

Classifier Evaluation

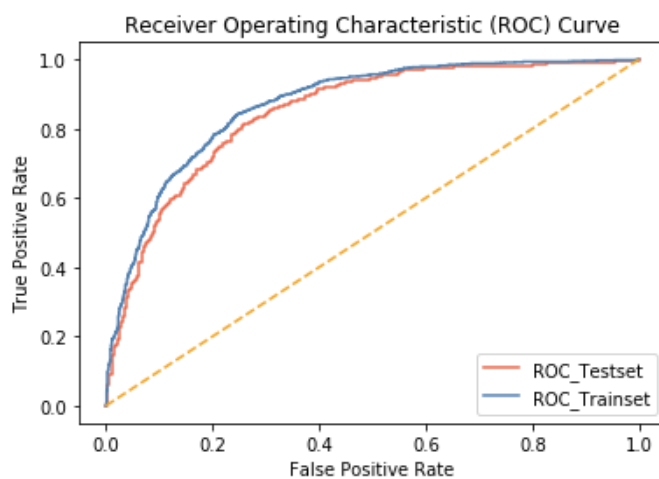
```
acc_scorer = make_scorer(accuracy_score)
GridSearchCV(model, parameters, cv = 5, scoring=acc_scorer)
```

1. Naive Bayes:

Before doing cross-validation:

AUC_Testset: 0.85

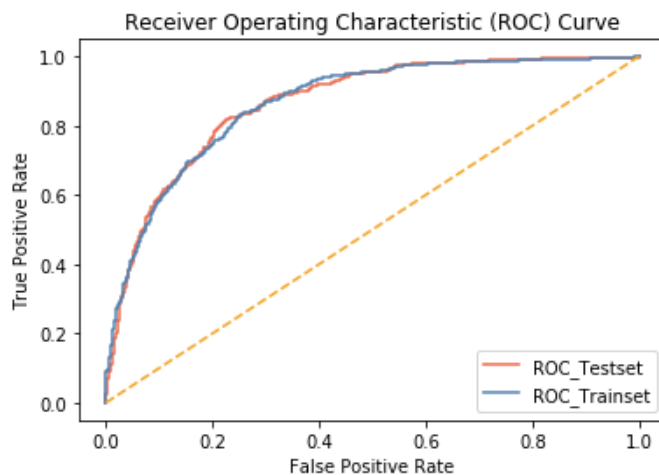
AUC_Trainset: 0.87



After doing cross-validation:

AUC_Testset: 0.87

AUC_Trainset: 0.86



best_score_: 0.7623690863127482

best_params_: { }

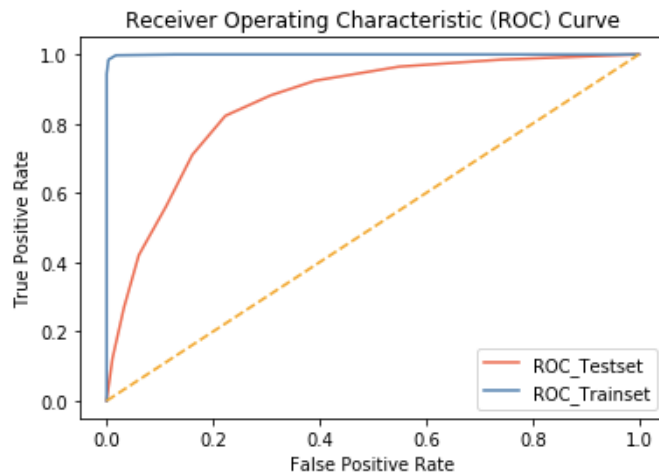
best_estimator_: GaussianNB(priors=None, var_smoothing=1e-09)

2. Random Forest:

Before doing cross-validation:

AUC_Testset: 0.86

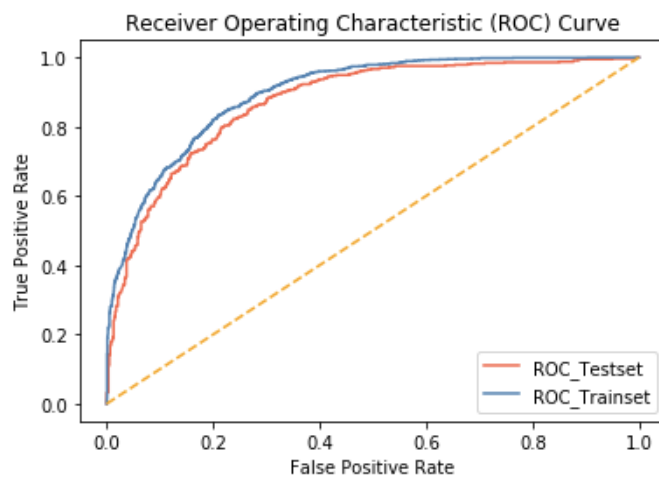
AUC_Trainset: 1.00



After doing cross-validation:

AUC_Testset: 0.87

AUC_Trainset: 0.90



best_score_: 0.7937883712531599

best_params_: {'criterion': 'entropy', 'max_depth': 5, 'max_features': 'sqrt',
'min_samples_leaf': 1, 'min_samples_split': 3, 'n_estimators': 9}

best_estimator_: RandomForestClassifier(bootstrap=True, class_weight=None,
criterion='entropy',

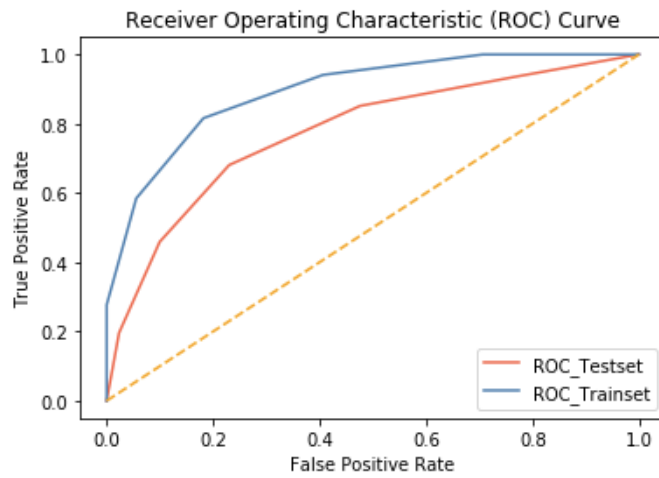
max_depth=5, max_features='sqrt', max_leaf_nodes=None,
min_impurity_decrease=0.0, min_impurity_split=None,
min_samples_leaf=1, min_samples_split=3,
min_weight_fraction_leaf=0.0, n_estimators=9,
n_jobs=None, oob_score=False, random_state=None,
verbose=0, warm_start=False)

3. KNN:

Before doing cross-validation(N=5):

AUC_Testset: 0.78

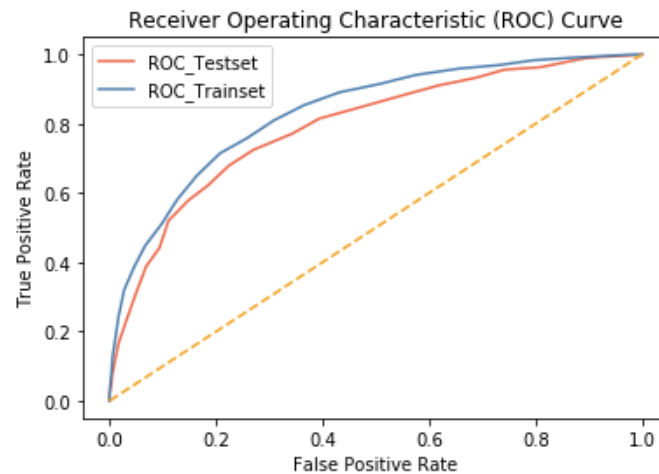
AUC_Trainset: 0.89



After doing cross-validation:

AUC_Testset: 0.79

AUC_Trainset: 0.83



best_score_: 0.7276995305164319

best_params_: {'n_neighbors': 28}

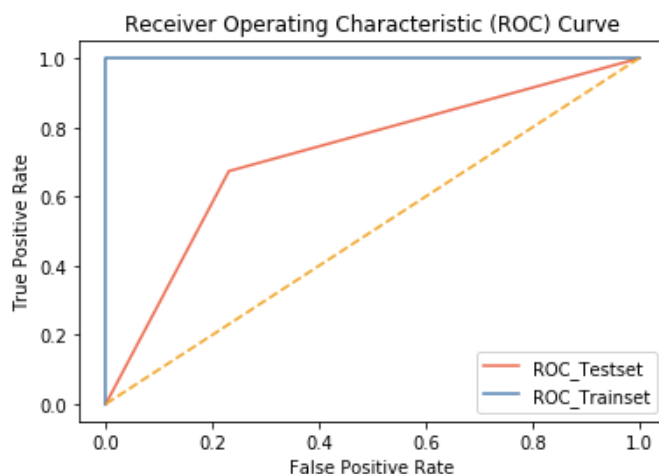
best_estimator_: KNeighborsClassifier(algorithm='auto', leaf_size=30,
metric='minkowski',
metric_params=None, n_jobs=None, n_neighbors=28, p=2,
weights='uniform')

4. Decision tree:

Before doing cross-validation:

AUC_Testset: 0.72

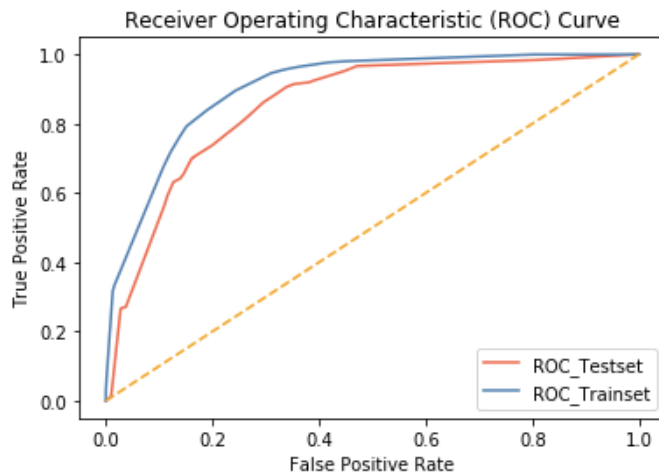
AUC_Trainset: 1.00



After doing cross-validation:

AUC_Testset: 0.85

AUC_Trainset: 0.90



best_score_: 0.7648970747562297

best_params_: {'criterion': 'entropy', 'max_depth': 6}

best_estimator_: DecisionTreeClassifier(class_weight=None, criterion='entropy',
max_depth=6,

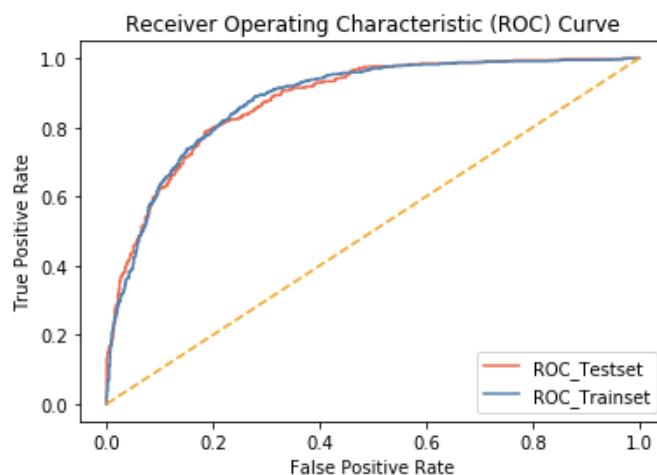
max_features=None, max_leaf_nodes=None,
min_impurity_decrease=0.0, min_impurity_split=None,
min_samples_leaf=1, min_samples_split=2,
min_weight_fraction_leaf=0.0, presort=False,
random_state=None, splitter='best')

5. SVM

Before doing cross-validation(N=5):

AUC_Testset: 0.88

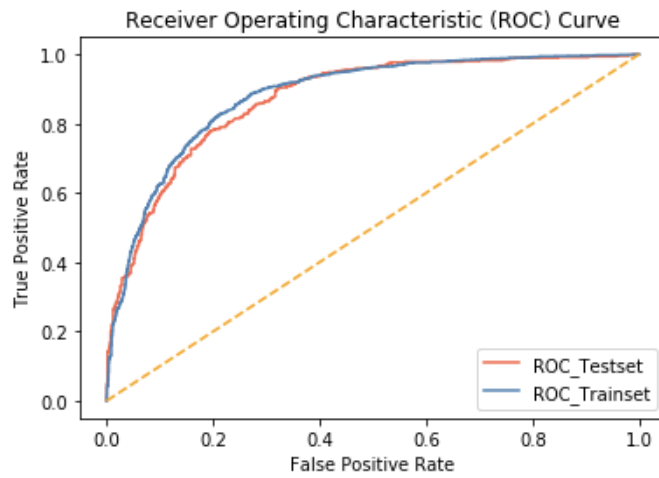
AUC_Trainset: 0.88



After doing cross-validation:

AUC_Testset: 0.87

AUC_Trainset: 0.88



best_score_: 0.800650054171181

best_params_: {'C': 10, 'gamma': 0.001}

best_estimator_: SVC(C=10, cache_size=200, class_weight=None, coef0=0.0, decision_function_shape='ovr', degree=3, gamma=0.001, kernel='linear', max_iter=-1, probability=True, random_state=None, shrinking=True, tol=0.001, verbose=False)