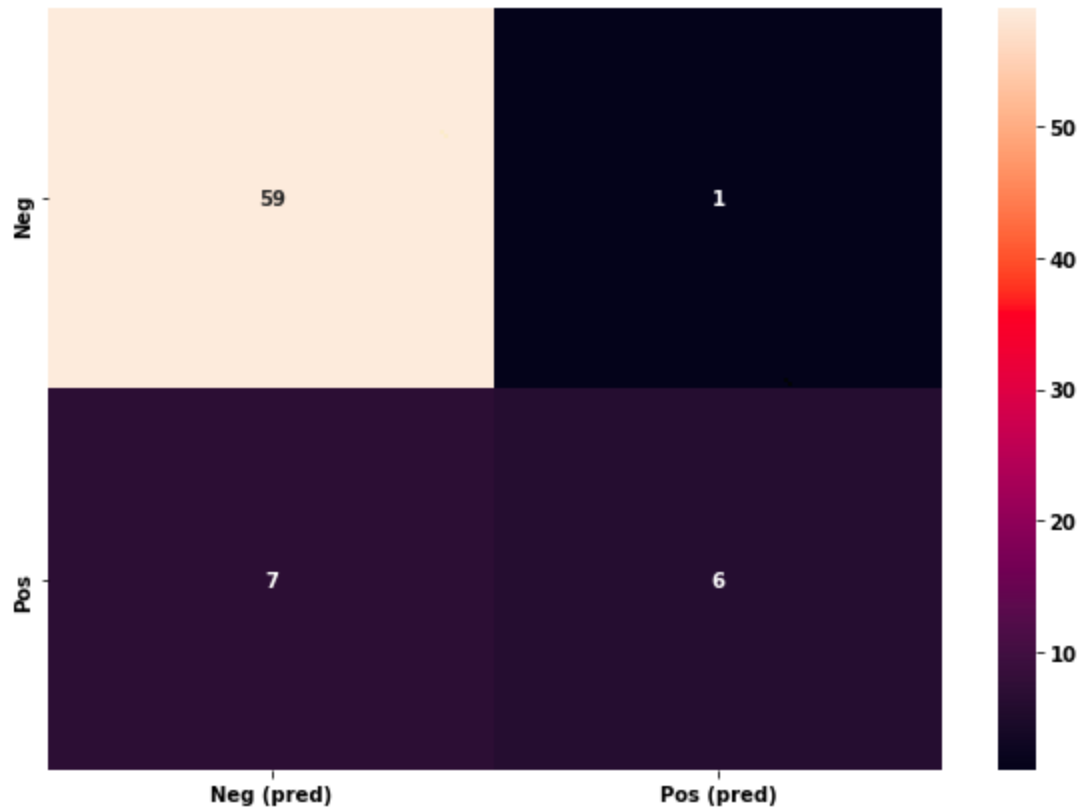
K-Nearest Neighbors

Notes:

- $k = 7$ had the highest accuracy

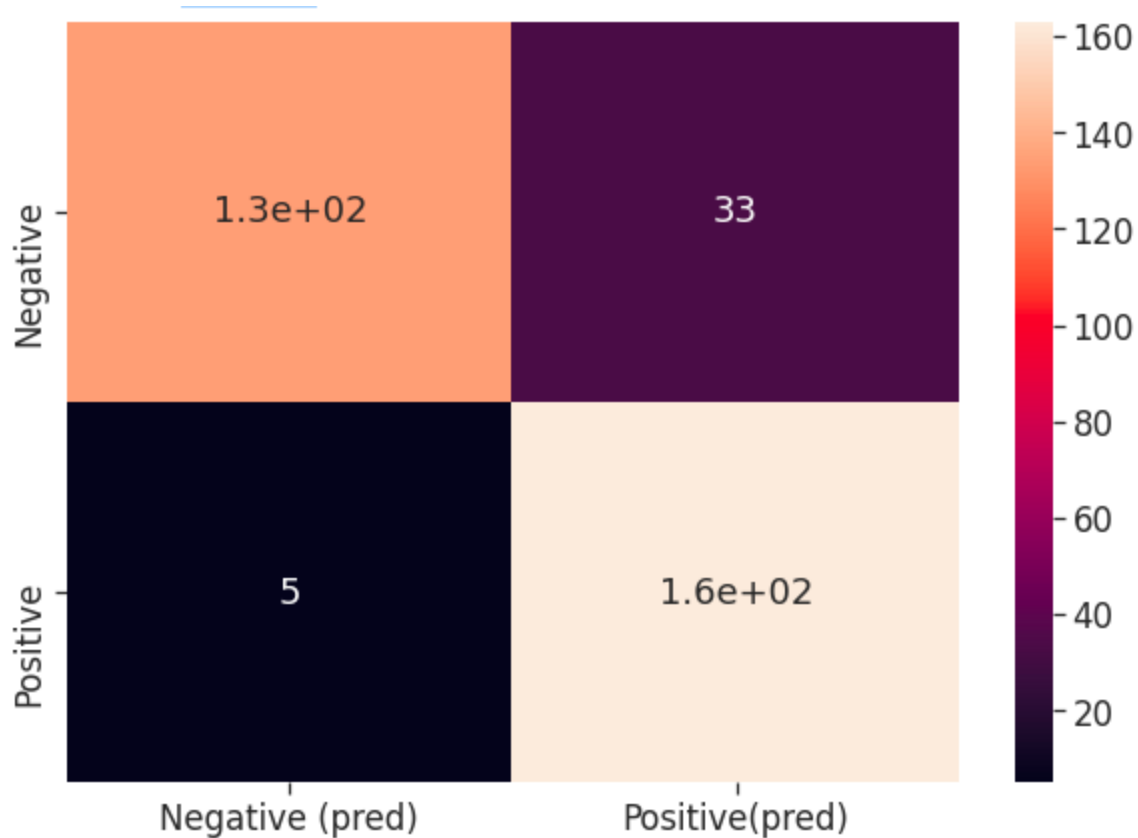
No NAN

Accuracy Score: 0.890411



KNN Filled Data

Accuracy Score: 0.886567



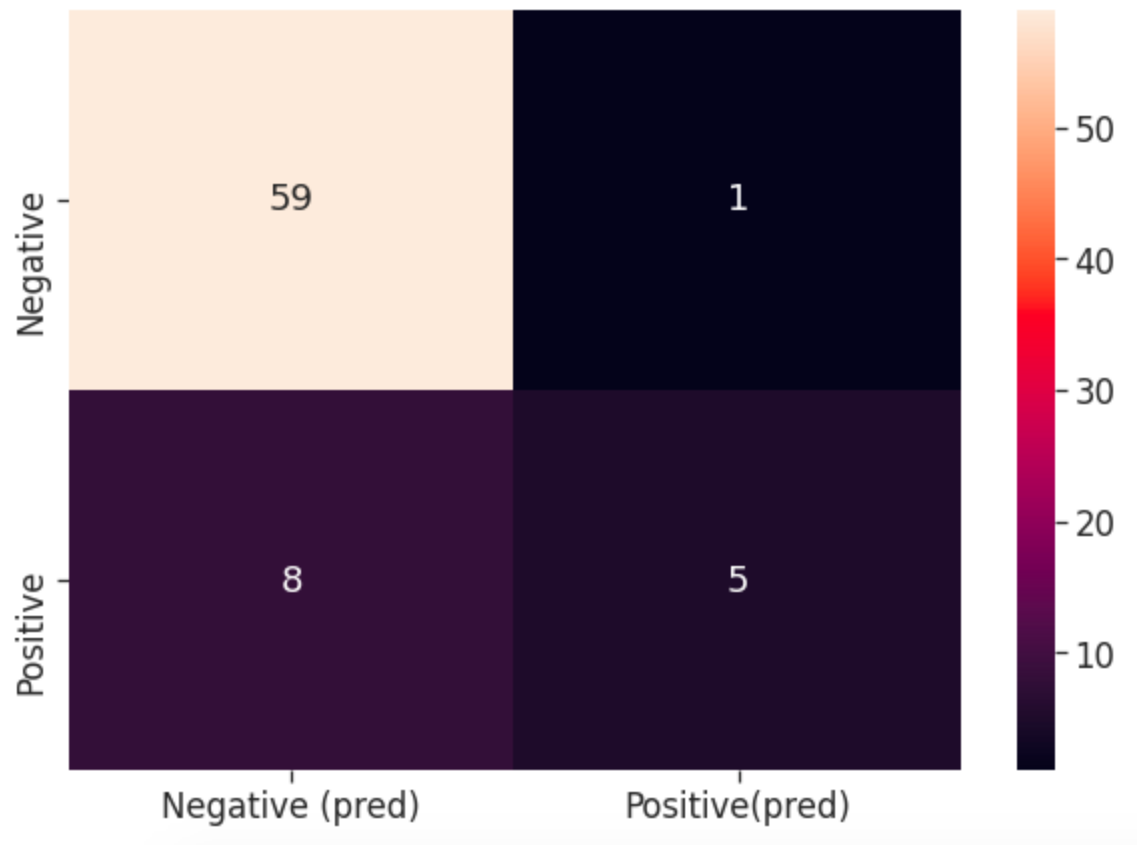
Random Forest Classifier

Notes:

- `n_estimators = 10` → “the number of trees you want to build before taking the maximum voting or averages of predictions; higher number of trees give you better performance but makes your code slower”

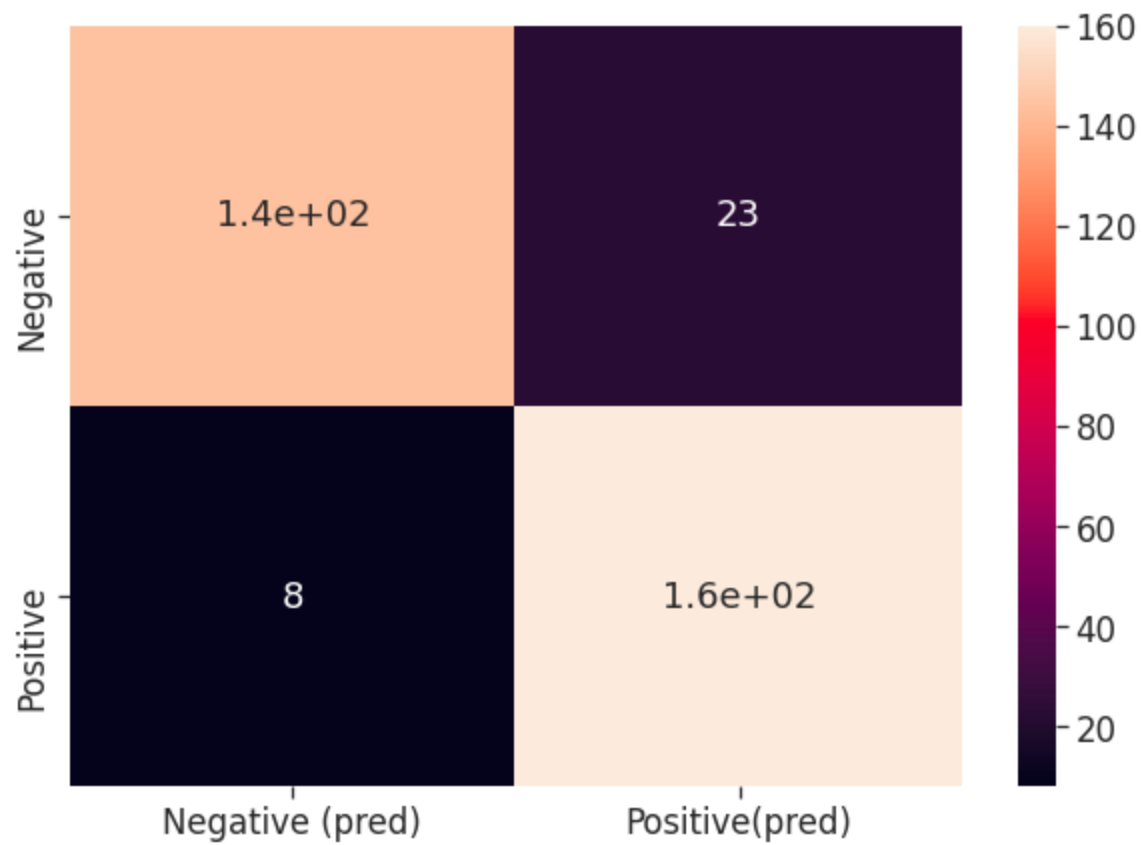
(<https://www.analyticsvidhya.com/blog/2015/06/tuning-random-forest-model/>)

Accuracy Score: 0.876712

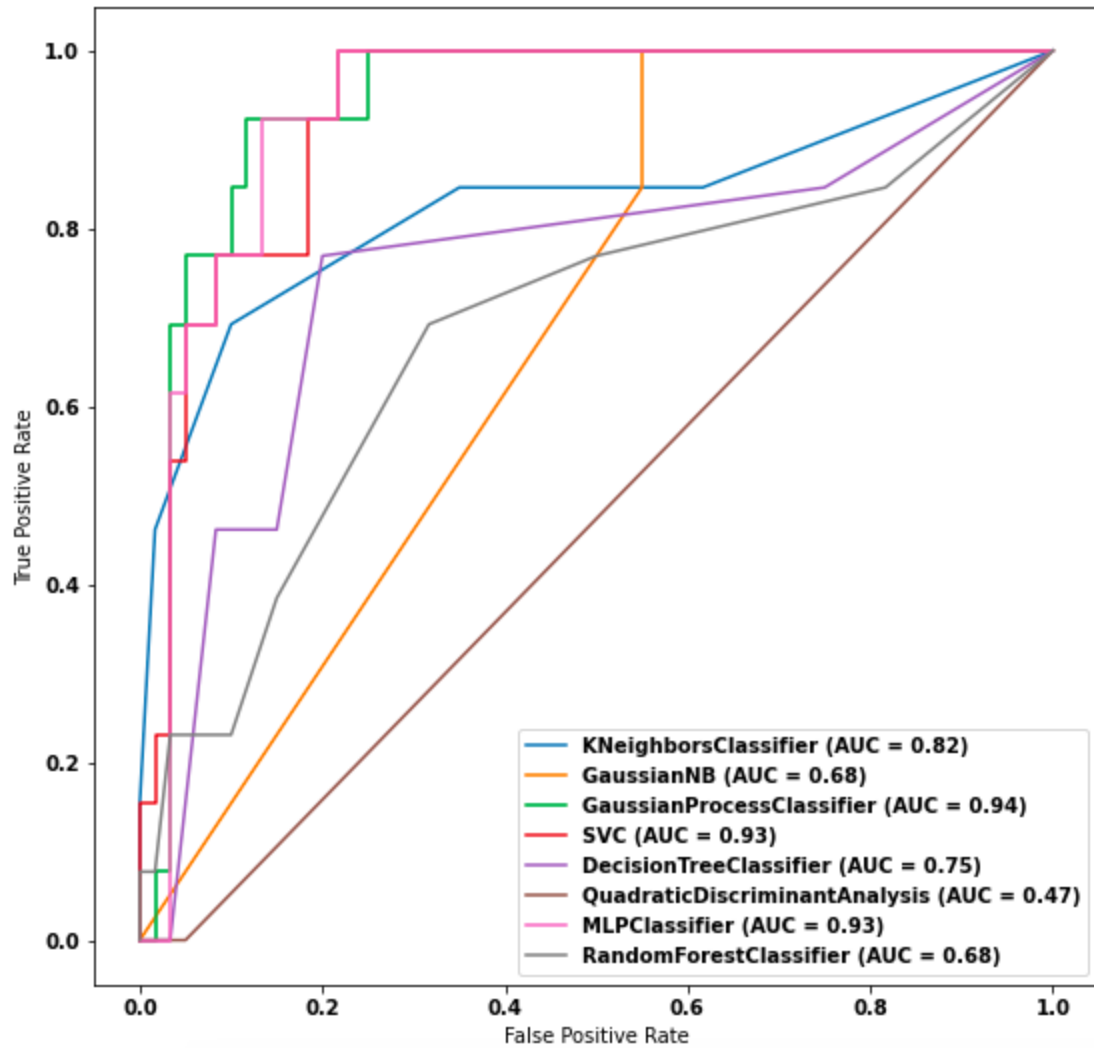


RFC Filled Data

Accuracy Score: 0.907463



ROC/AUC Compilation! (NoNAN below)



Filled Data ROC/AUC:

