

```
In [1]: import numpy as np # Linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import os
from sklearn import tree
```

```
In [2]: from sklearn.model_selection import train_test_split
from xgboost import XGBClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import StandardScaler
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from sklearn.gaussian_process.kernels import RBF
from sklearn.tree import DecisionTreeClassifier

from sklearn.metrics import accuracy_score
from sklearn.metrics import confusion_matrix

from sklearn import preprocessing
from sklearn import metrics
import matplotlib.pyplot as plt
from xgboost import plot_tree
```

```
In [3]: data = pd.read_csv('Myocardial.csv')
```

```
In [4]: models = {
    'LogisticRegression': LogisticRegression(max_iter=3000),
    'XGBoostClassifier': XGBClassifier(num_class = 1),
    'KNN': KNeighborsClassifier(),
    'DecisionTree': DecisionTreeClassifier(max_depth=6)
}

models2 = {
    'LogisticRegression': LogisticRegression(max_iter=3000),
    'XGBoostClassifier': XGBClassifier(),
    'KNN': KNeighborsClassifier(),
    'Linear SVM': SVC(kernel="linear", C=0.025),
    'RBF SVM': SVC(gamma=2, C=1),
    'DecisionTree': DecisionTreeClassifier(max_depth=6)
}
```

```
In [5]: mean_missing = data.isna().mean()
data = data.loc[:, mean_missing < .1]
[target in y.columns for target in data.loc[:, mean_missing > .1].columns]
```

```
Out[5]: []
```

```
In [6]: X = data.iloc[:,1:95] # input
y = data.iloc[:,95:107] # target
y.loc[y['LET_IS'] > 0, 'LET_IS'] = 1 ## recode y as binary. 0 = alive, 1 = dead, cause

from sklearn.impute import SimpleImputer
imp = SimpleImputer(missing_values=np.nan, strategy='mean')

X = imp.fit_transform(X)

from sklearn import preprocessing
```

```

scaler = StandardScaler()
# scaler = preprocessing.MinMaxScaler()

X = scaler.fit_transform(X)

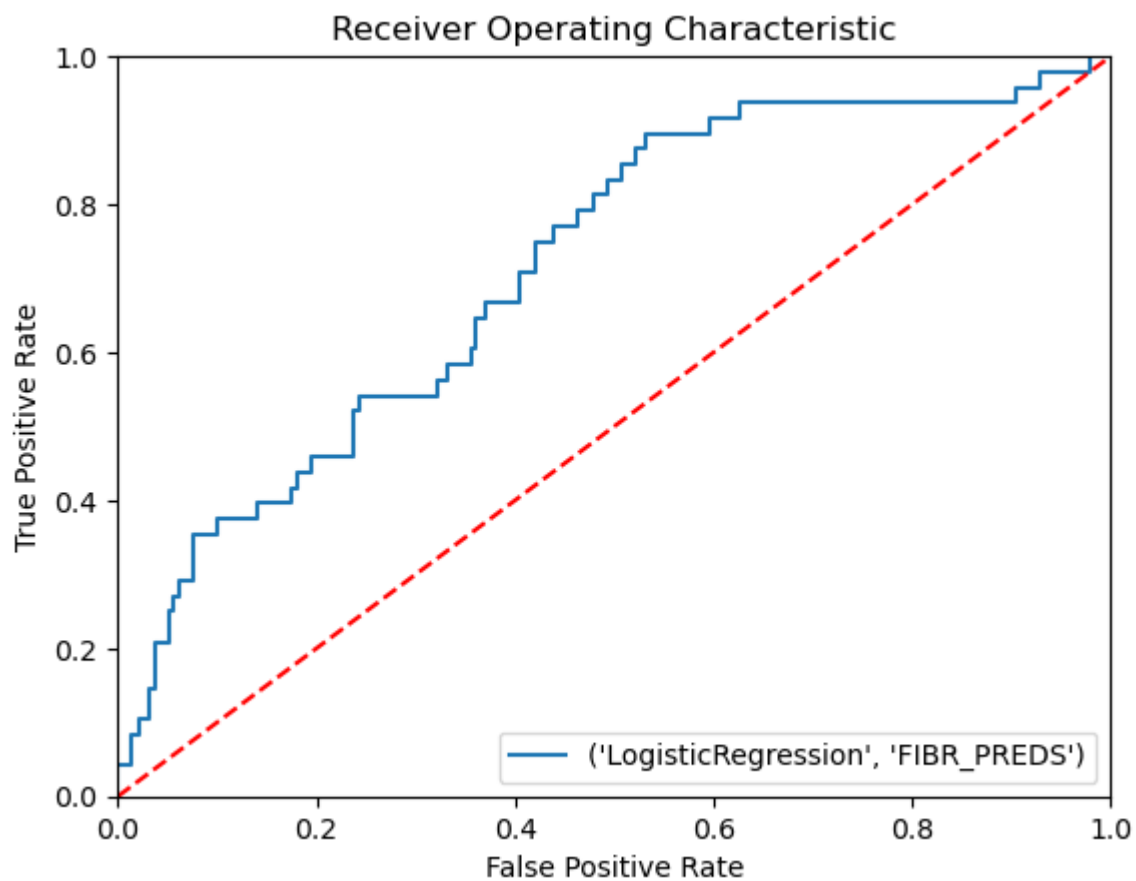
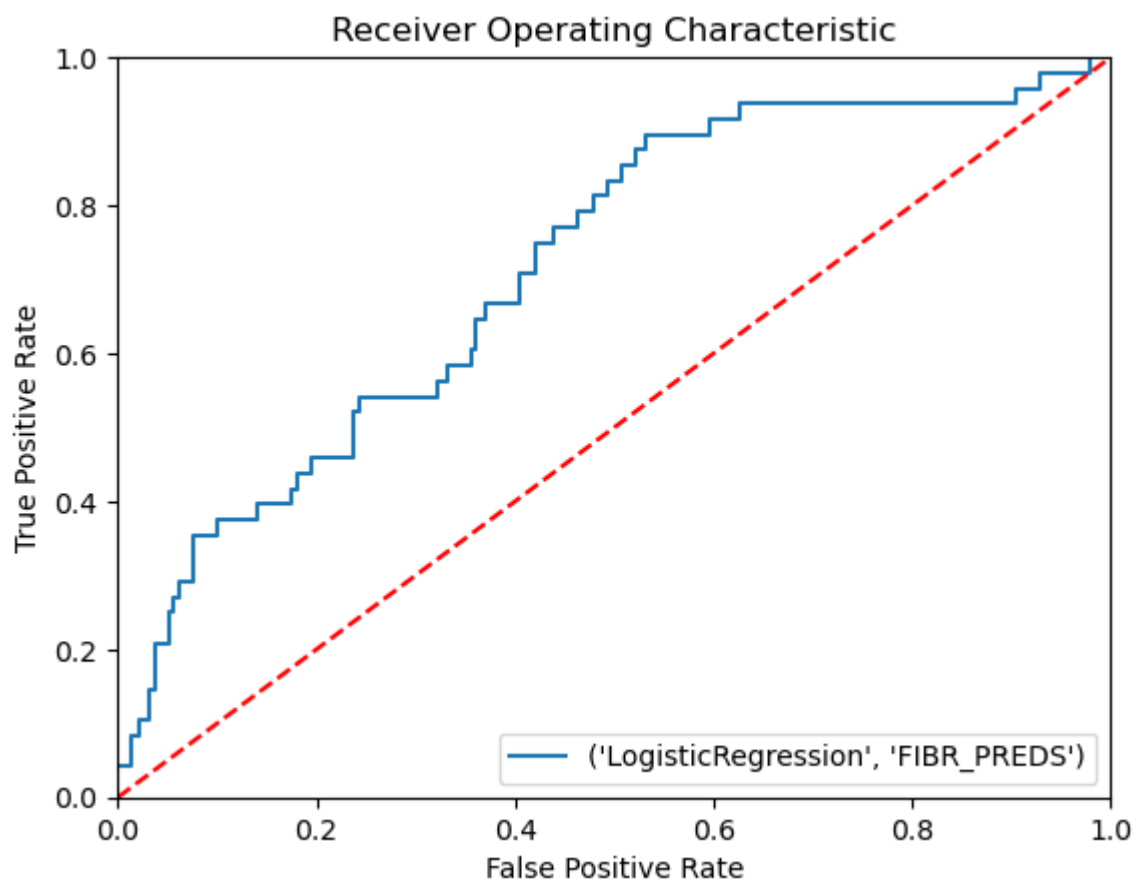
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=

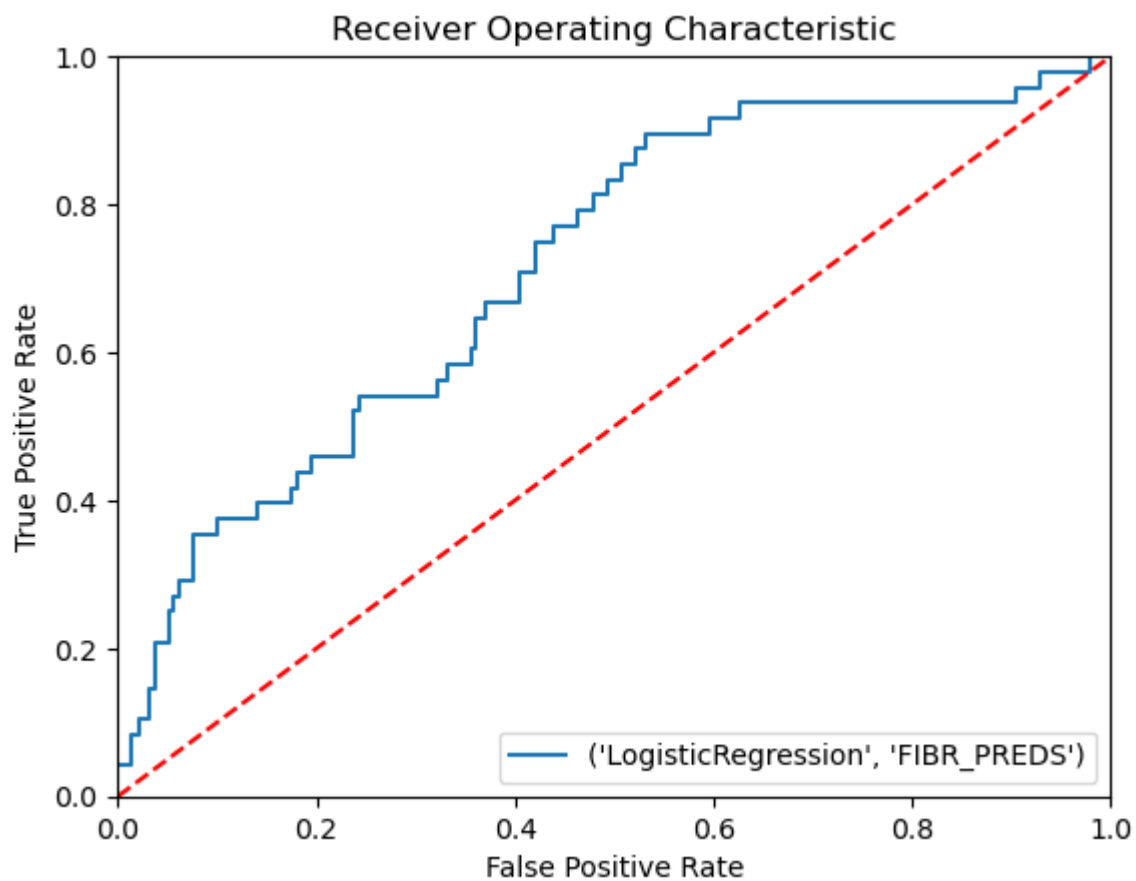
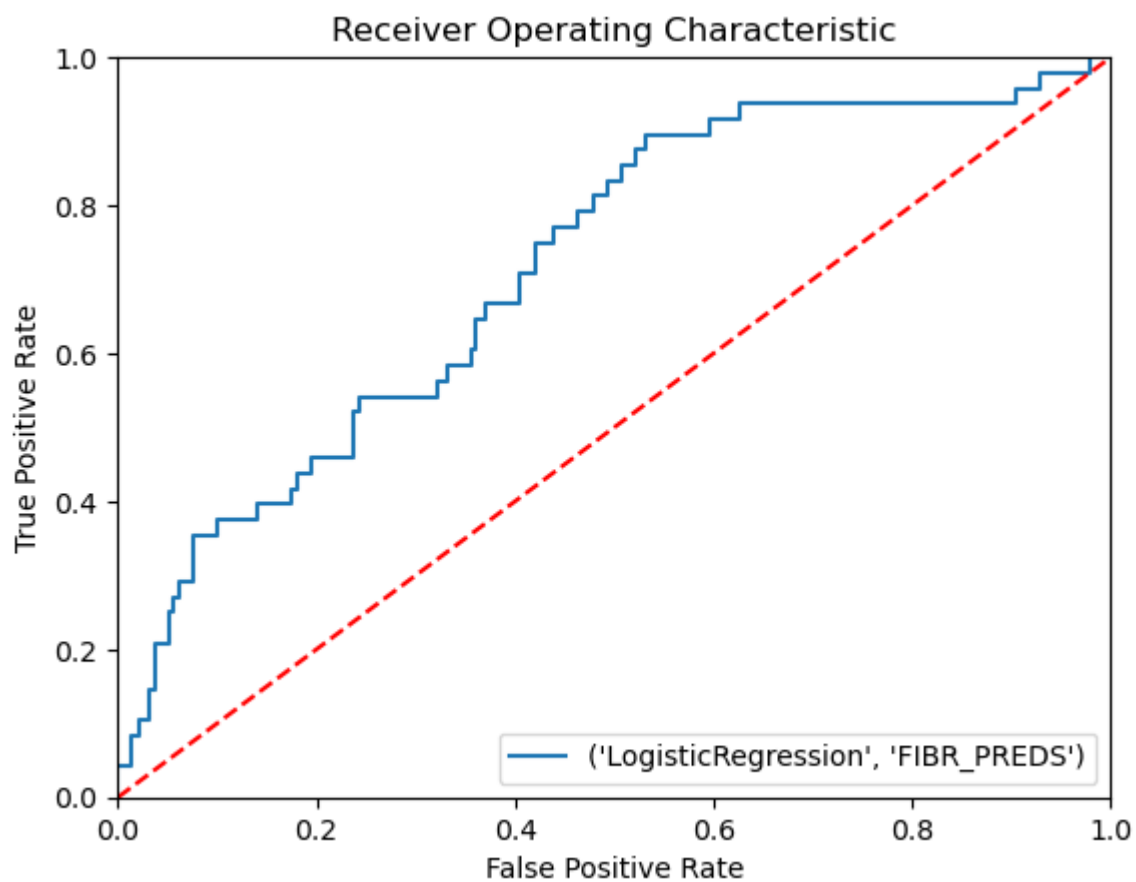
```

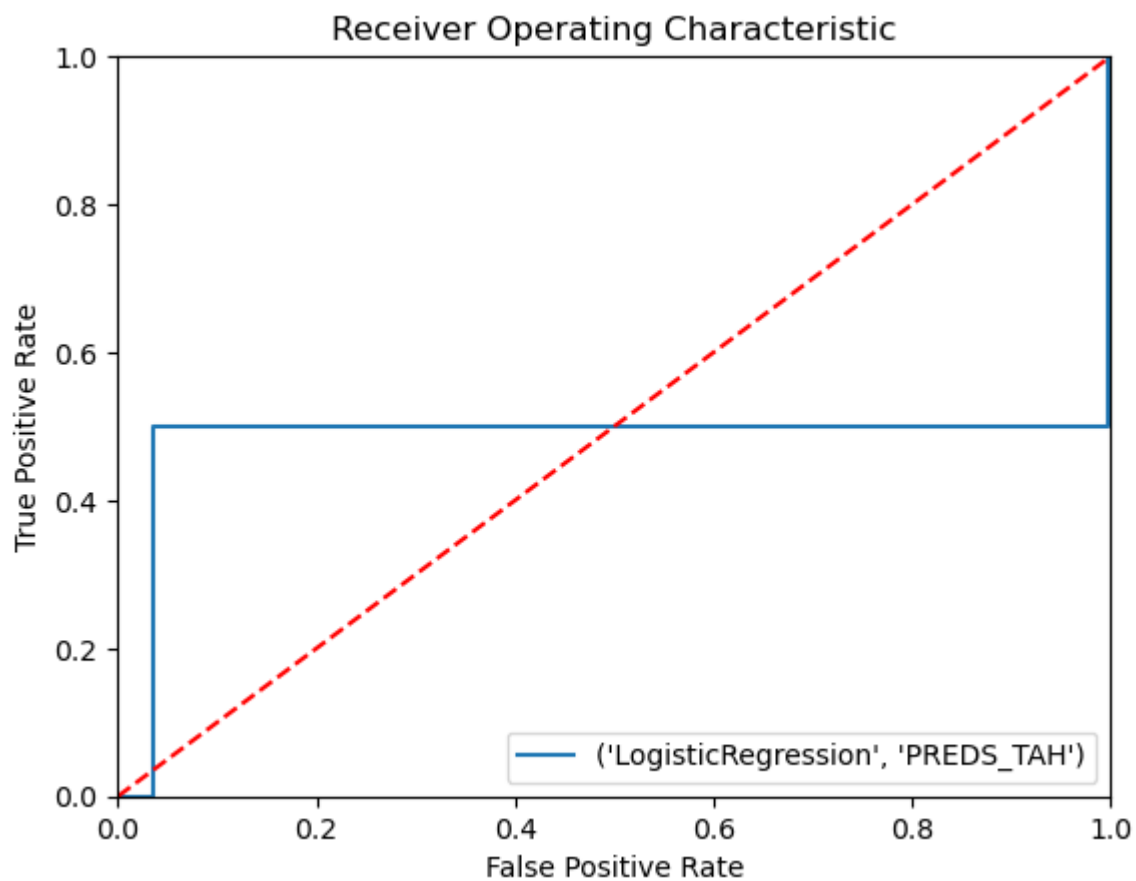
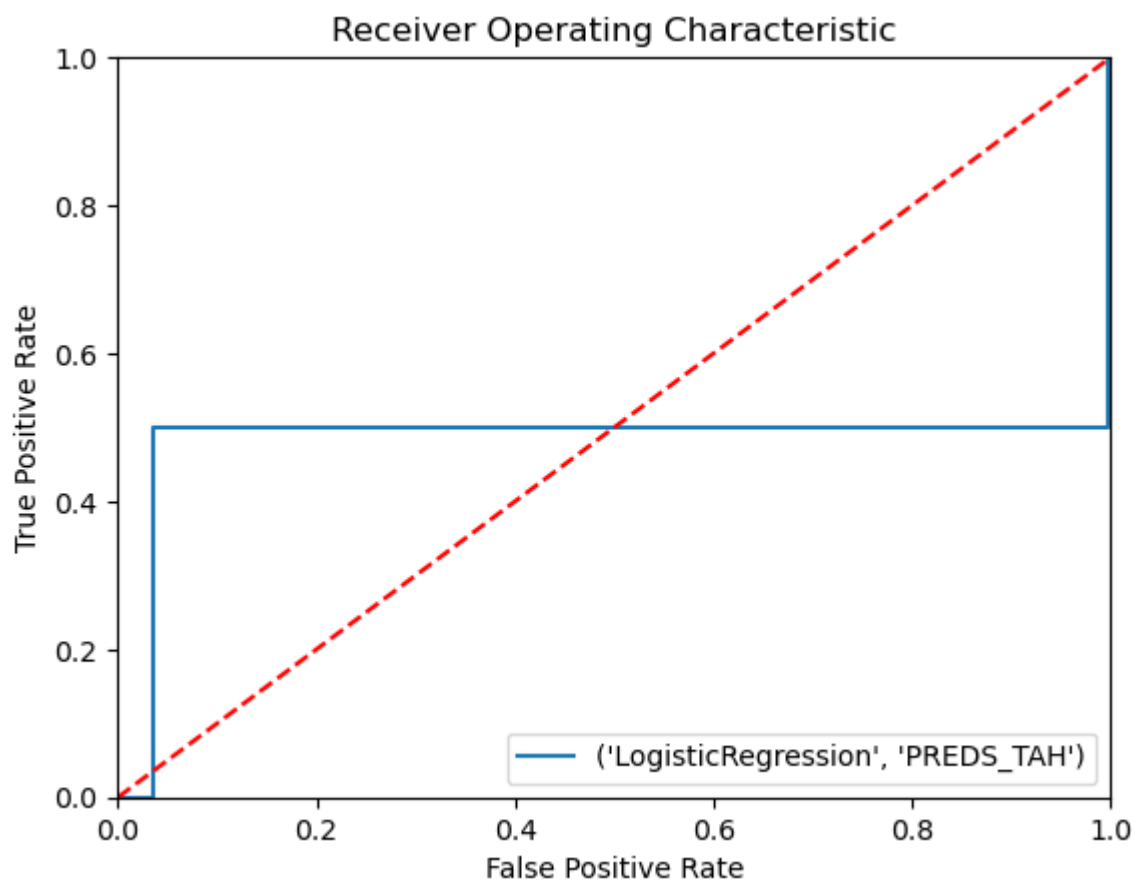
```

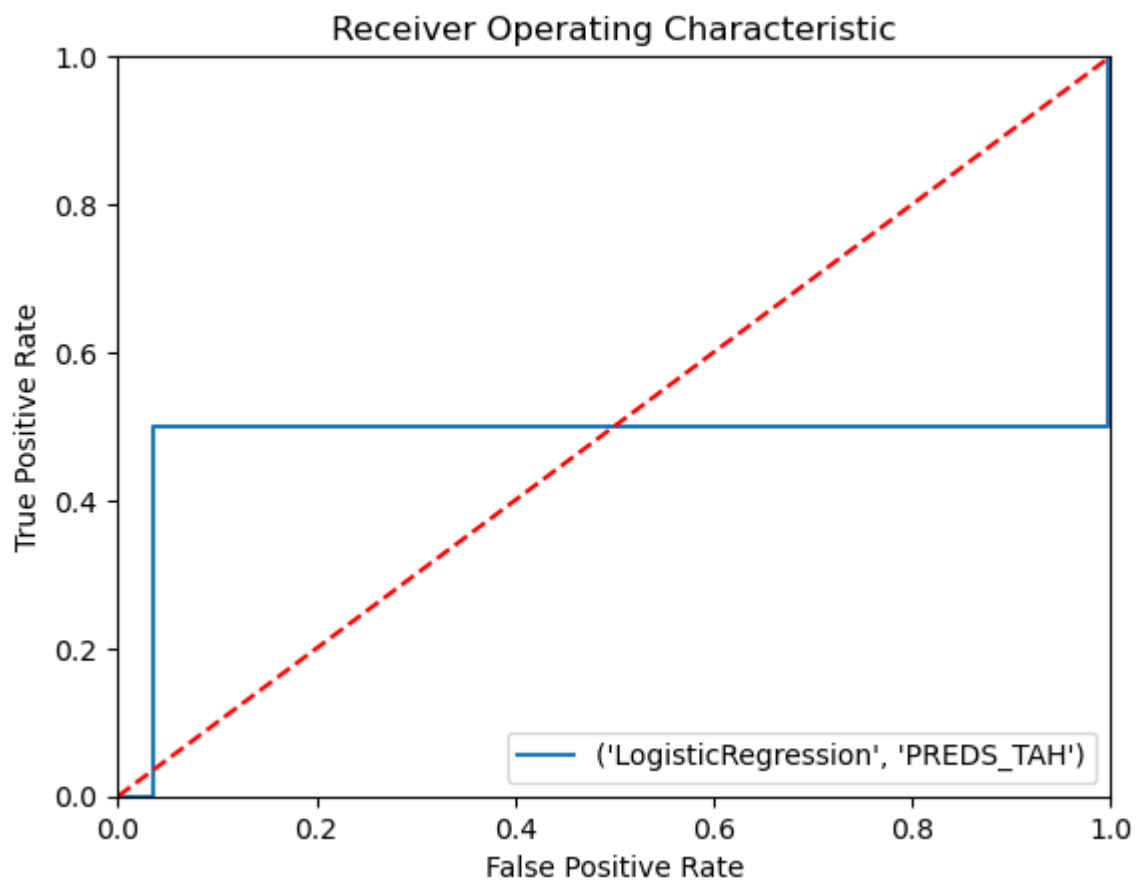
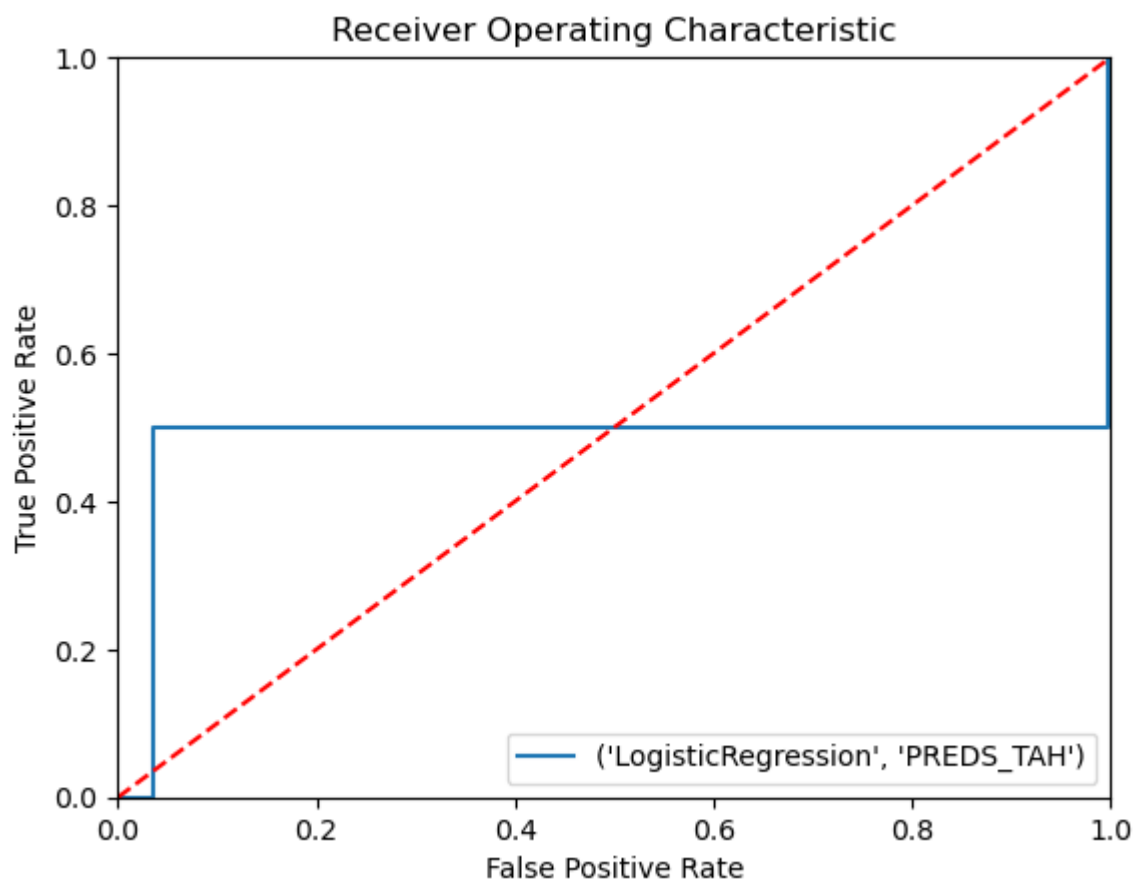
In [7]: acc=[]
models_names = models.keys()
models_names = list(models_names)
cols = list(y_train.columns.values)
for model_name, model in models.items():
    clf = model
    acc_model=[]
    for i in range(0, 12):
        clf.fit(X_train, y_train.iloc[:, i])
        acc_feature = round(clf.score(X_test, y_test.iloc[:, i]) * 100, 5)
        acc_model.append(acc_feature)
        y_pred = clf.predict(X_test)
        fpr, tpr, thresholds = metrics.roc_curve(y_test.iloc[:,i], clf.predict_proba(X_test).iloc[:,i])
        auc = metrics.auc(fpr, tpr)
        for b in range(0,4):
            plt.plot(fpr, tpr, label=(model_name, cols[i]))
            plt.title('Receiver Operating Characteristic')
            plt.legend(loc = 'lower right')
            plt.plot([0, 1], [0, 1], 'r--')
            plt.xlim([0, 1])
            plt.ylim([0, 1])
            plt.ylabel('True Positive Rate')
            plt.xlabel('False Positive Rate')
            plt.show()
    acc.append(acc_model)

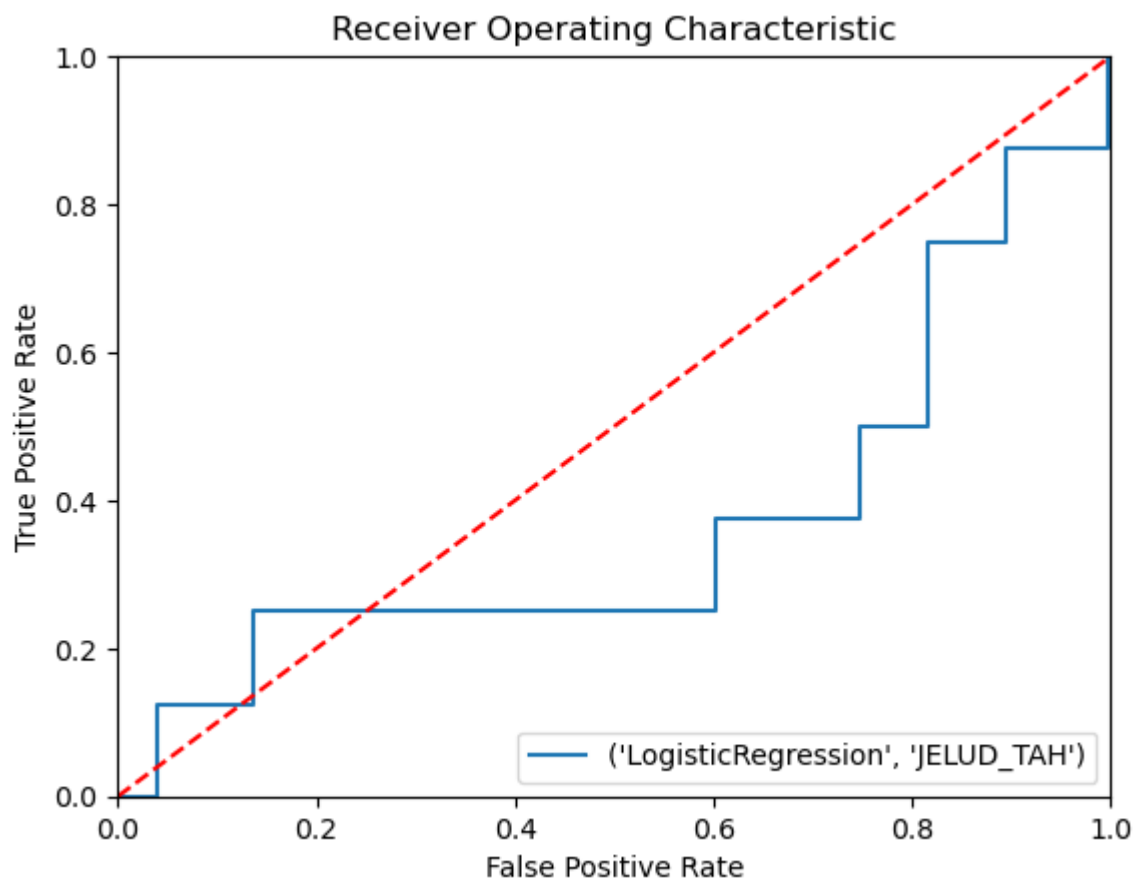
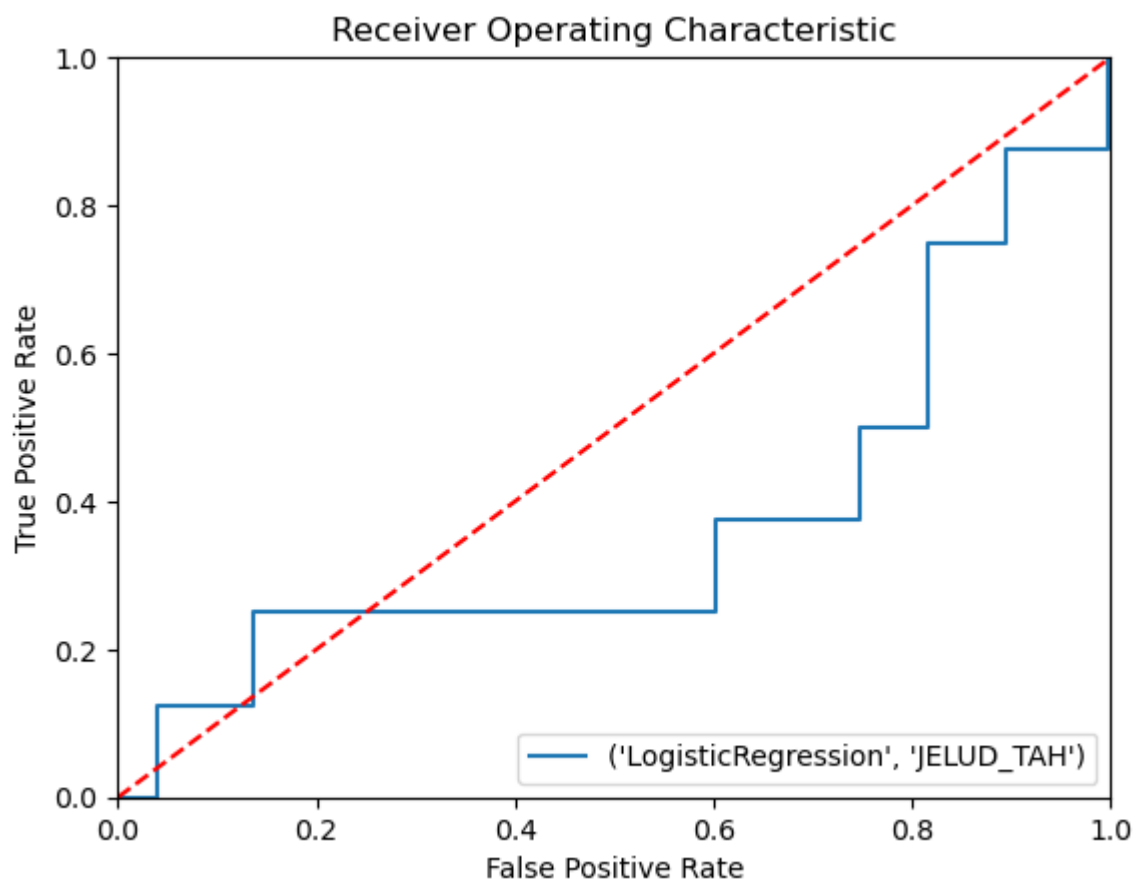
```

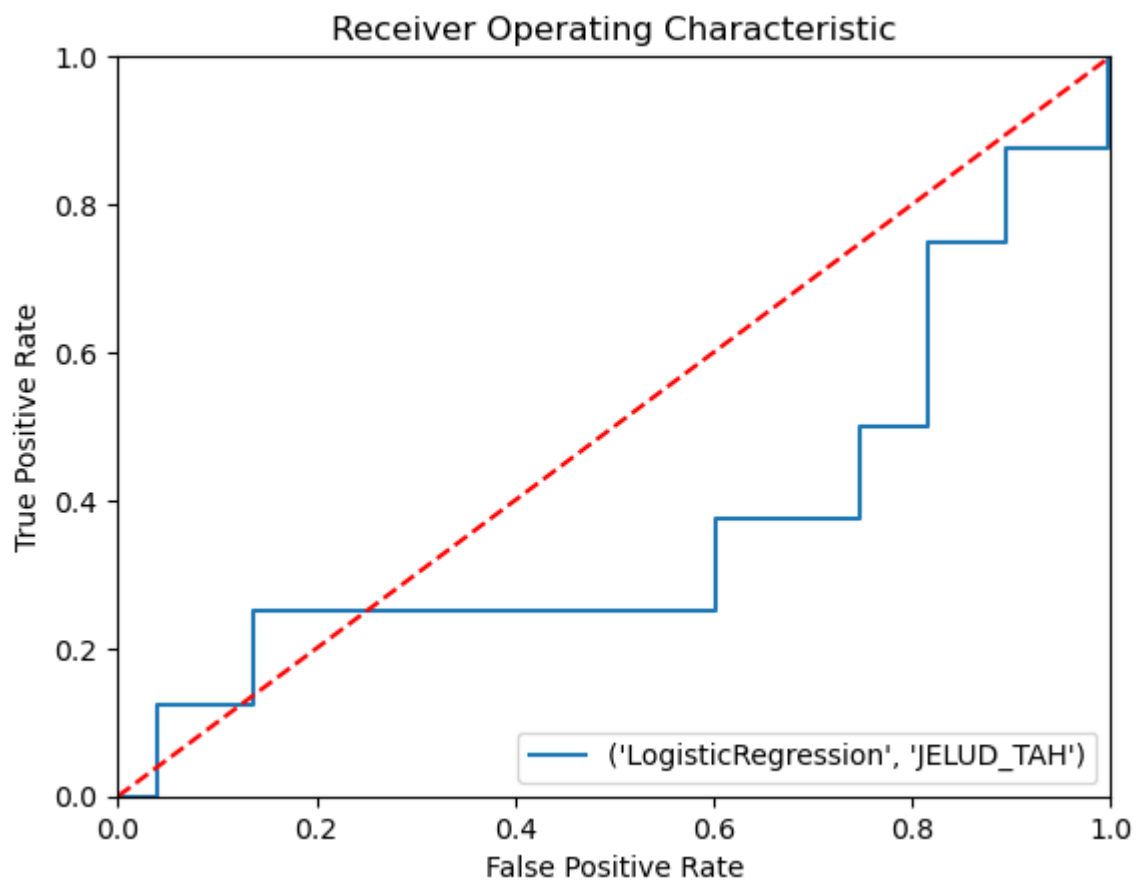
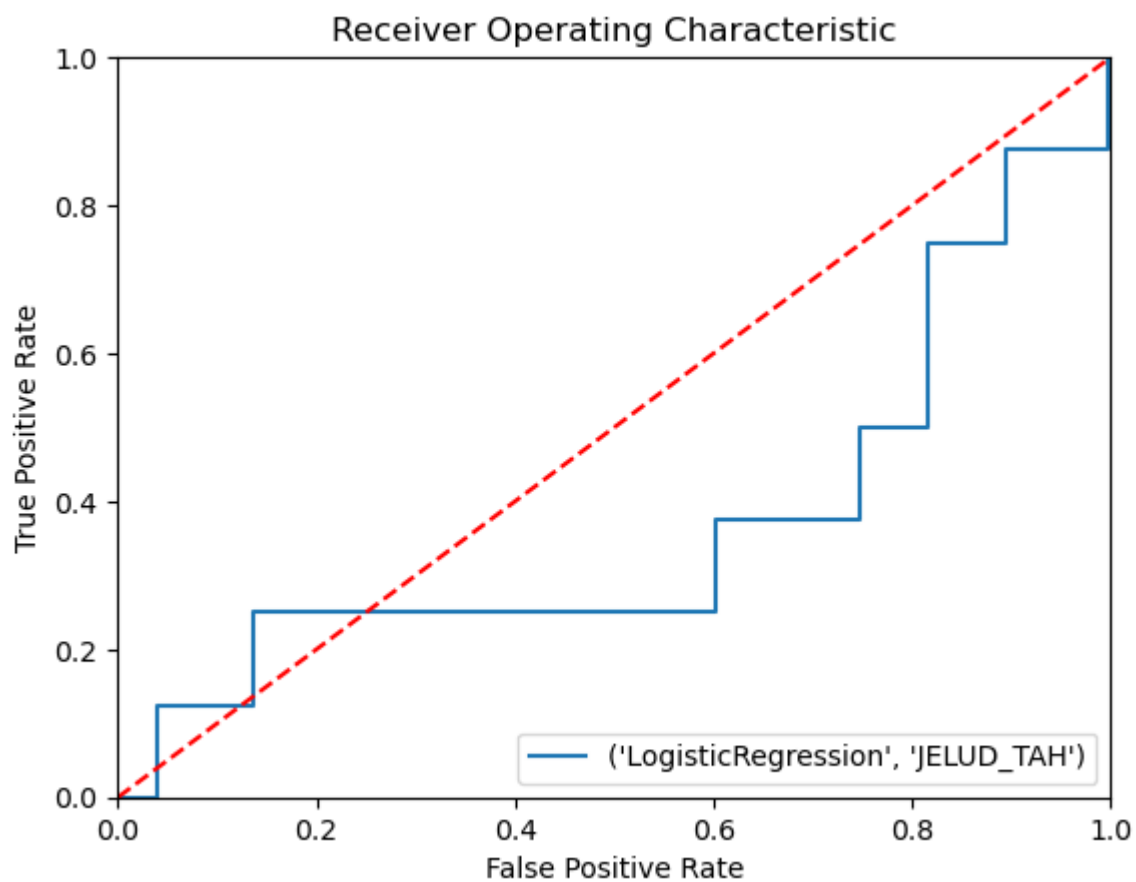


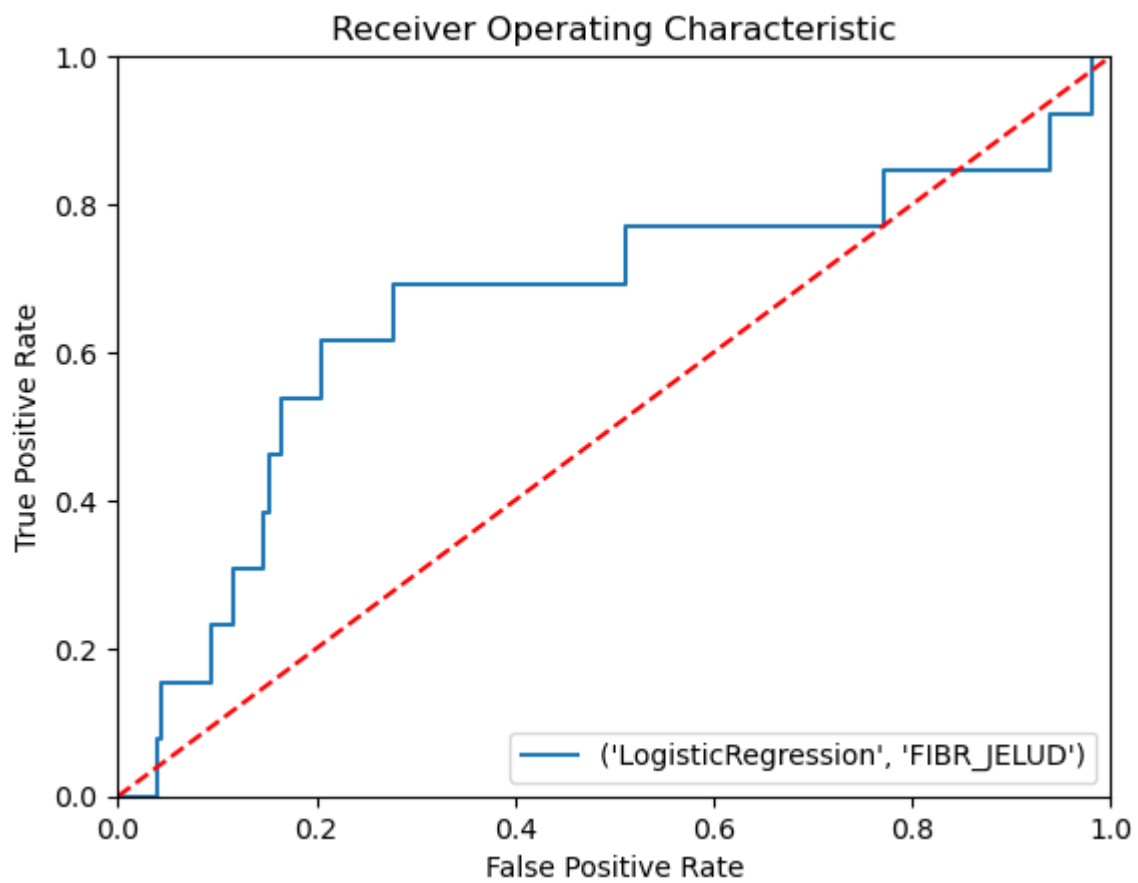
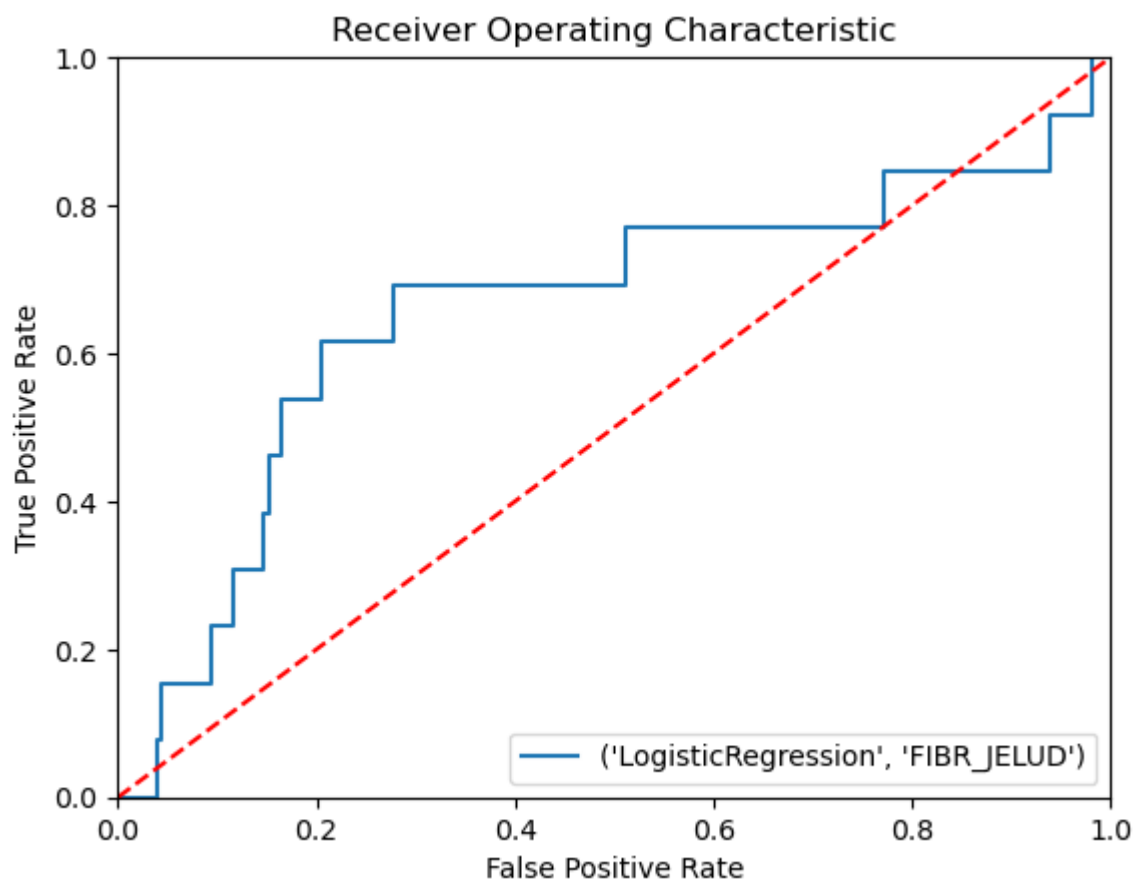


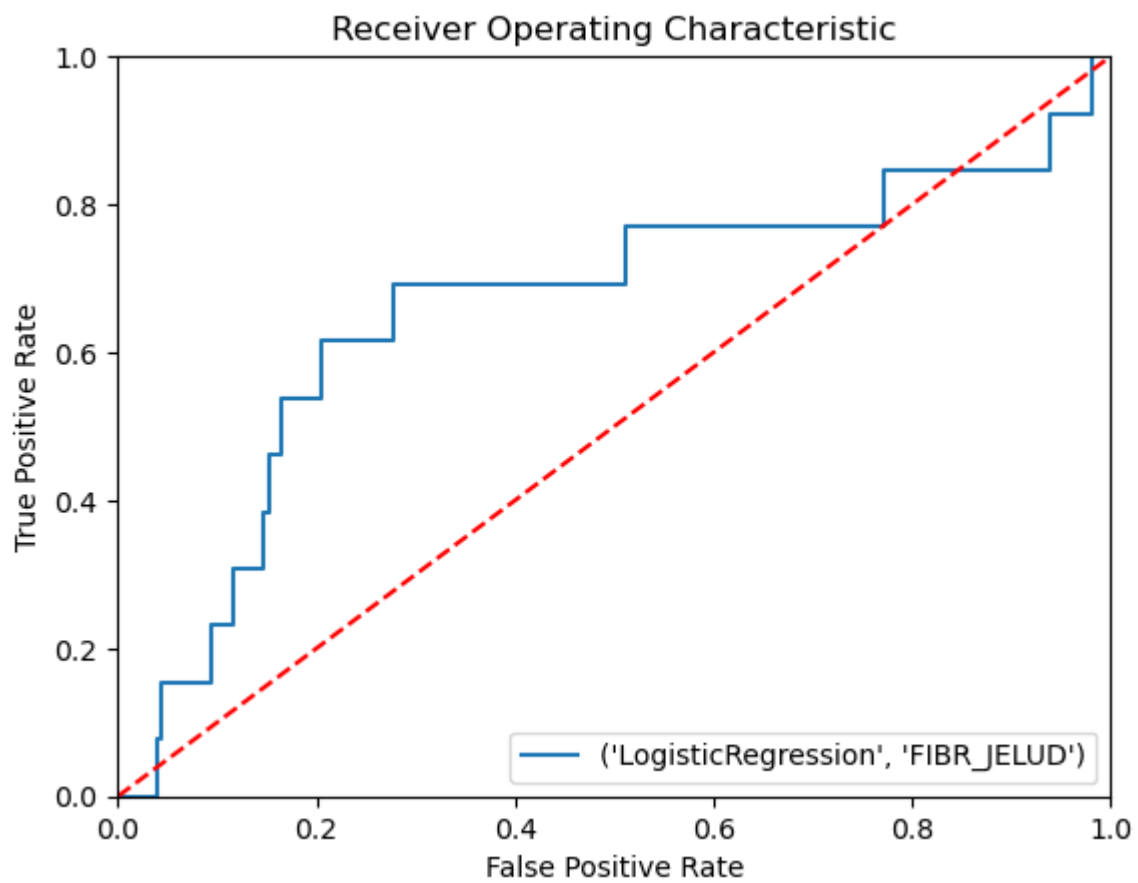
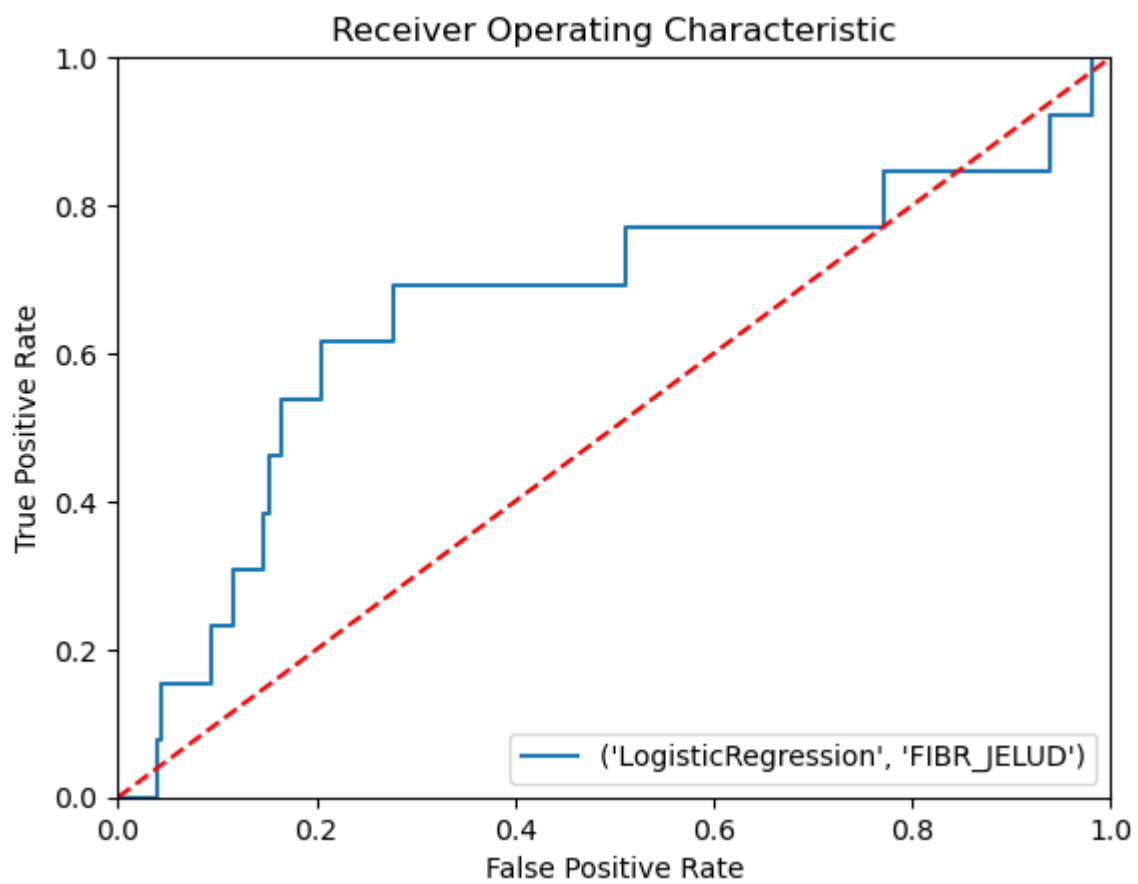


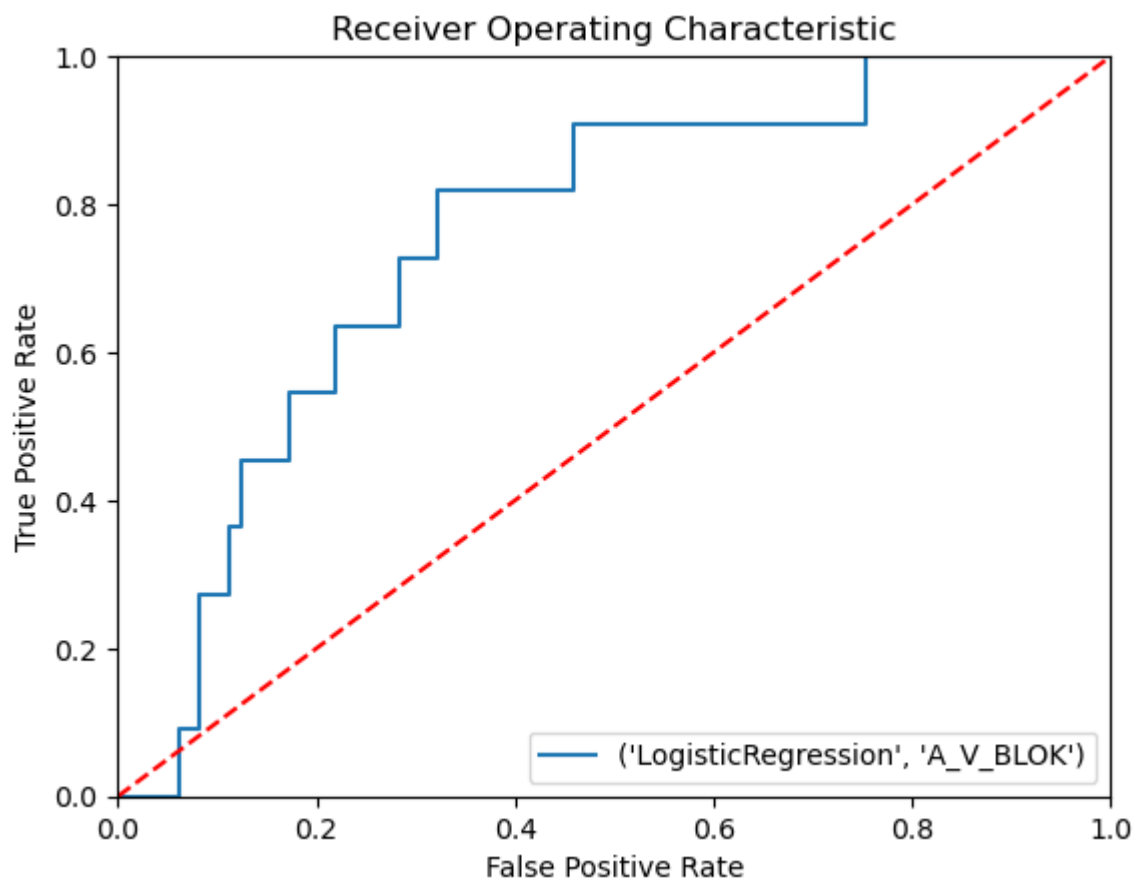
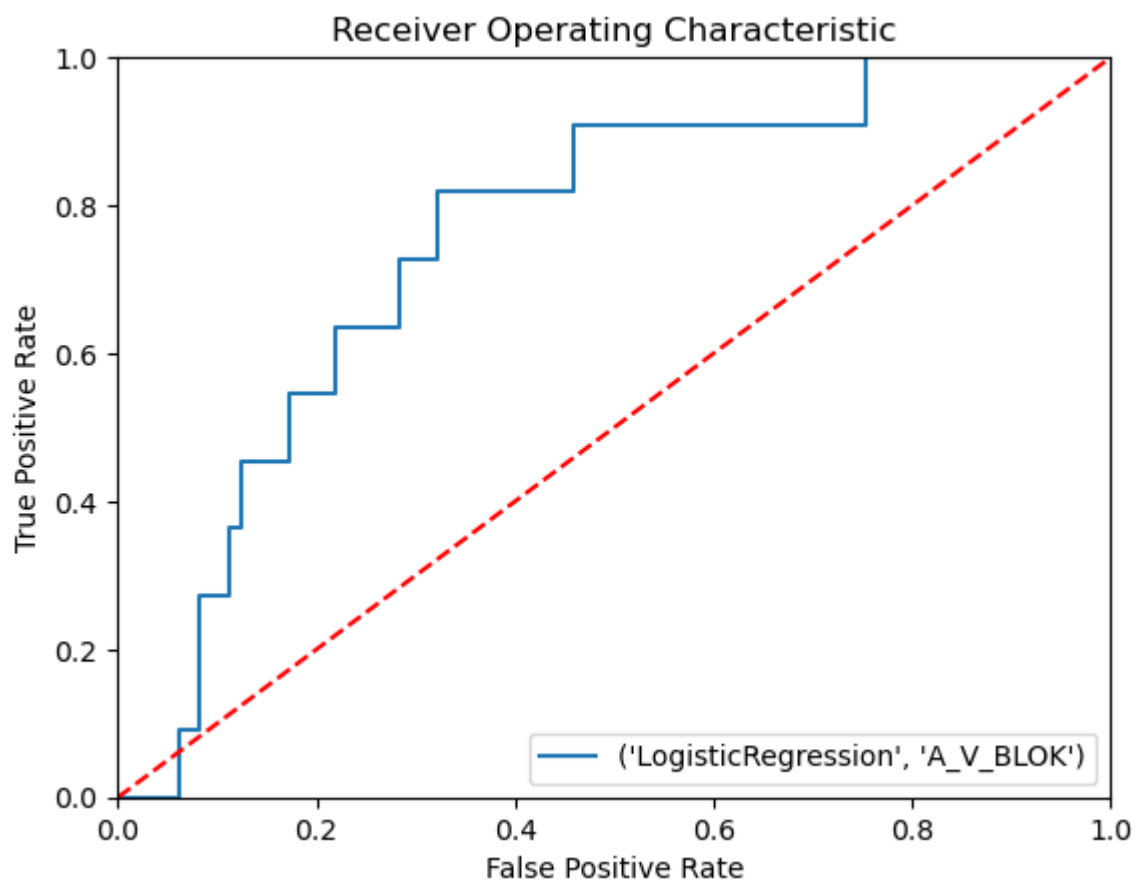


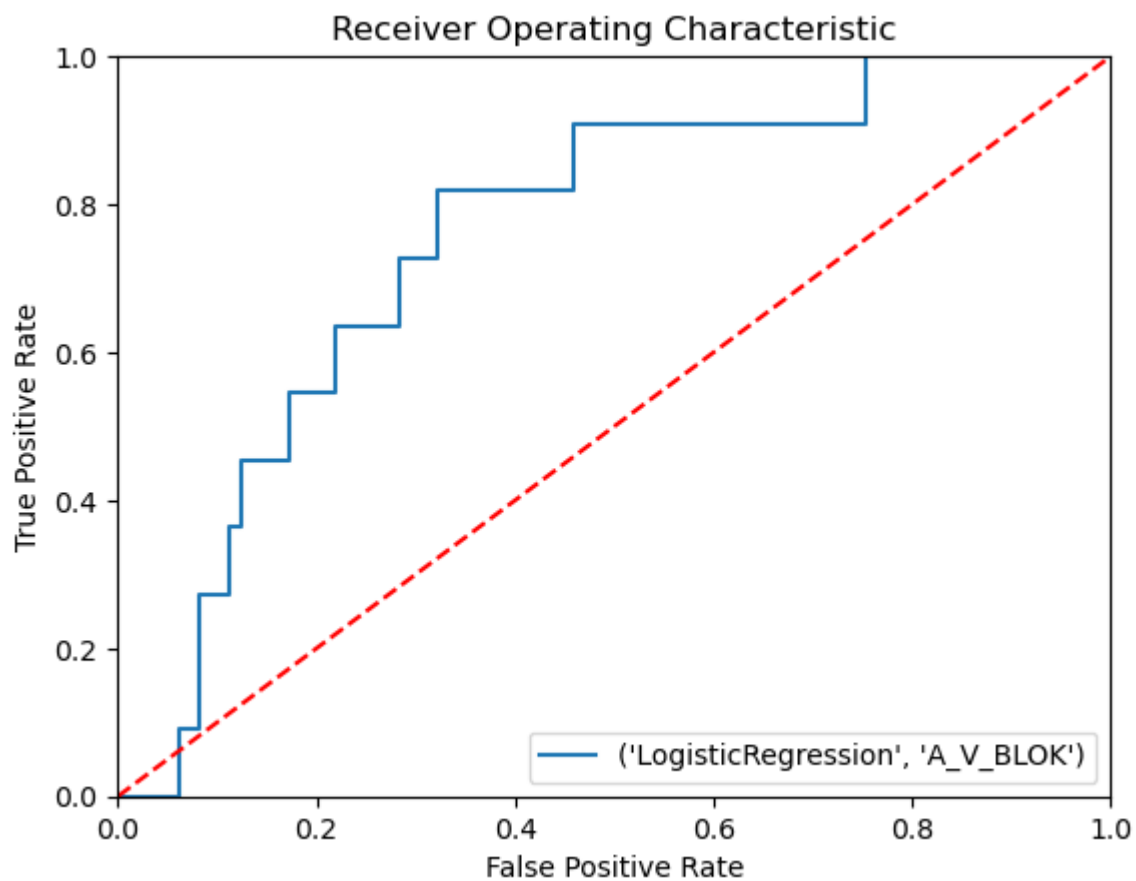
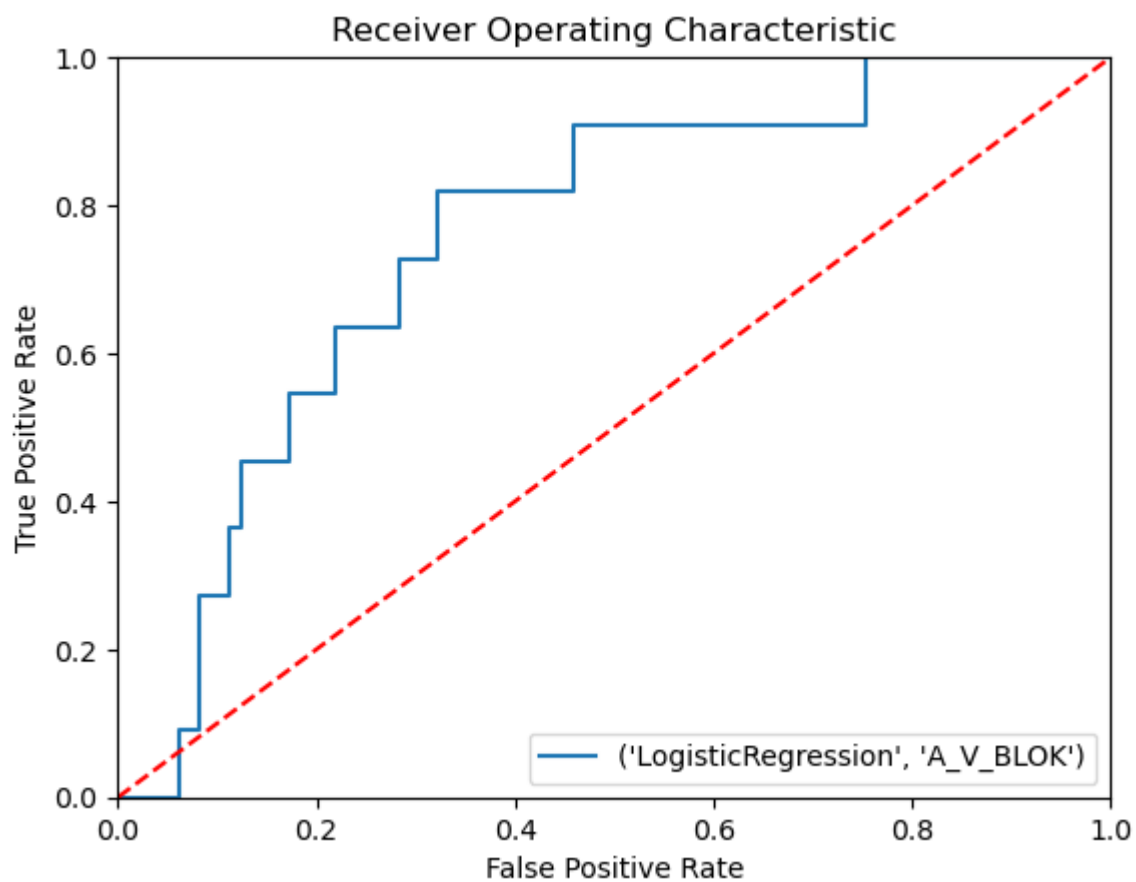


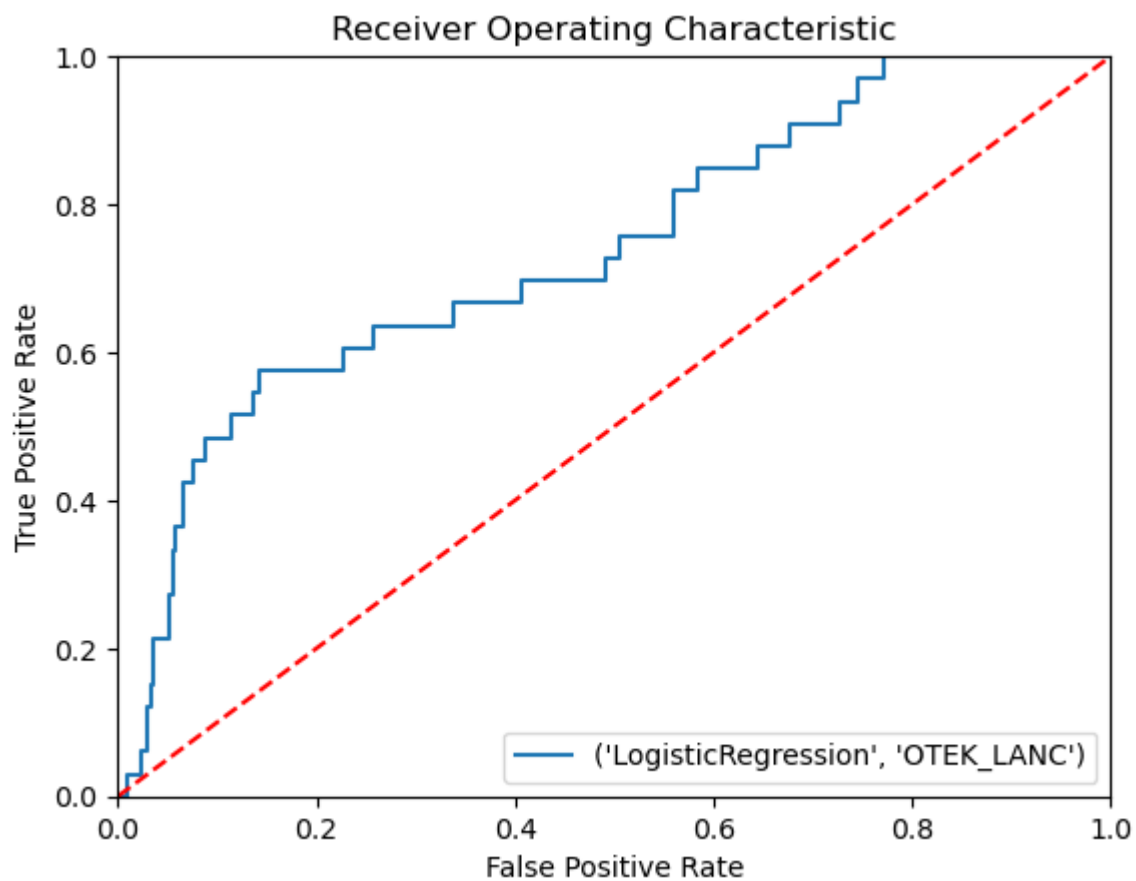
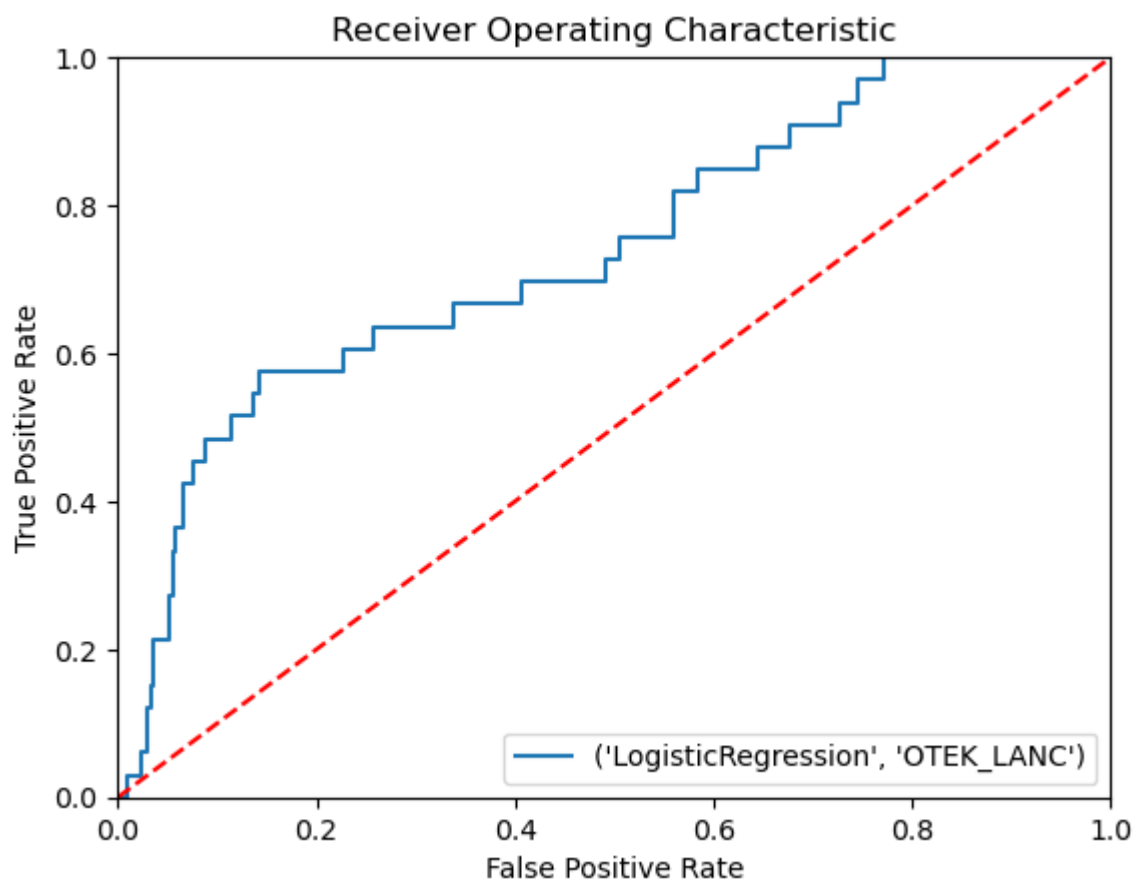


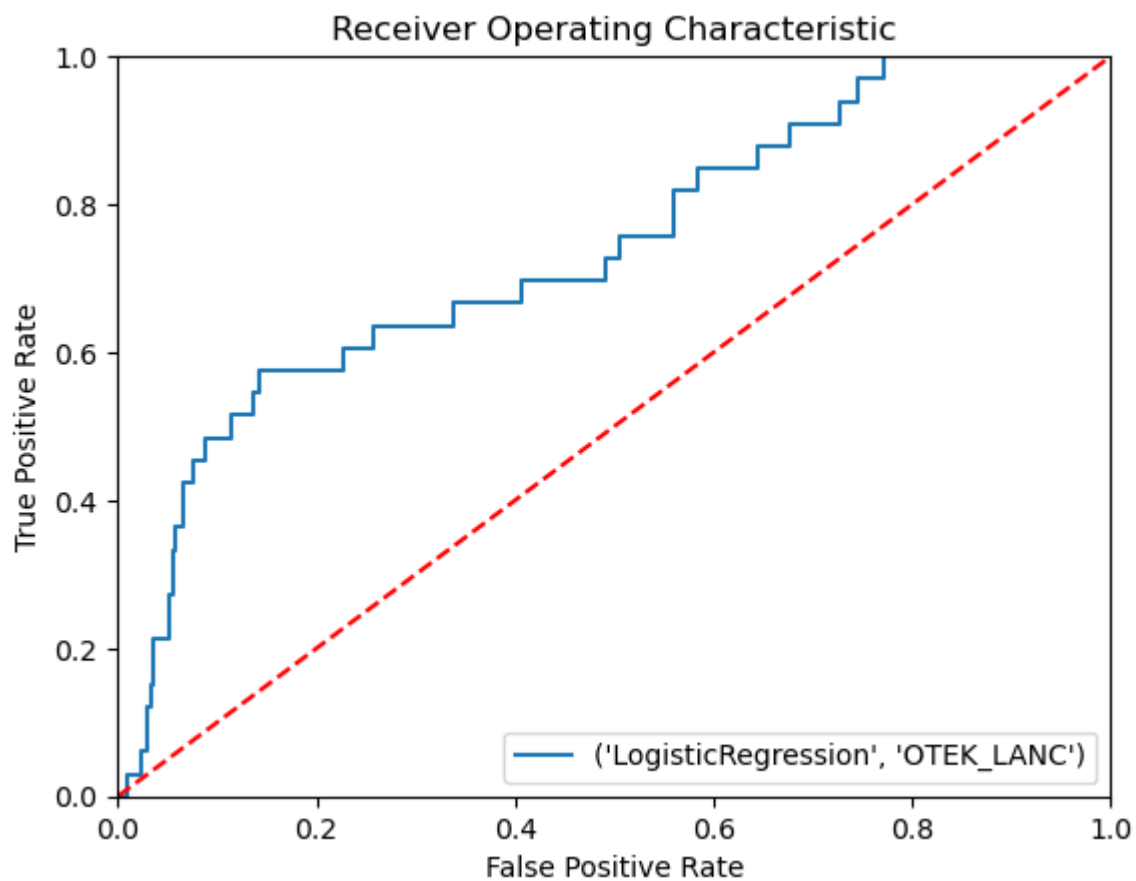
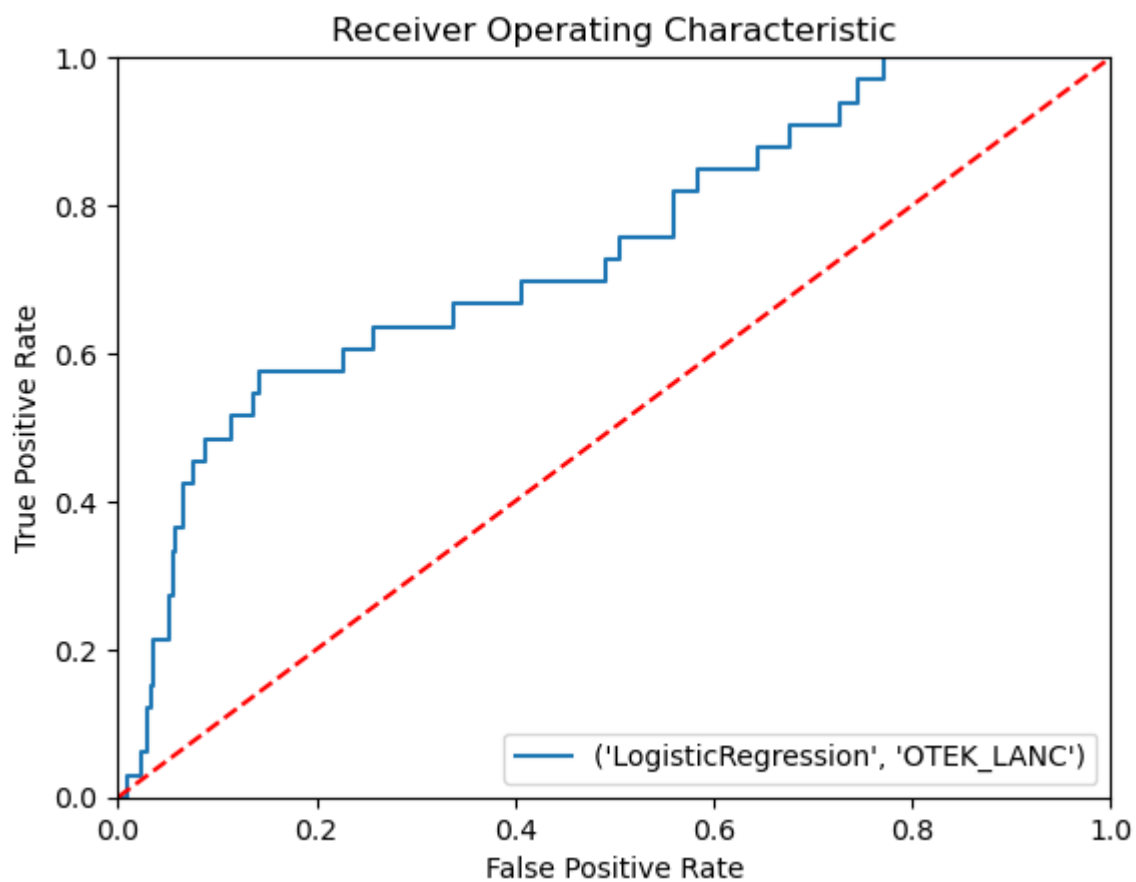


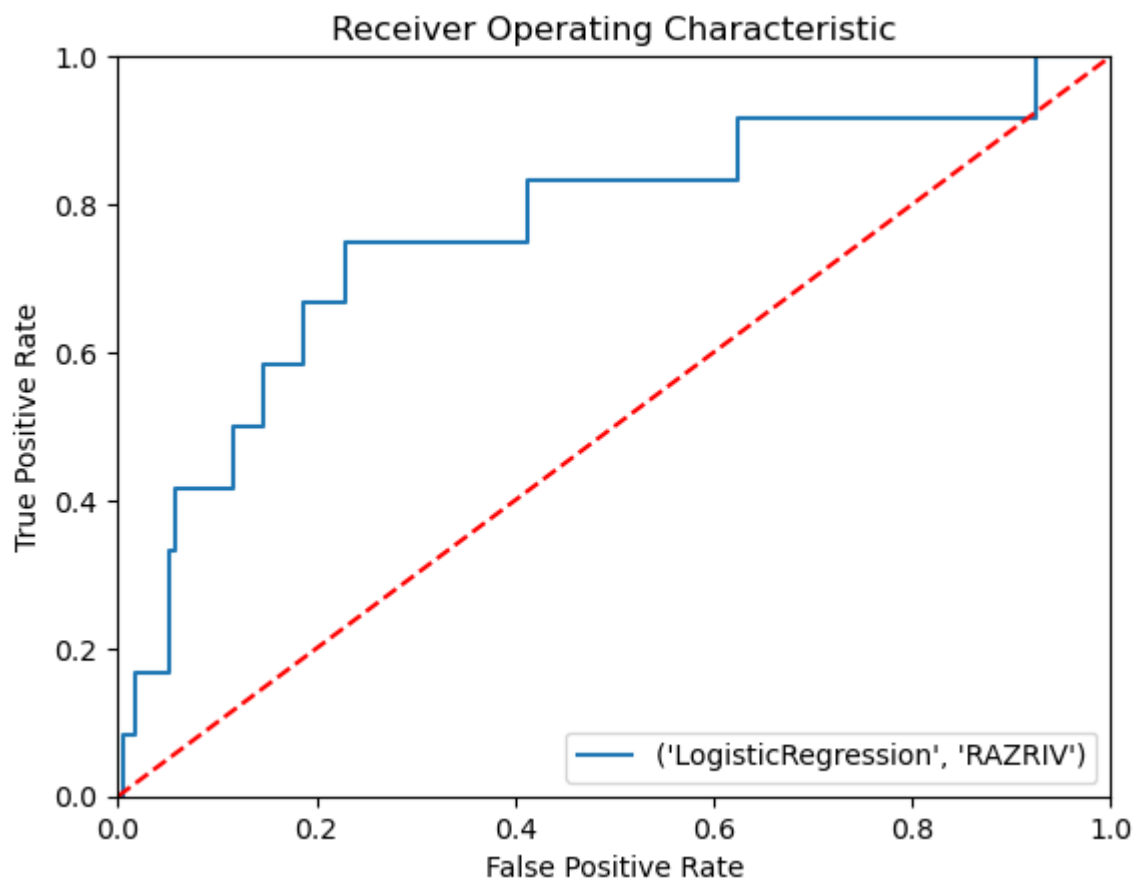
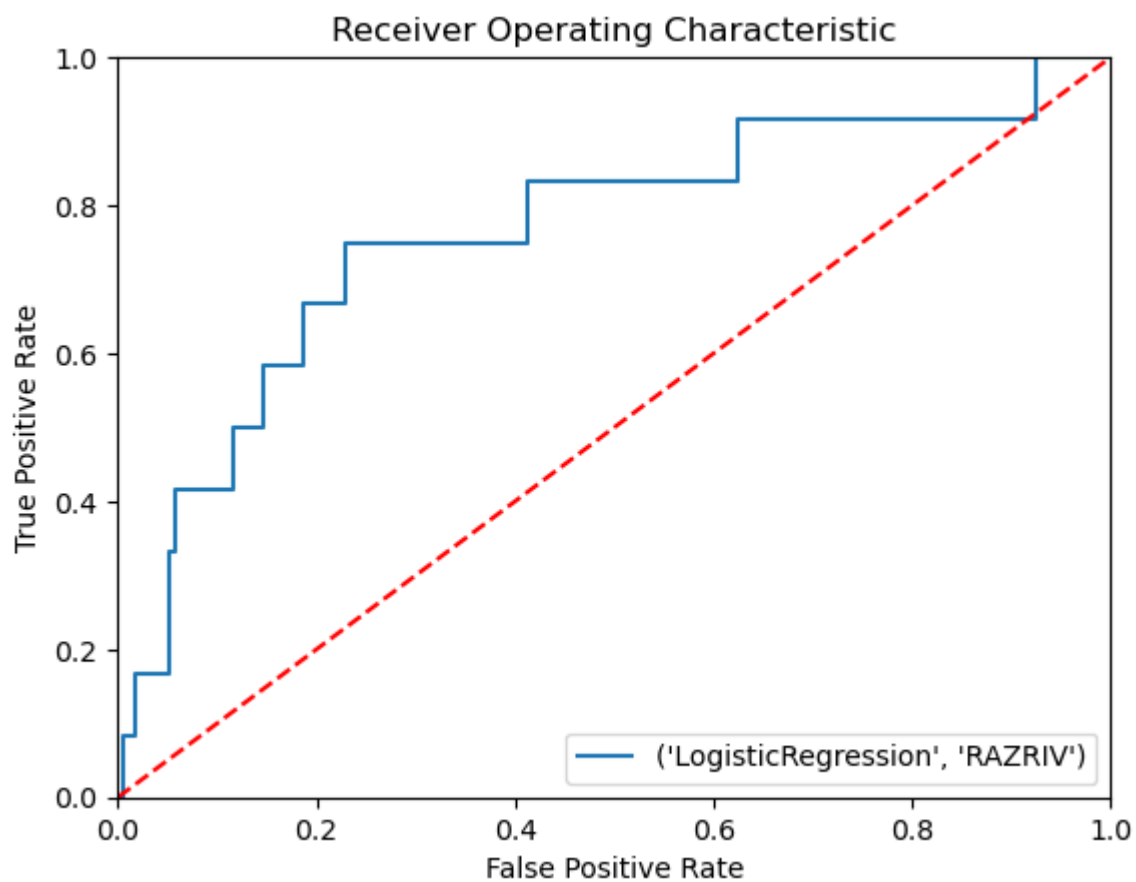


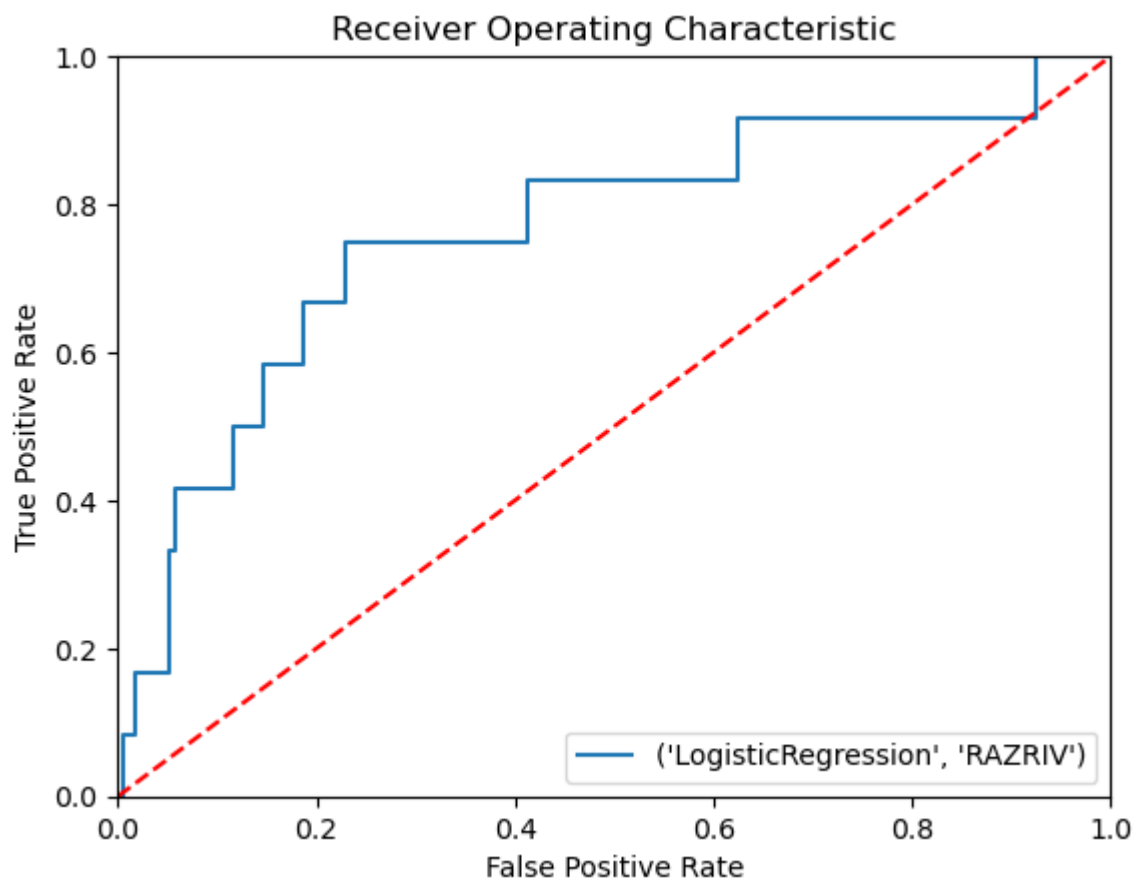
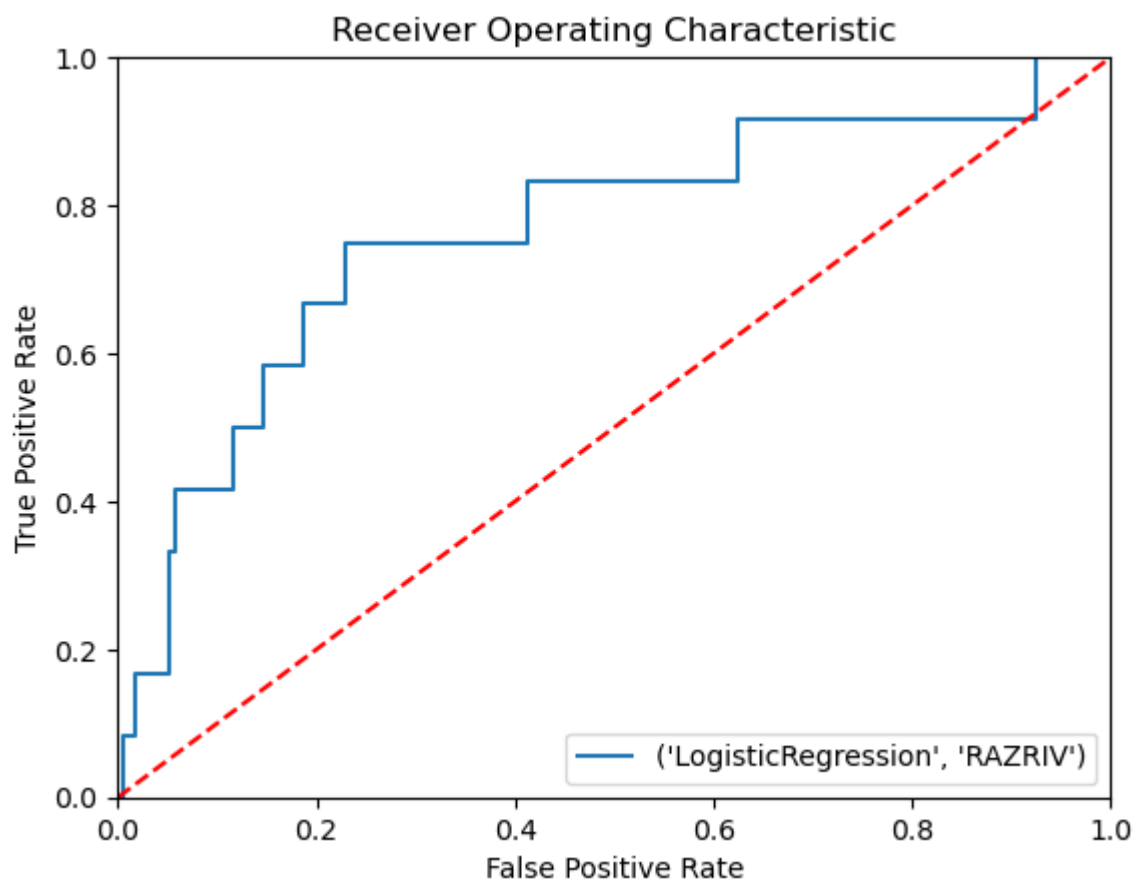


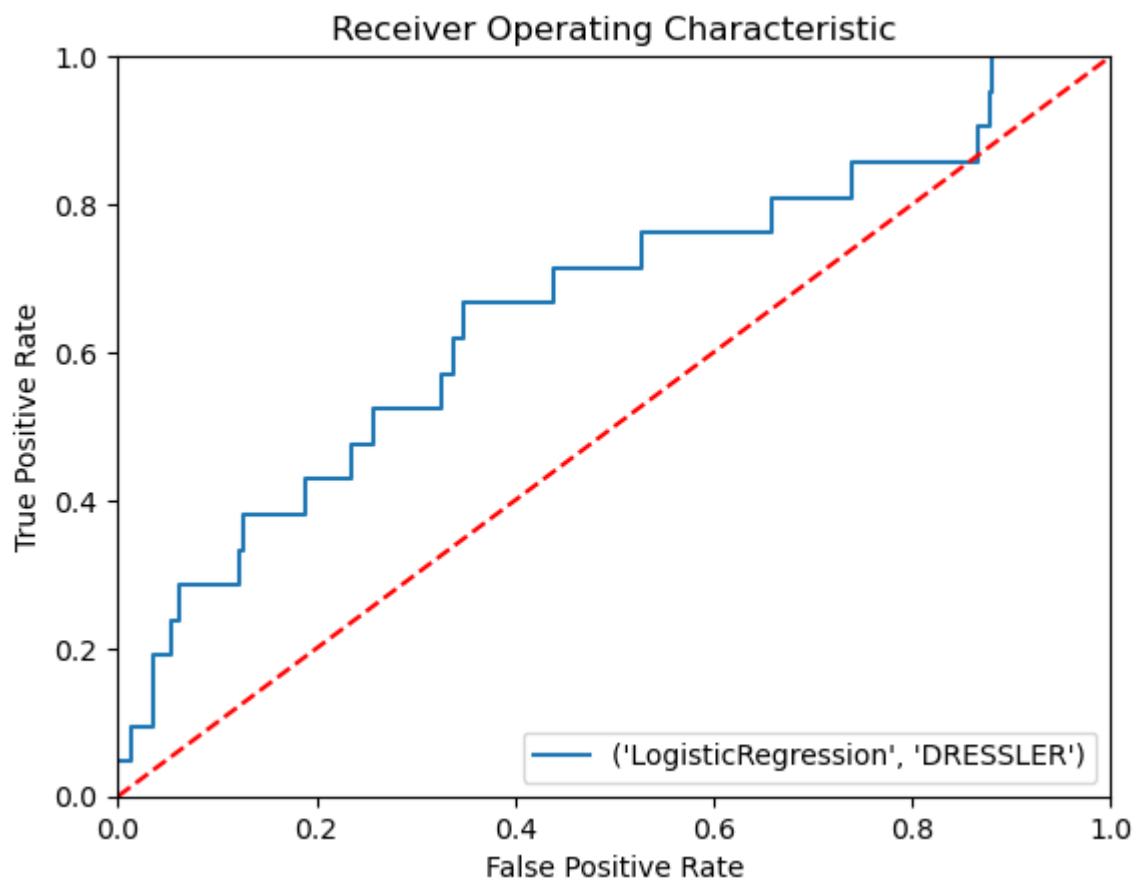
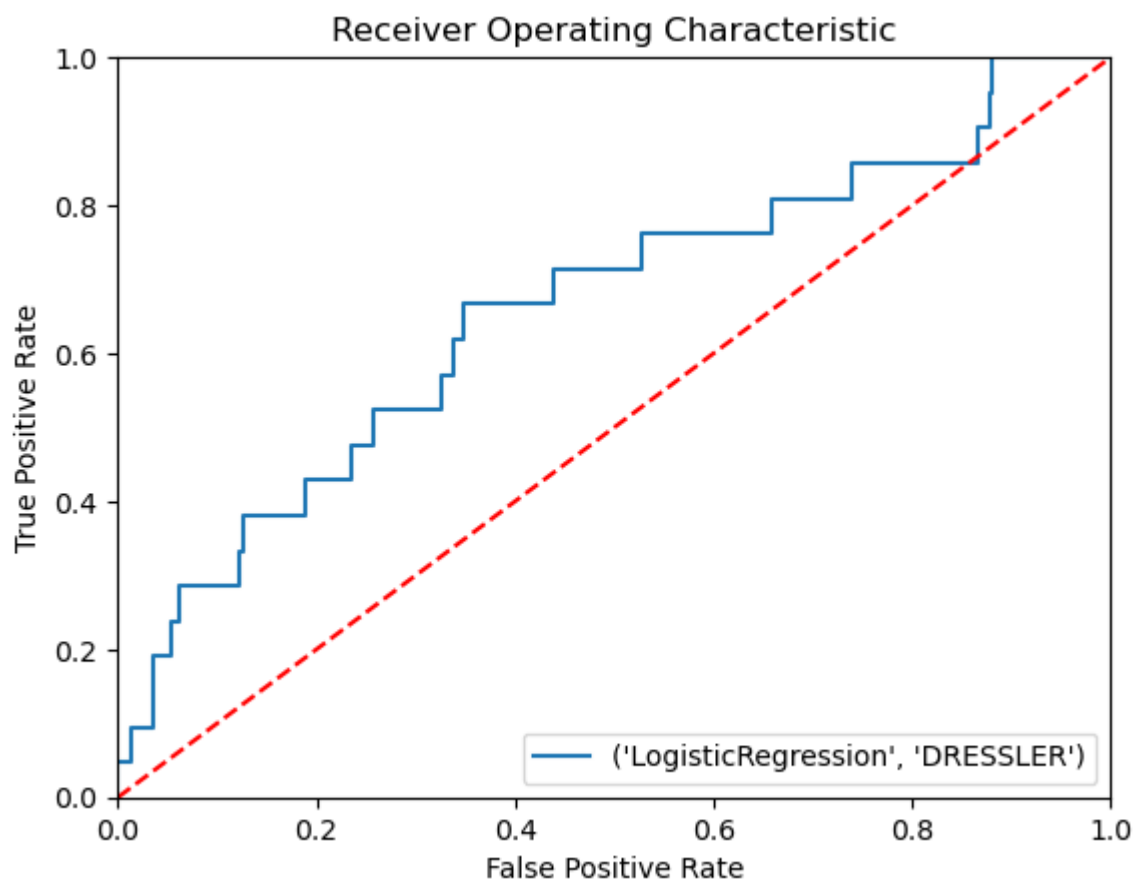


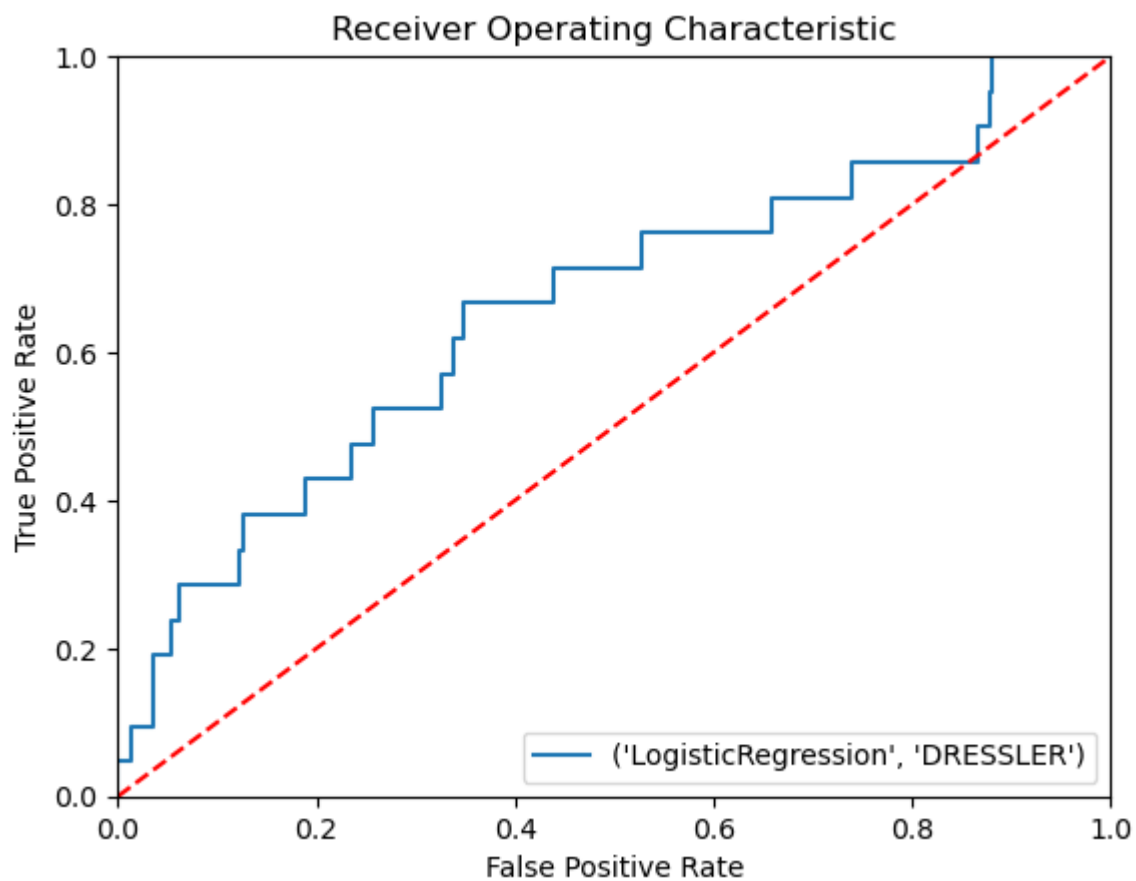
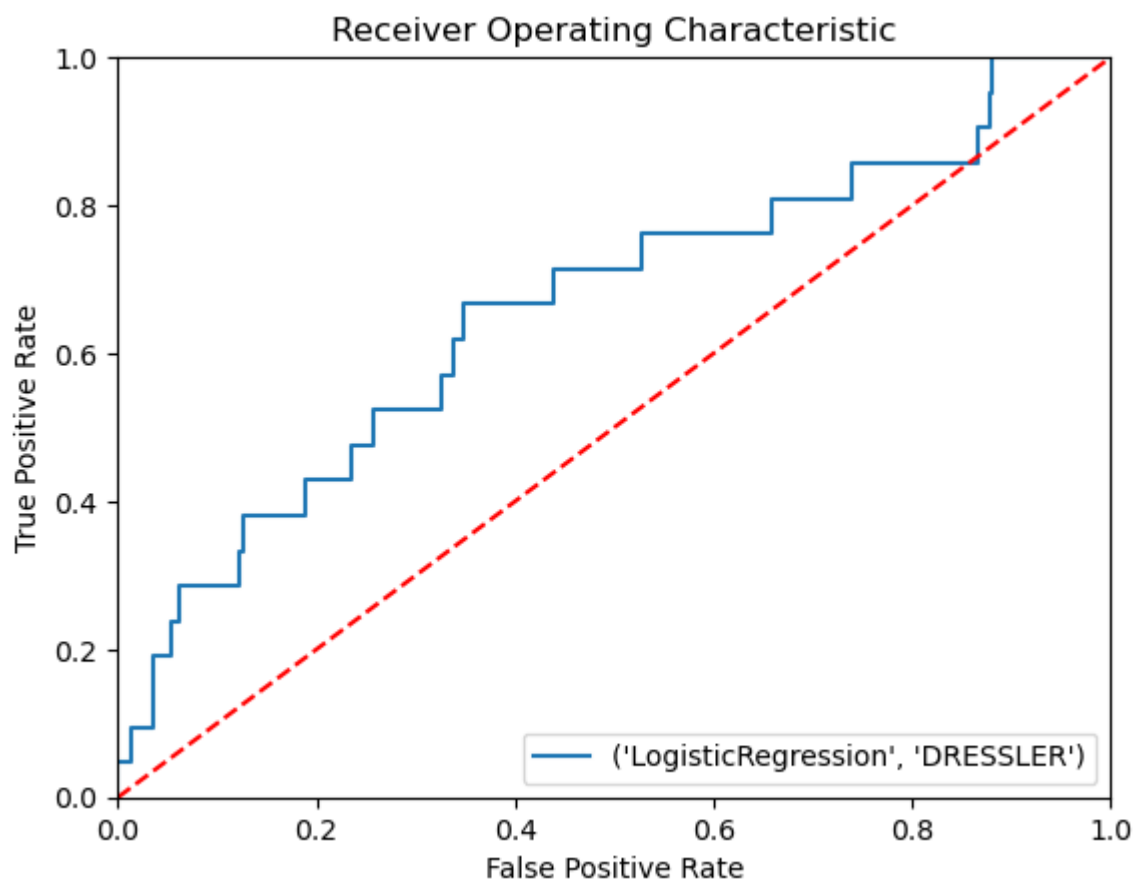


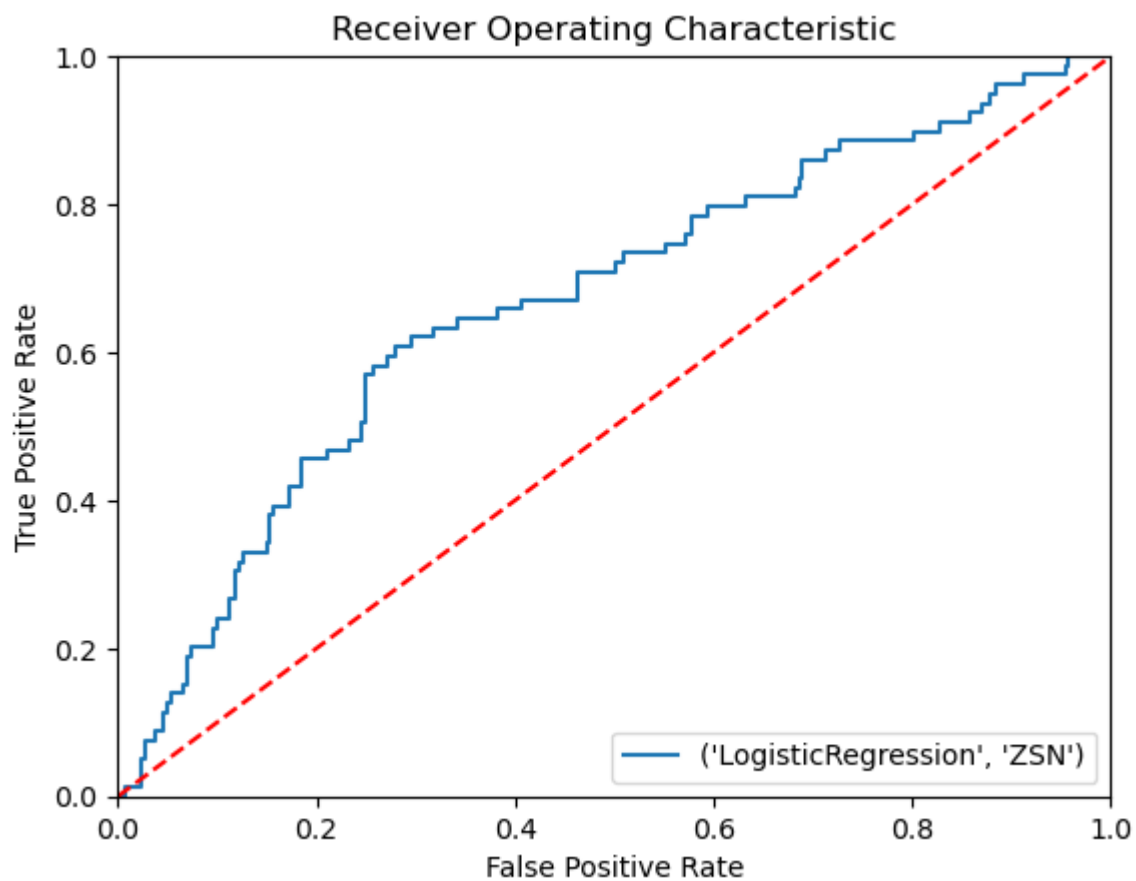
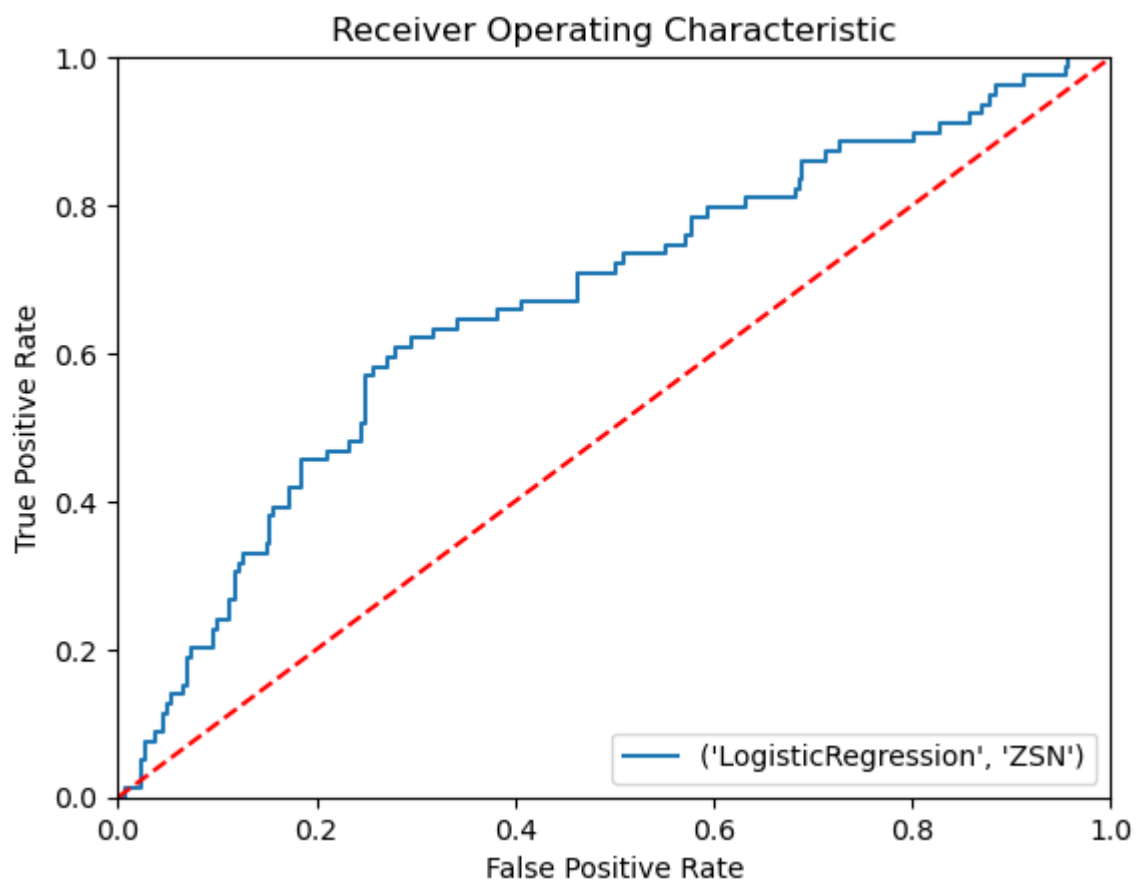


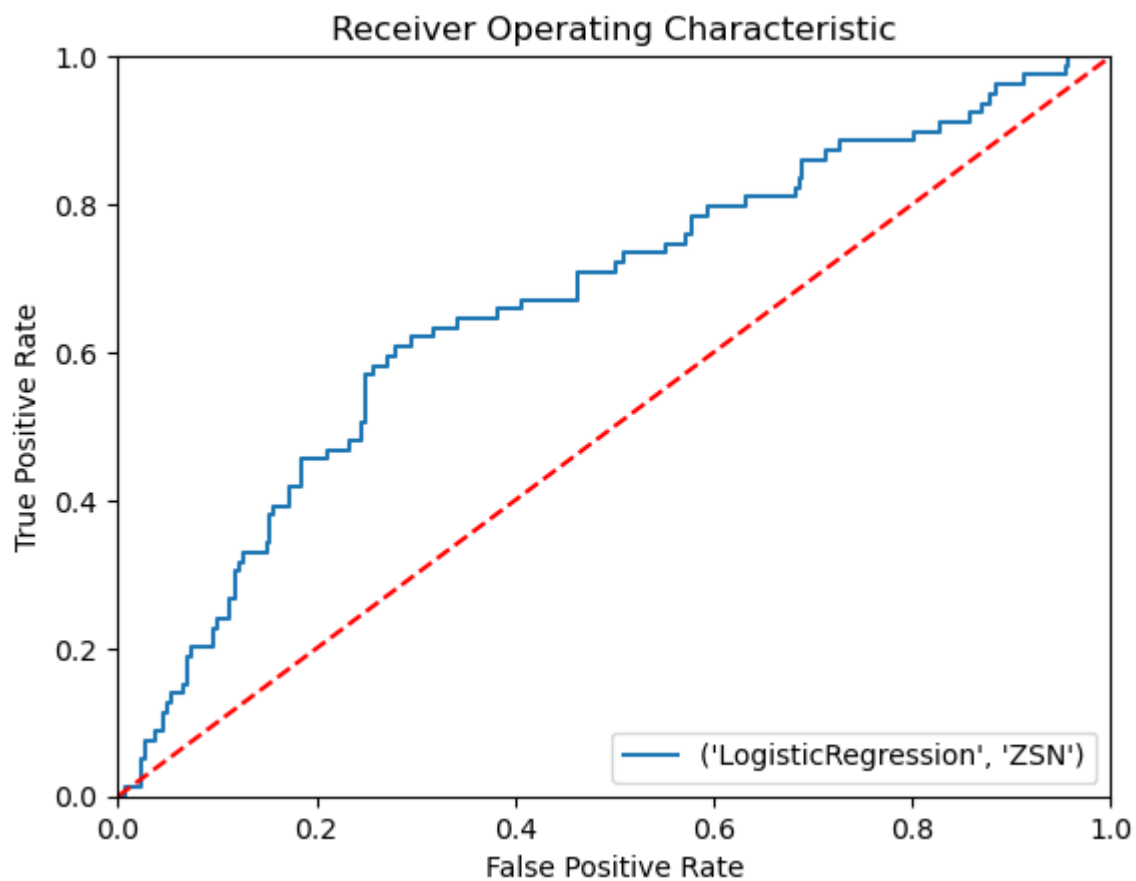
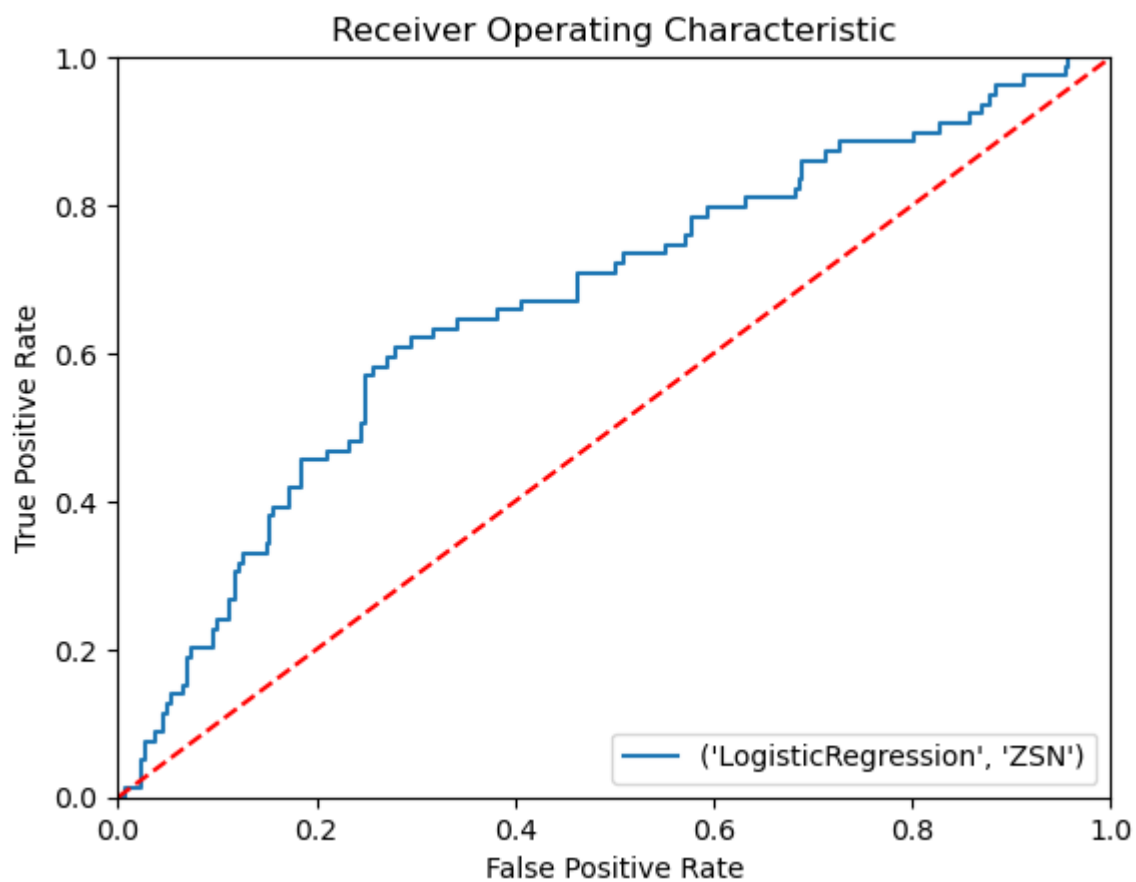


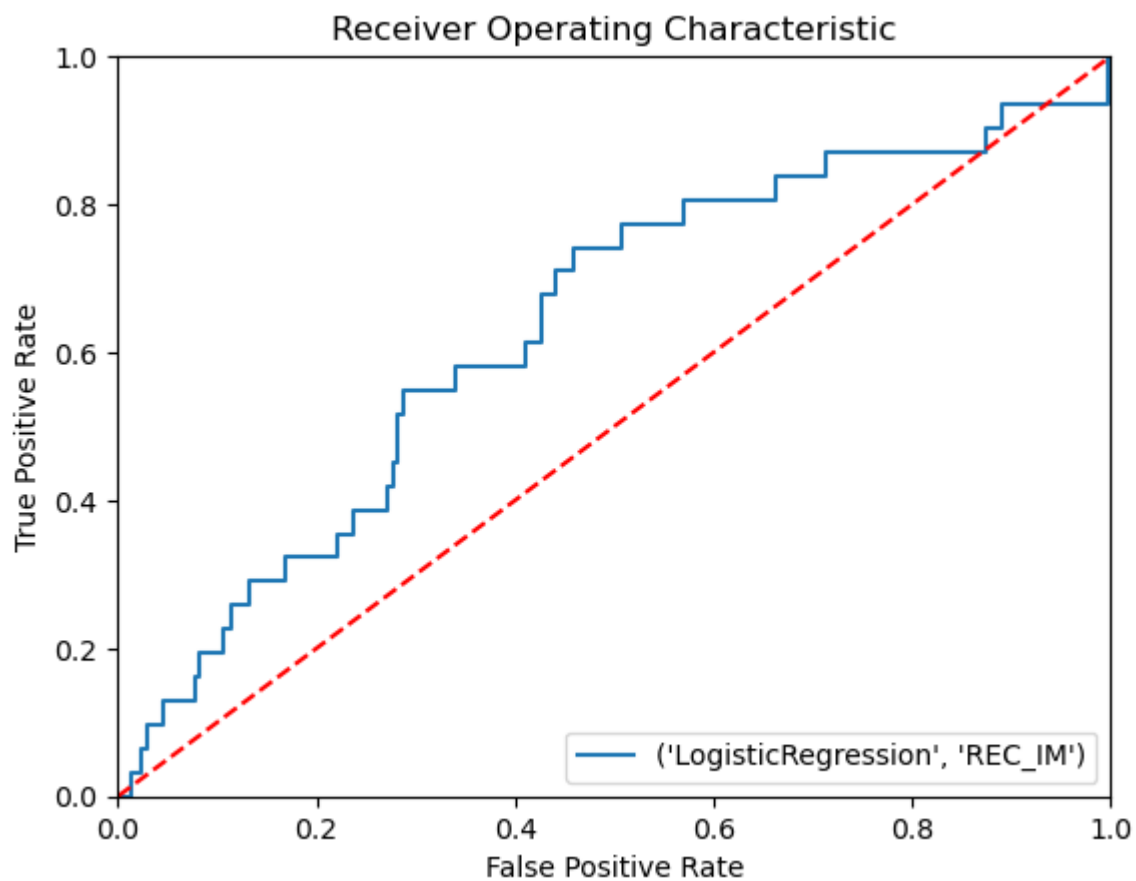
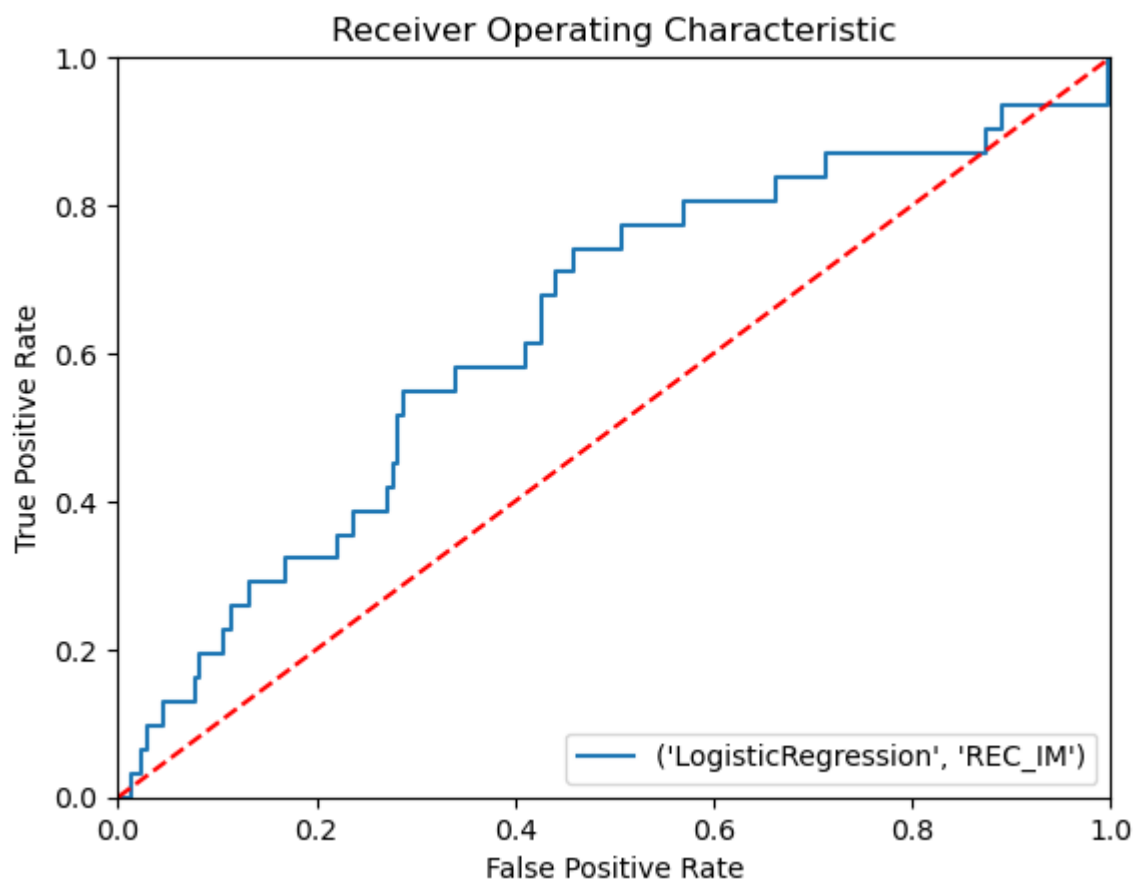


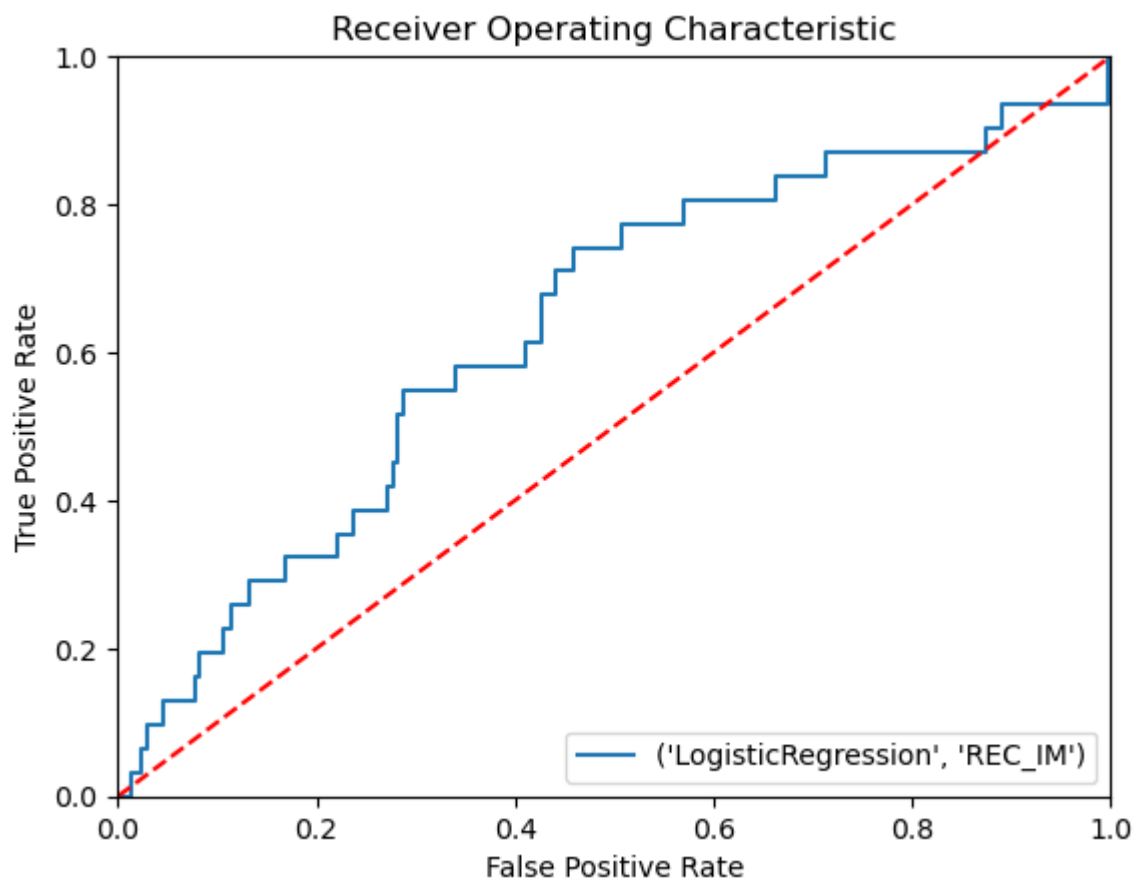
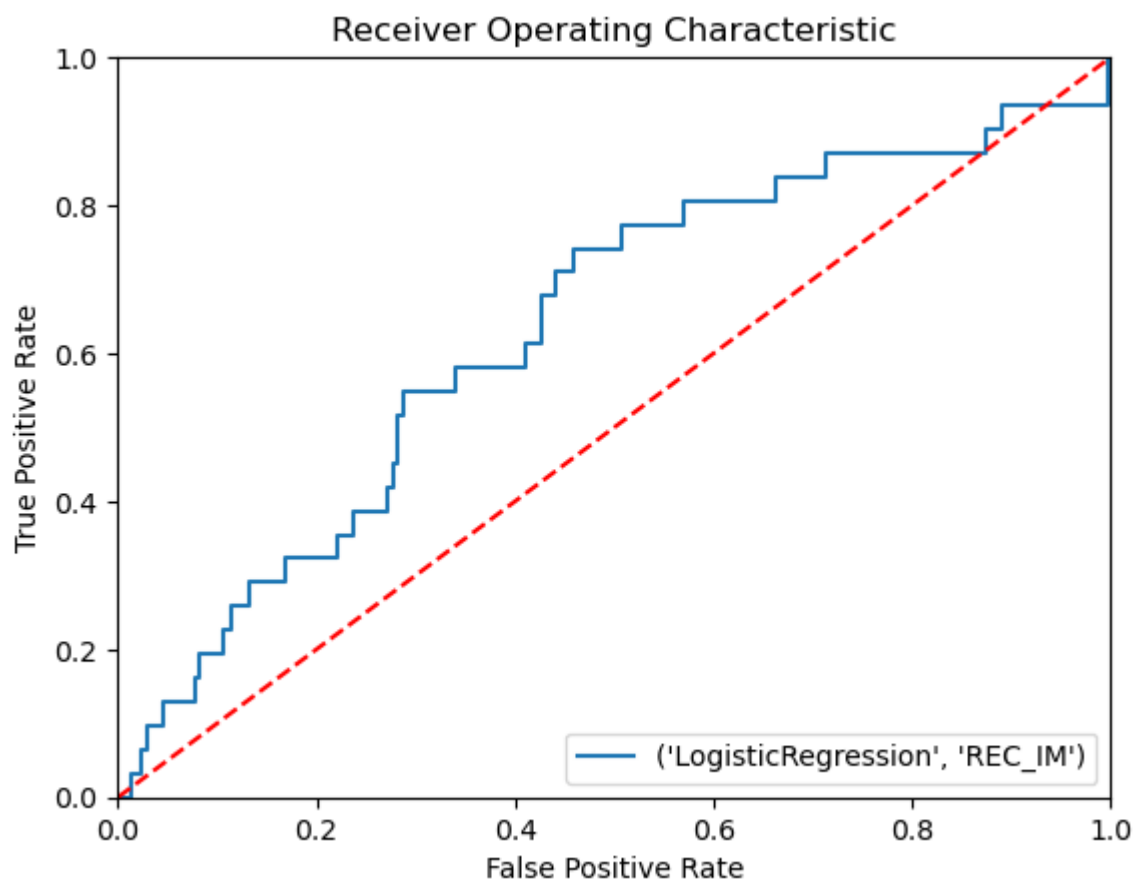


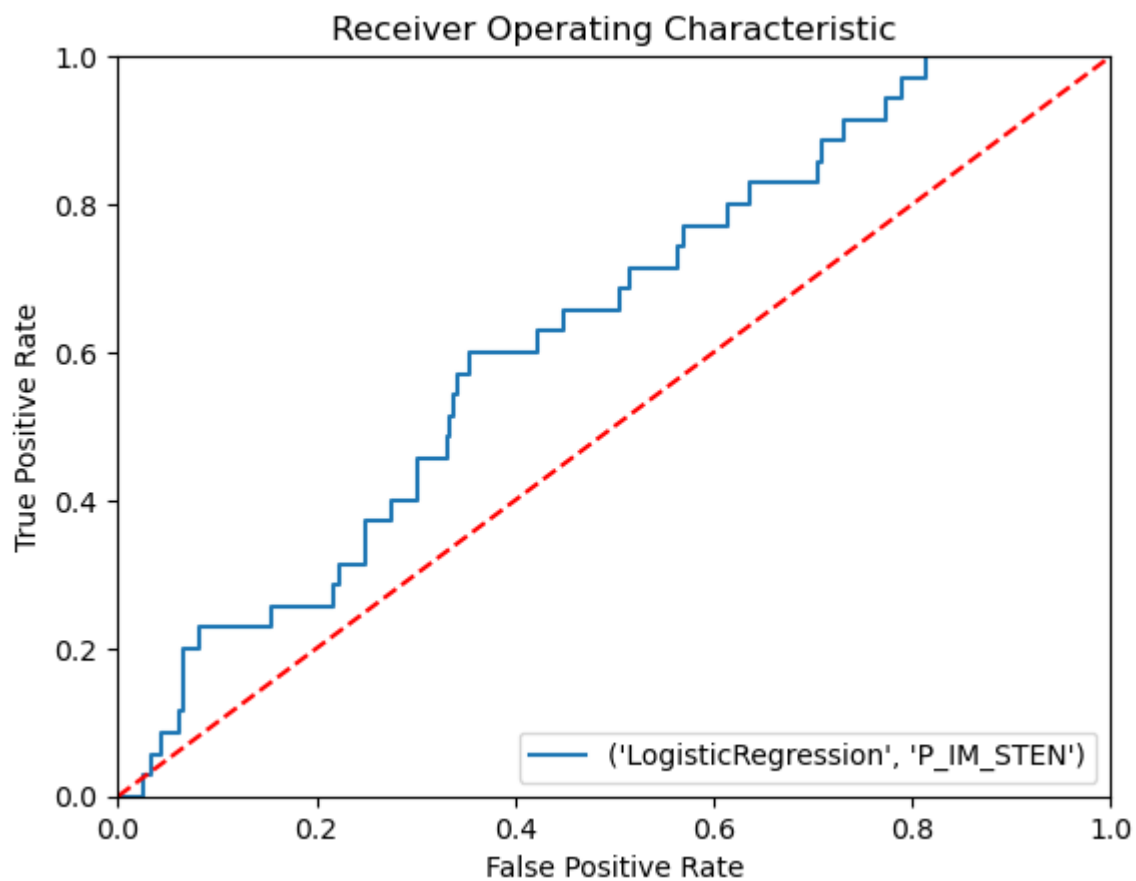
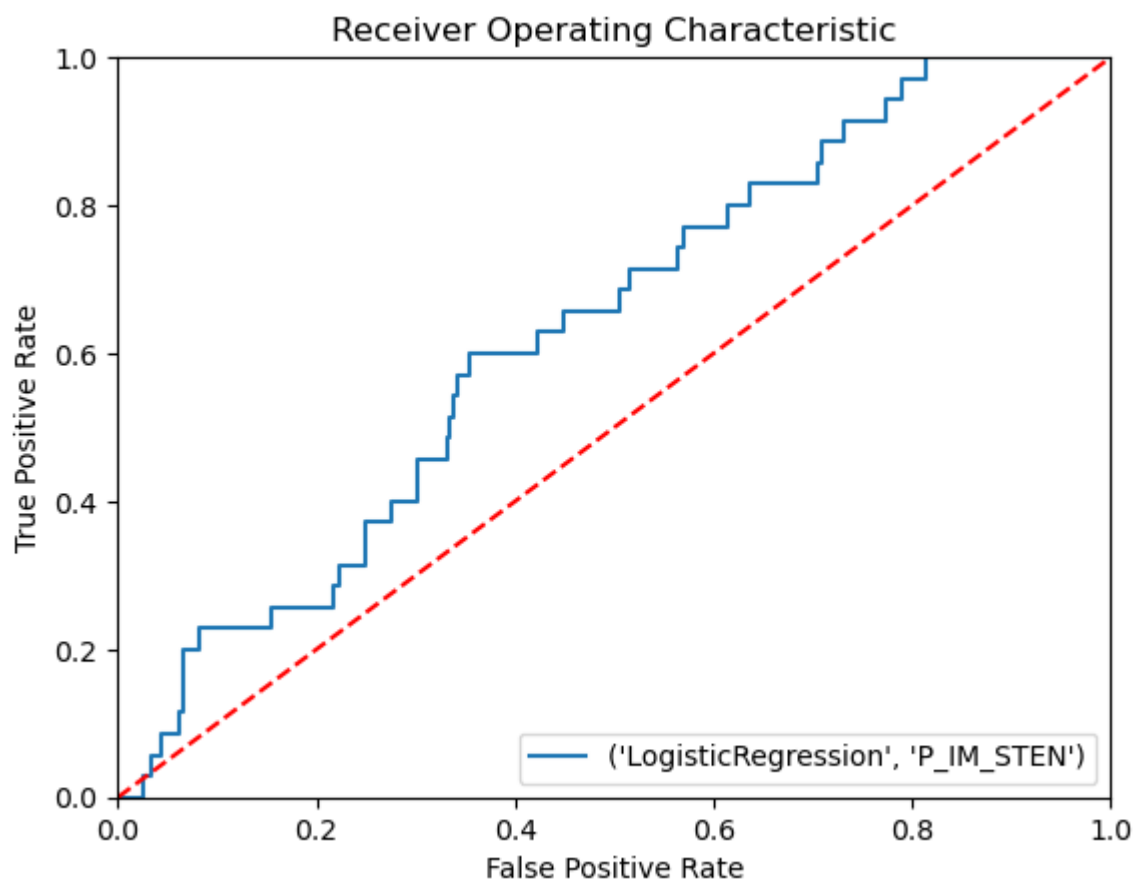


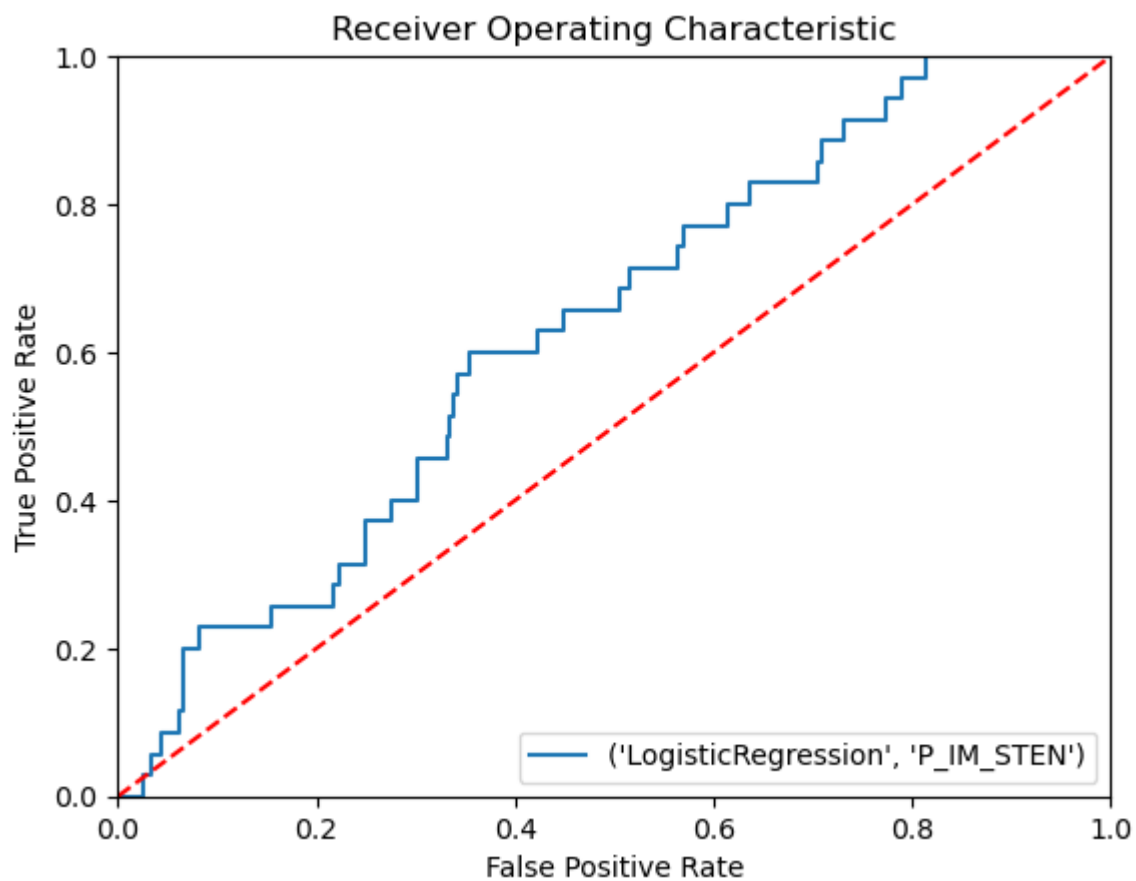
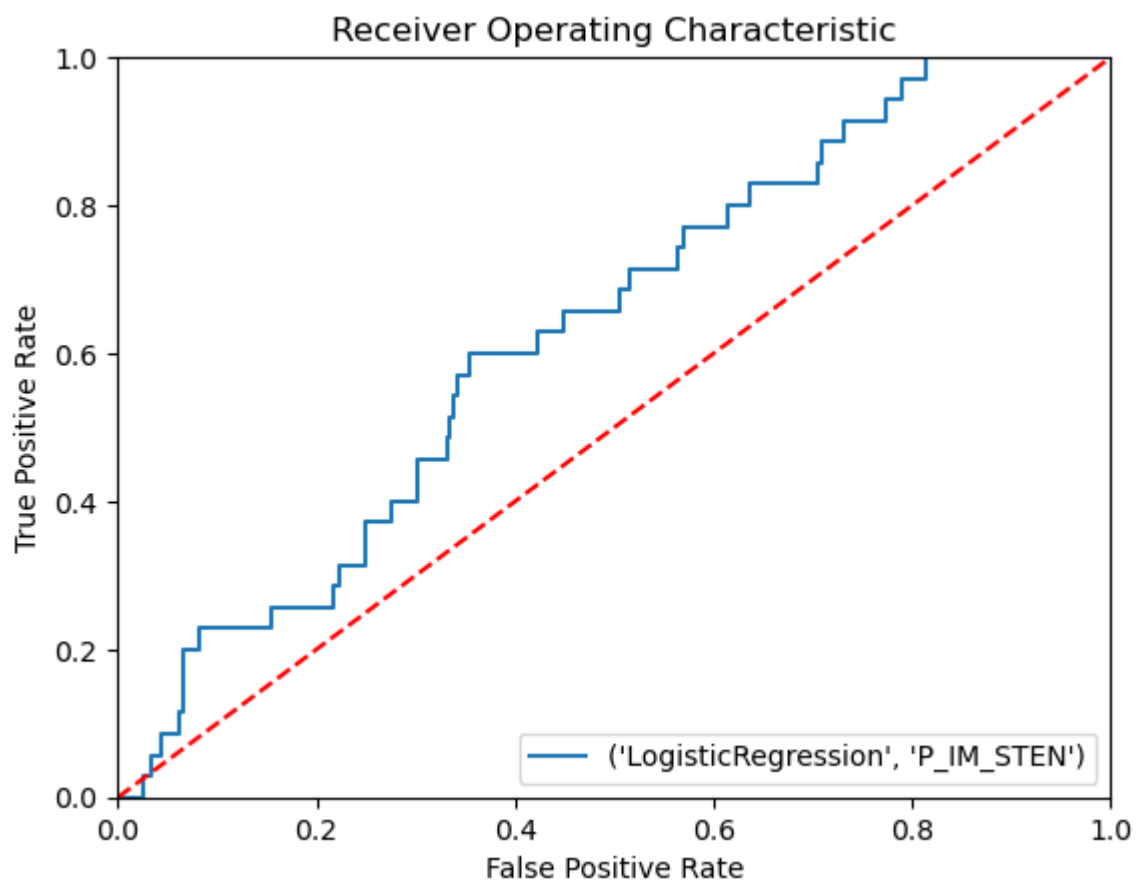


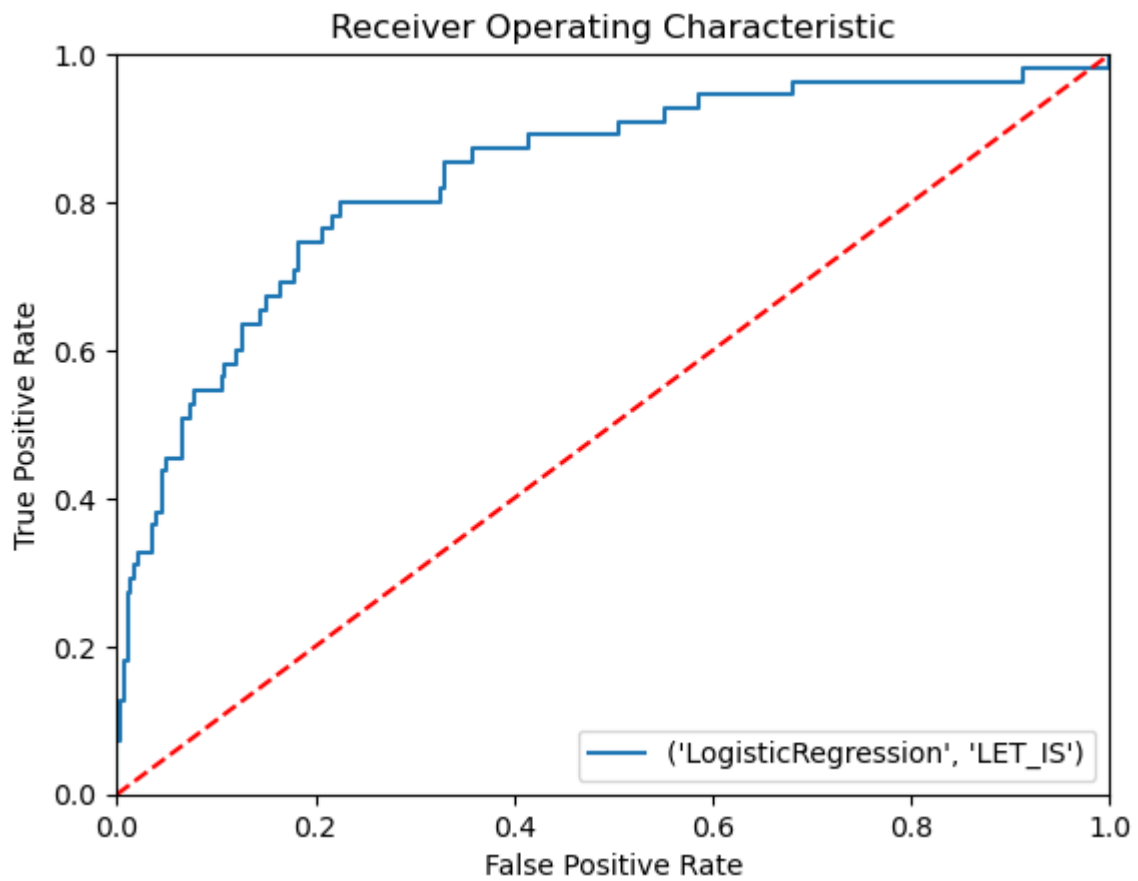
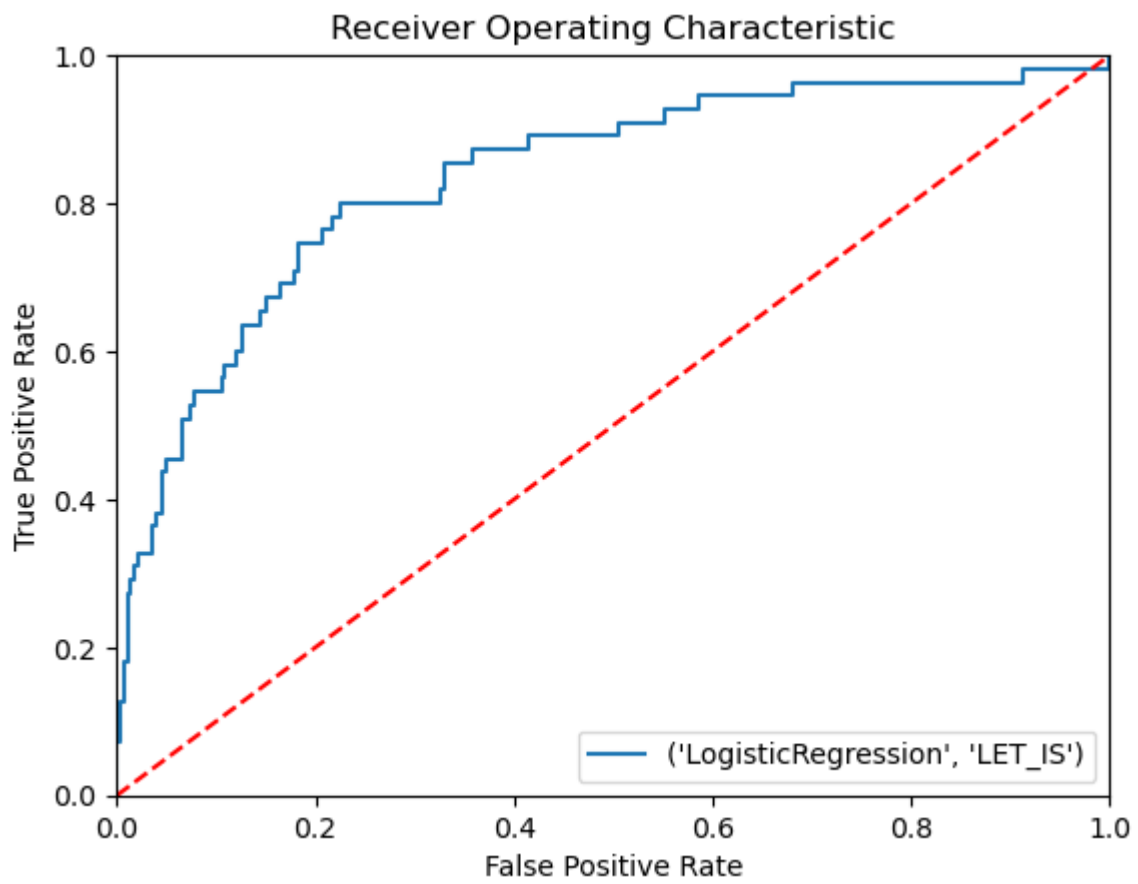


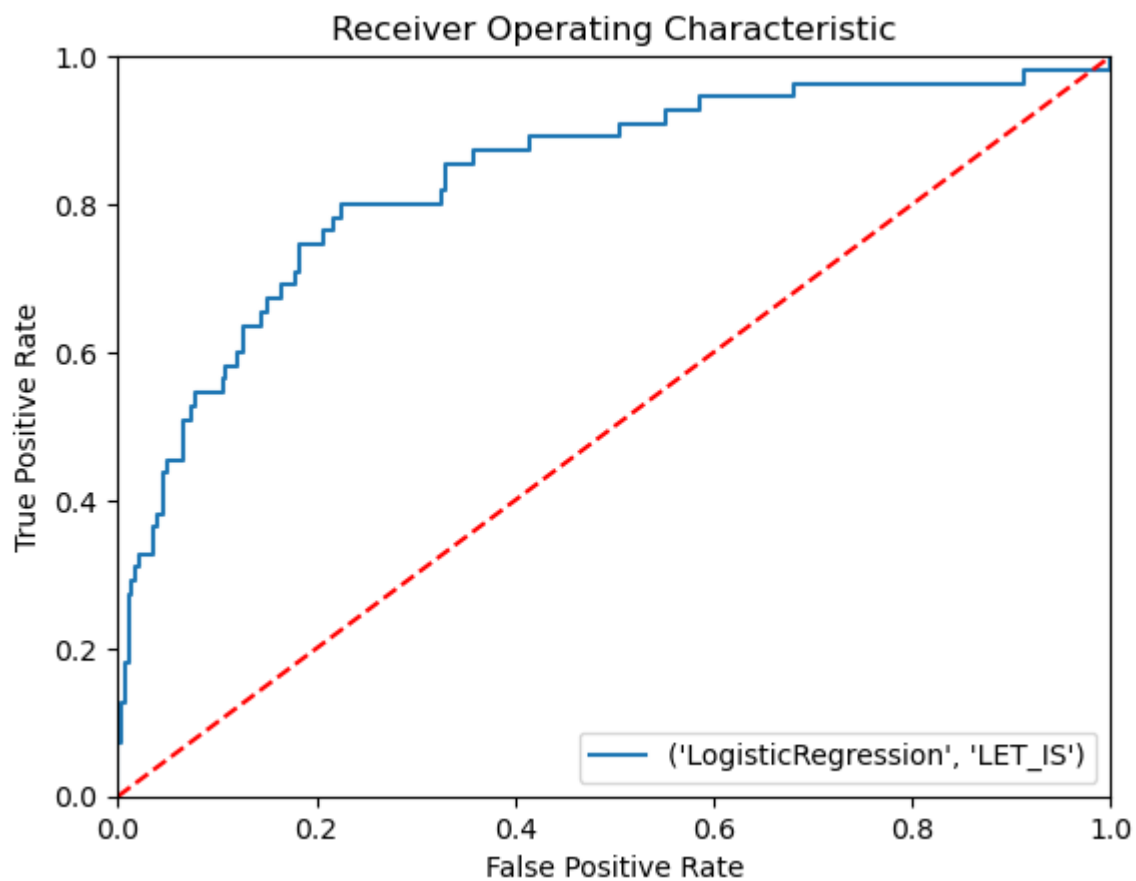
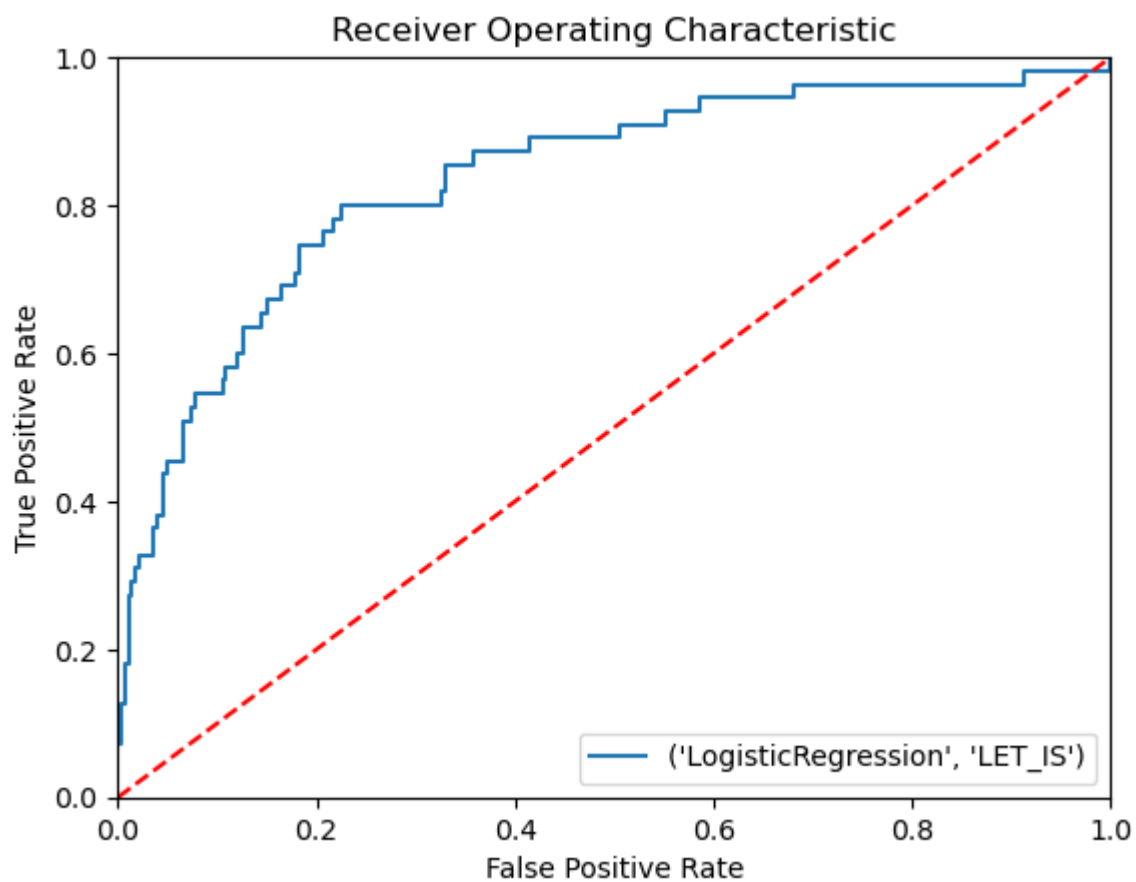


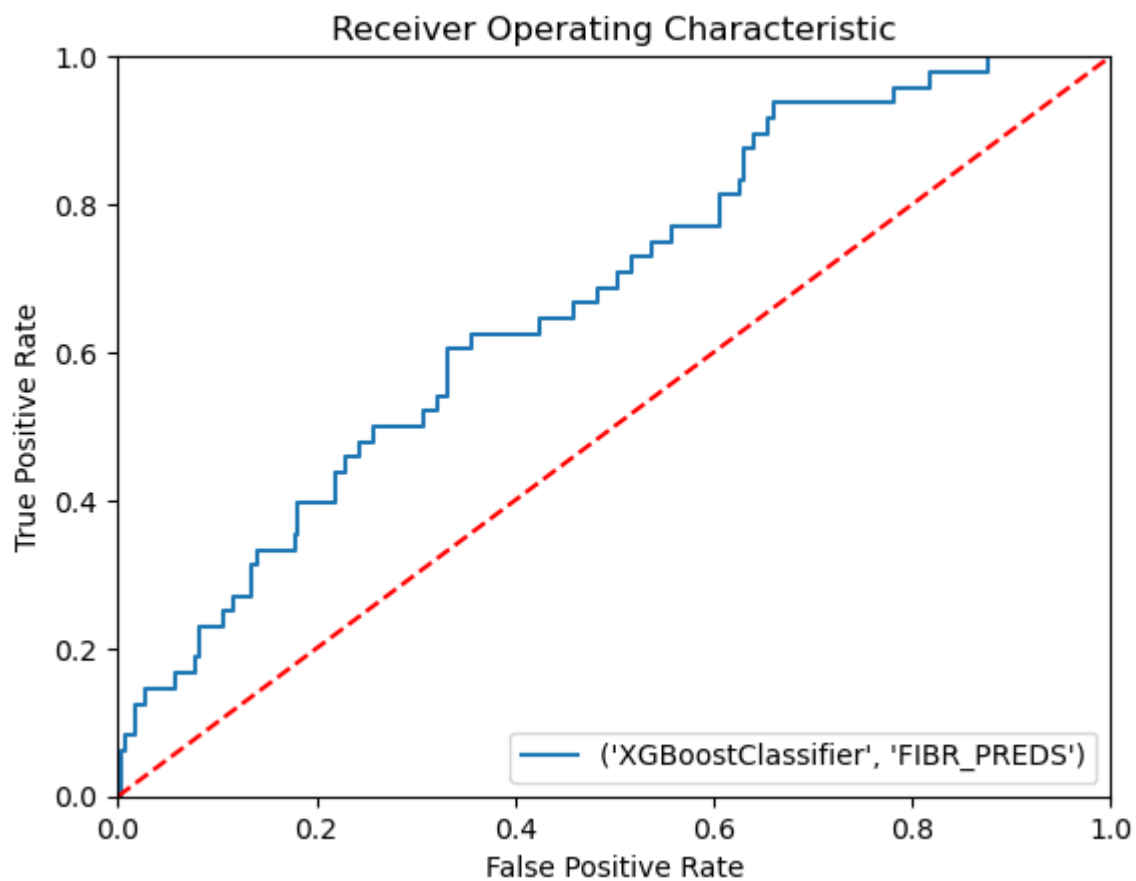
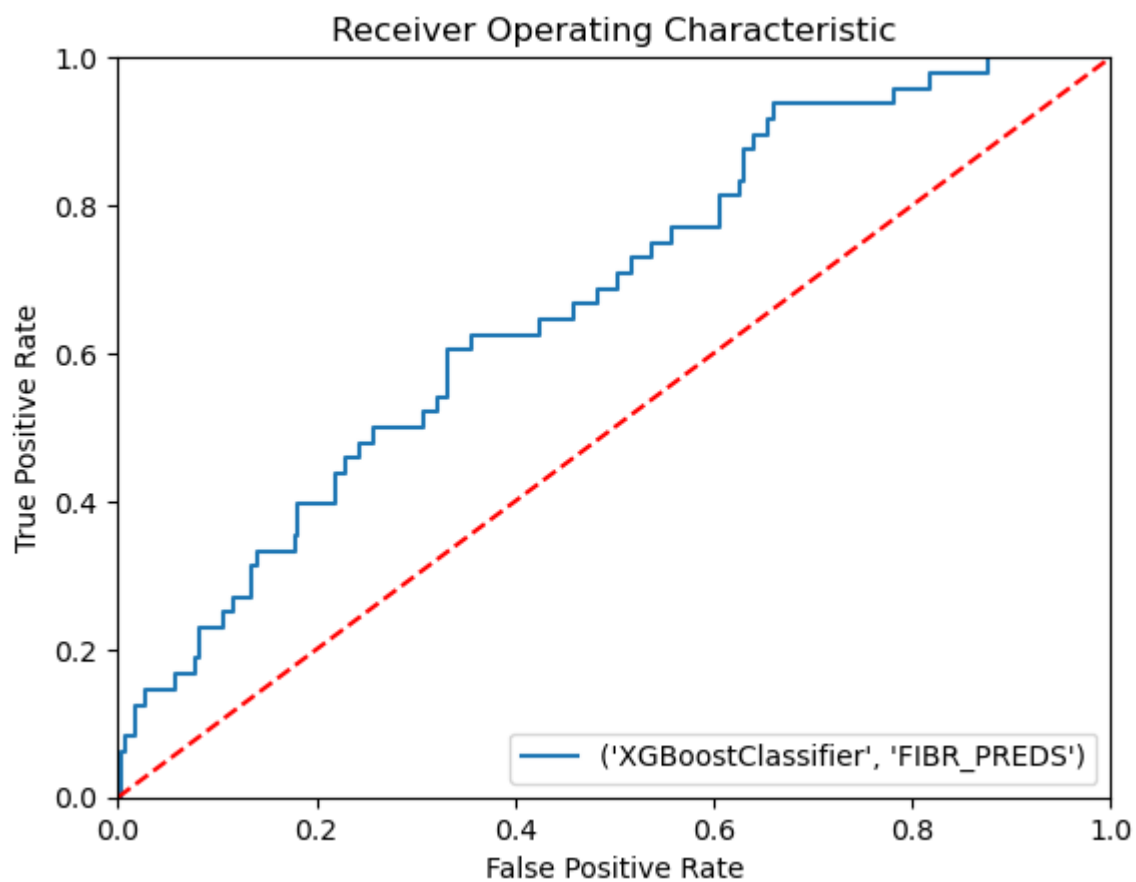


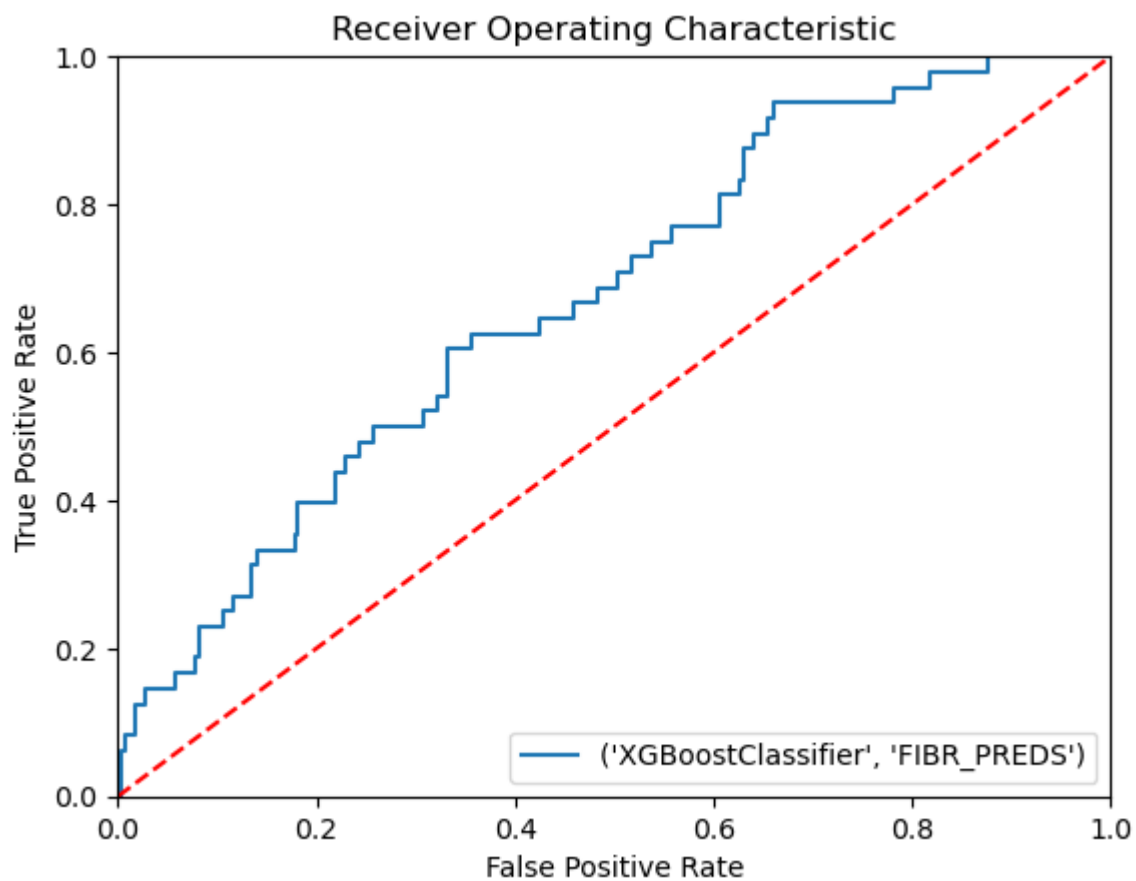
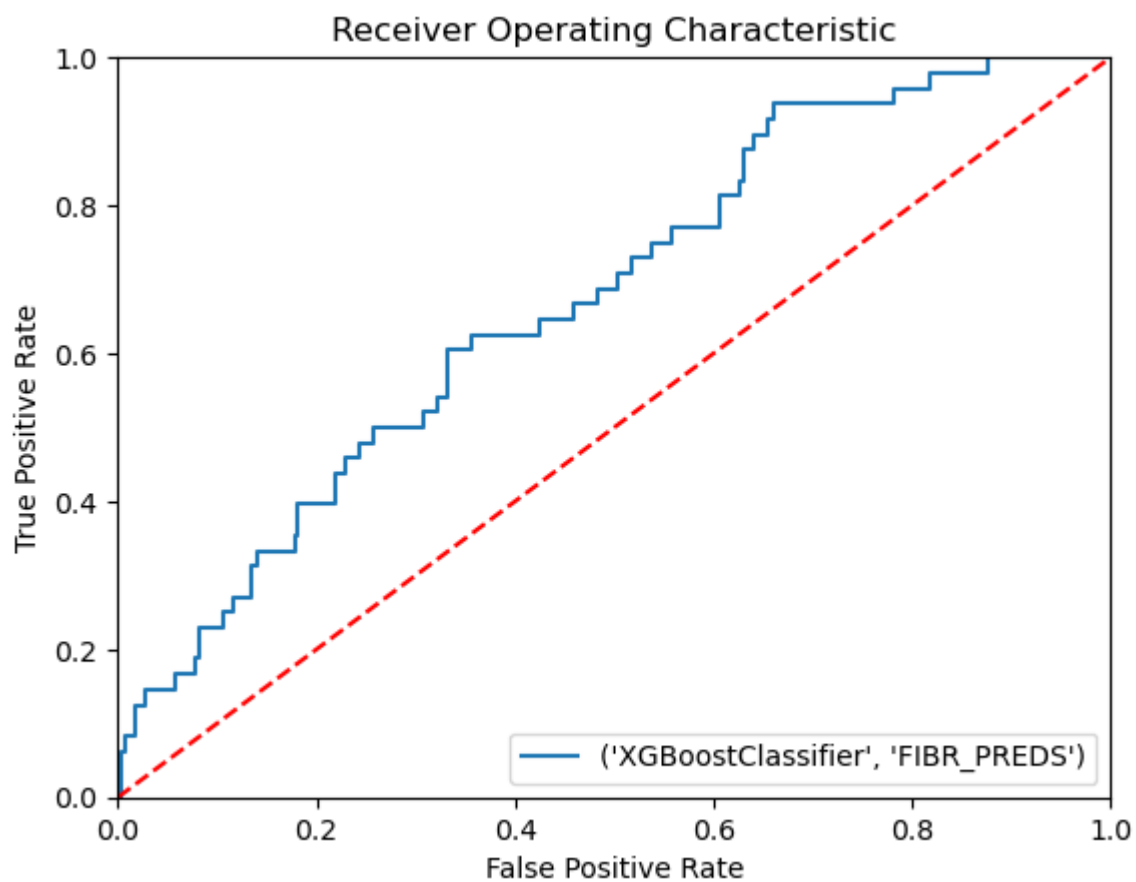


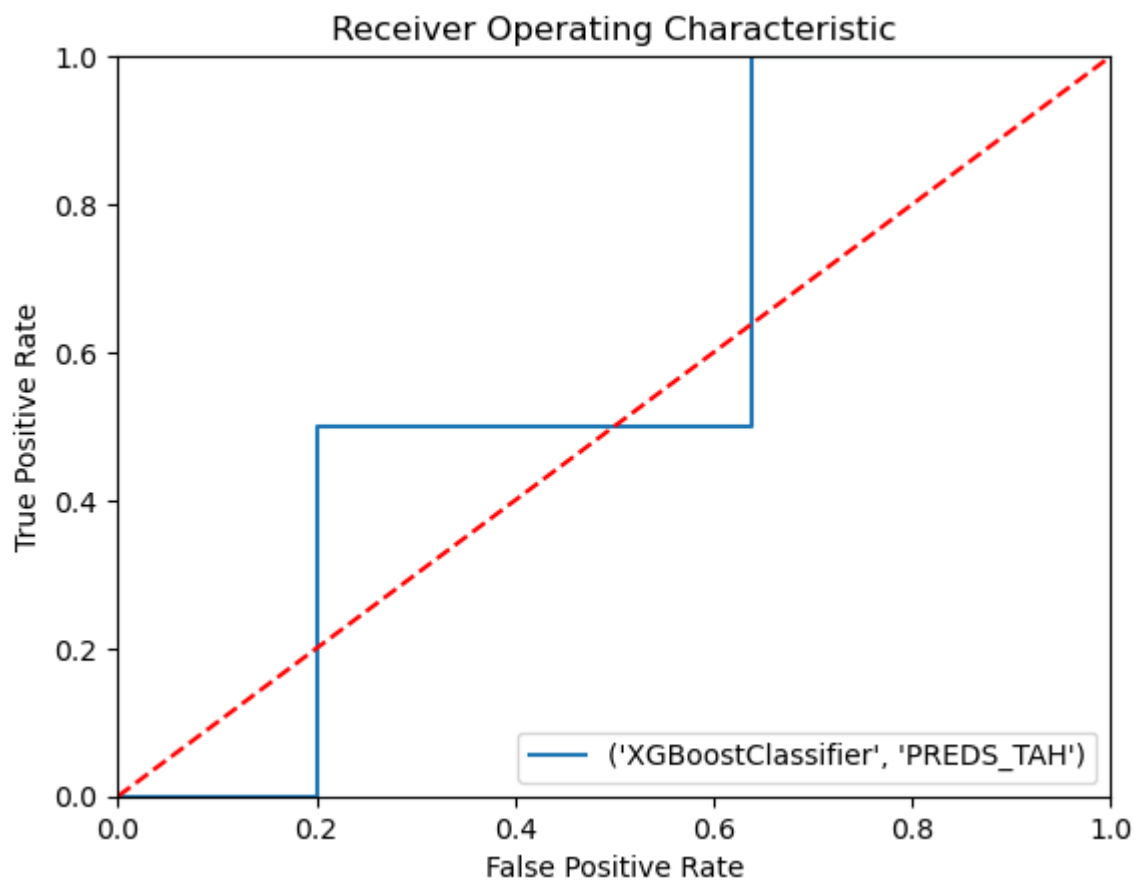
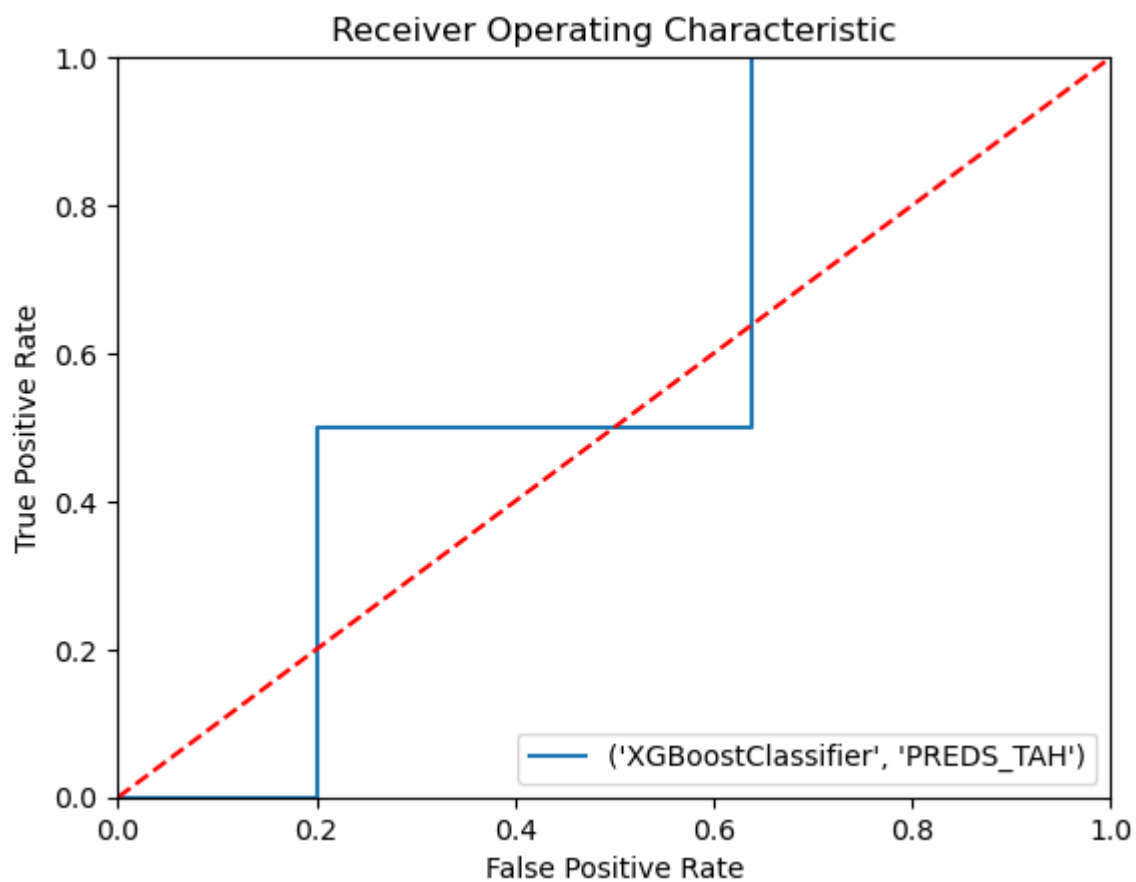


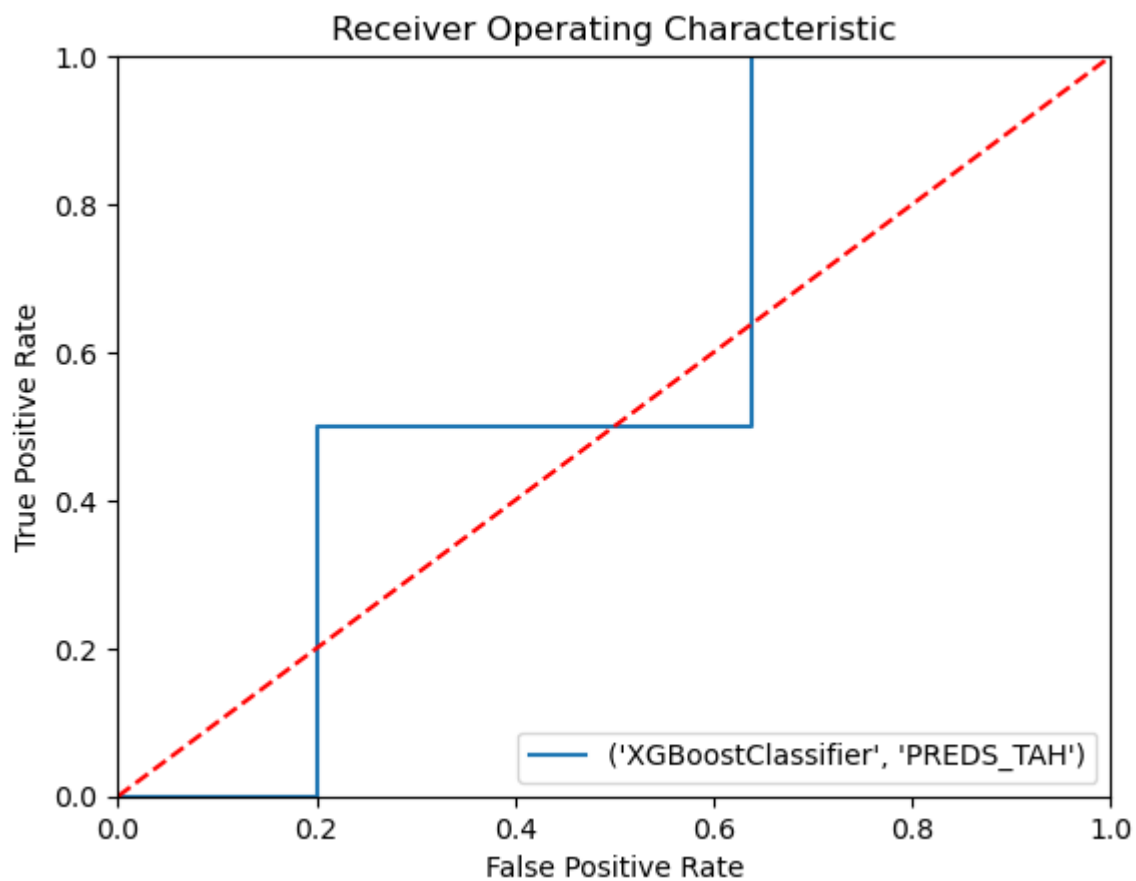
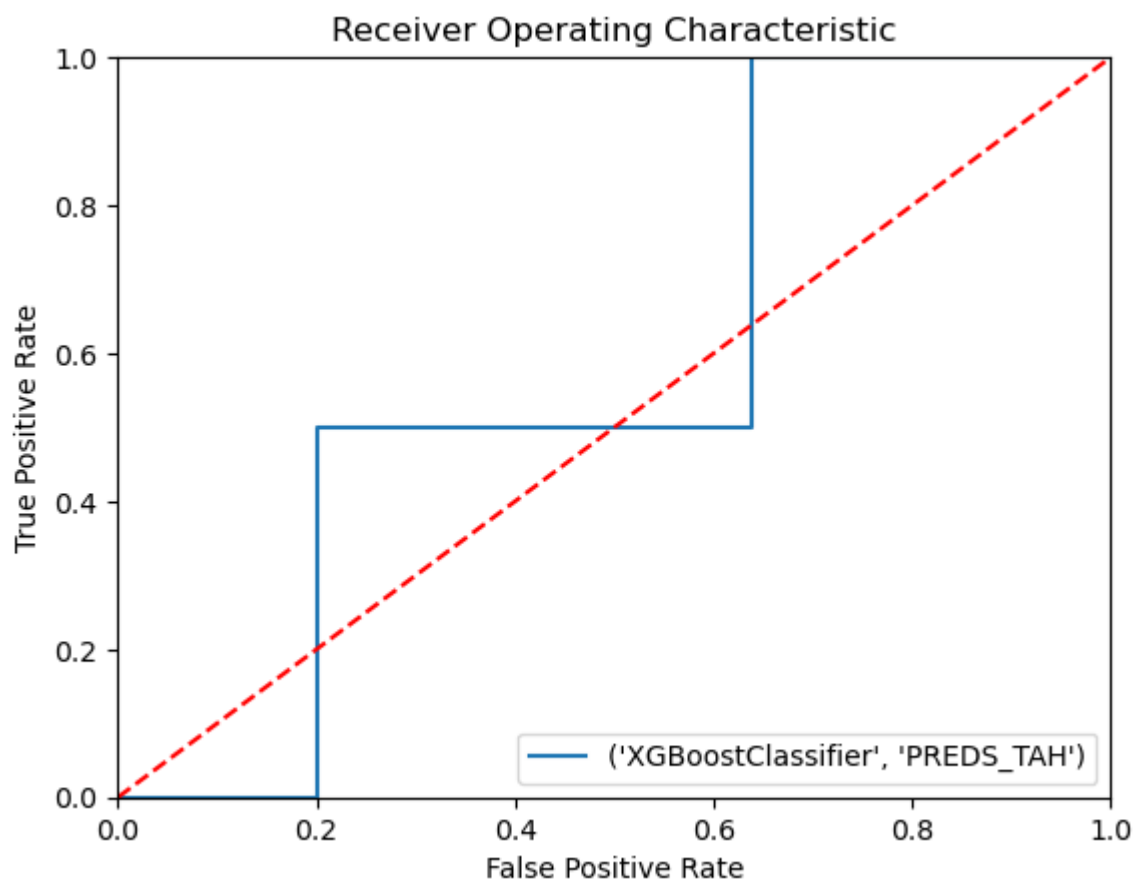


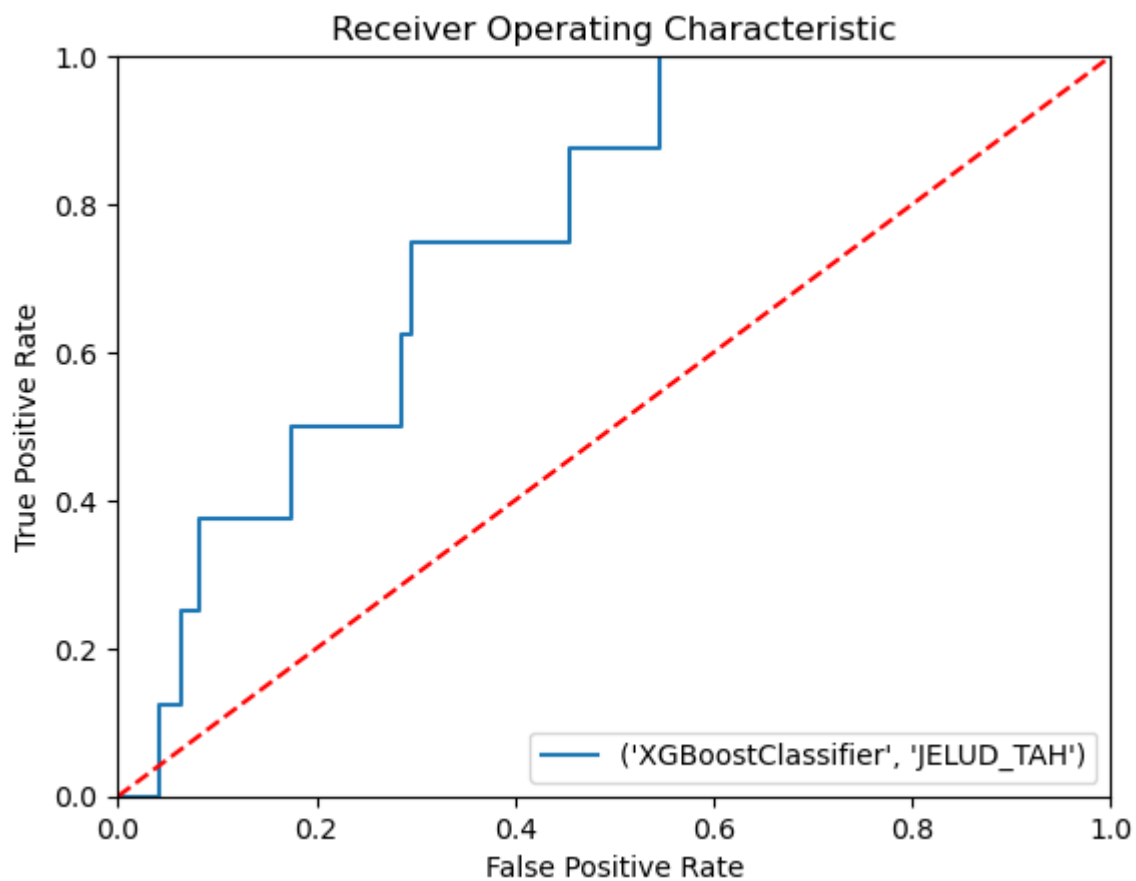
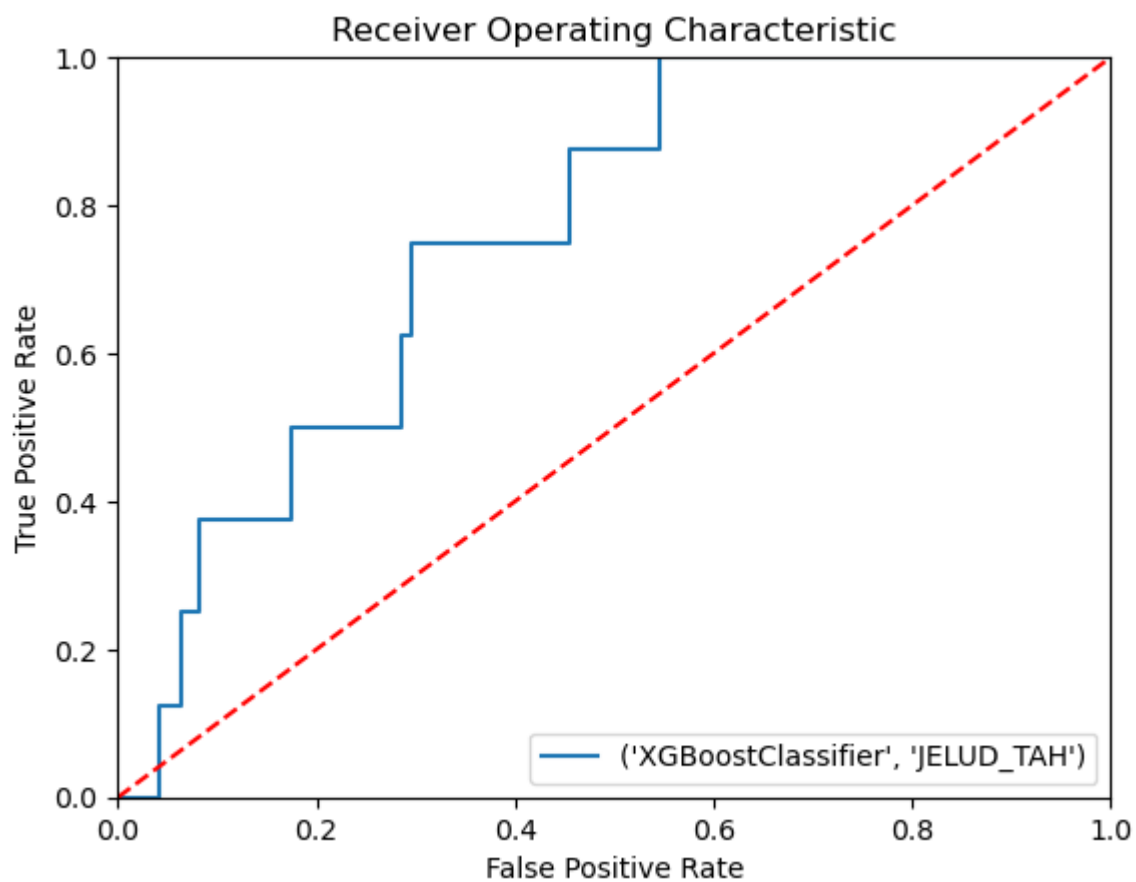


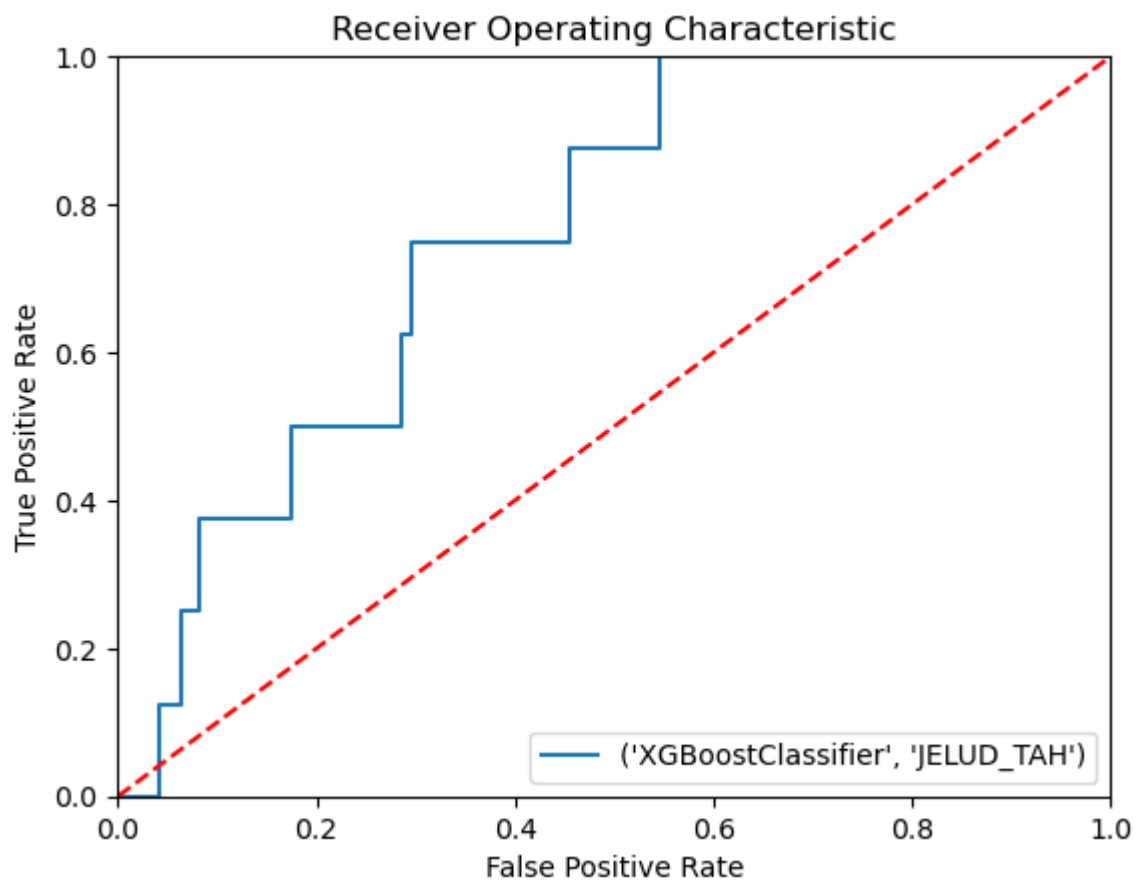
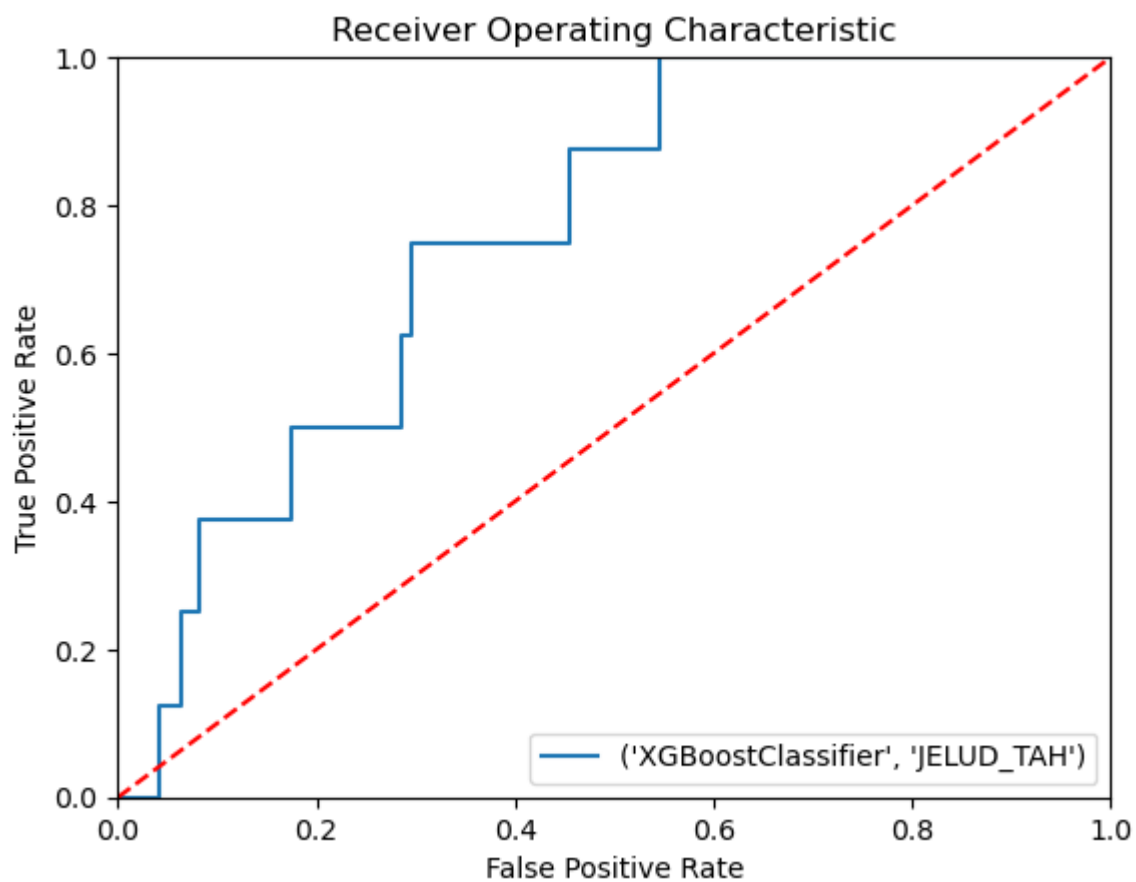


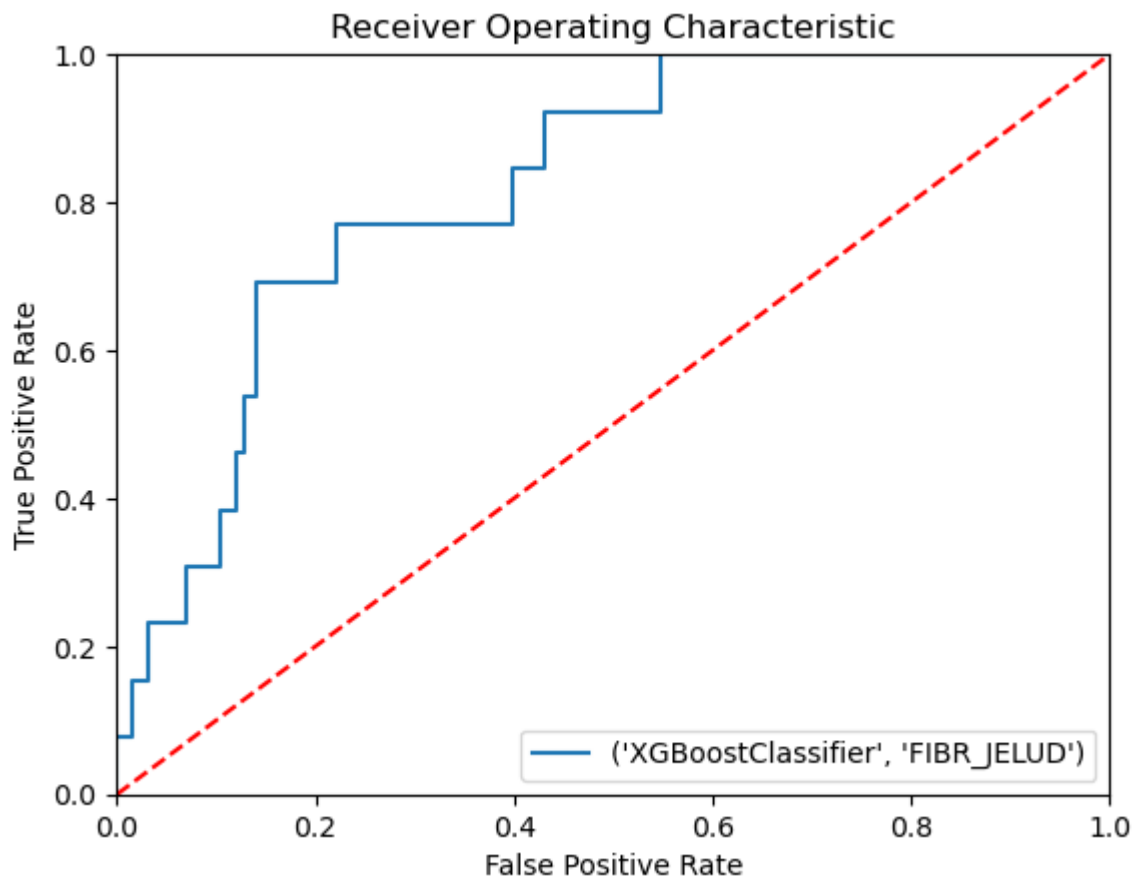
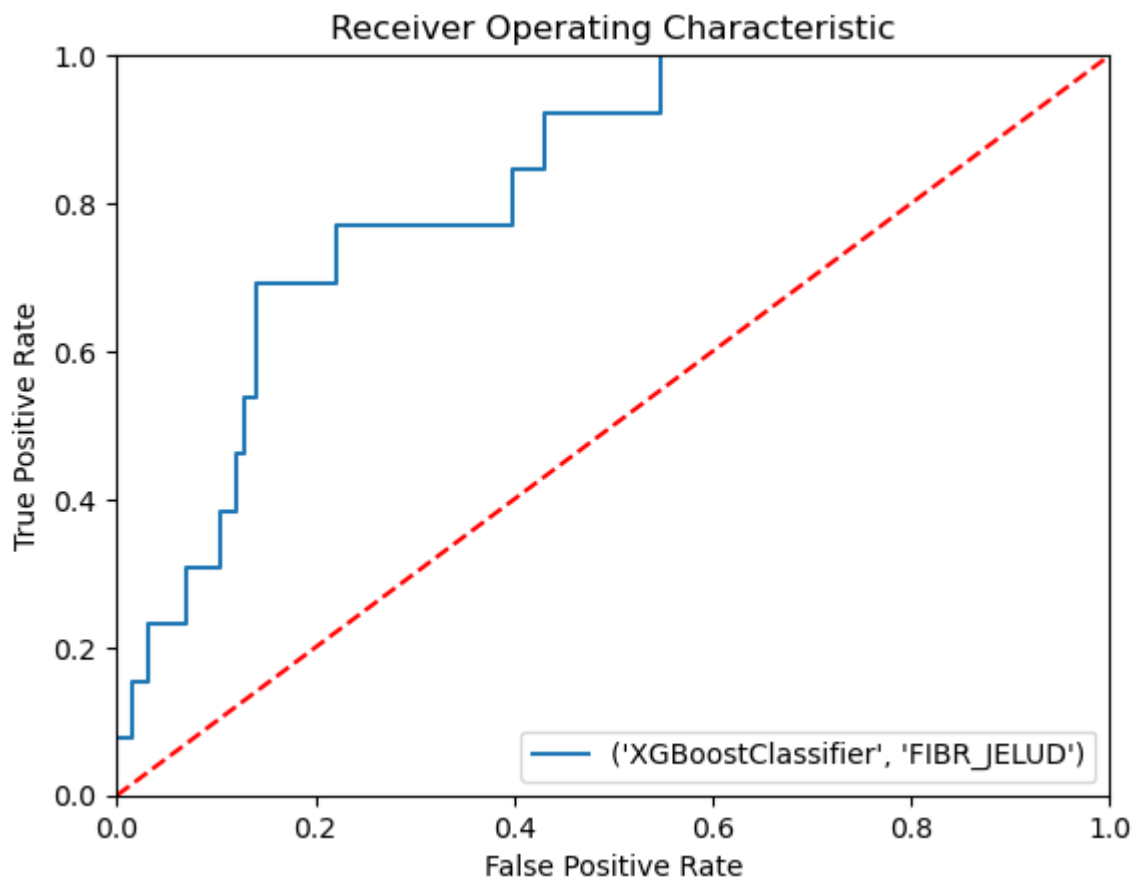


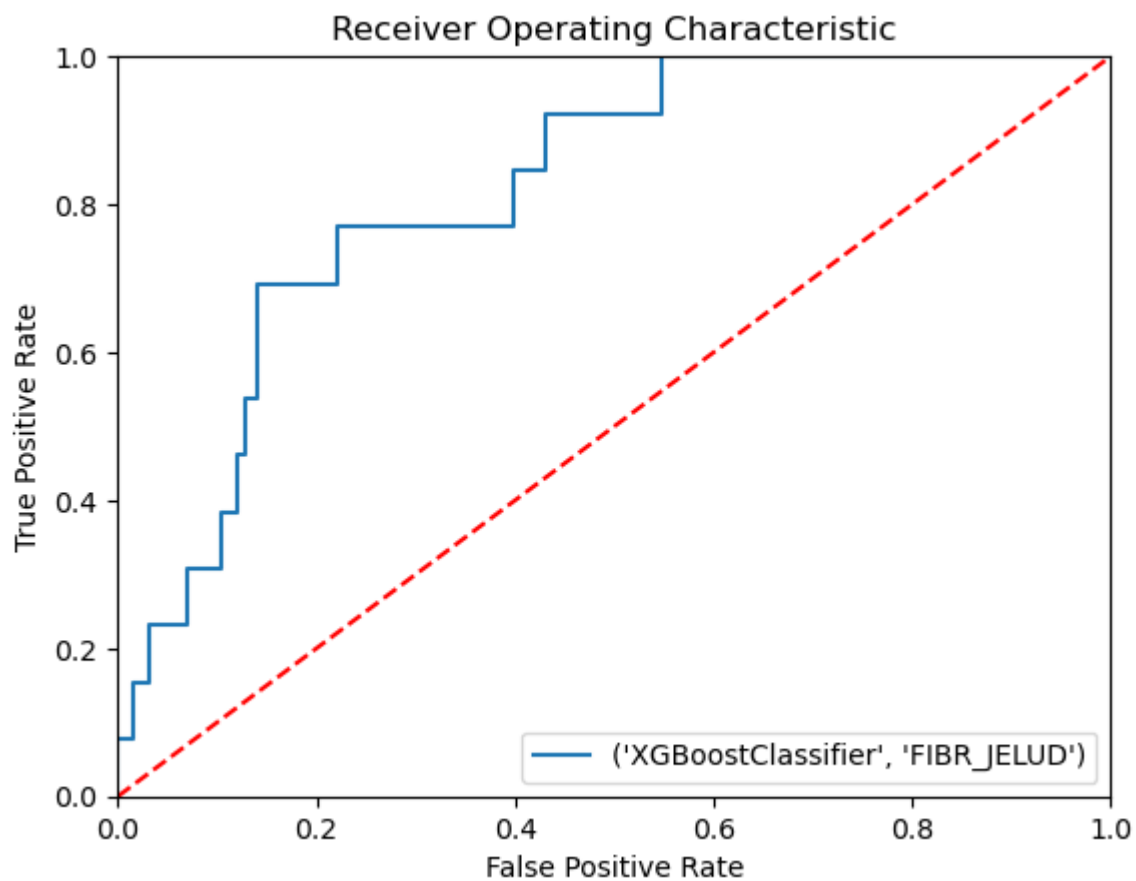
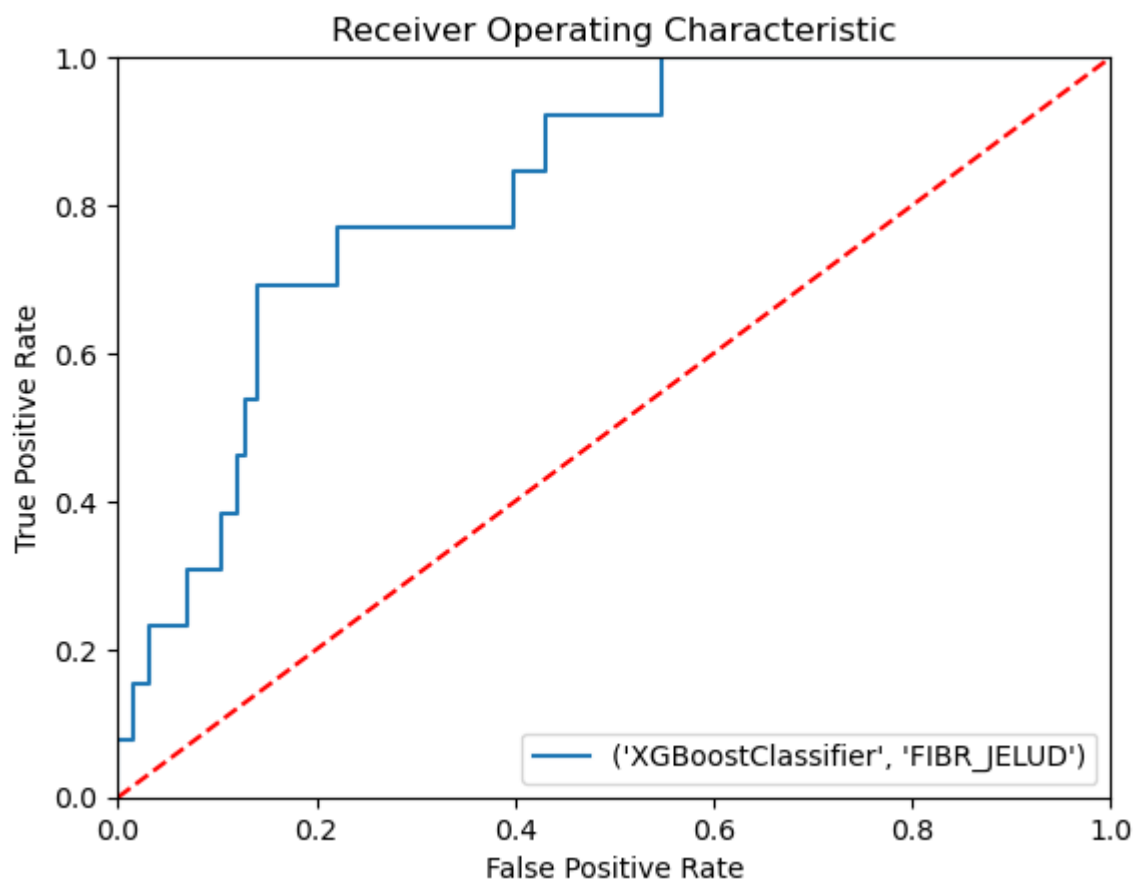


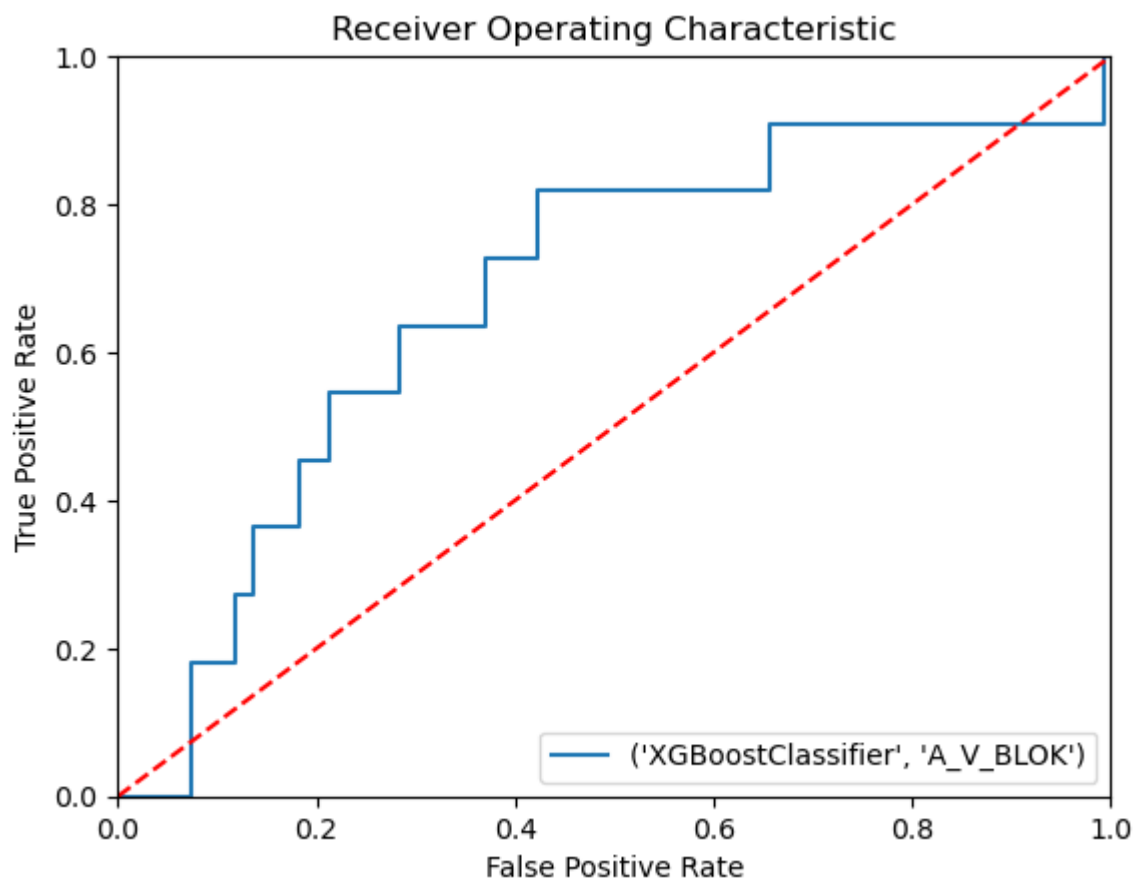
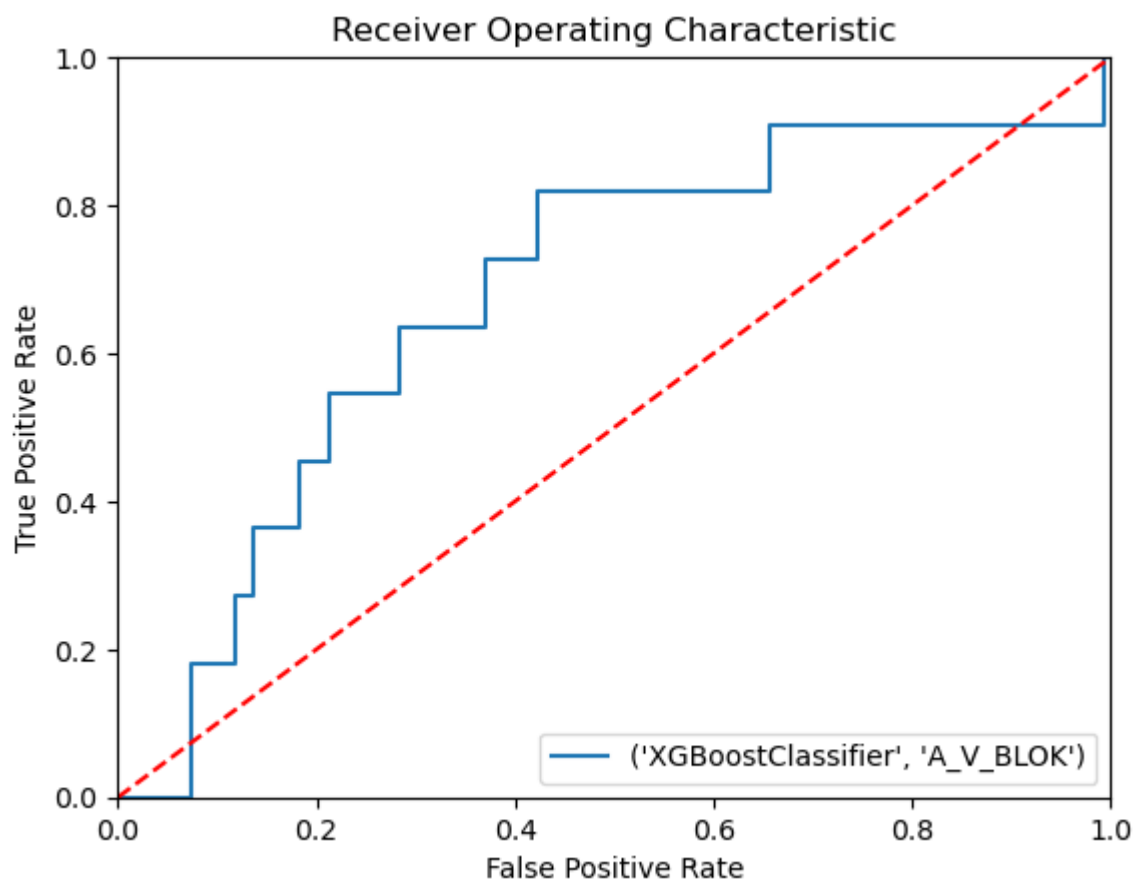


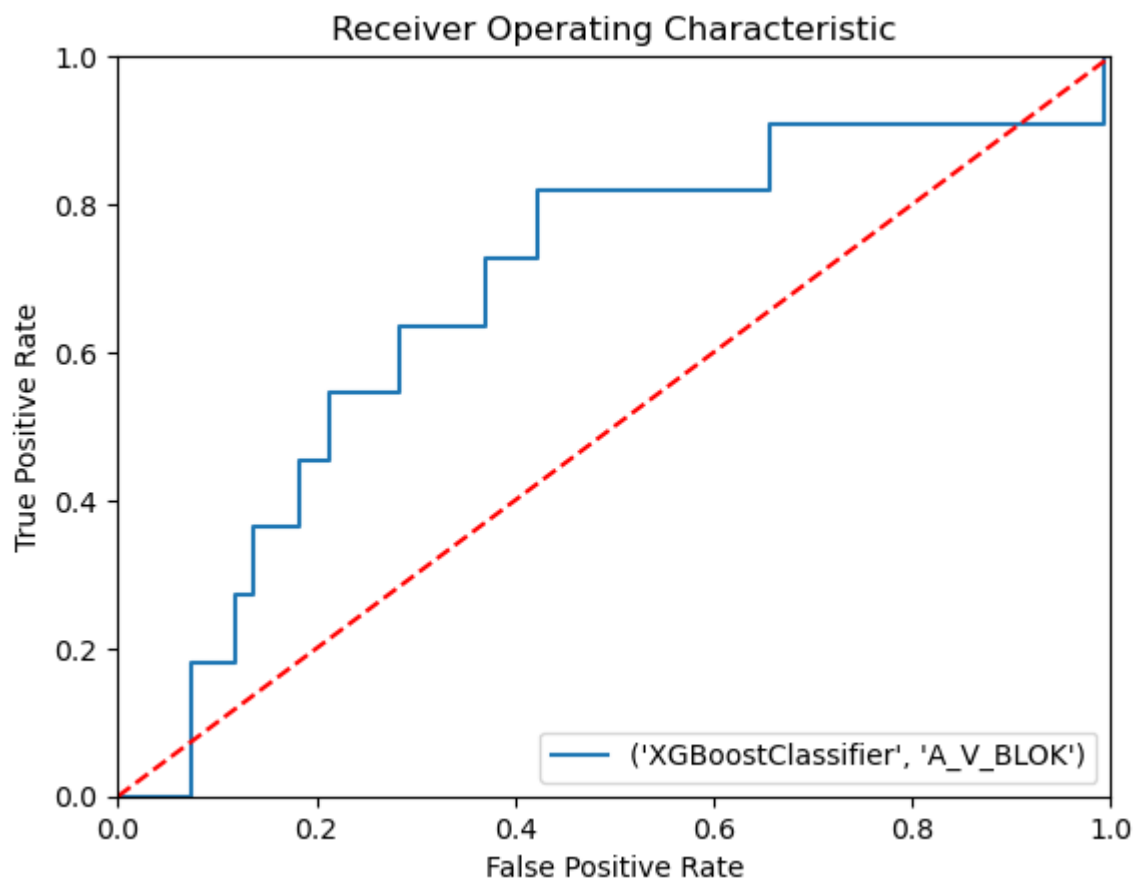
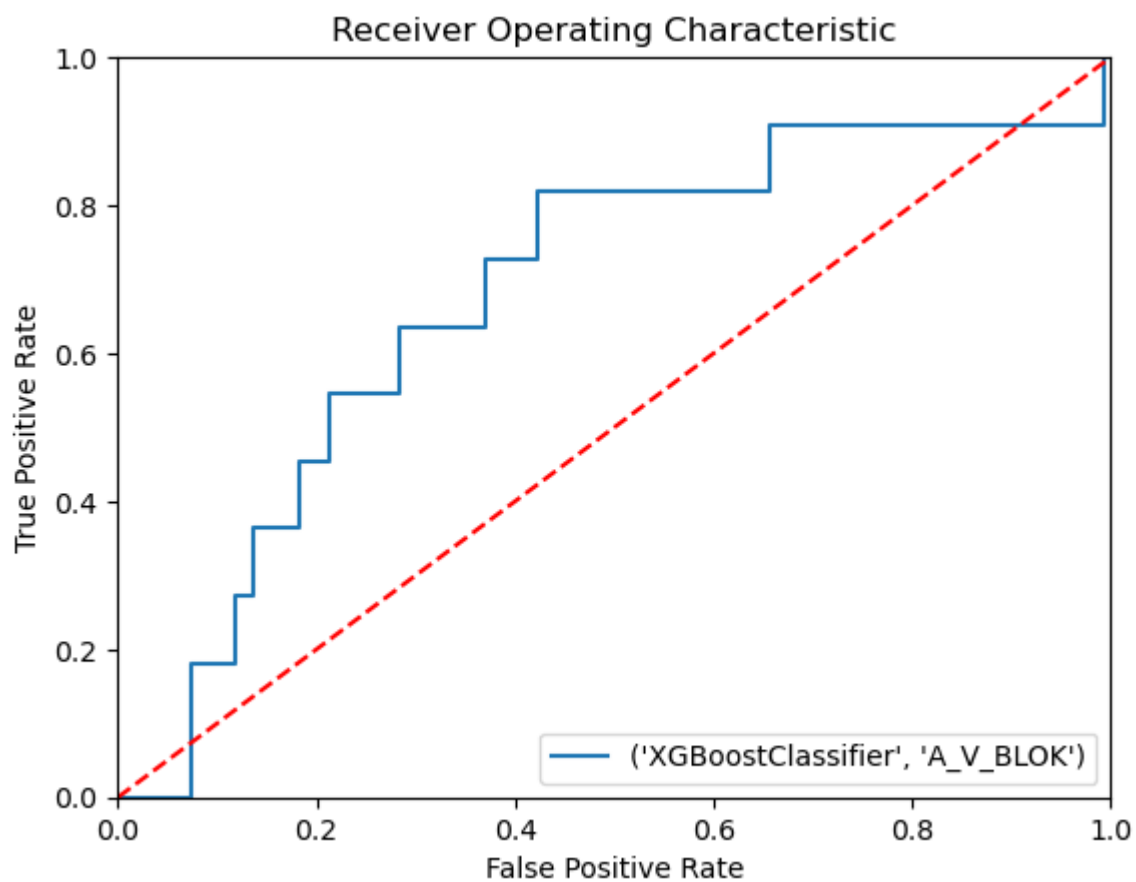


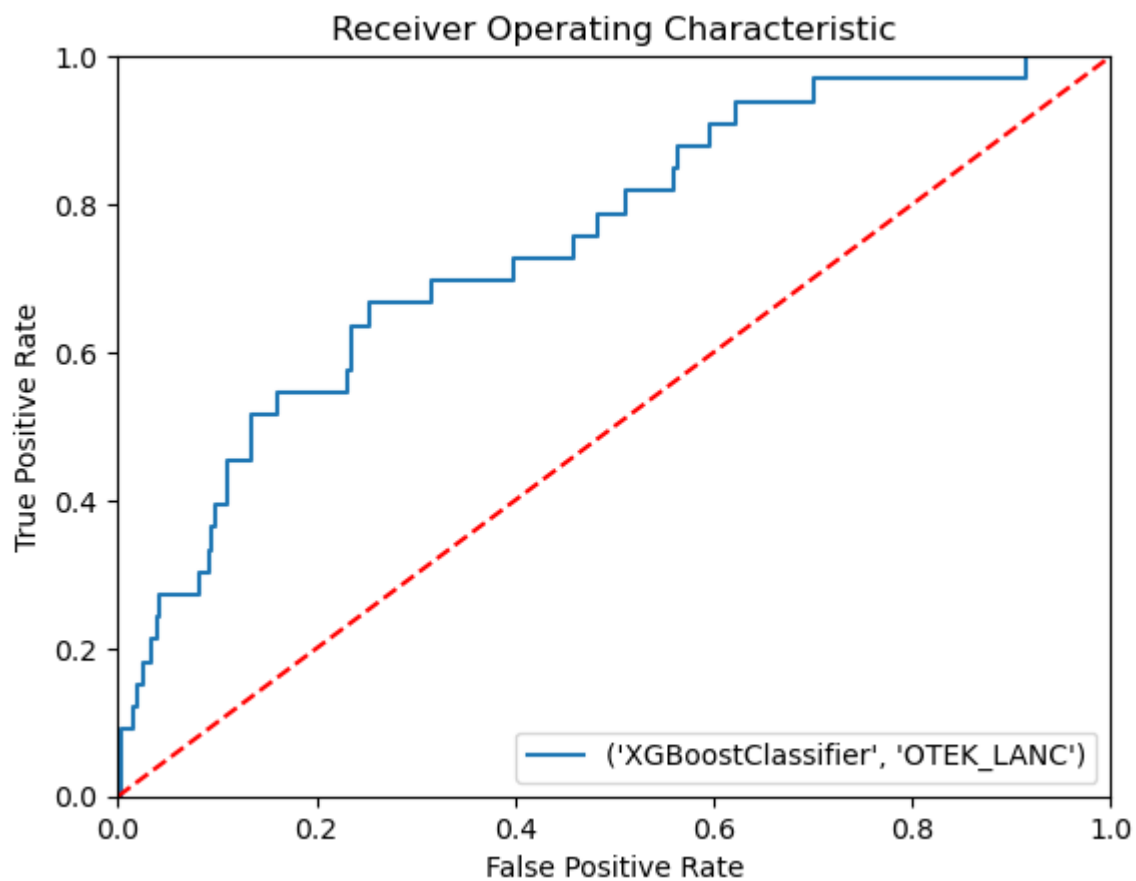
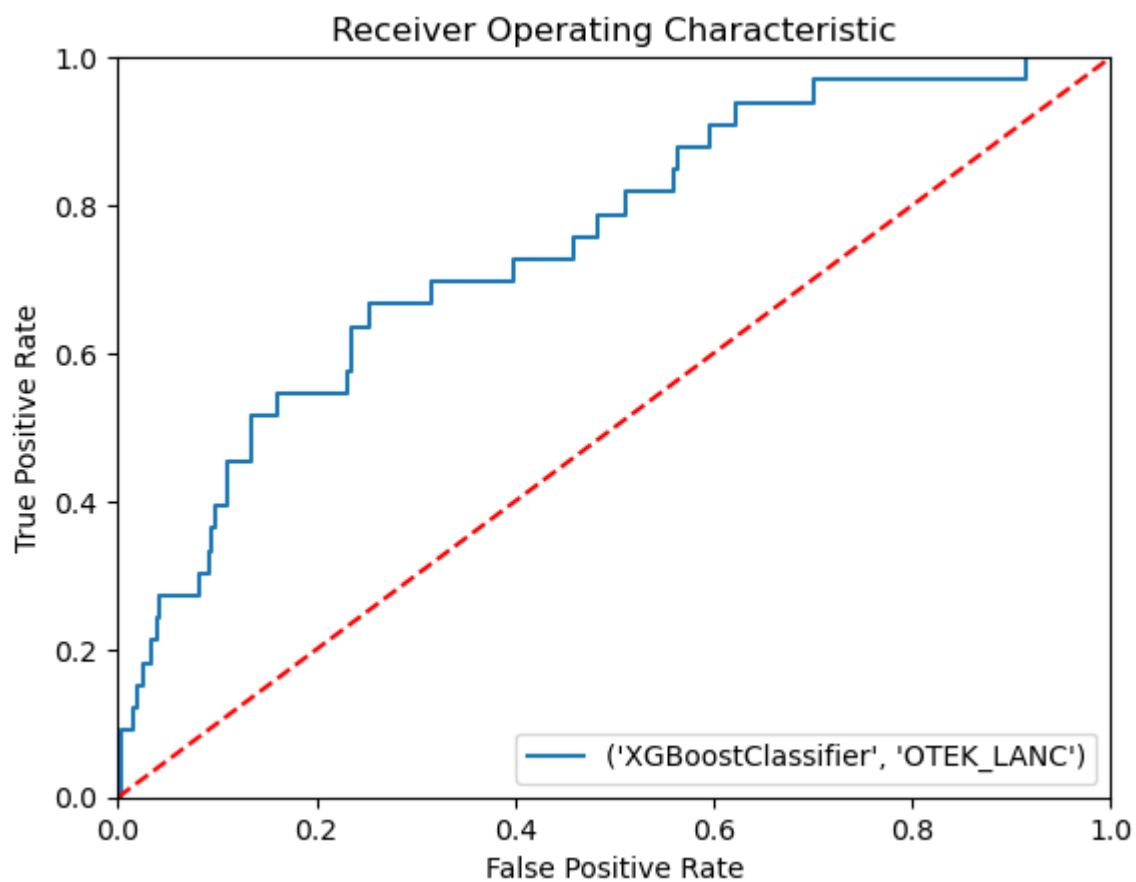


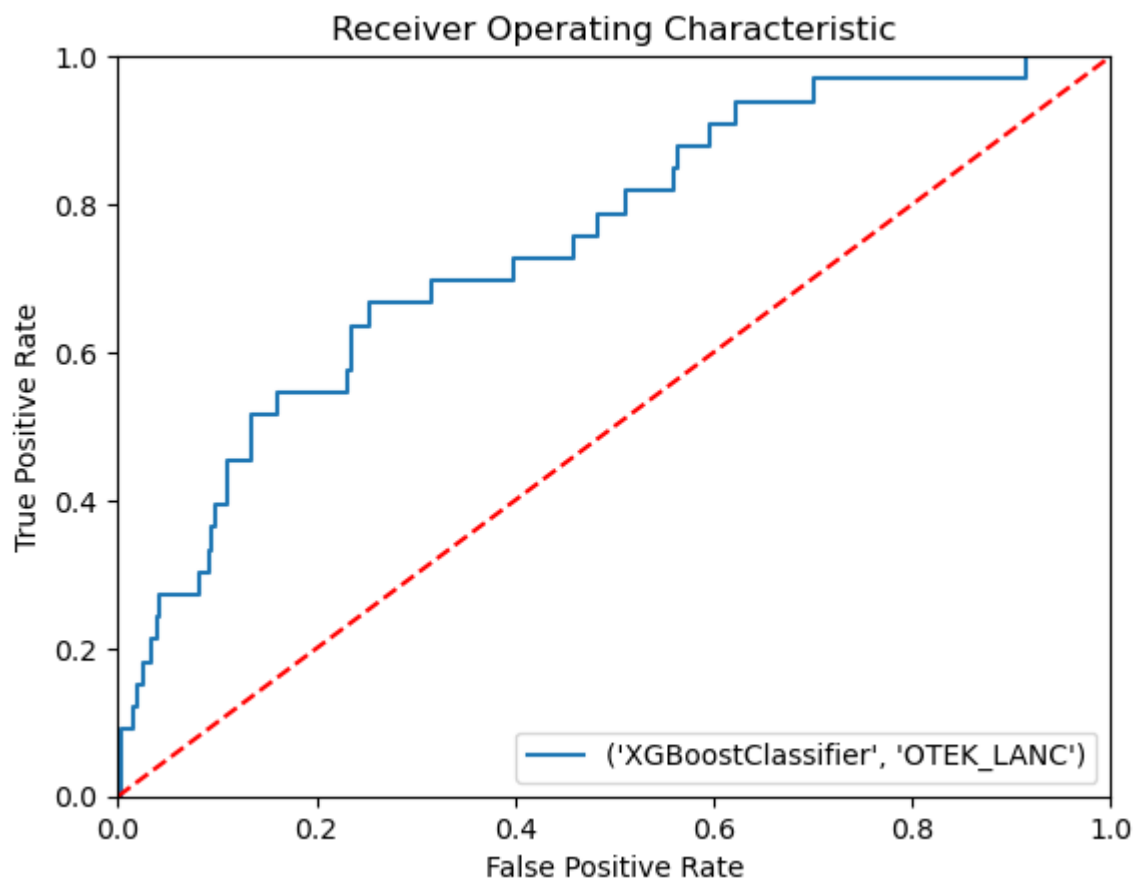
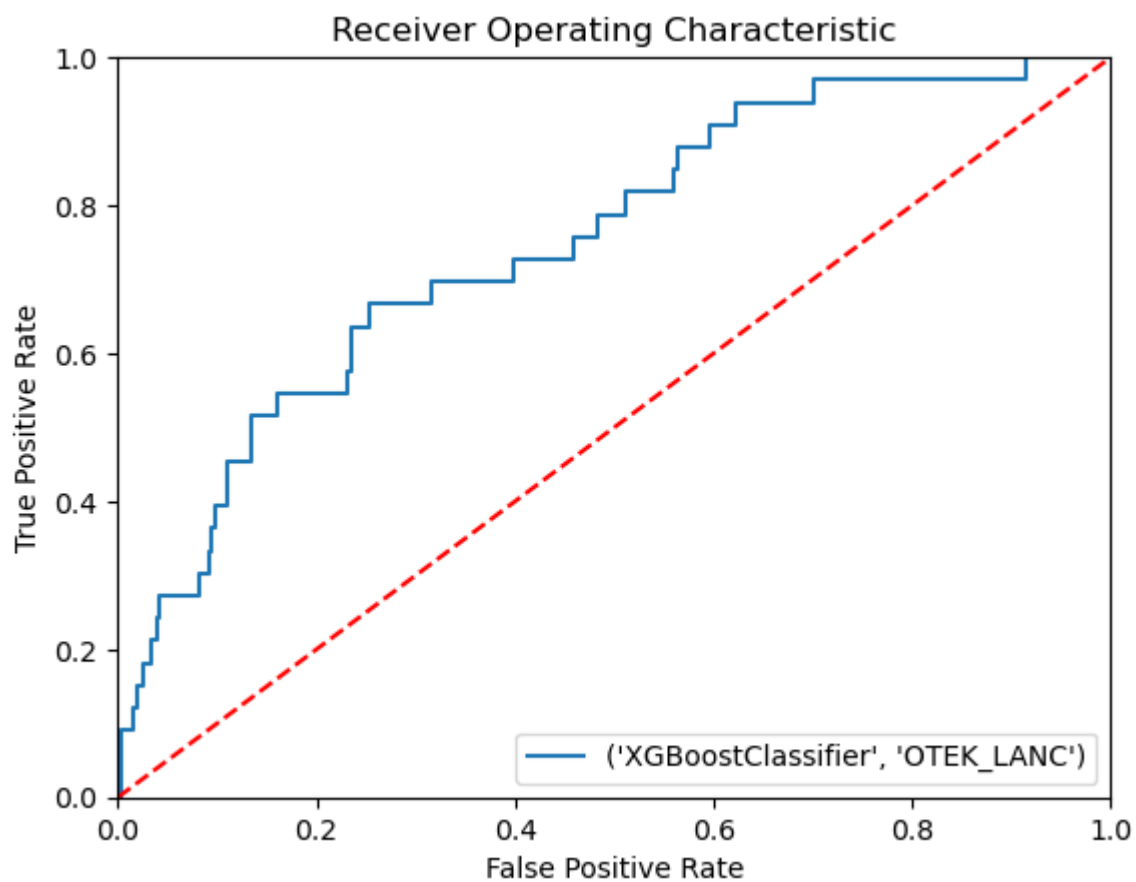


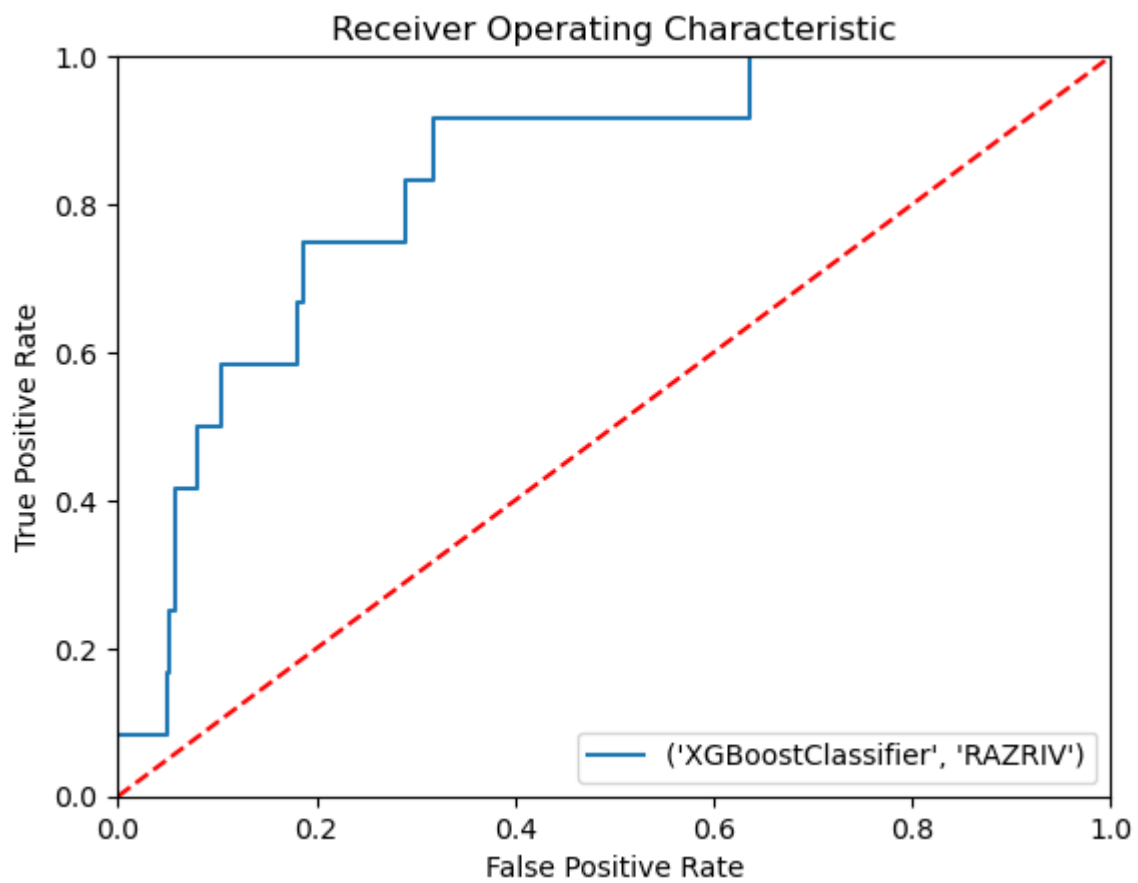
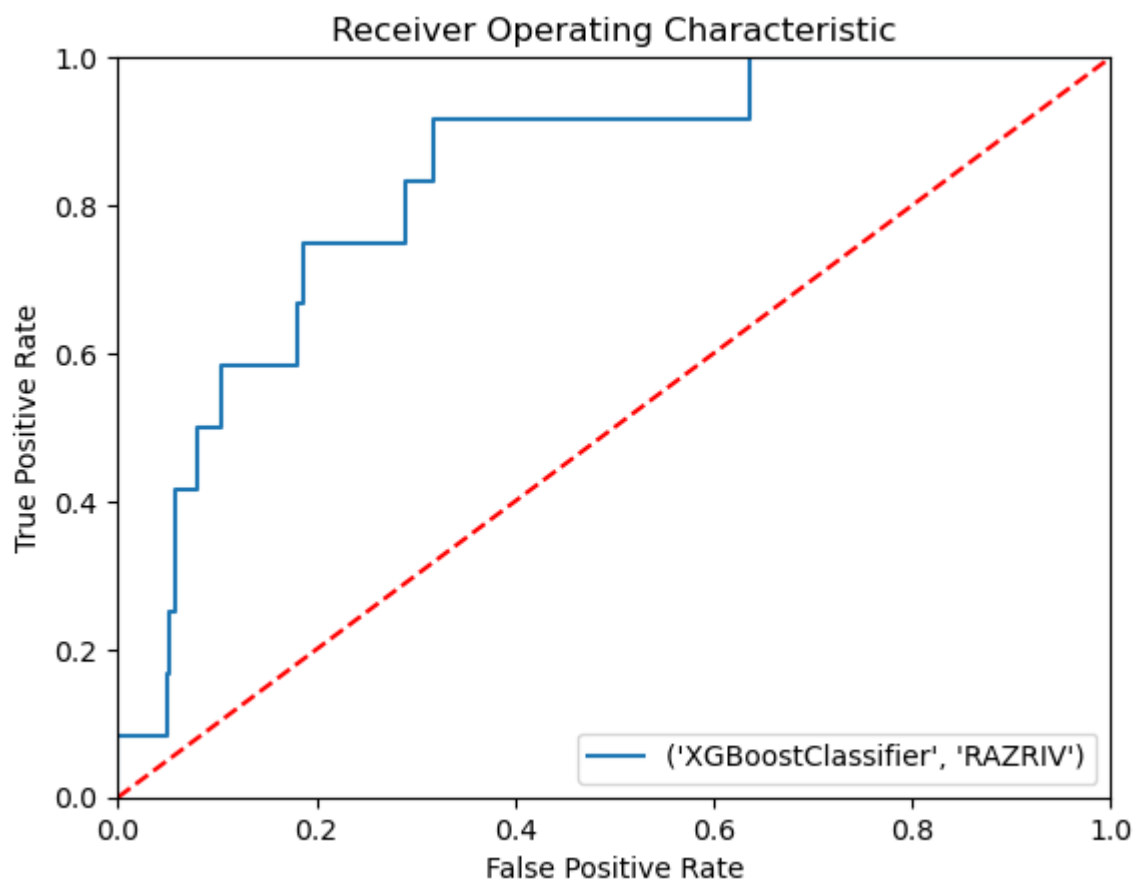


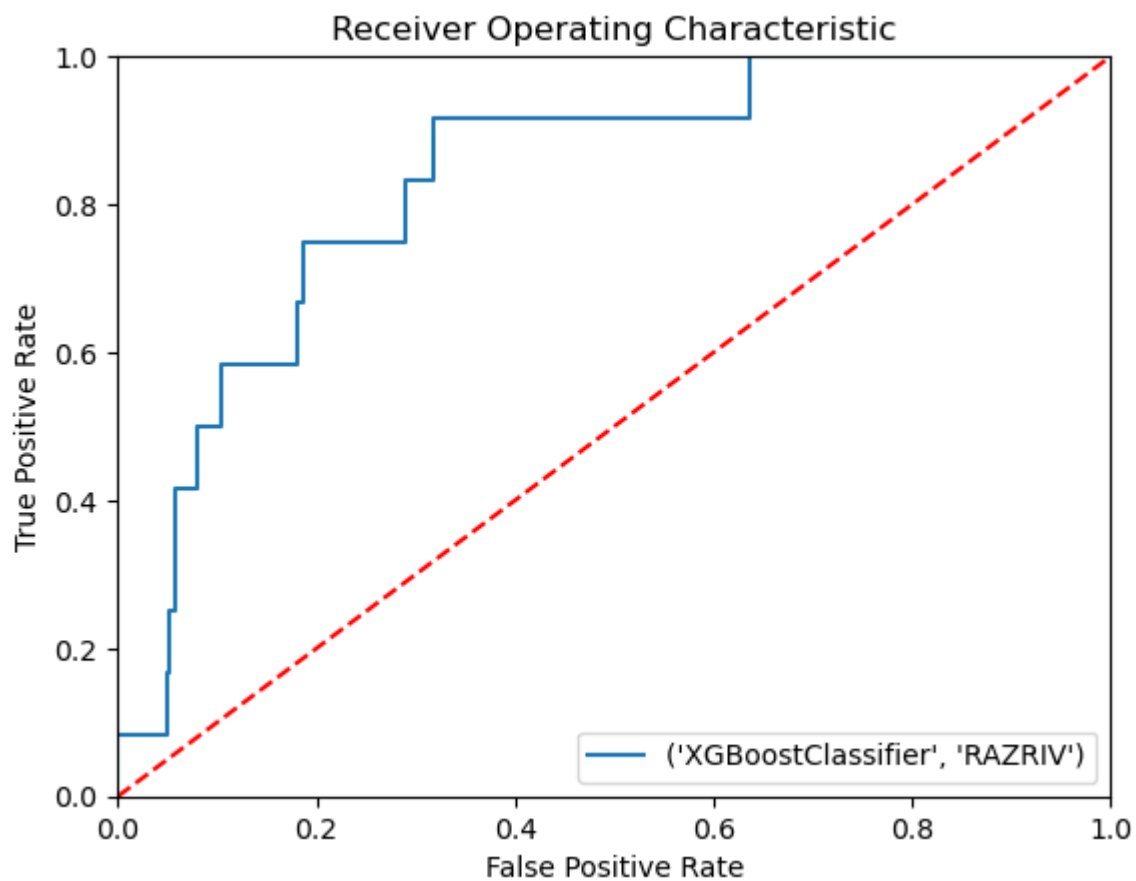
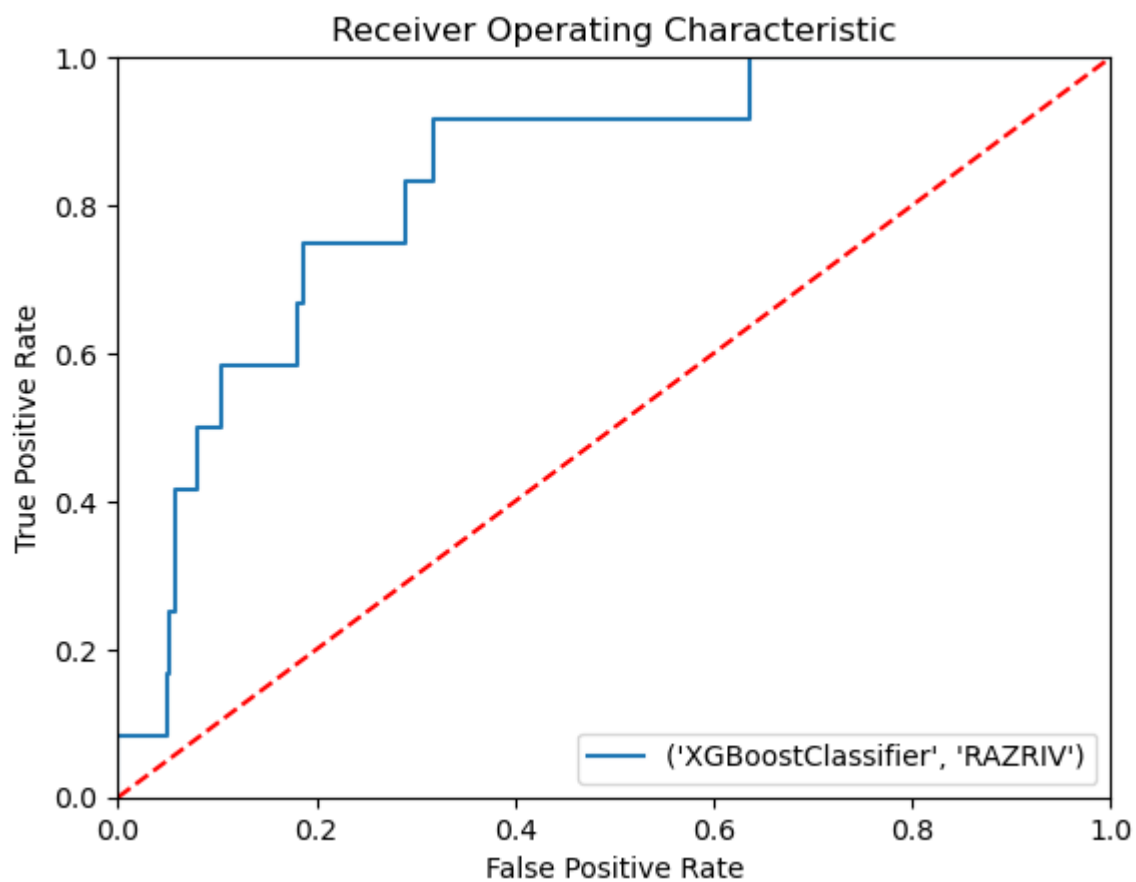


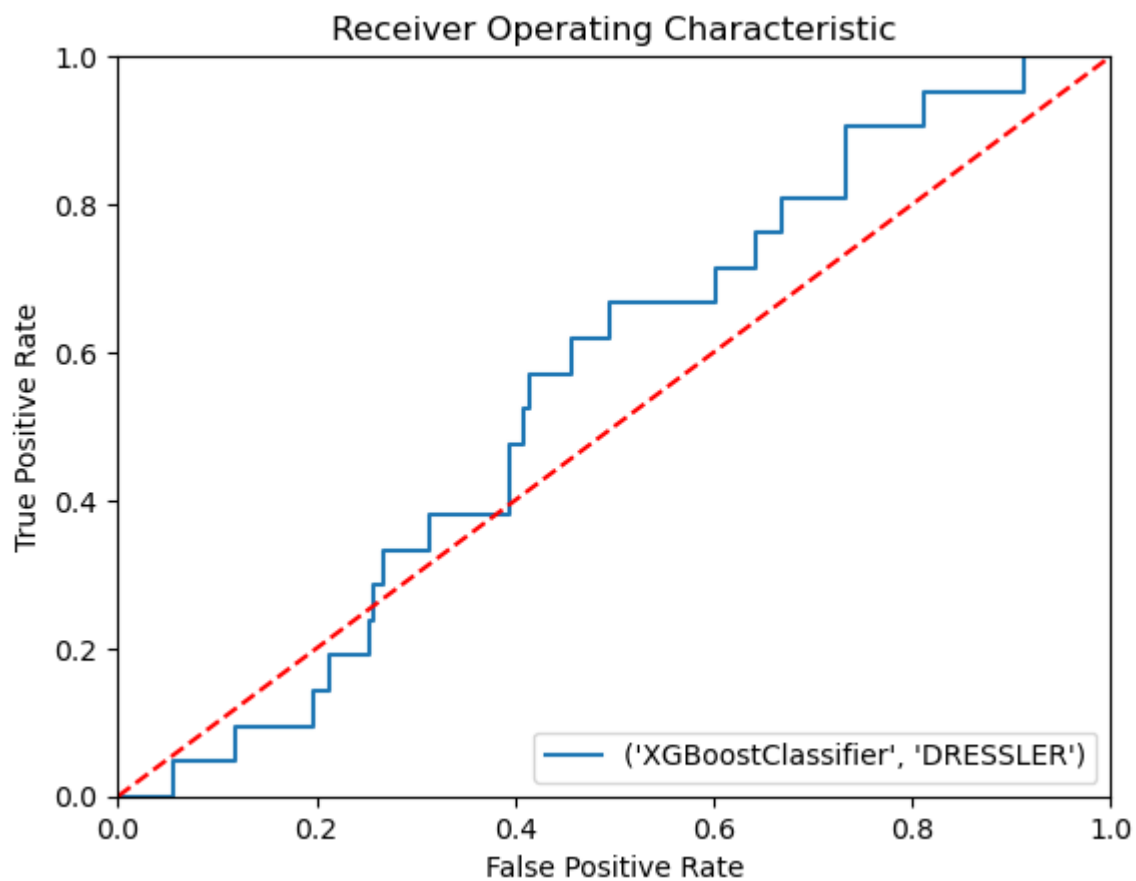
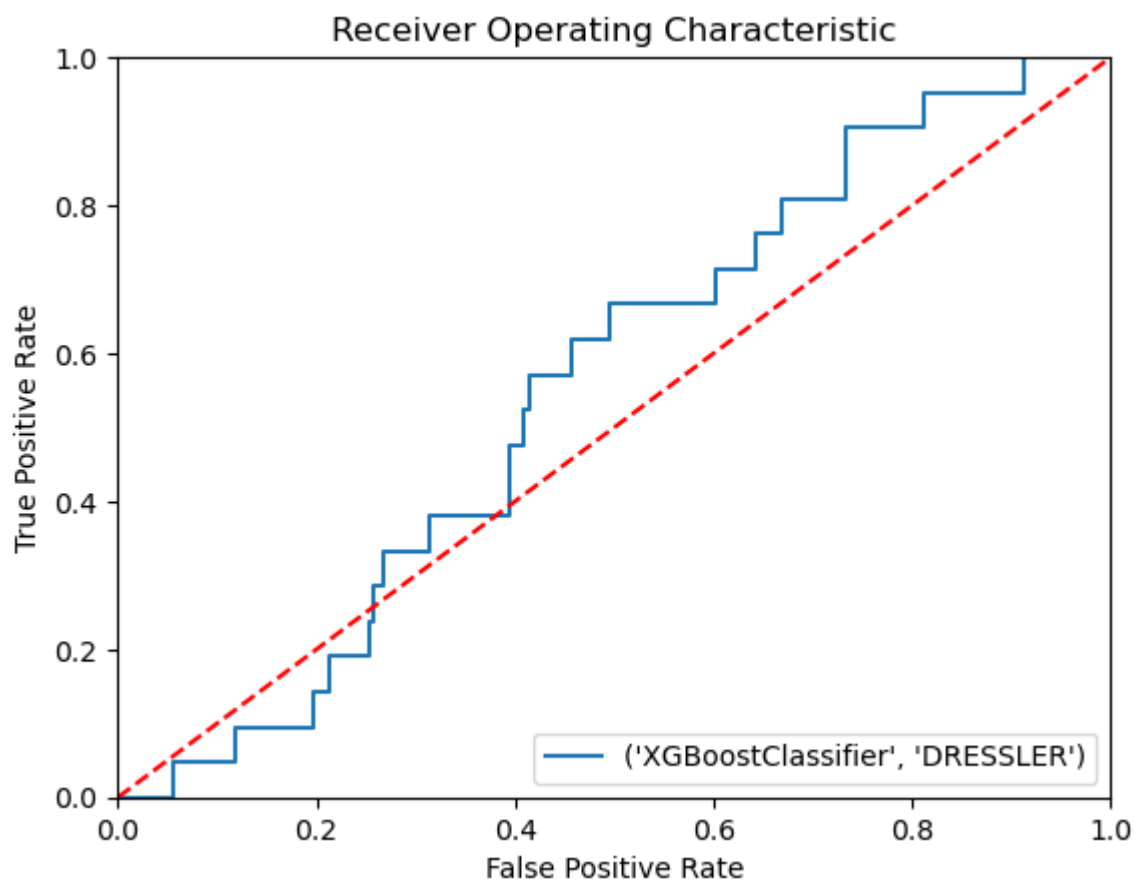


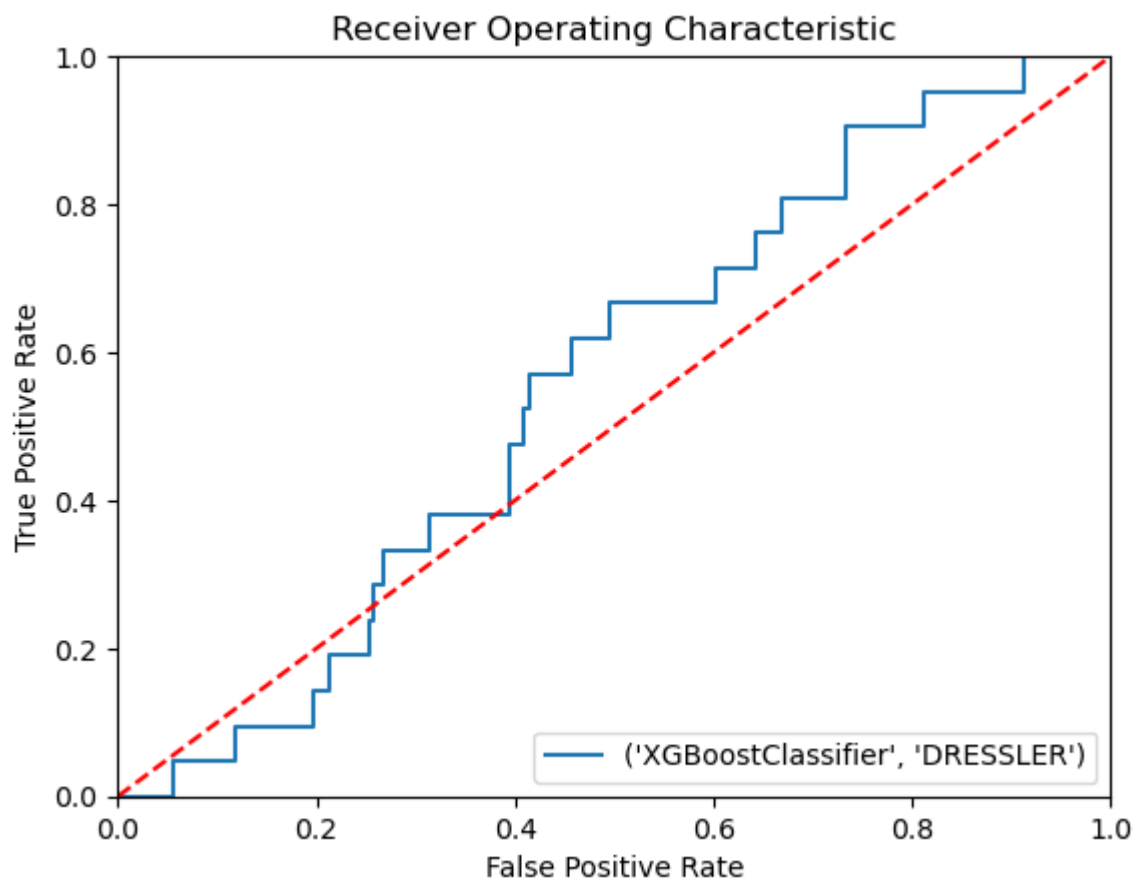
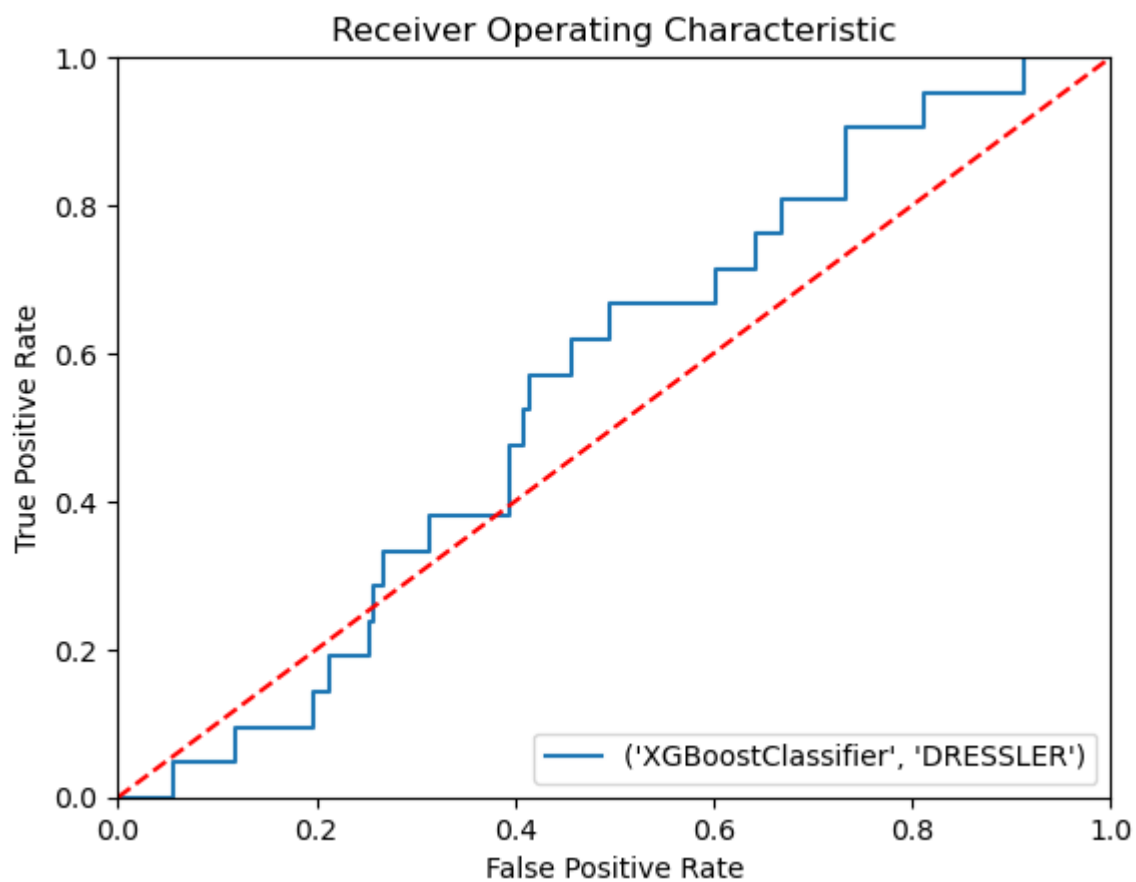


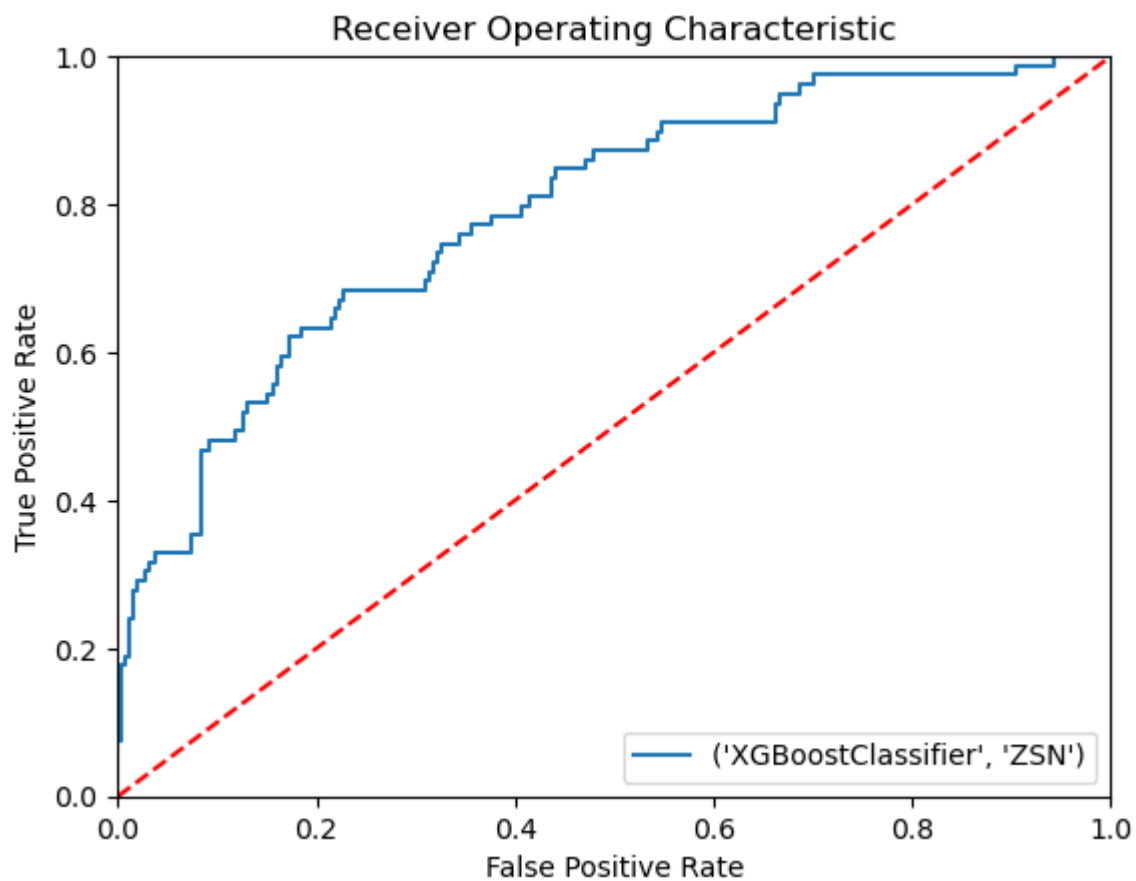
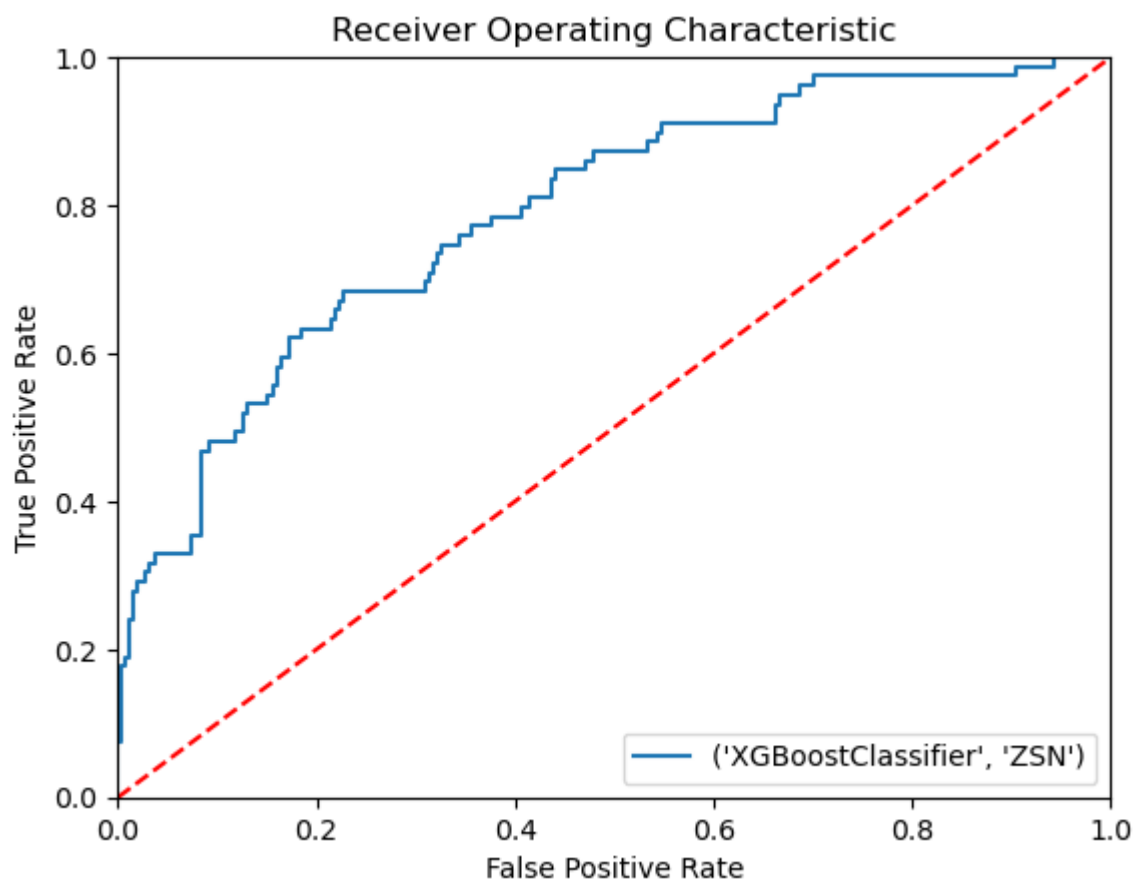


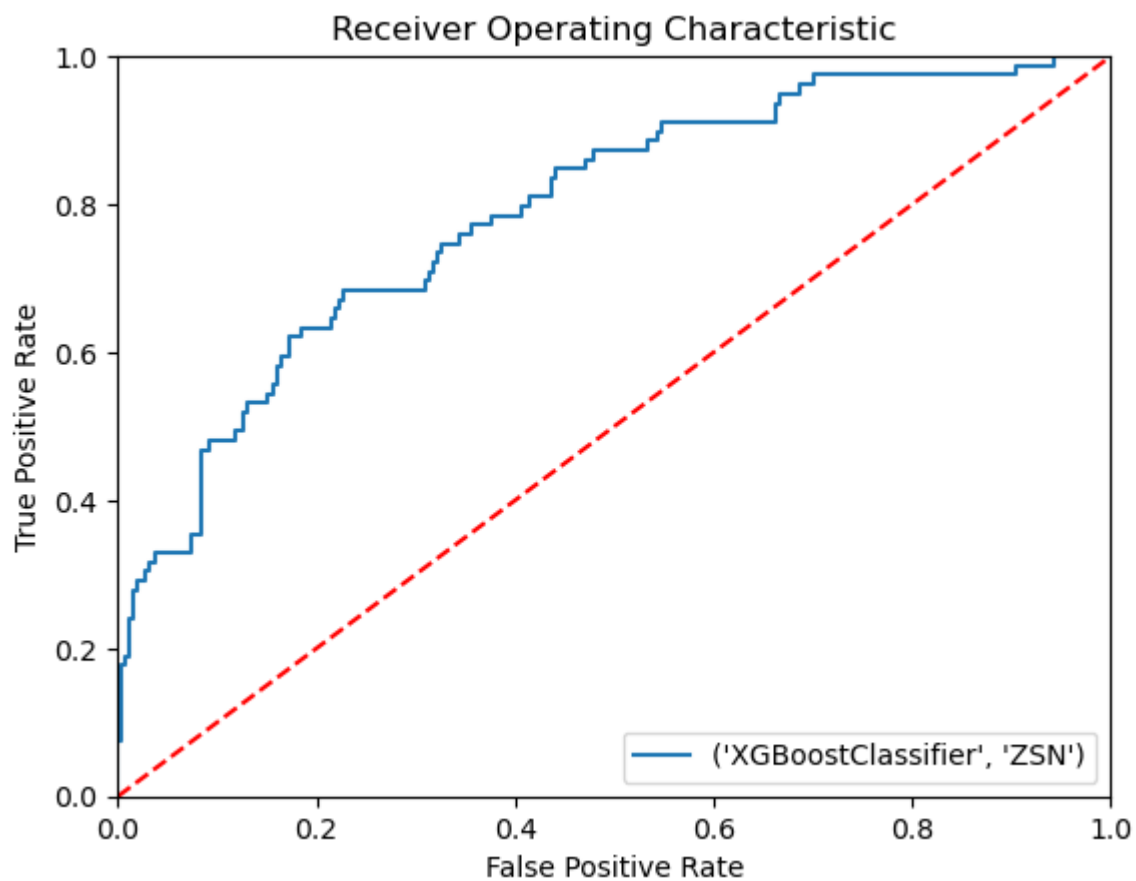
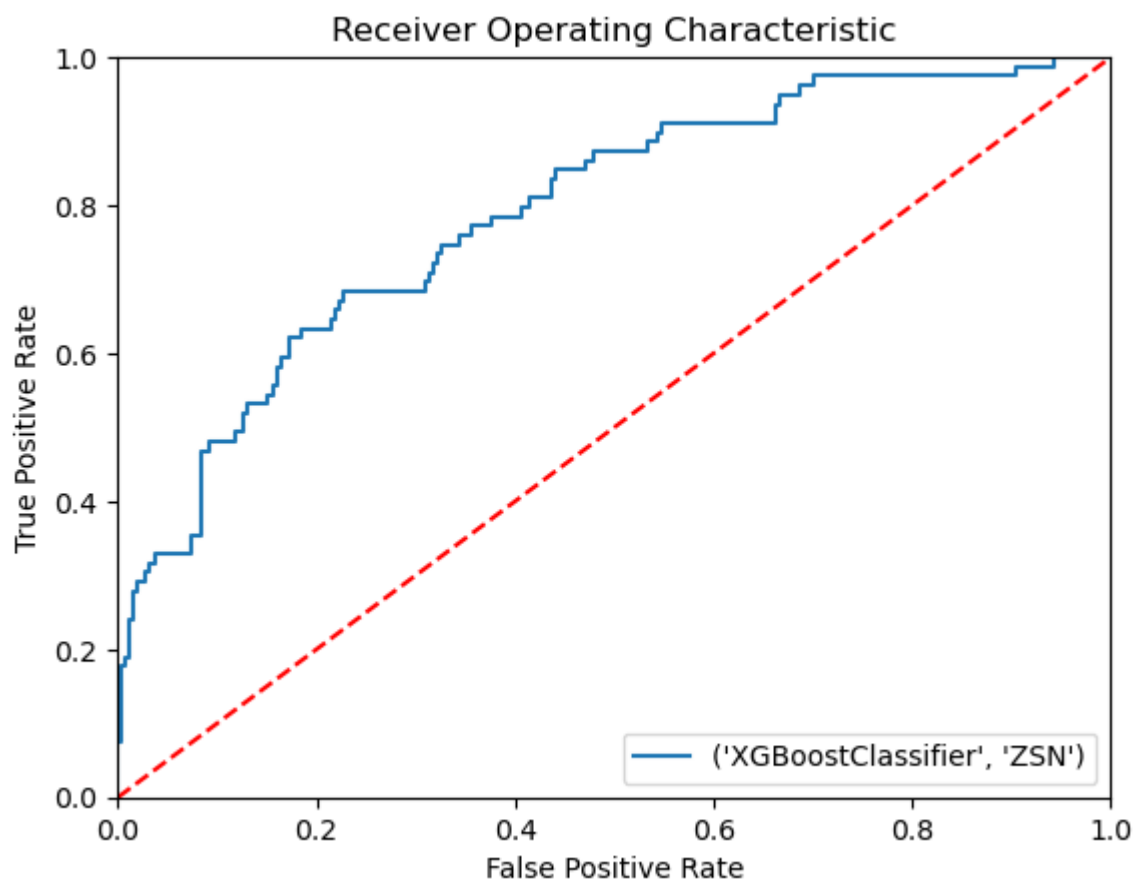


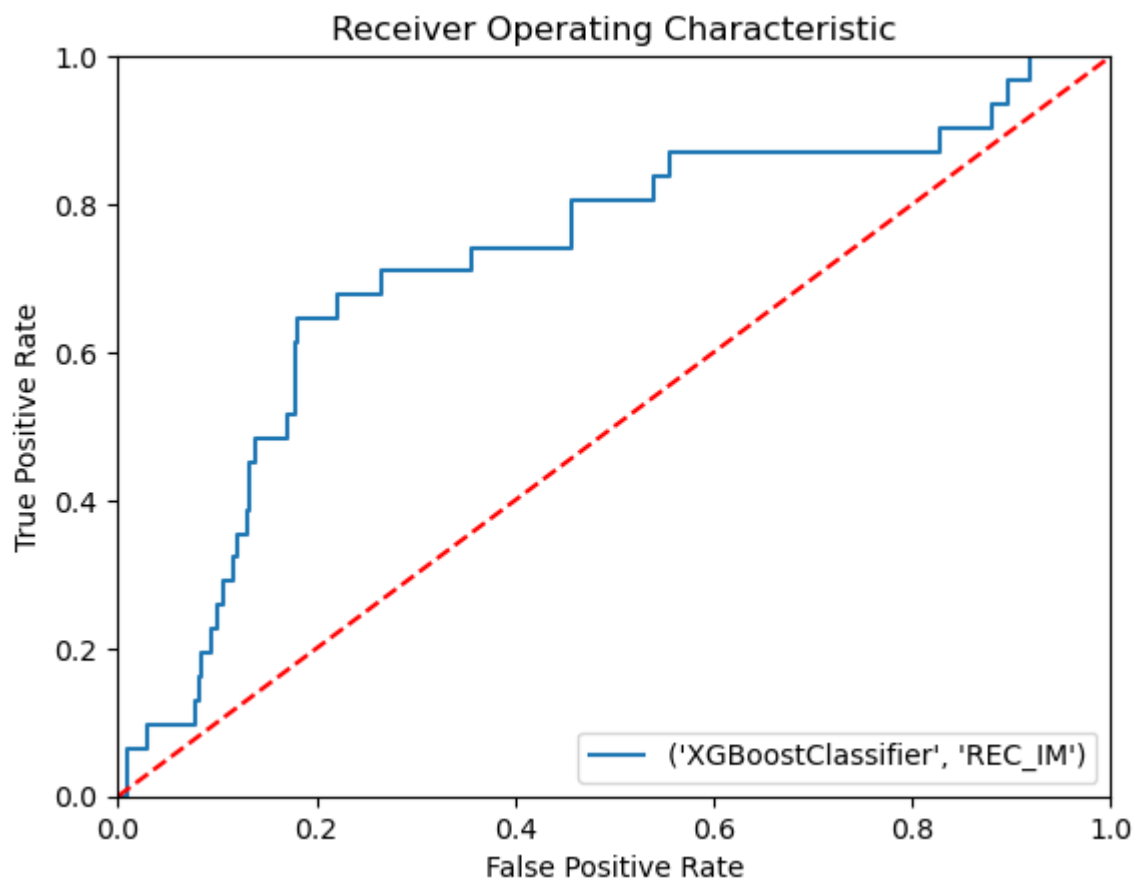
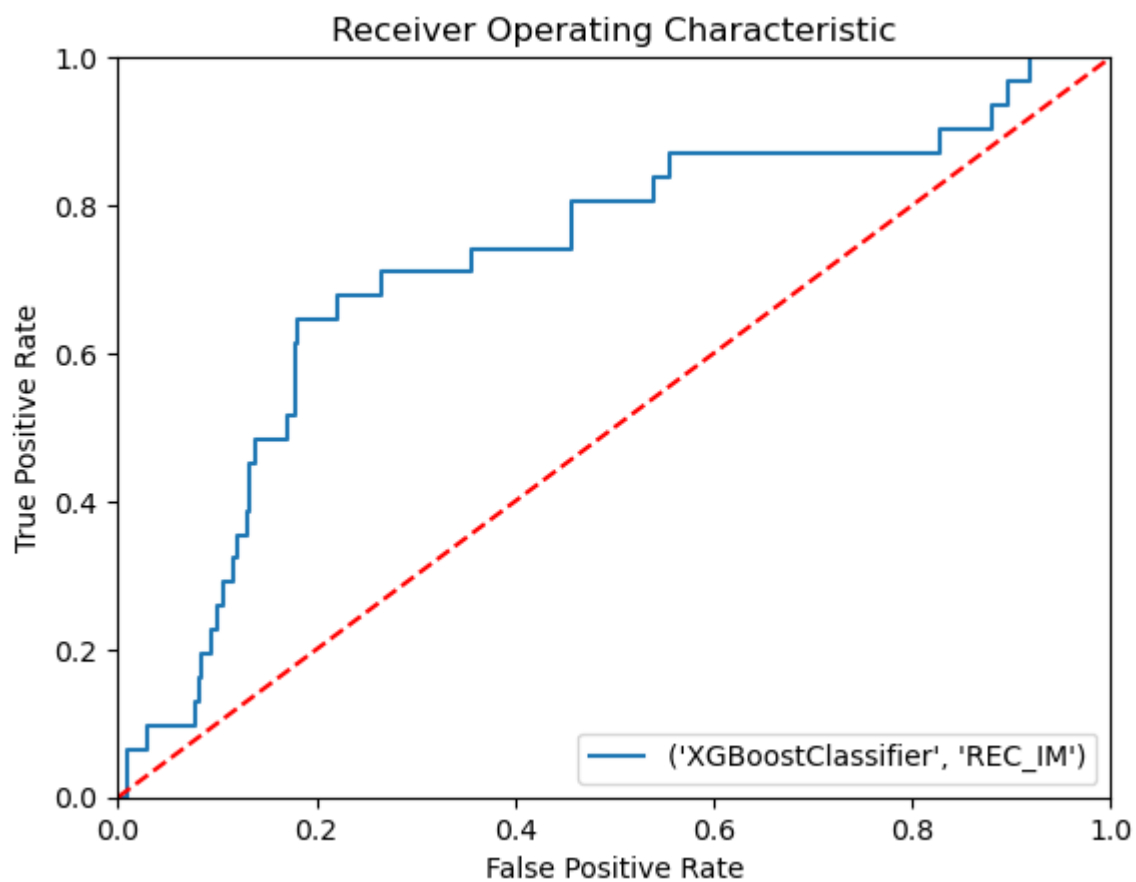


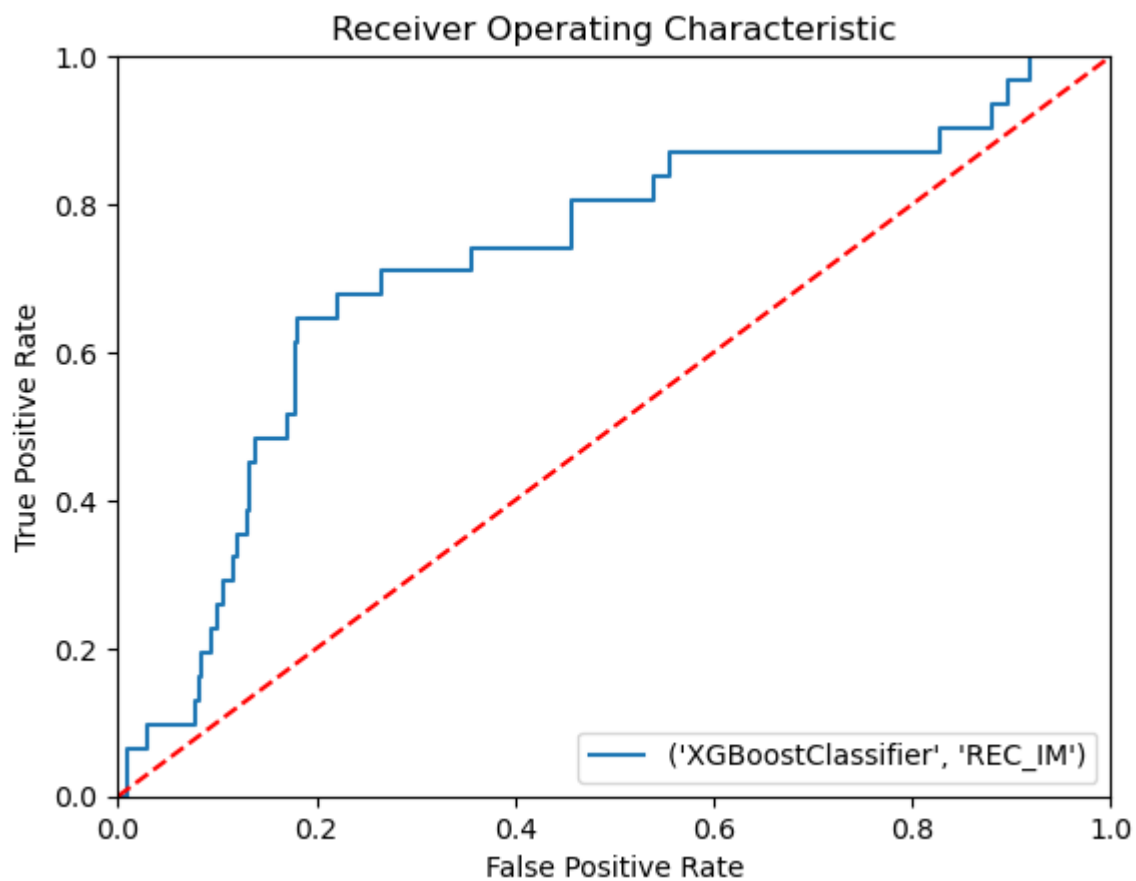
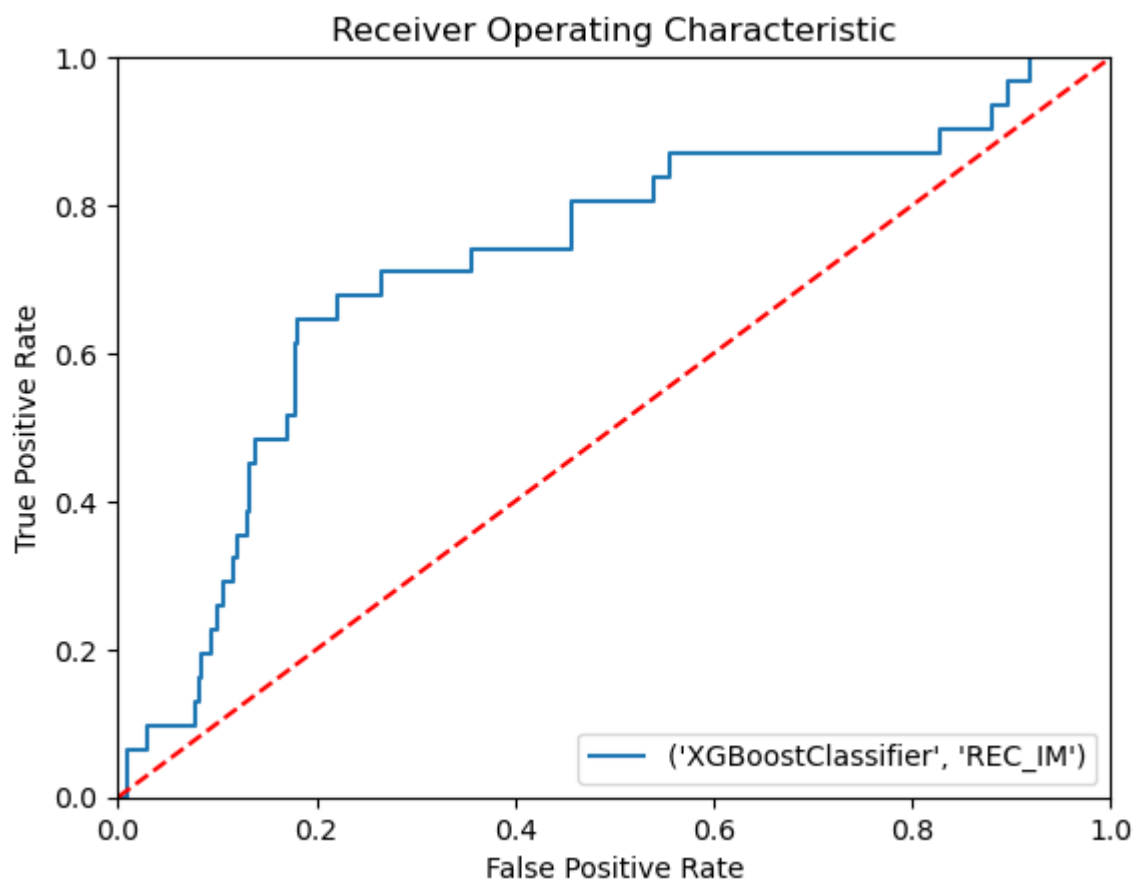


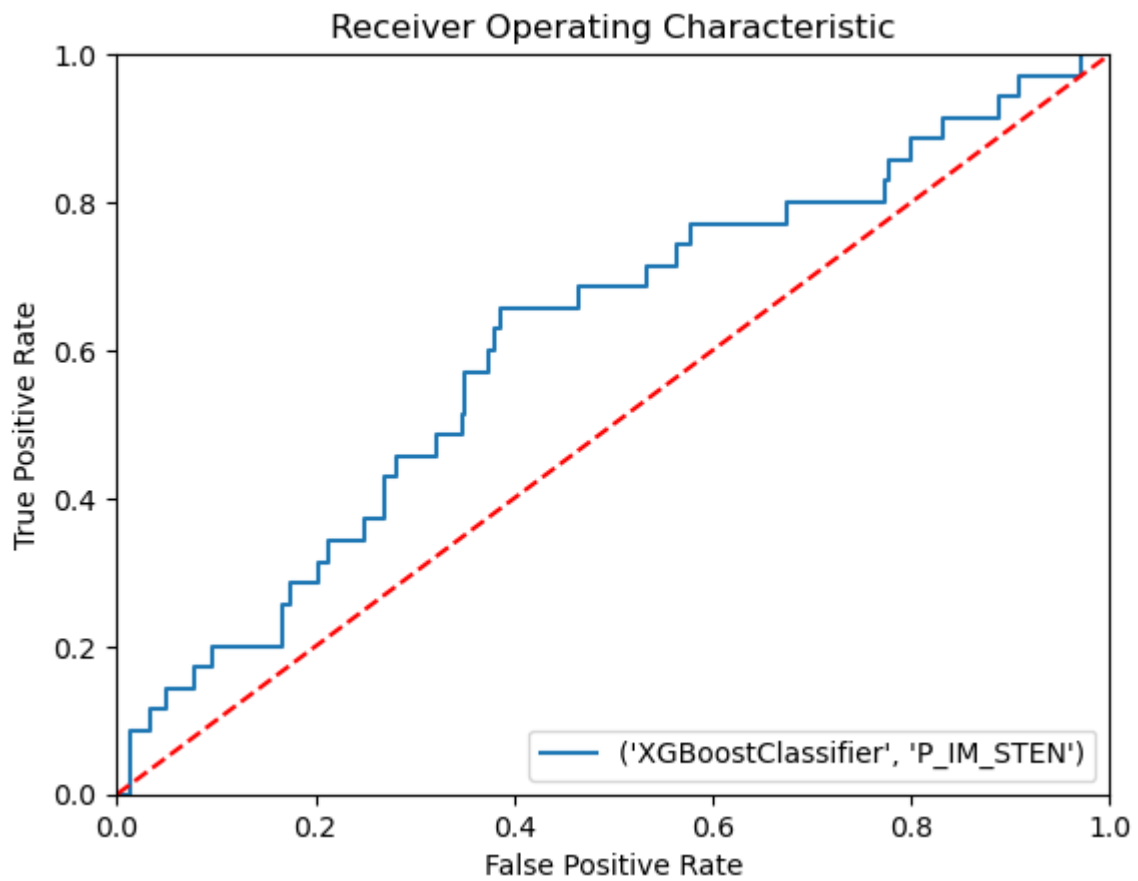
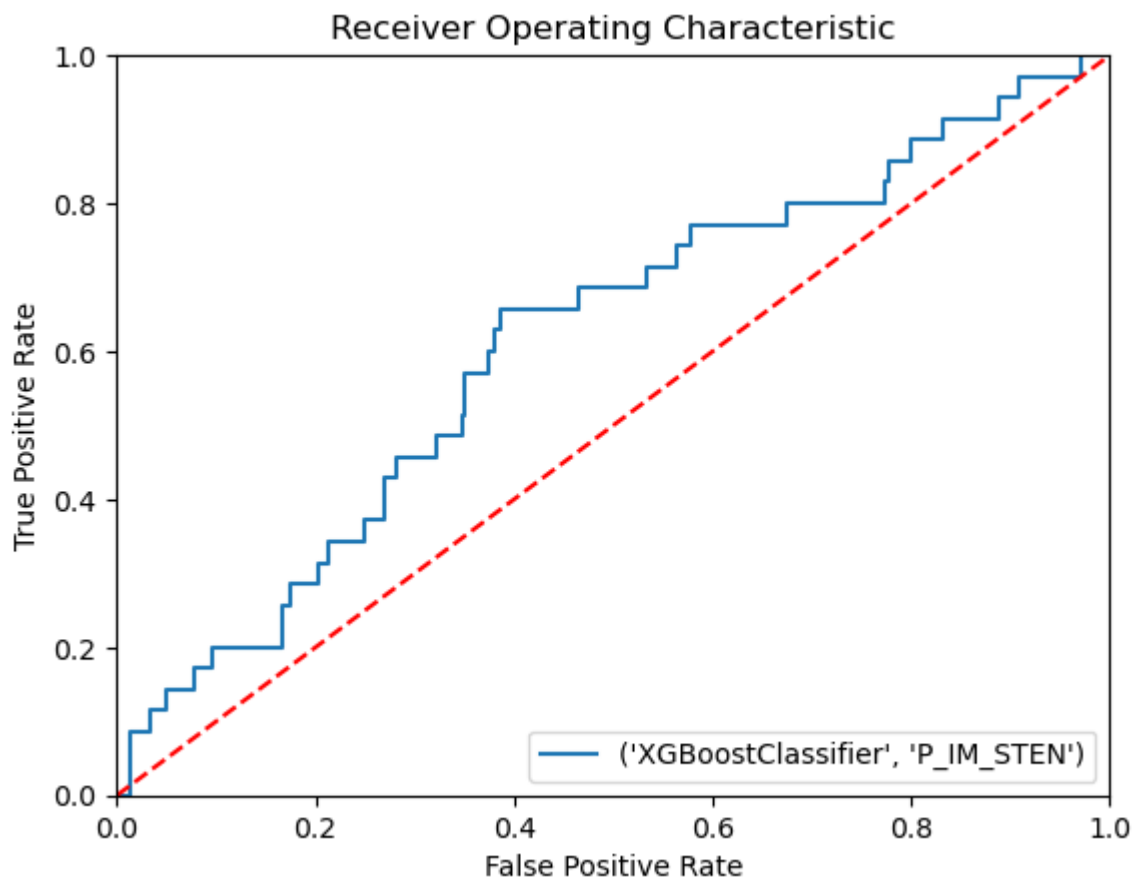


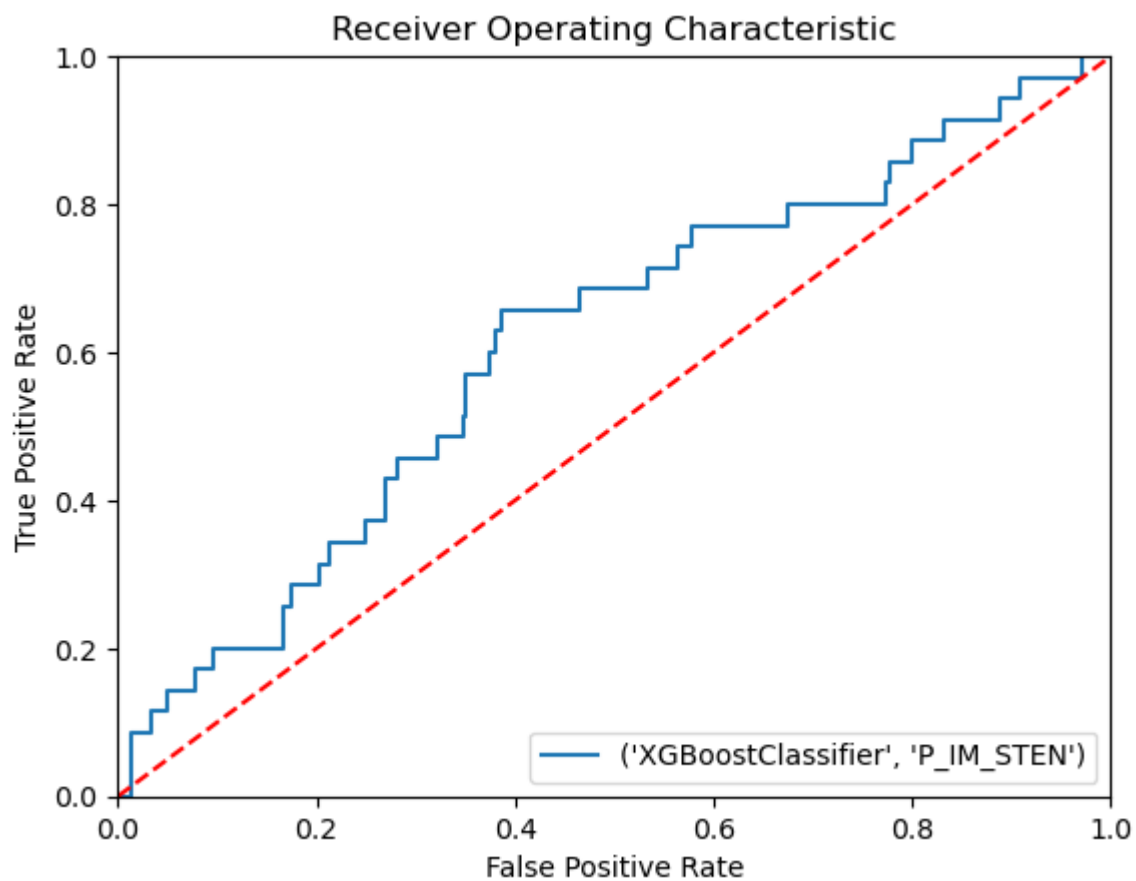
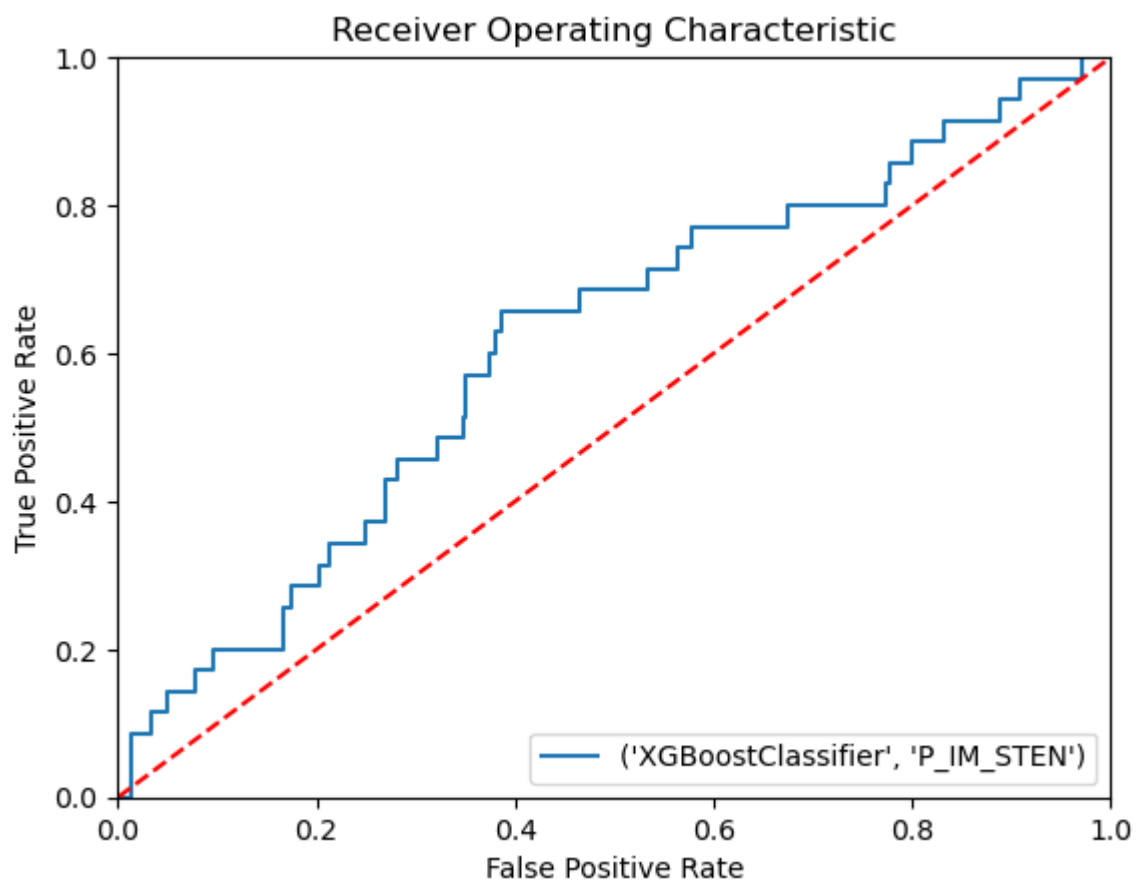


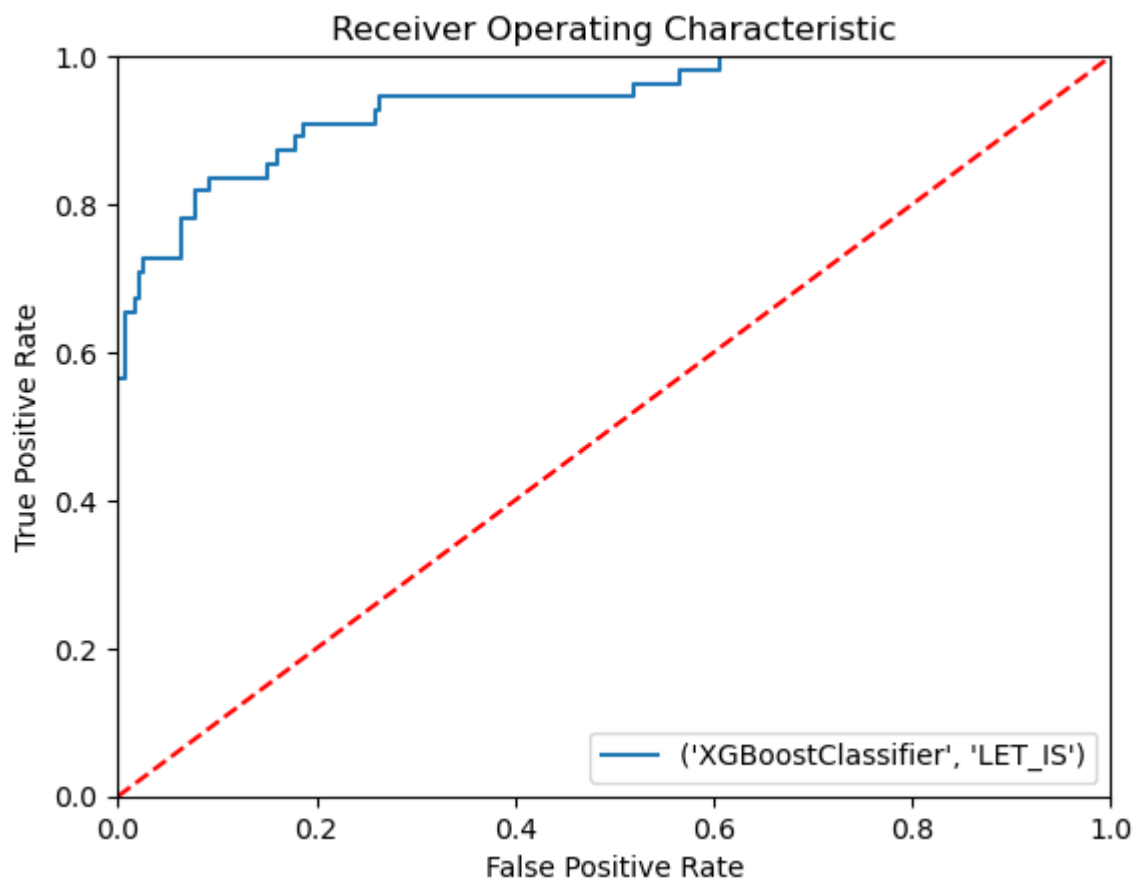
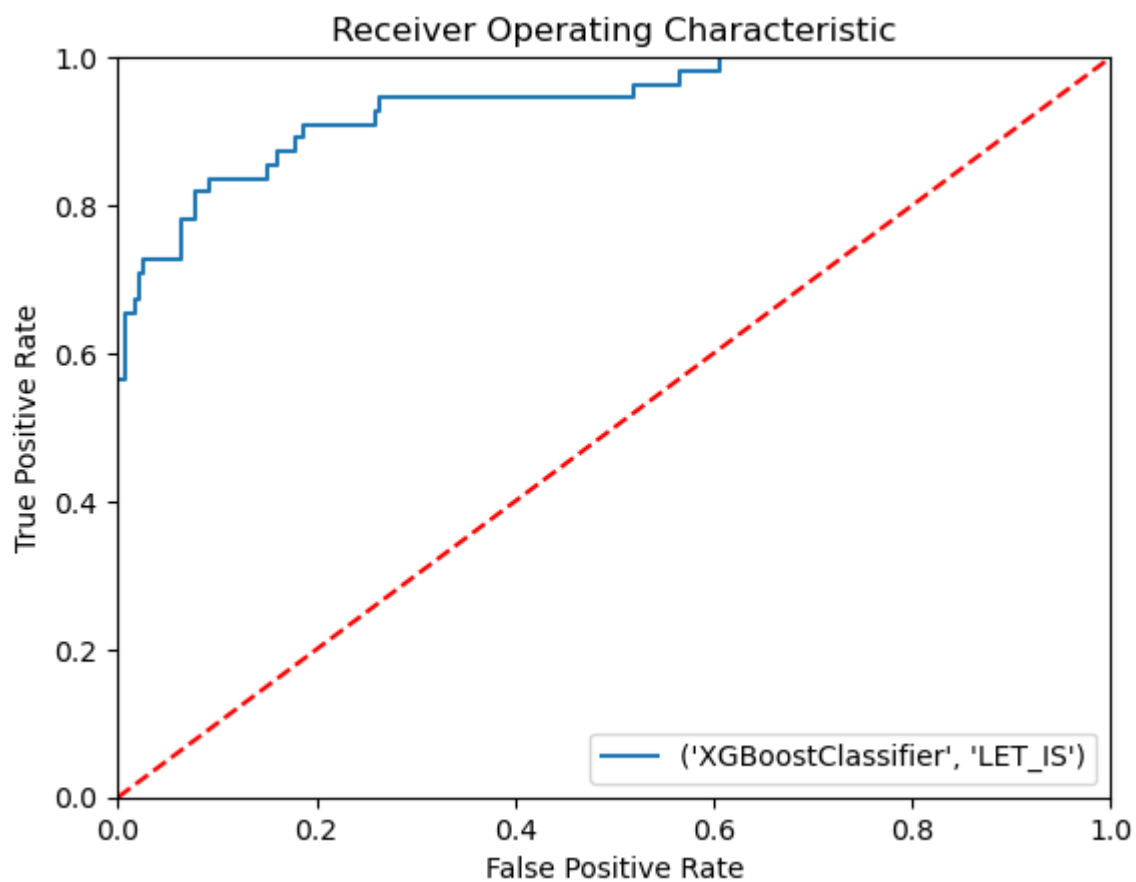


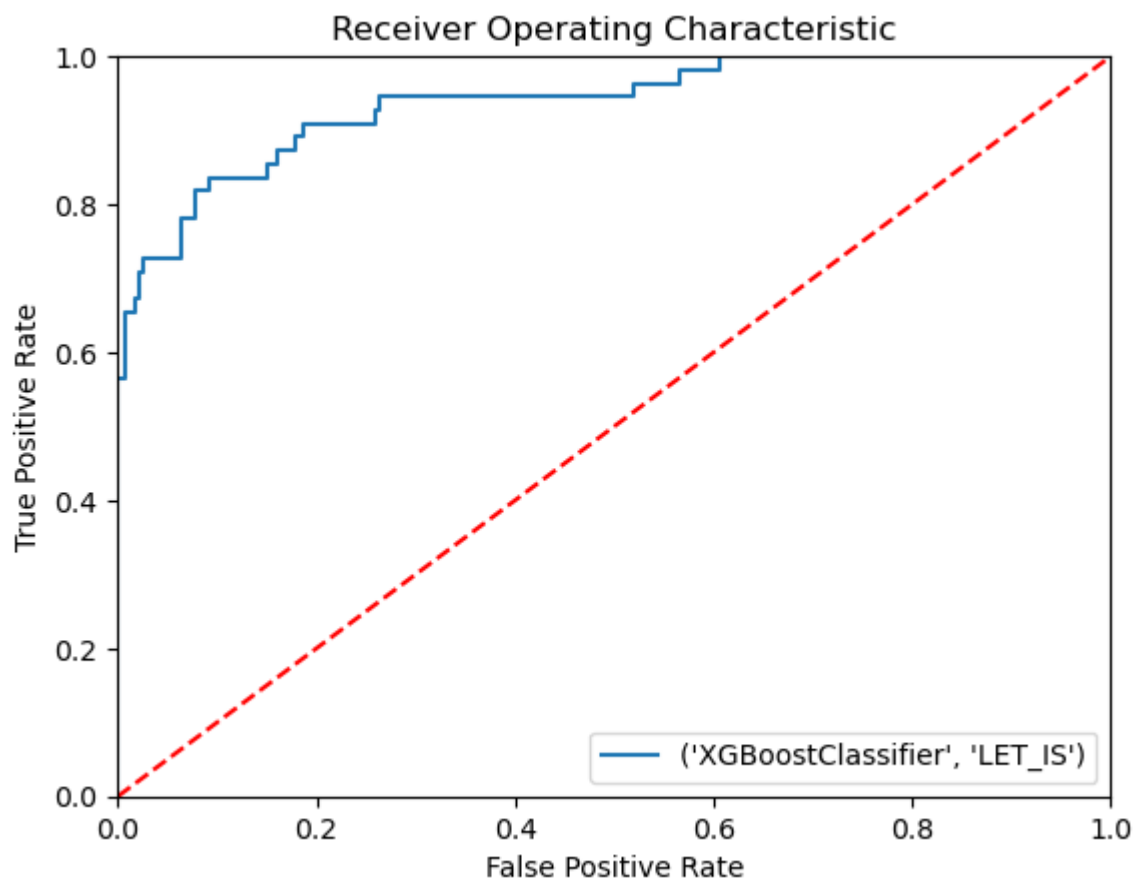
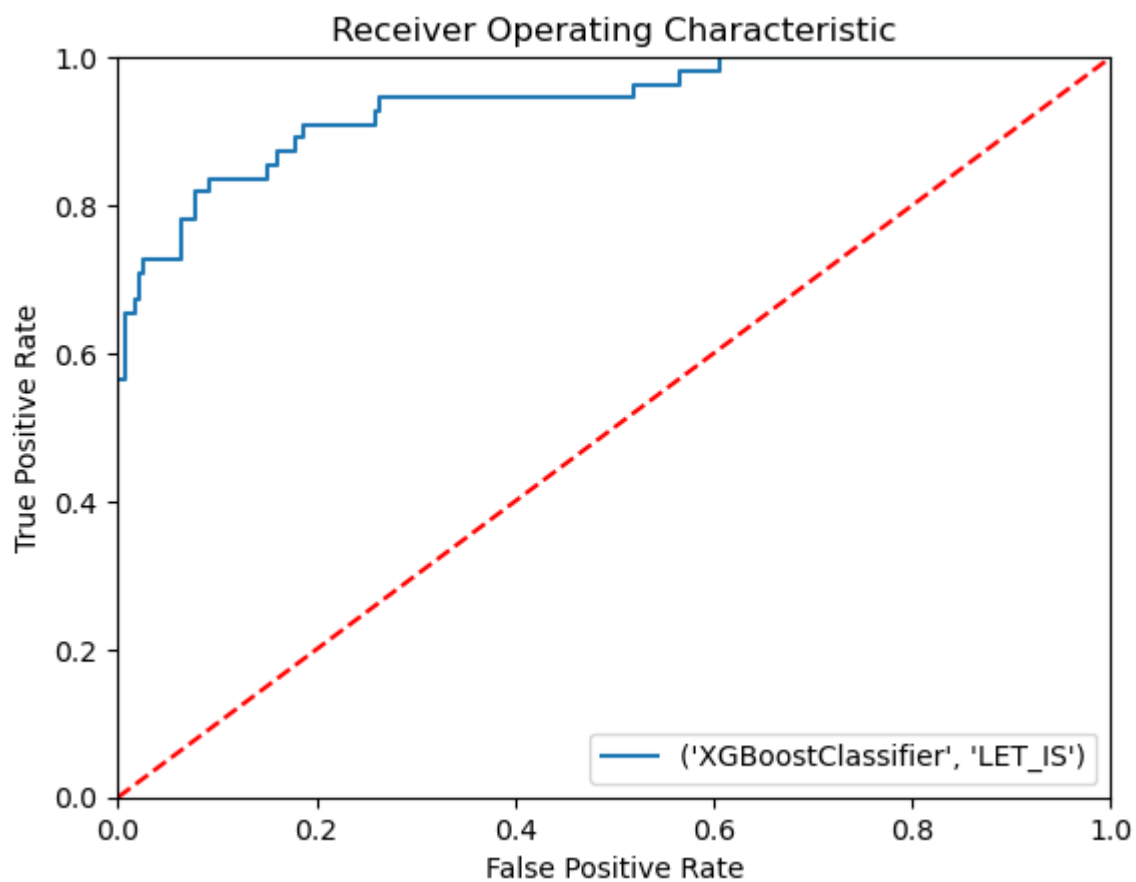


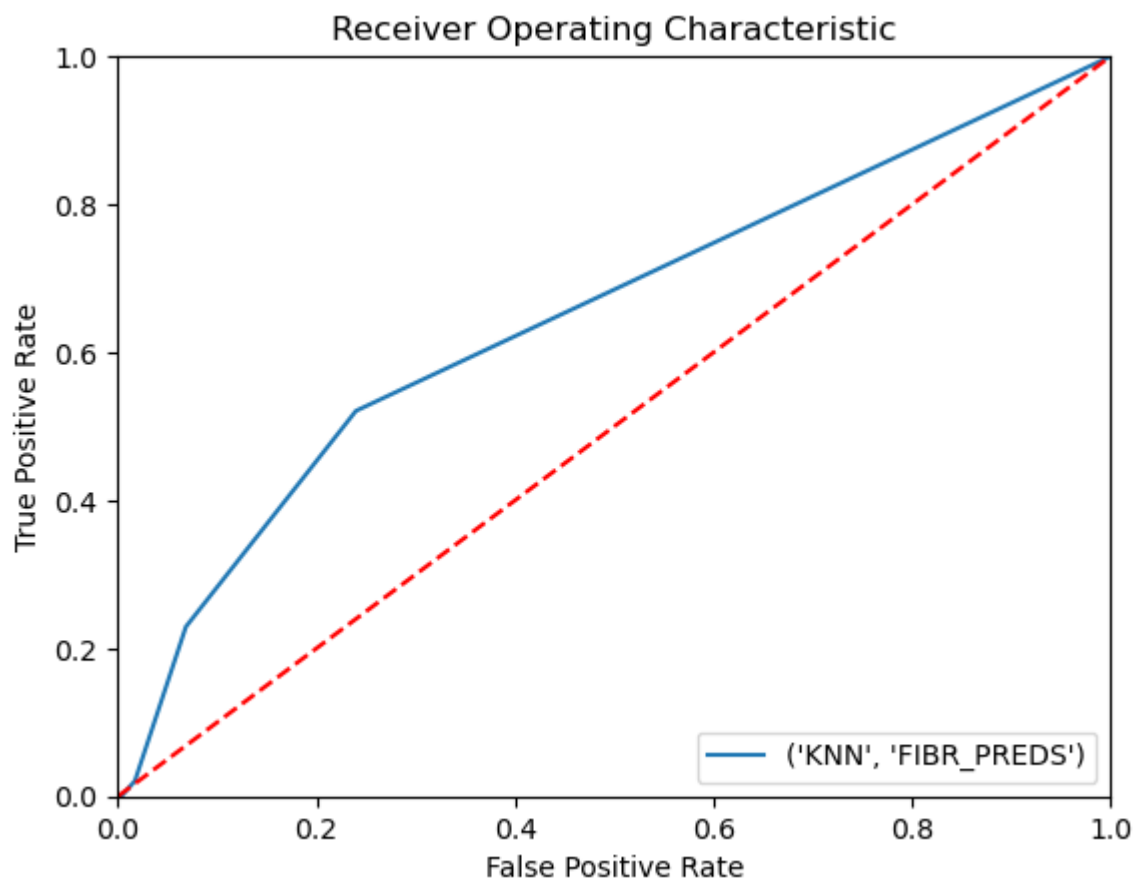
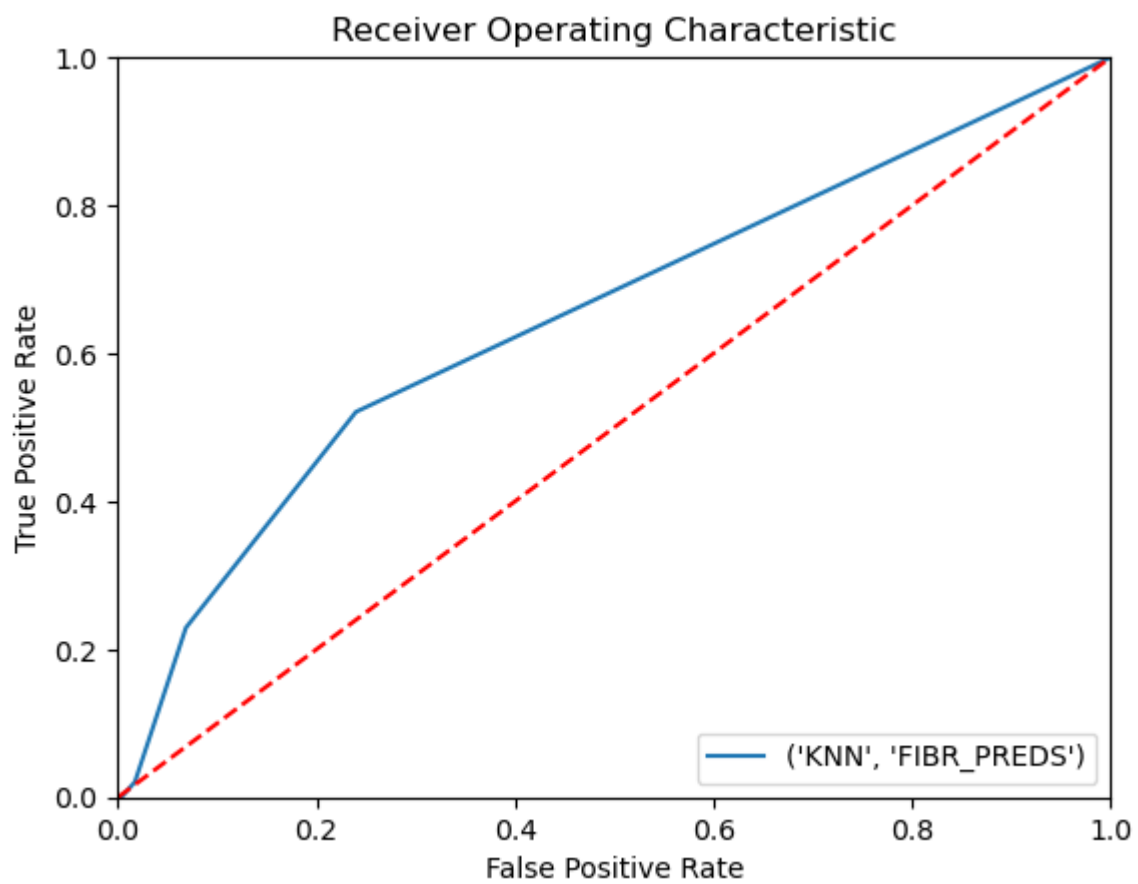


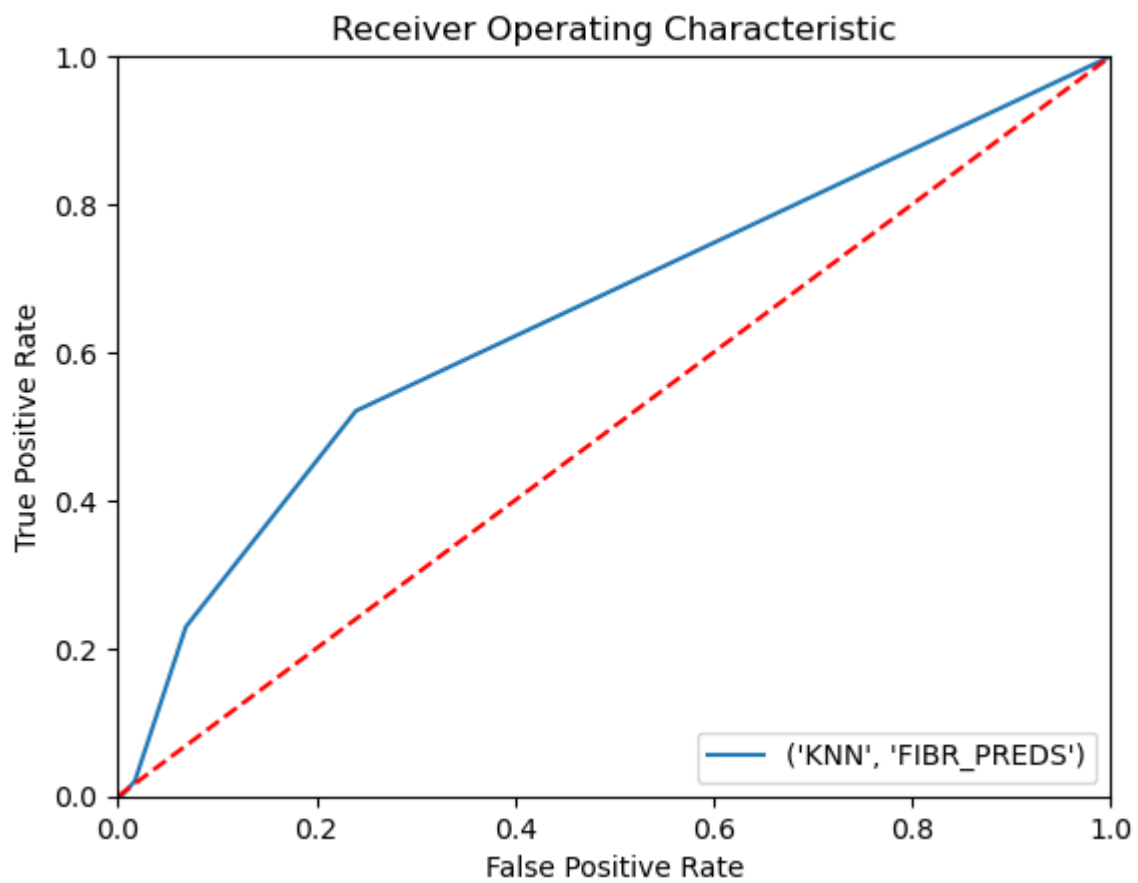
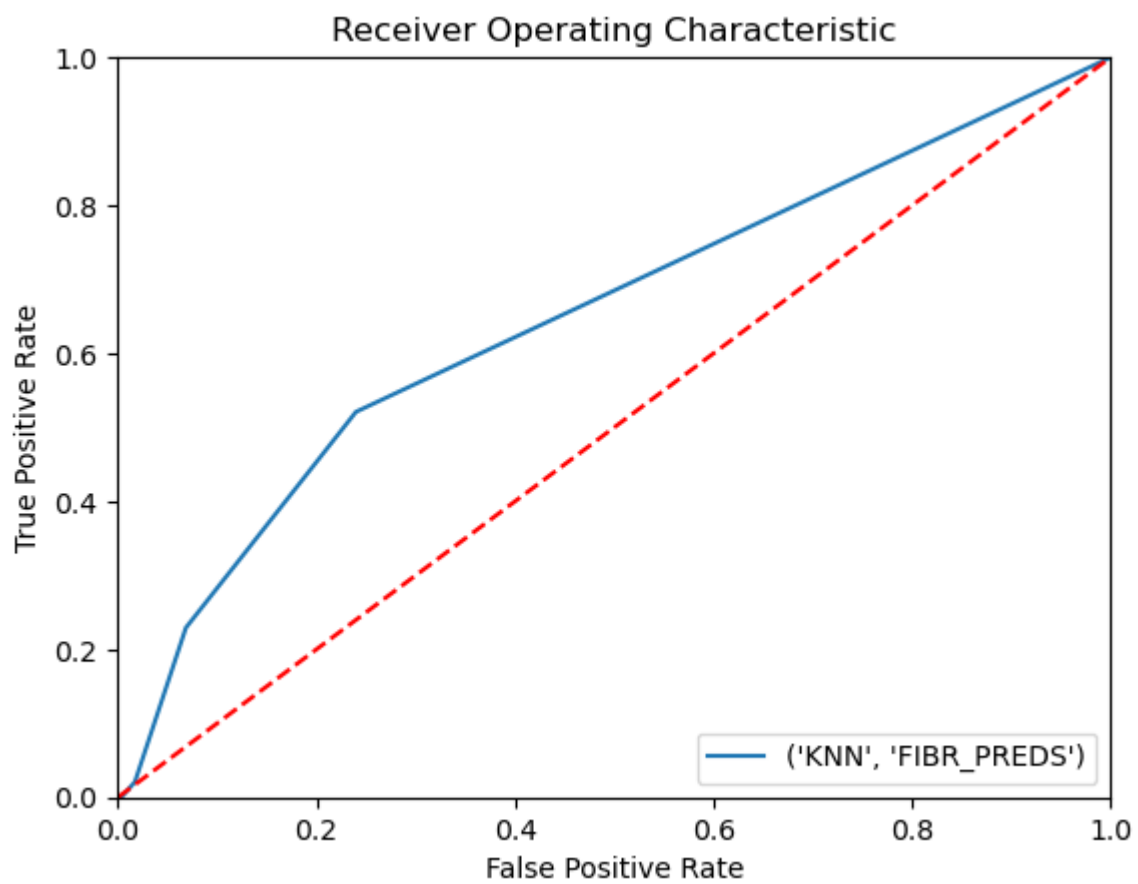


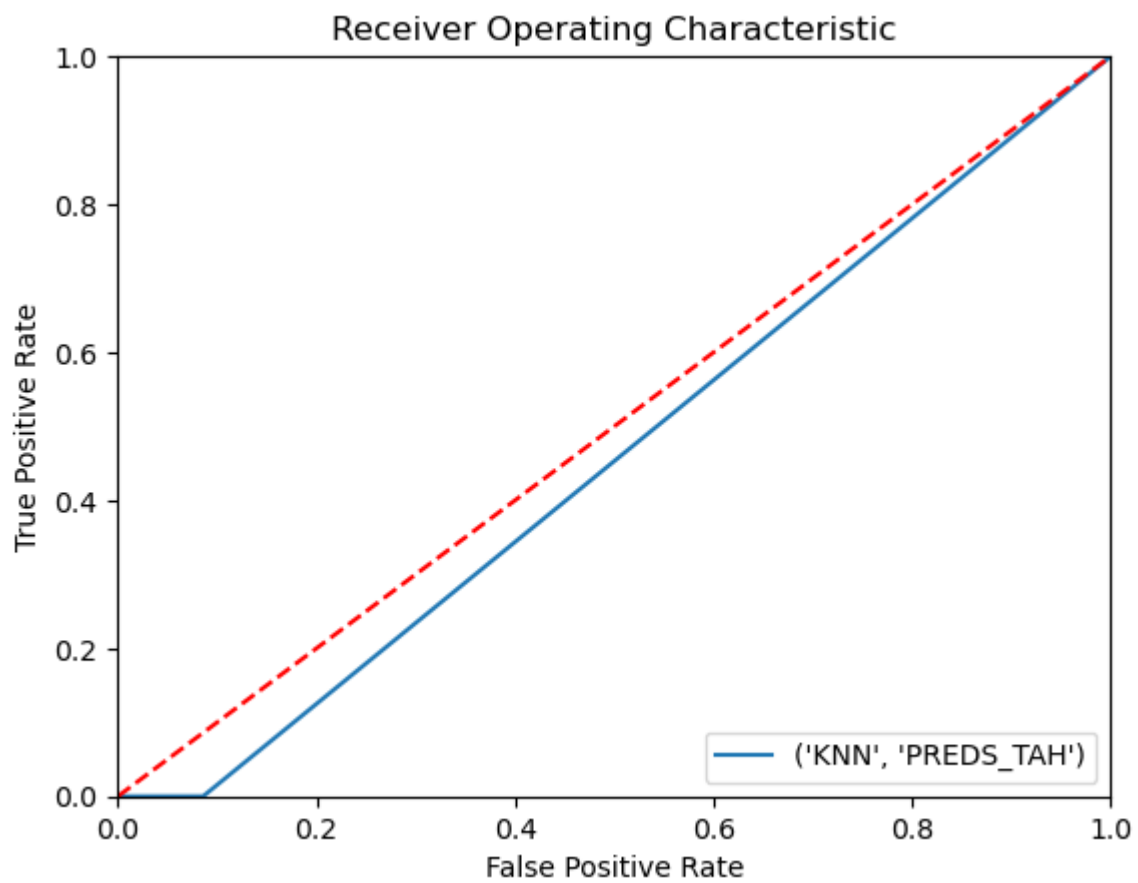
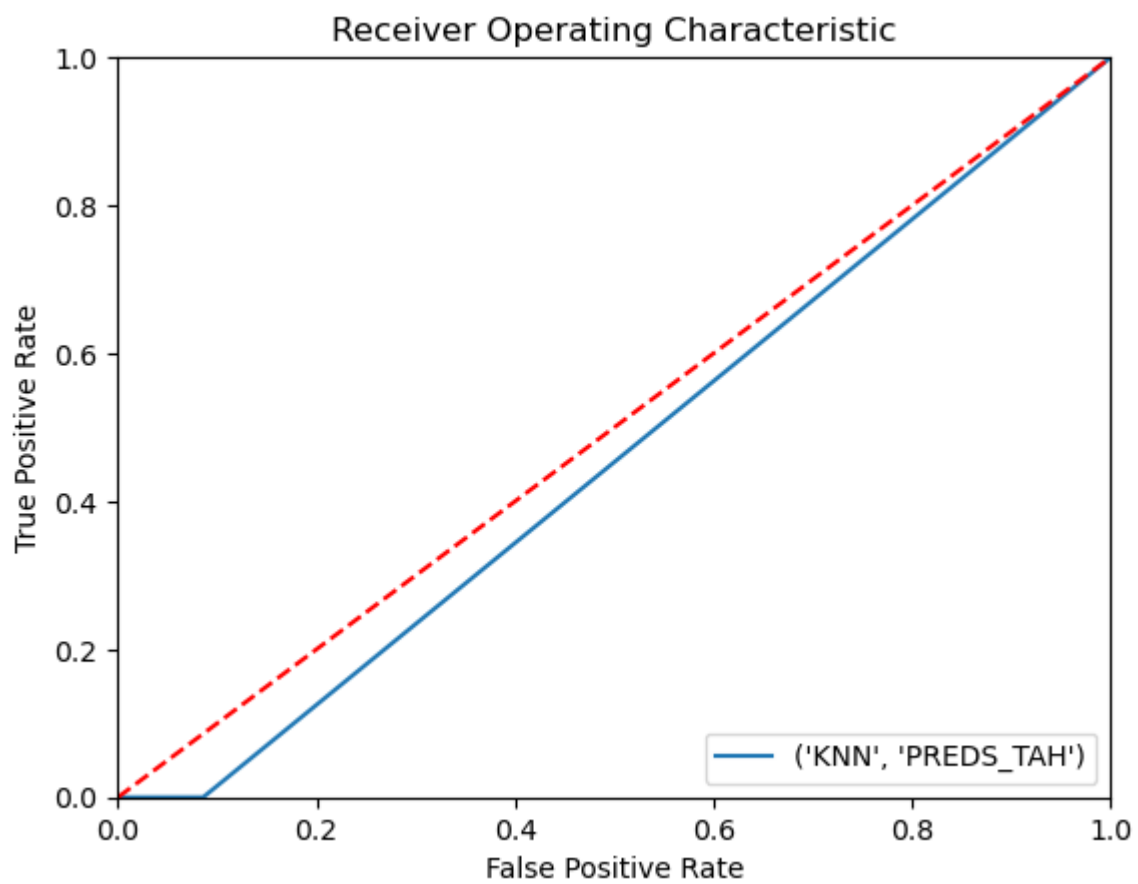


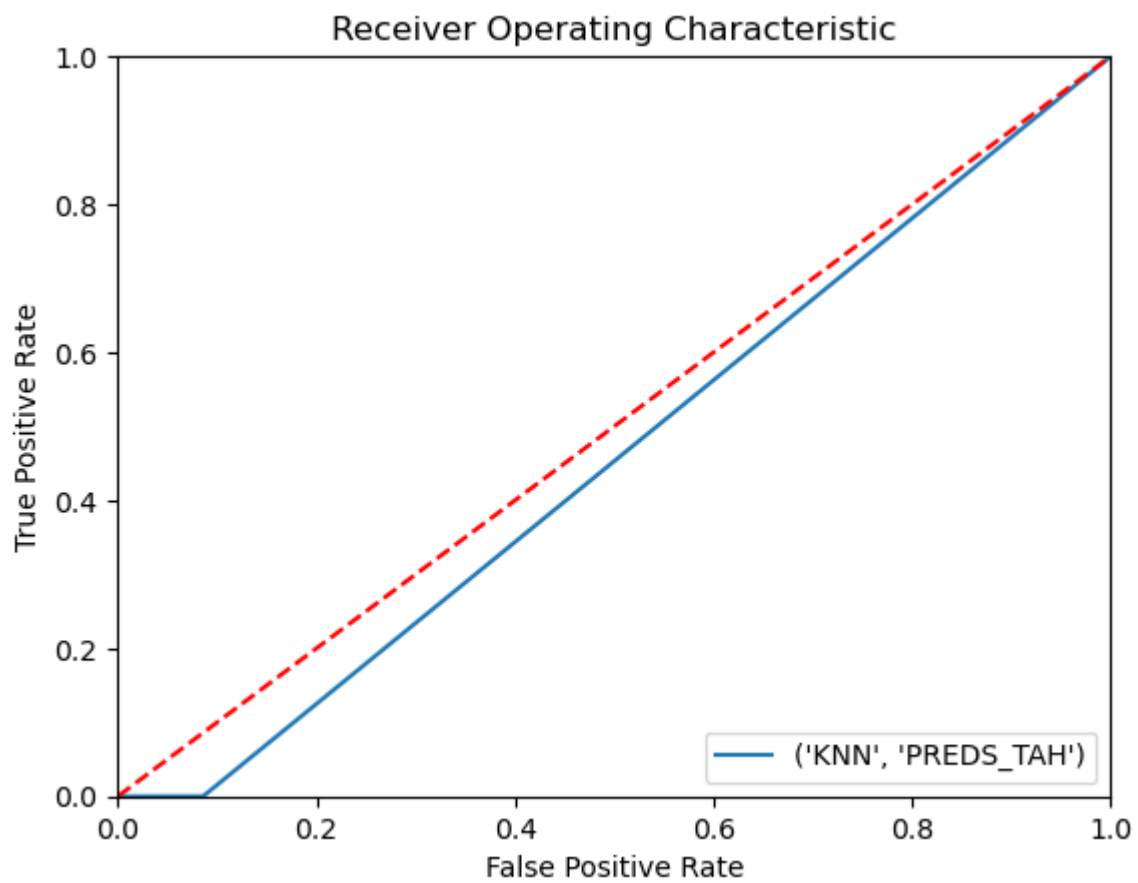
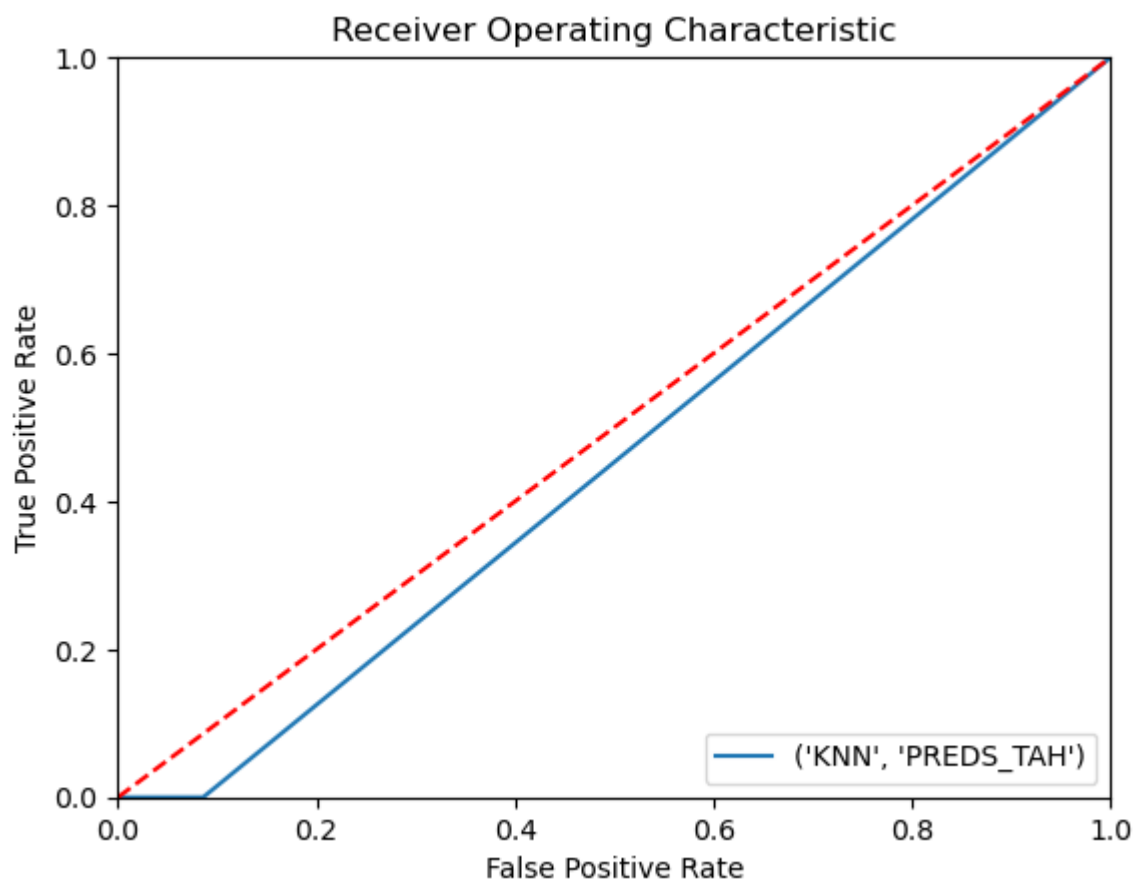


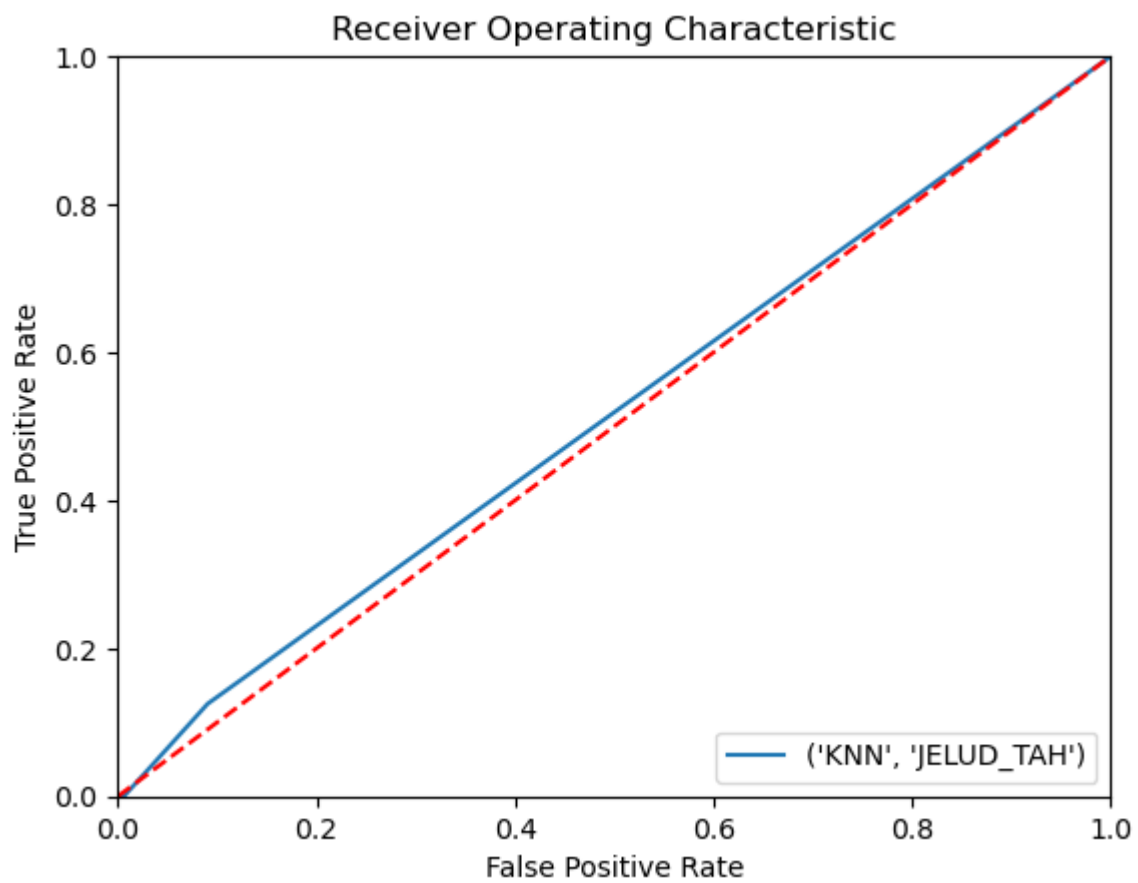
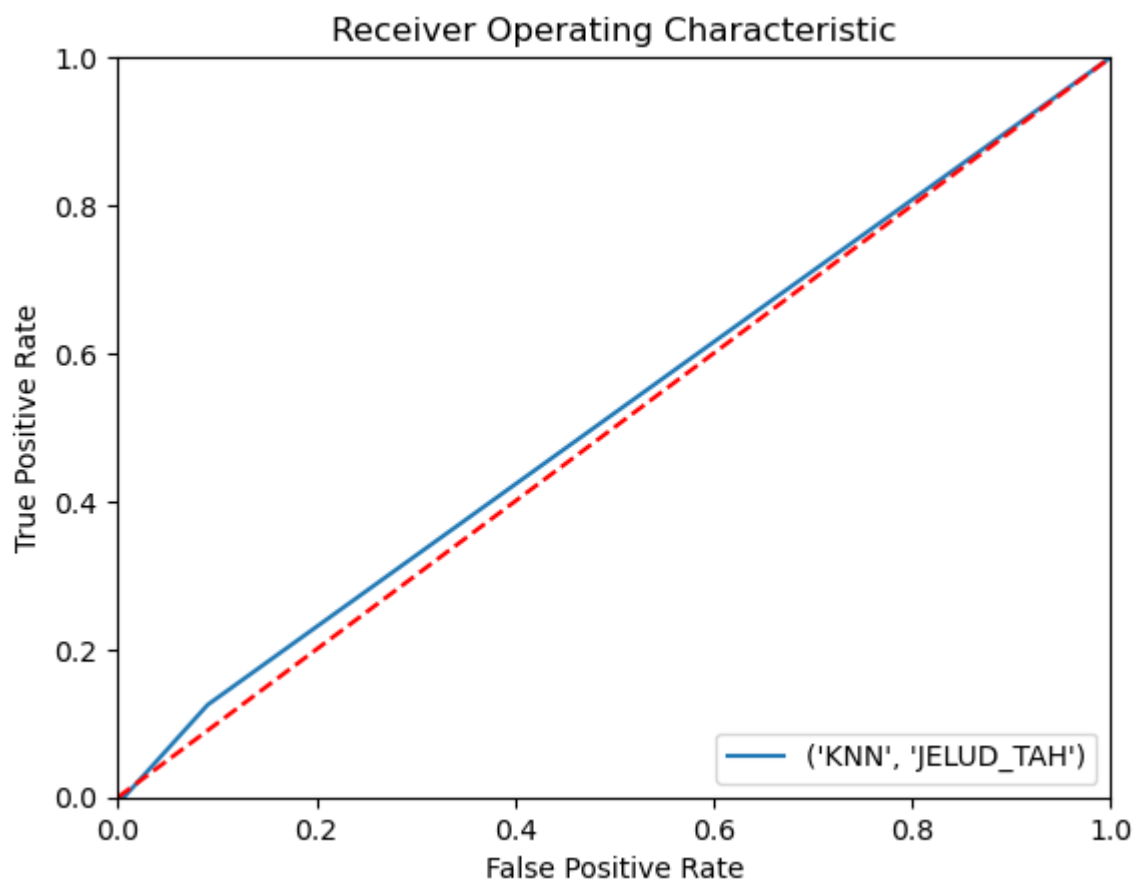


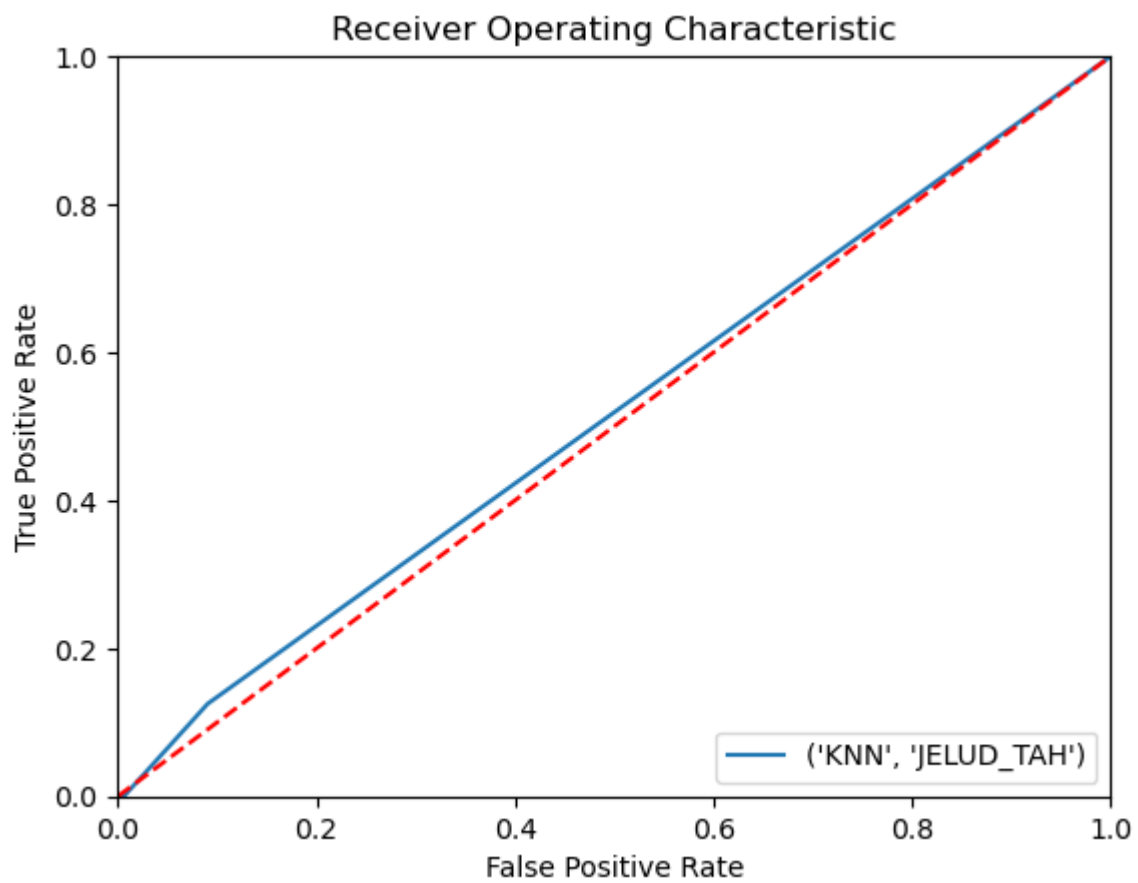
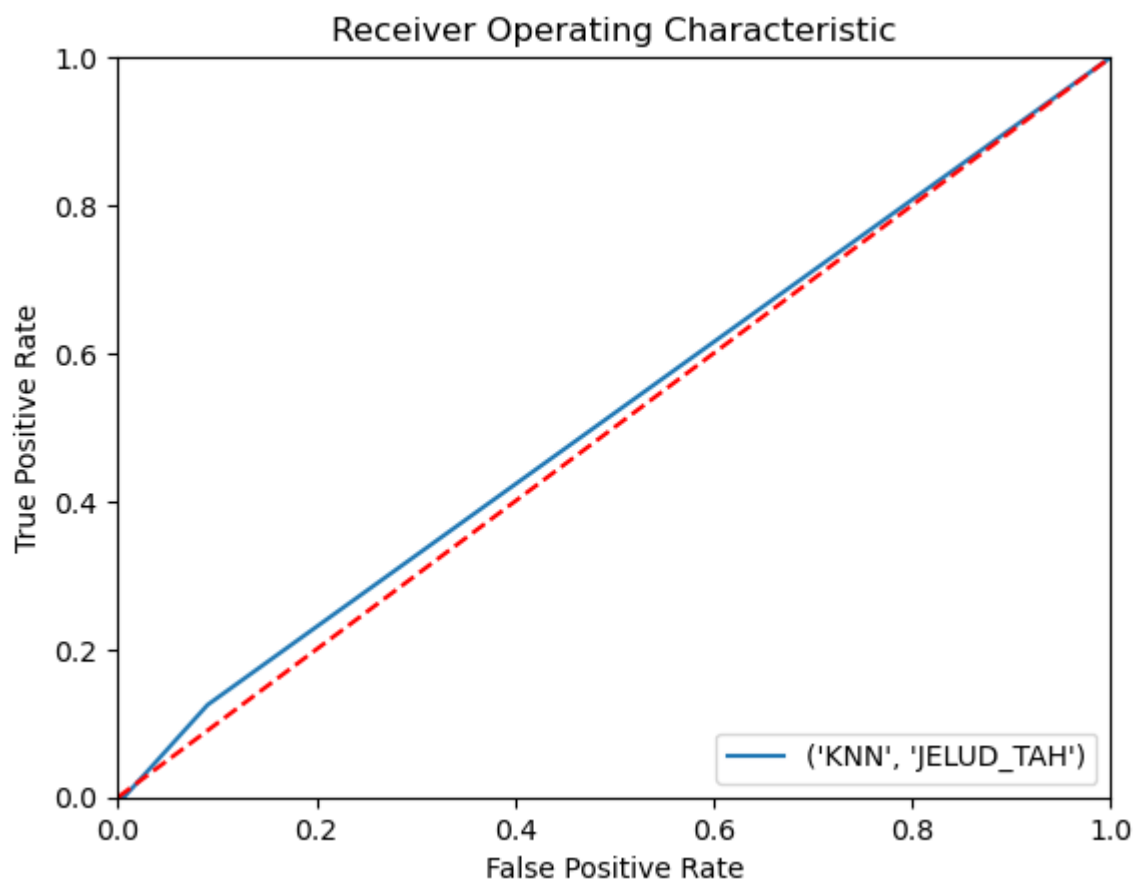


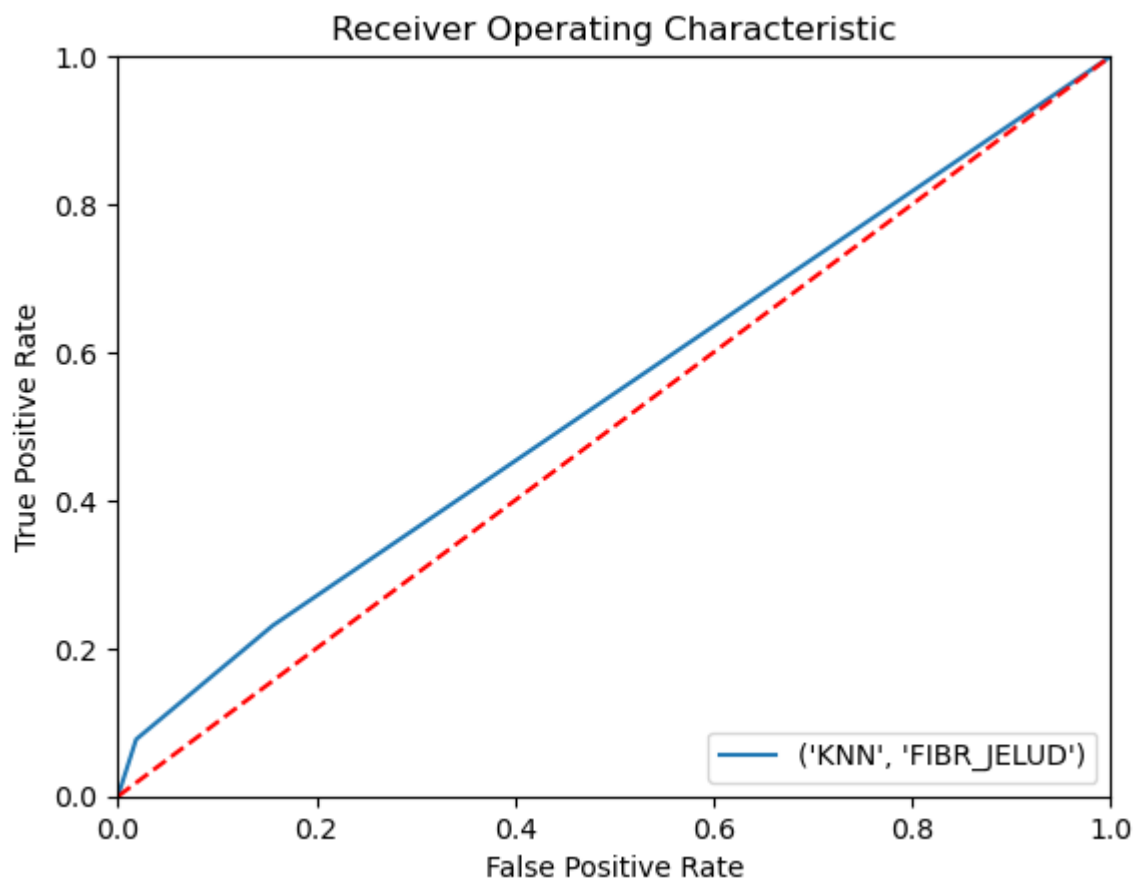
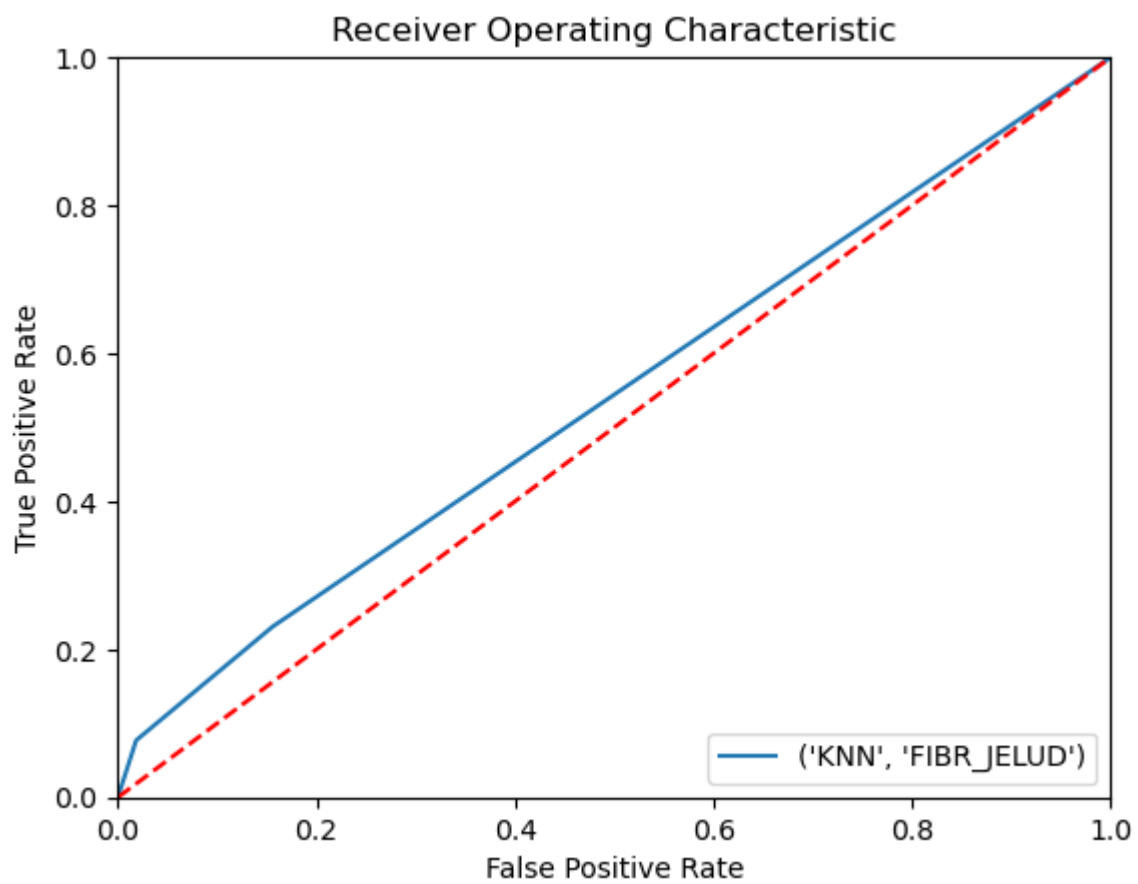


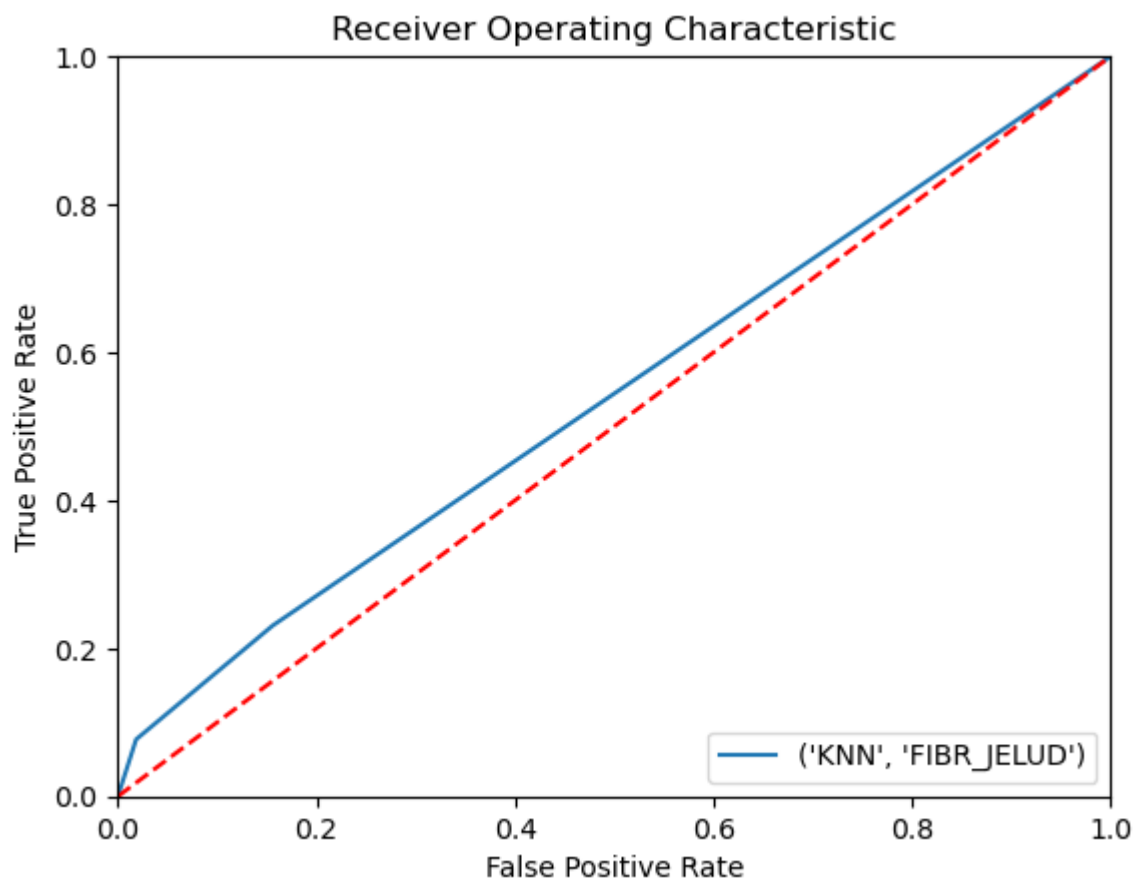
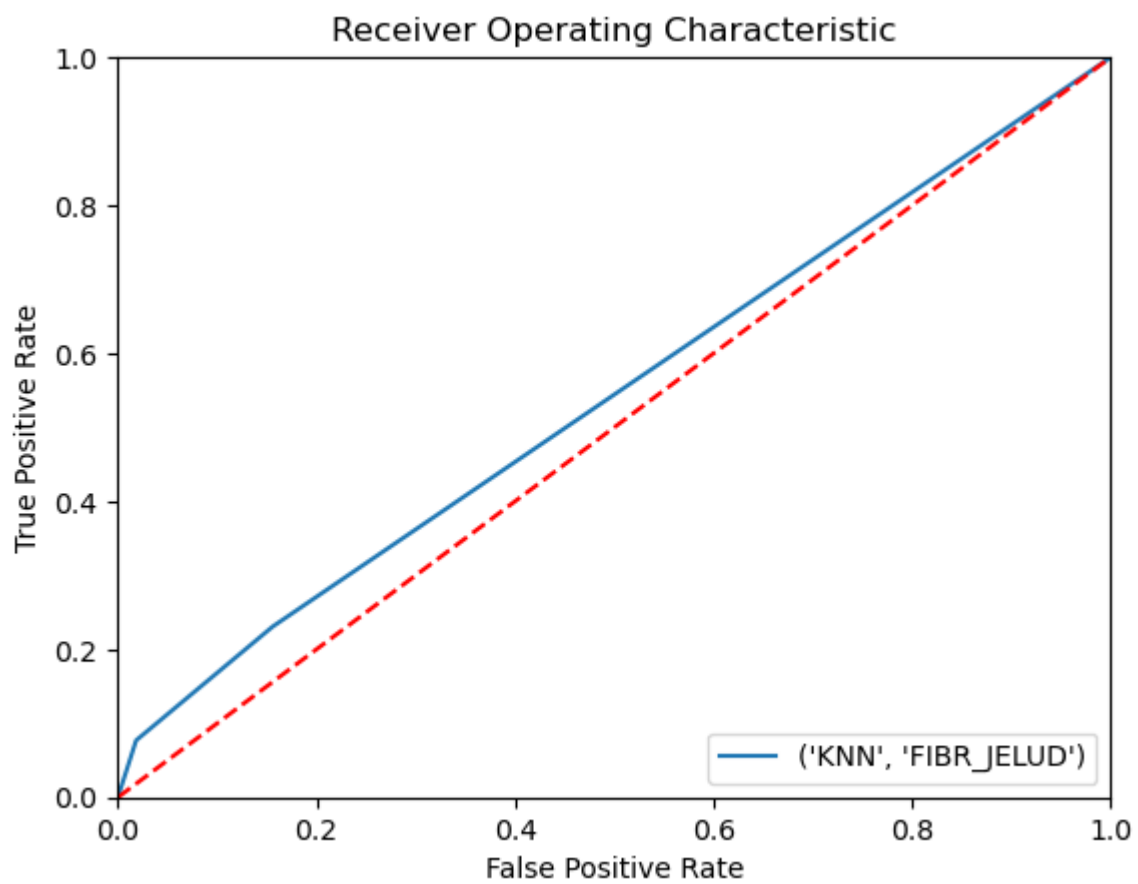


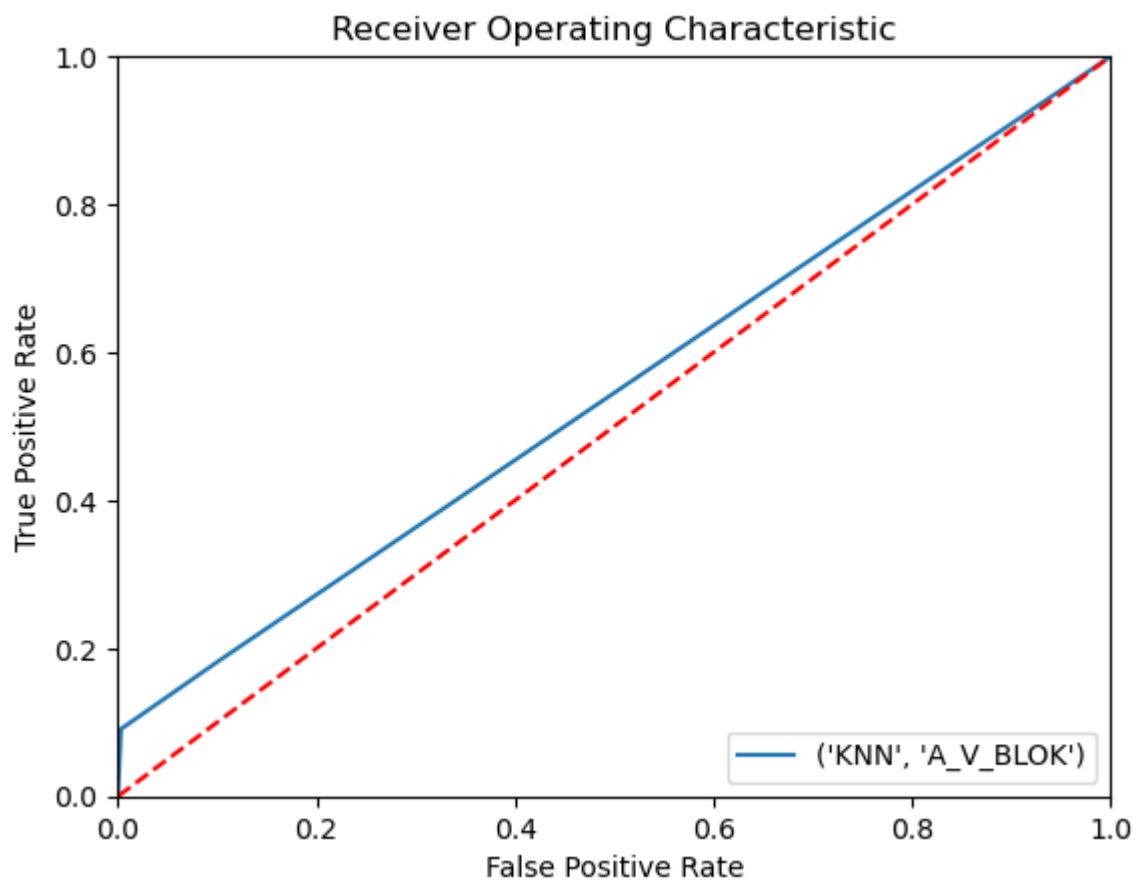
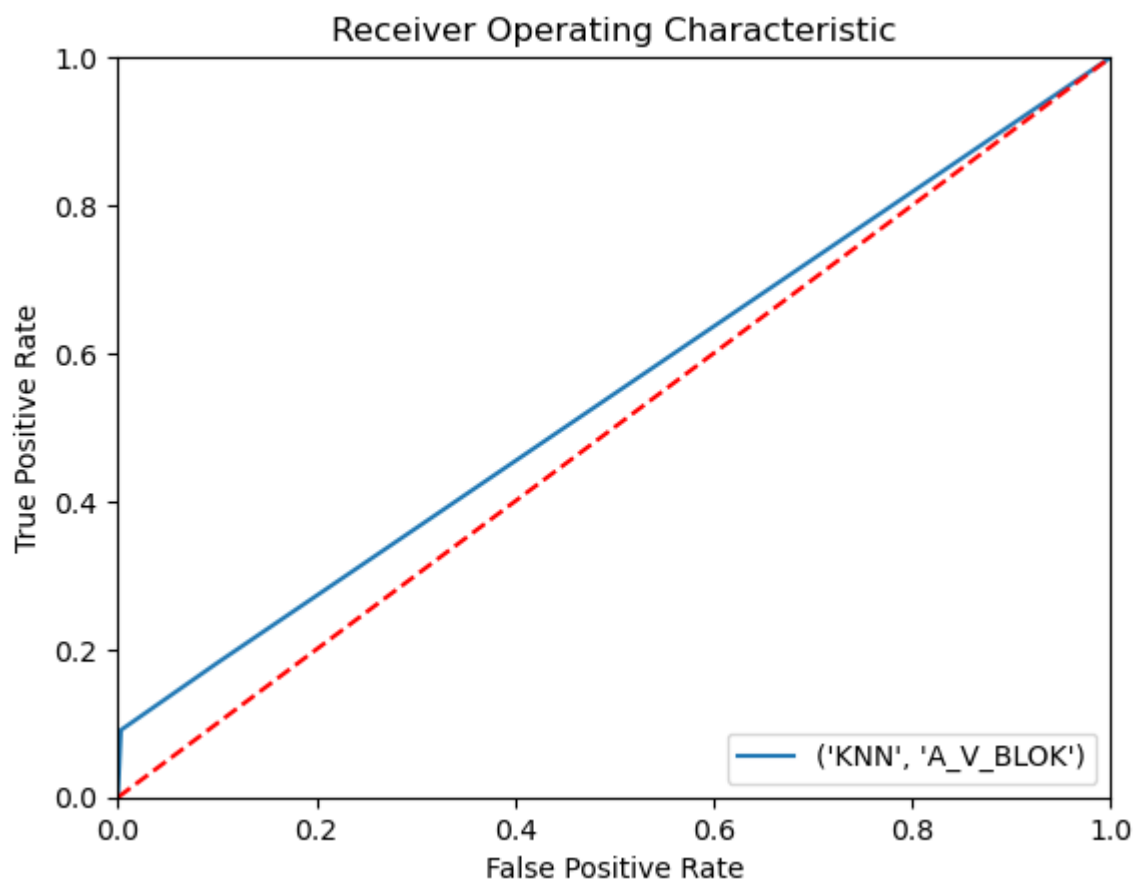


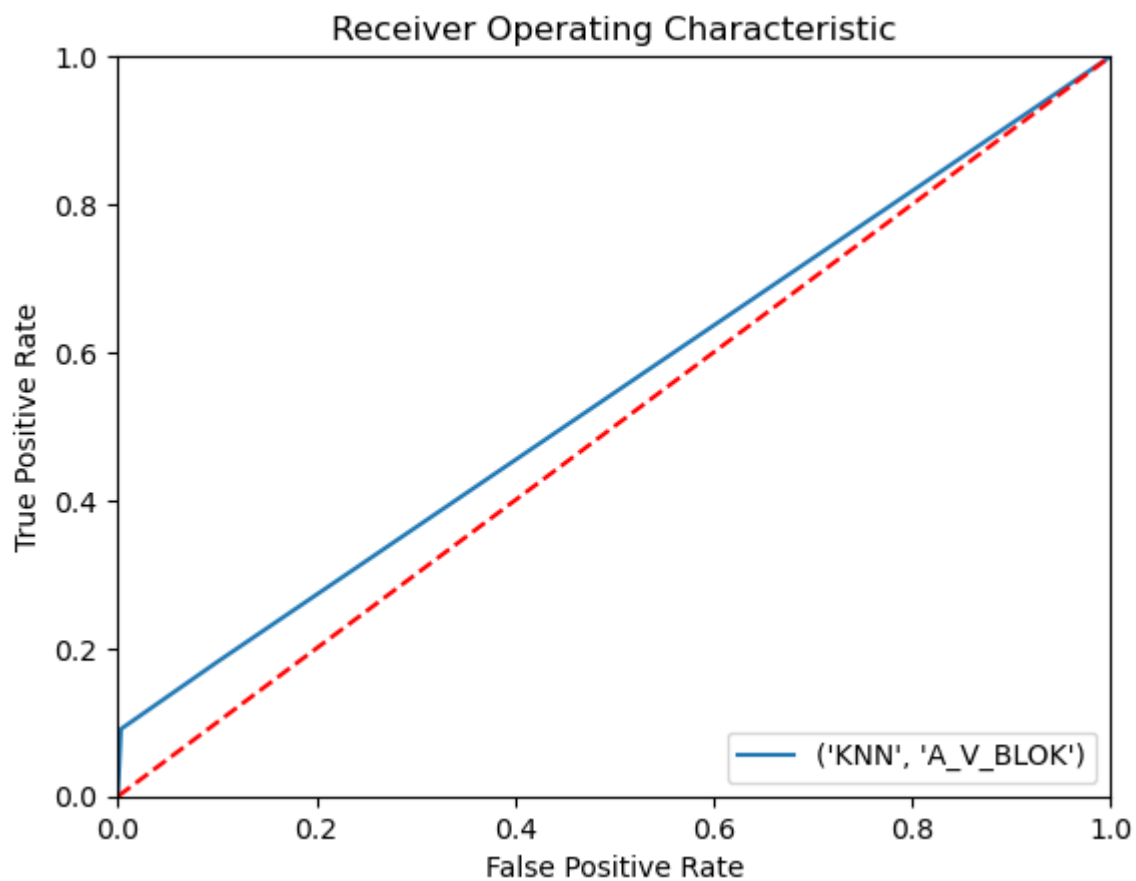
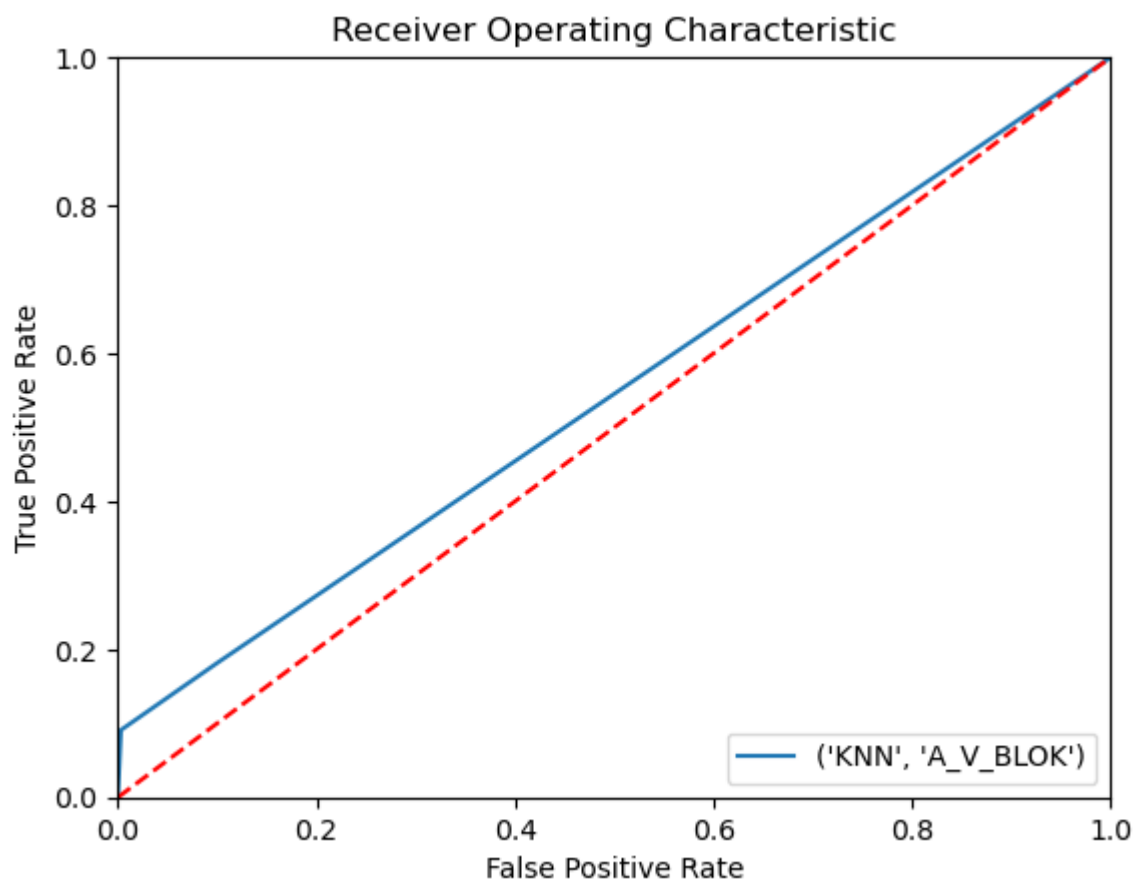


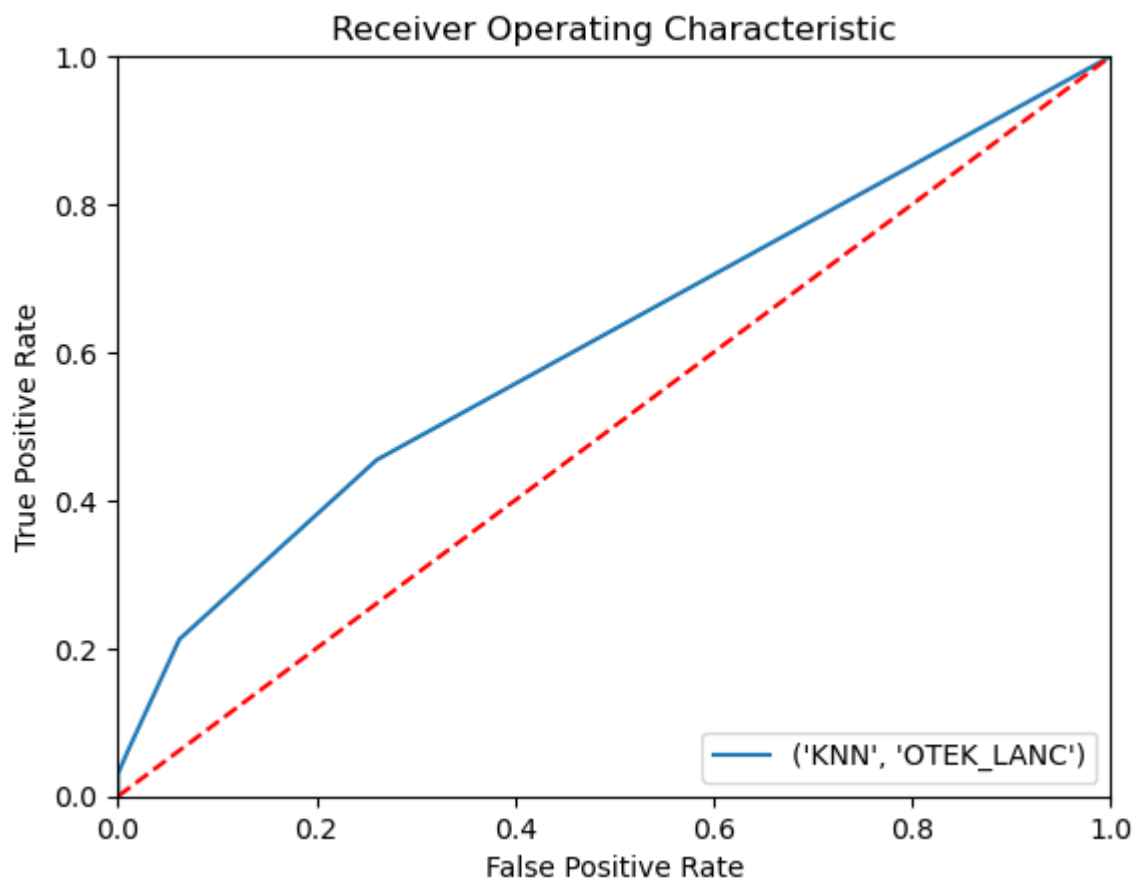
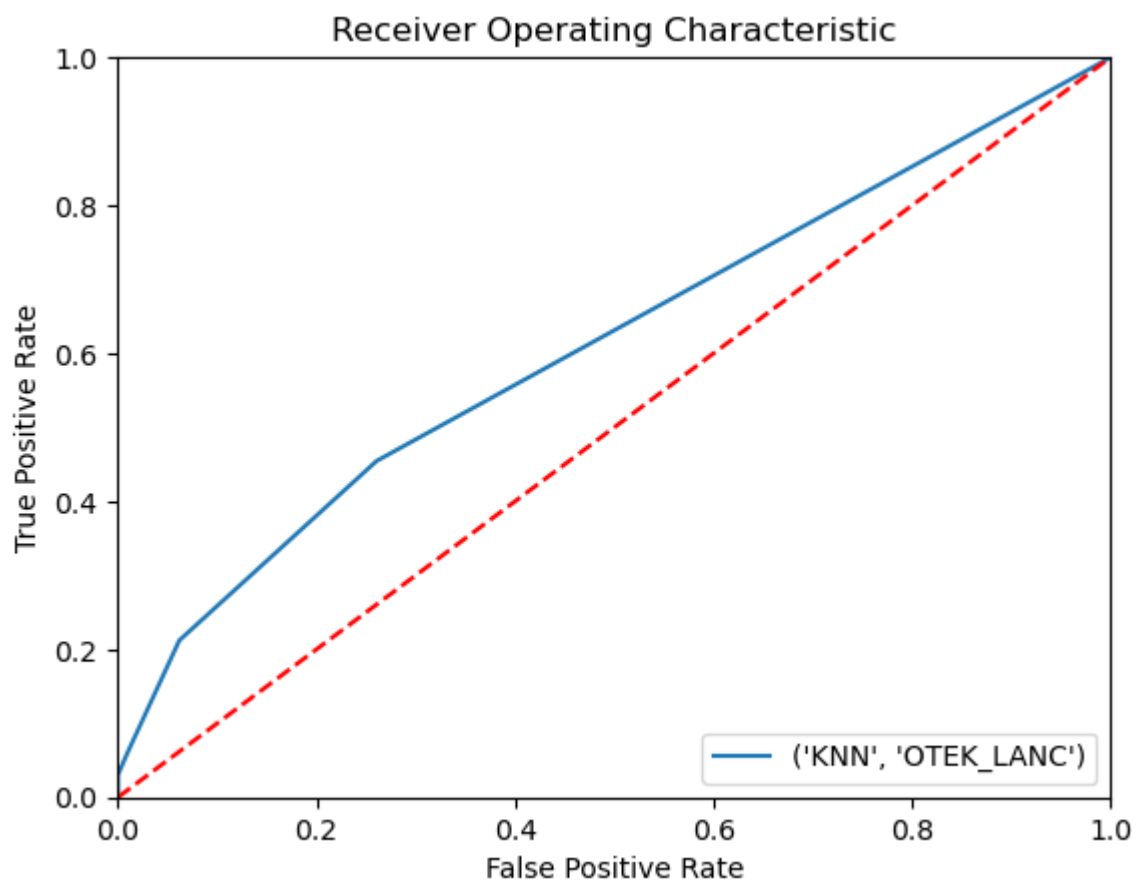


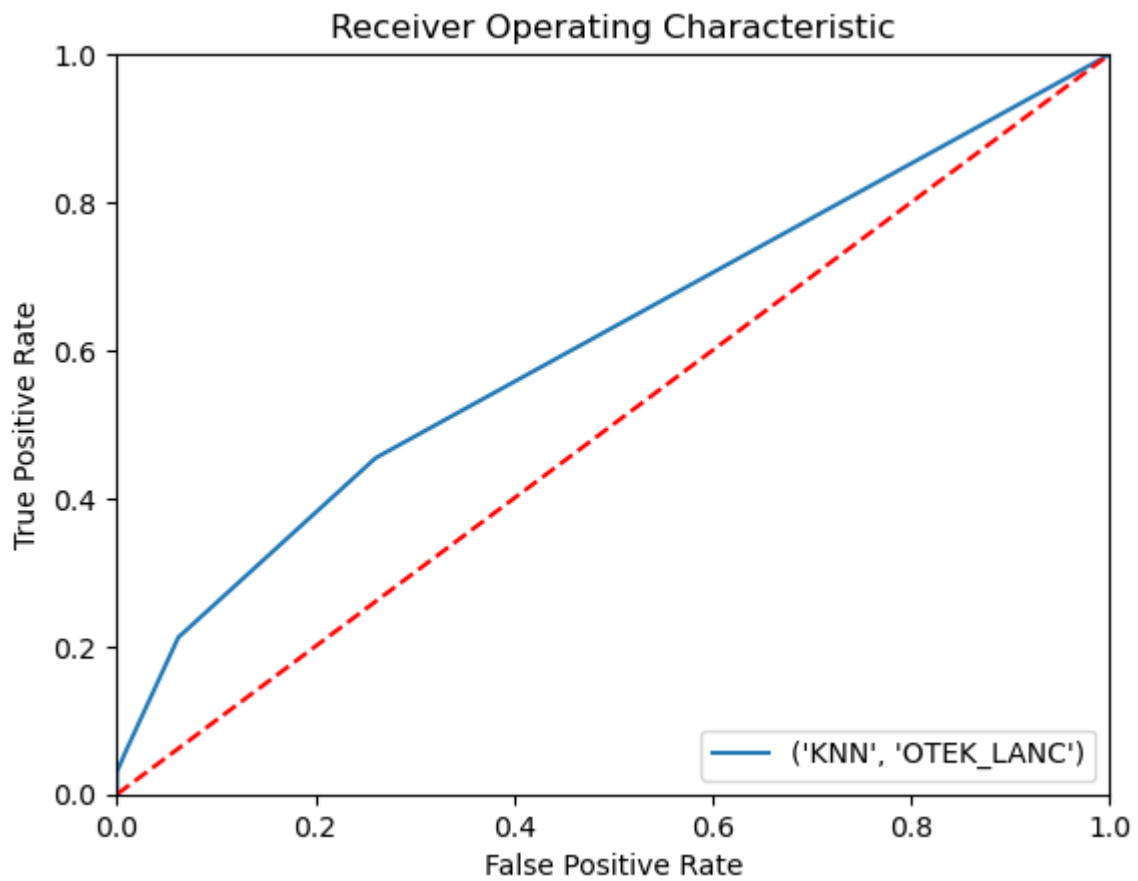
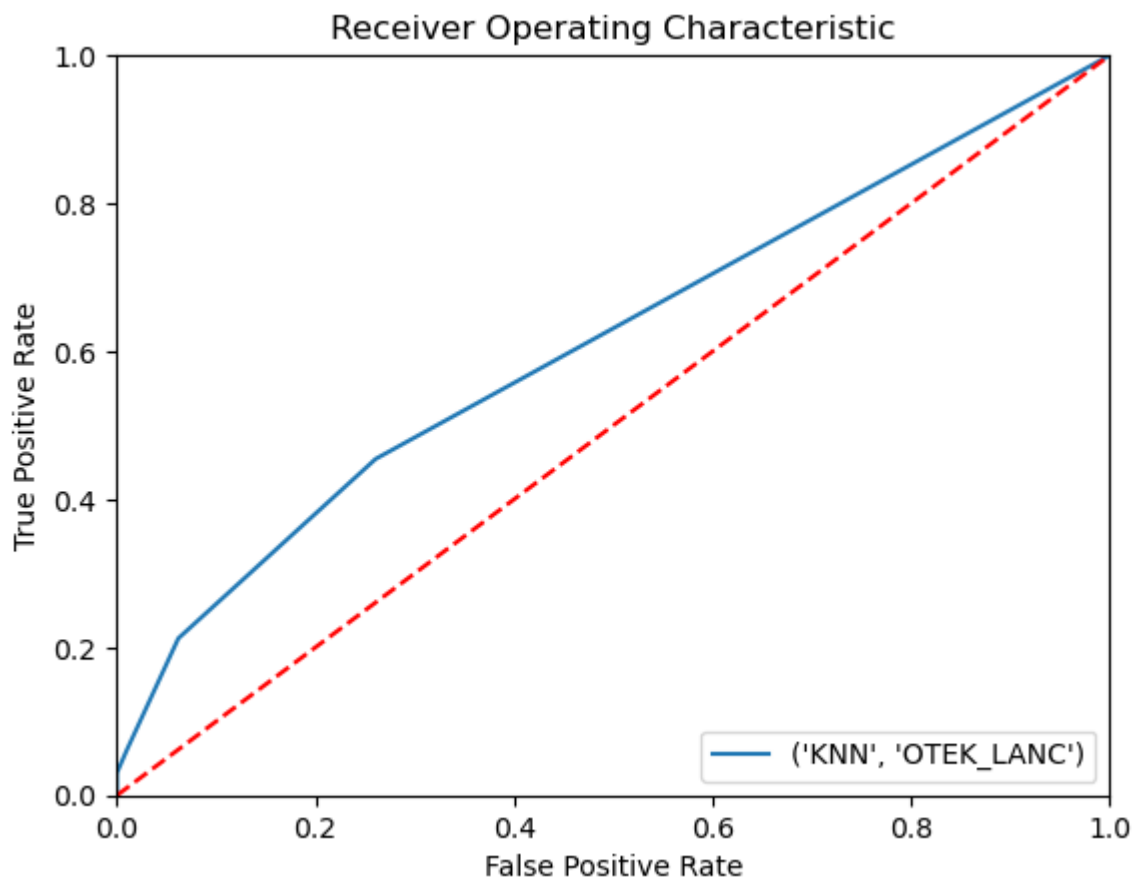


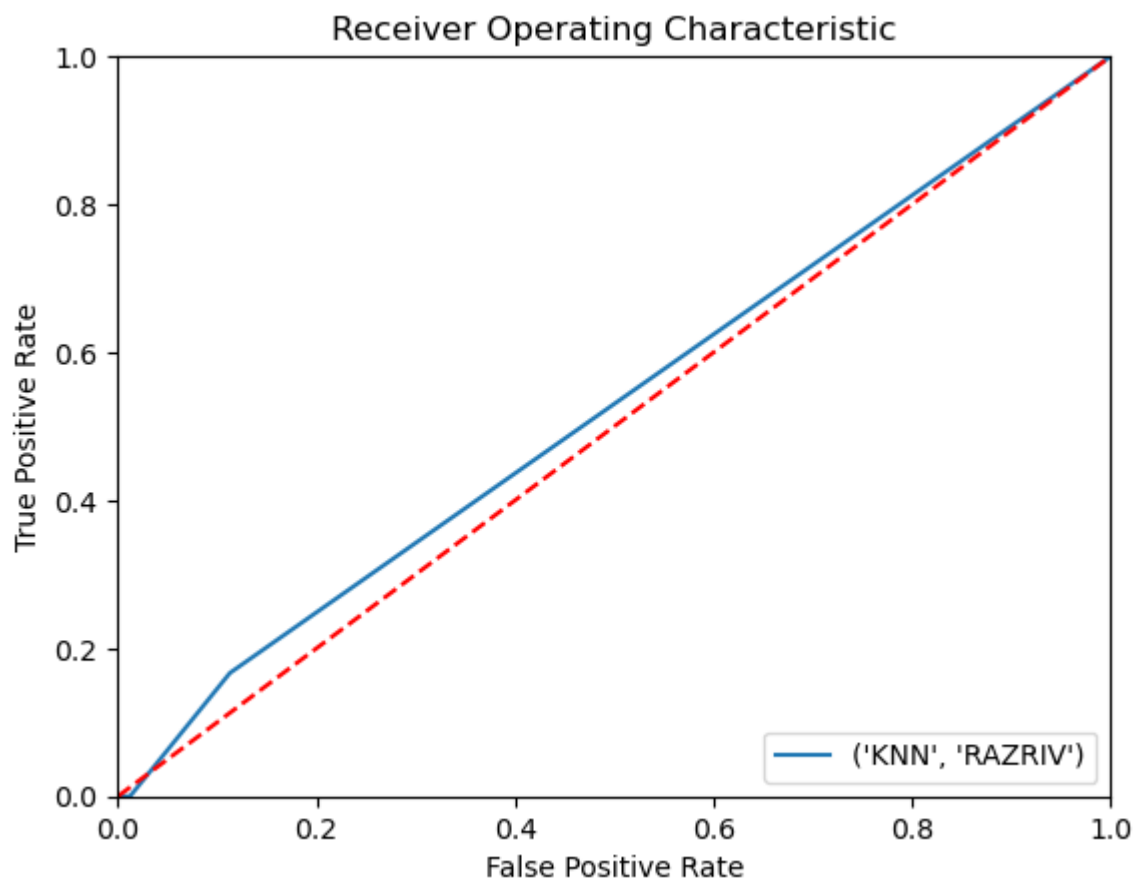
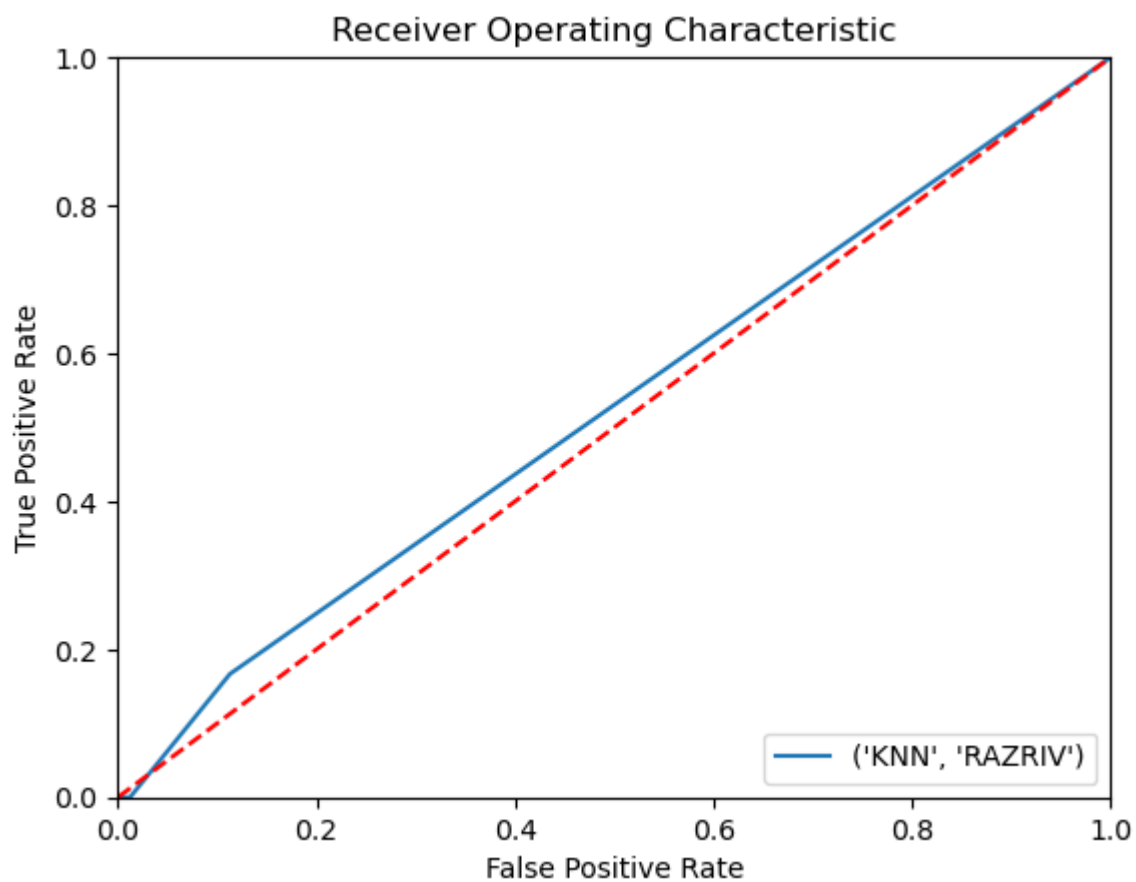


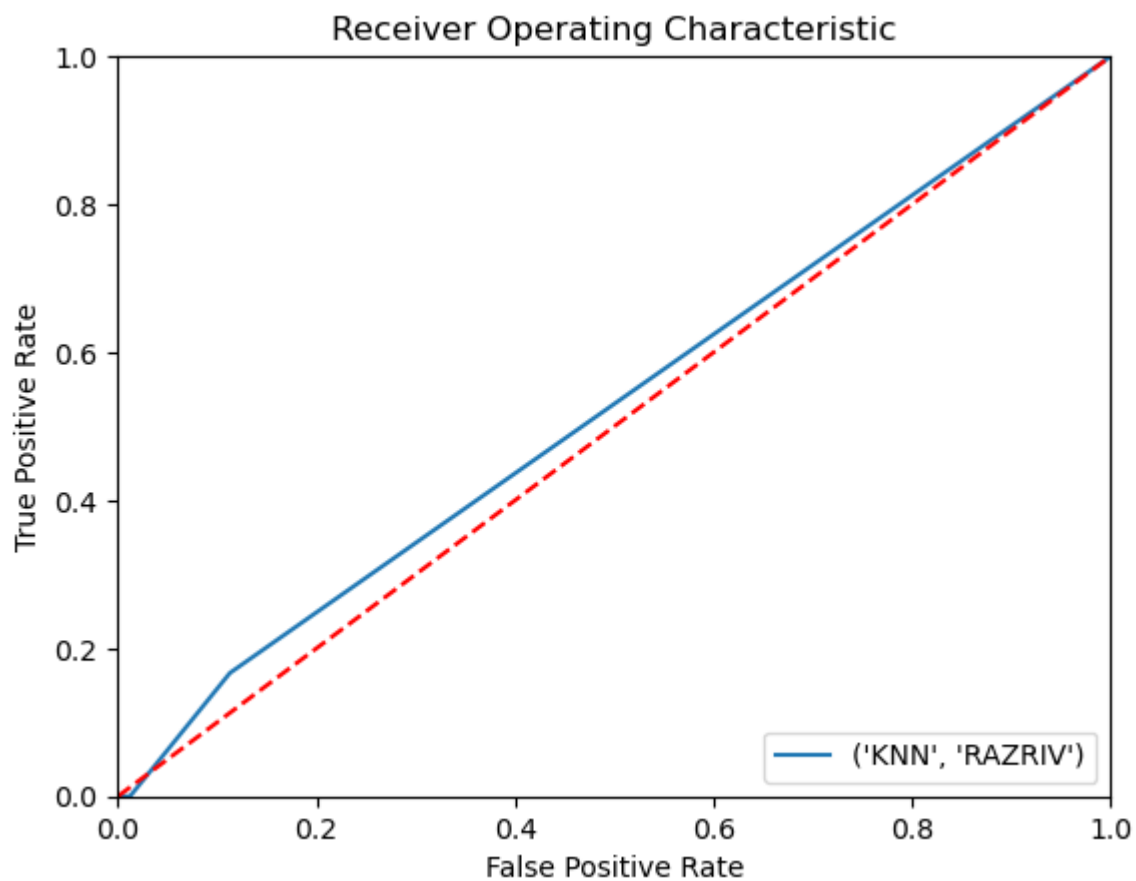
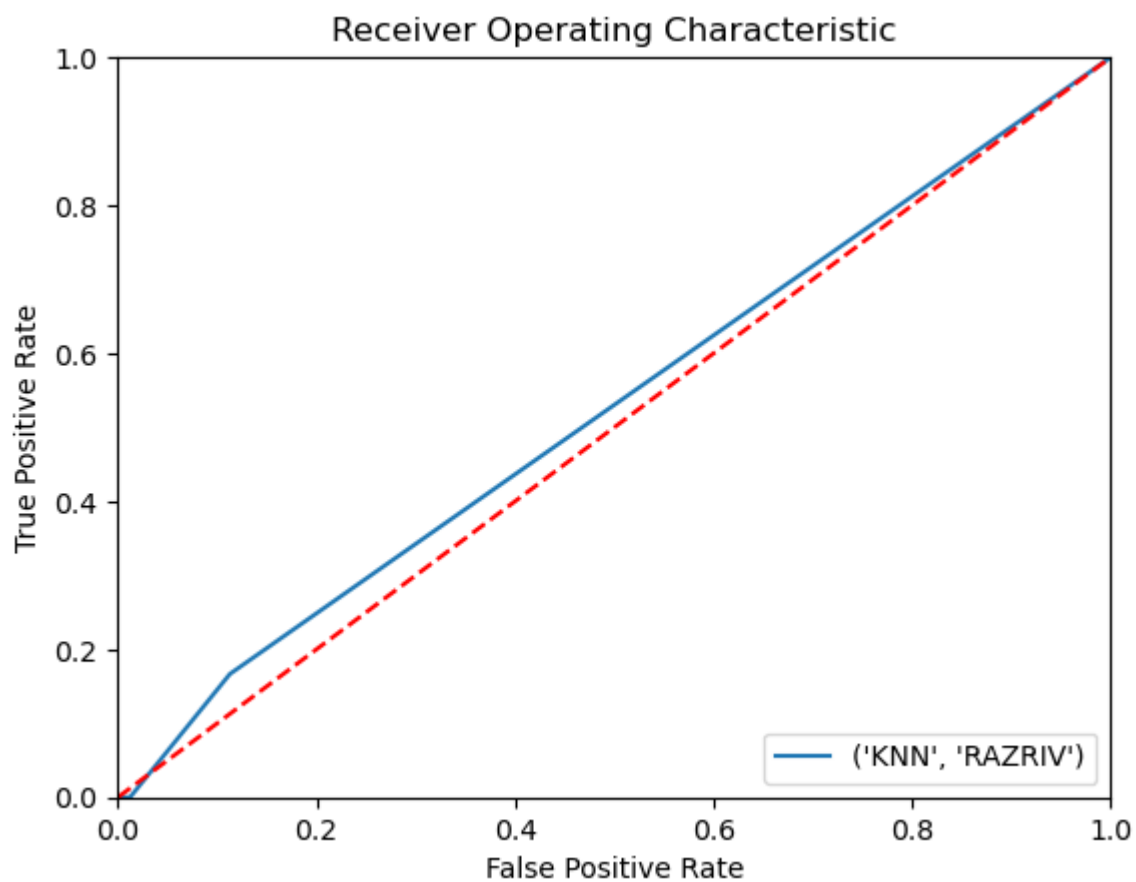


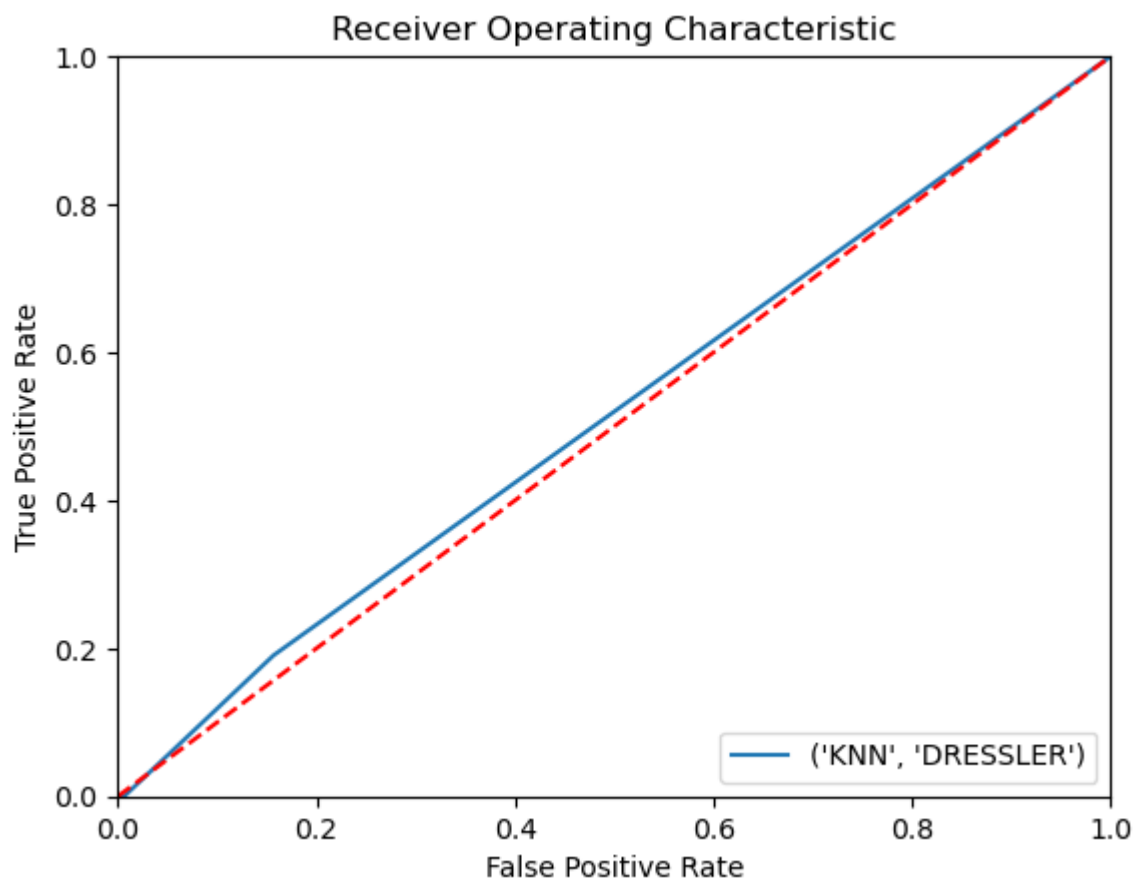
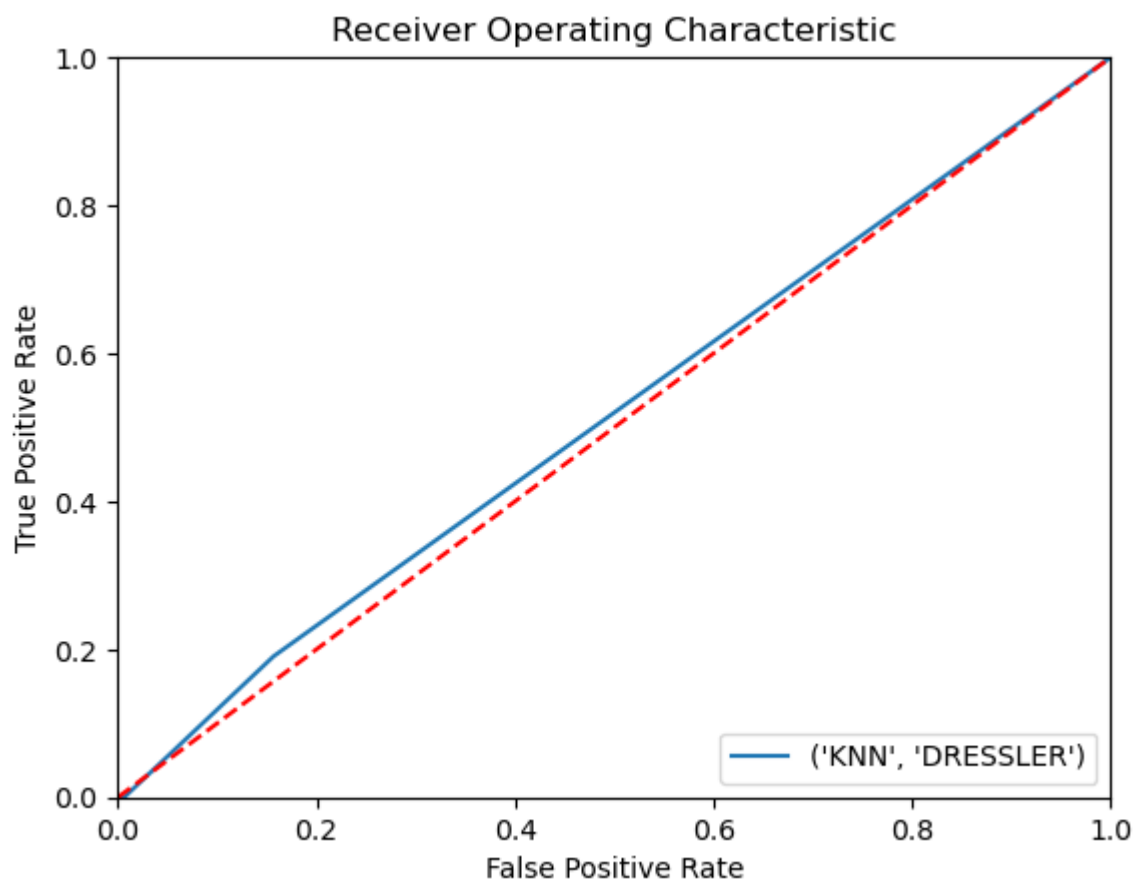


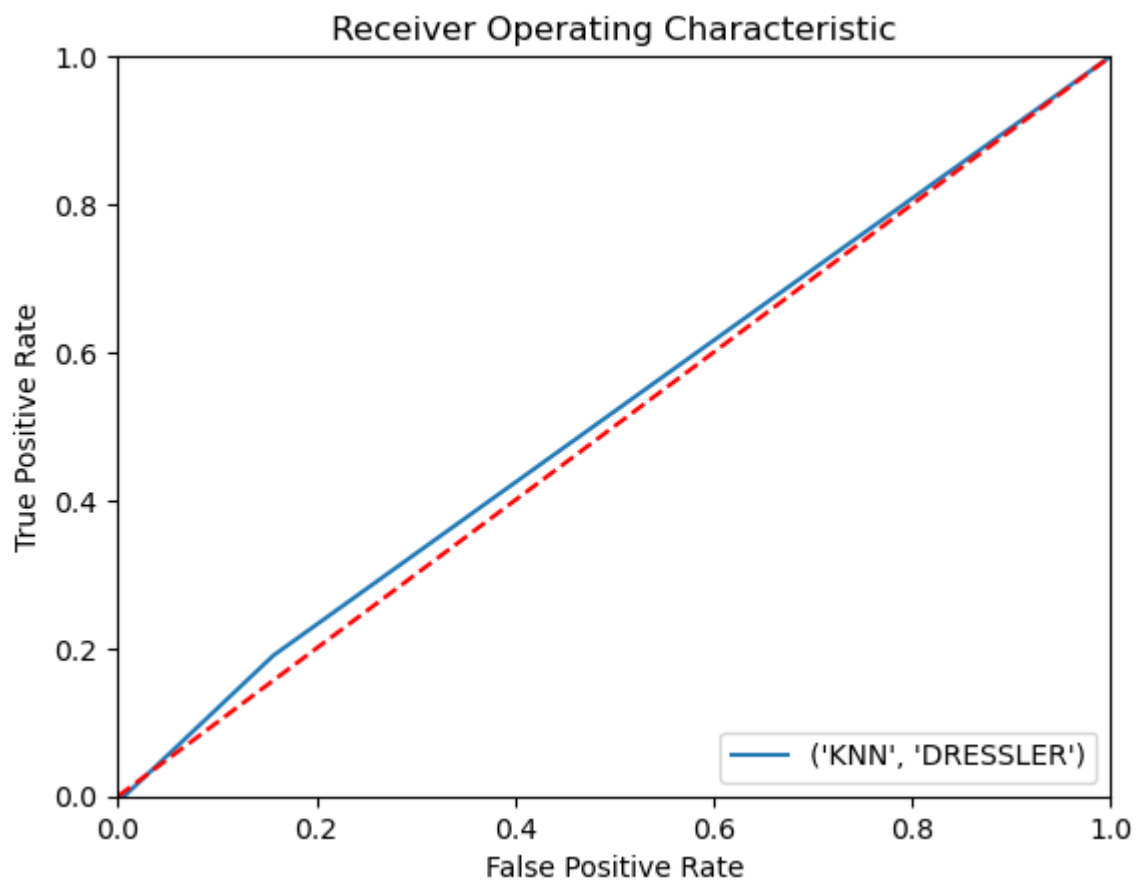
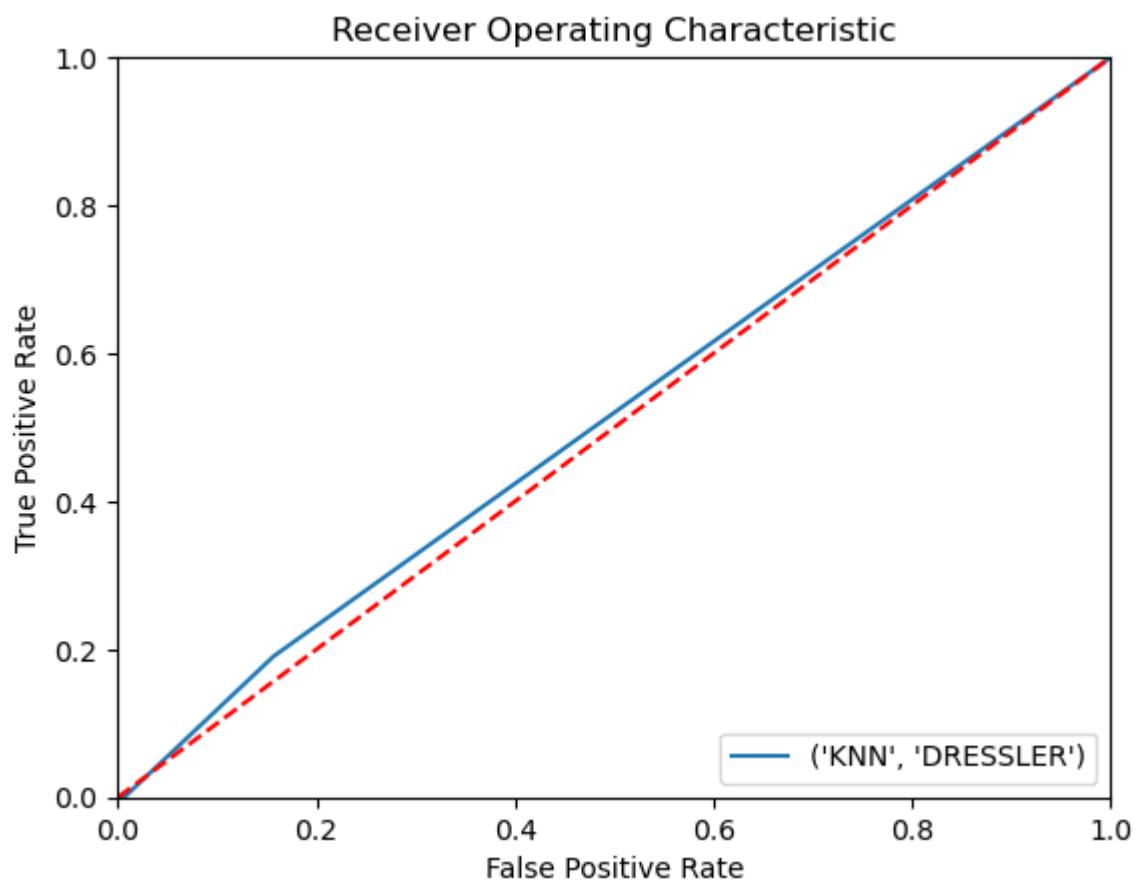


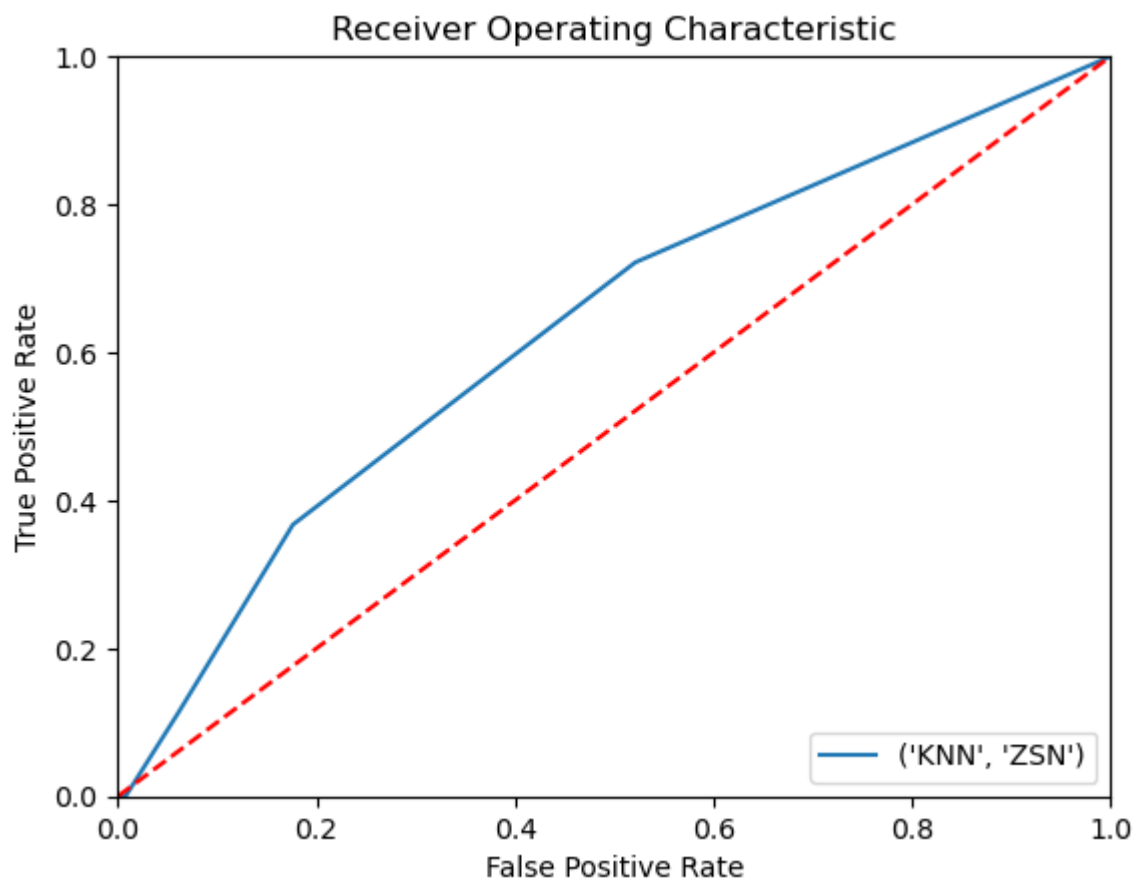
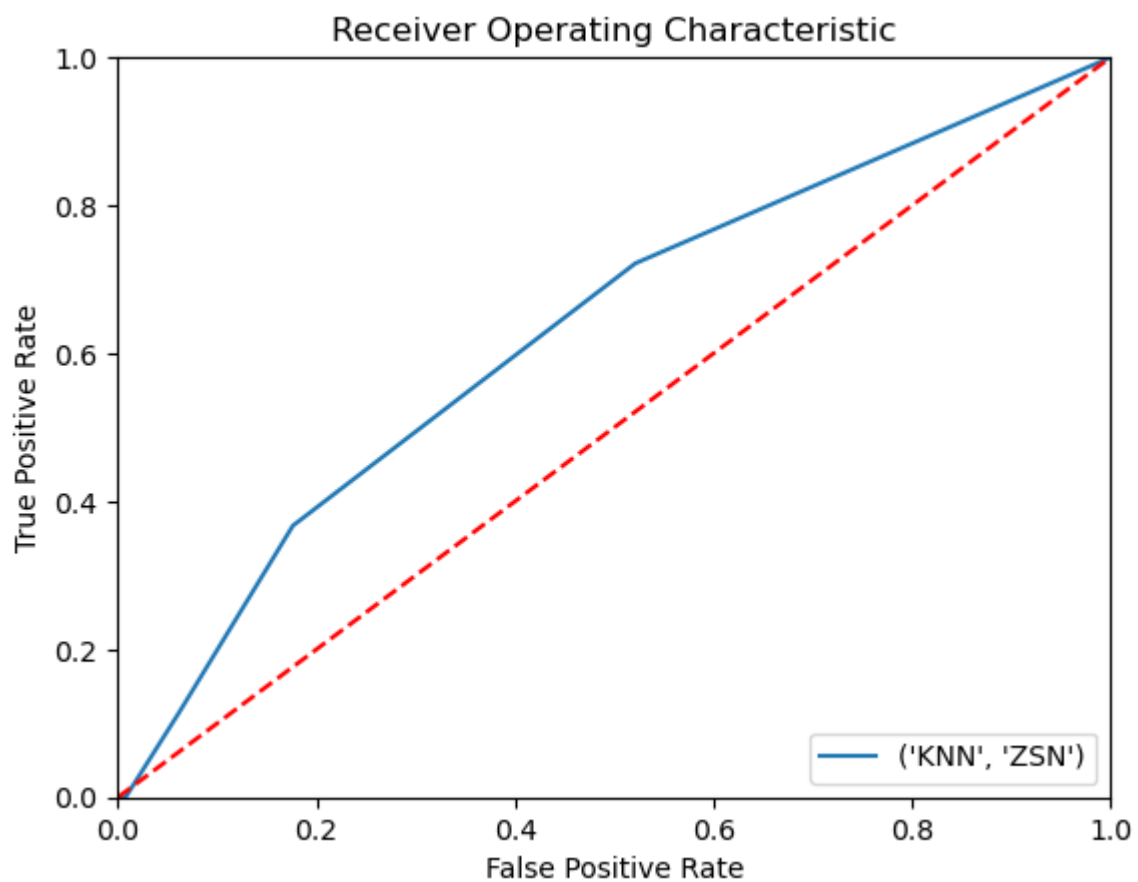


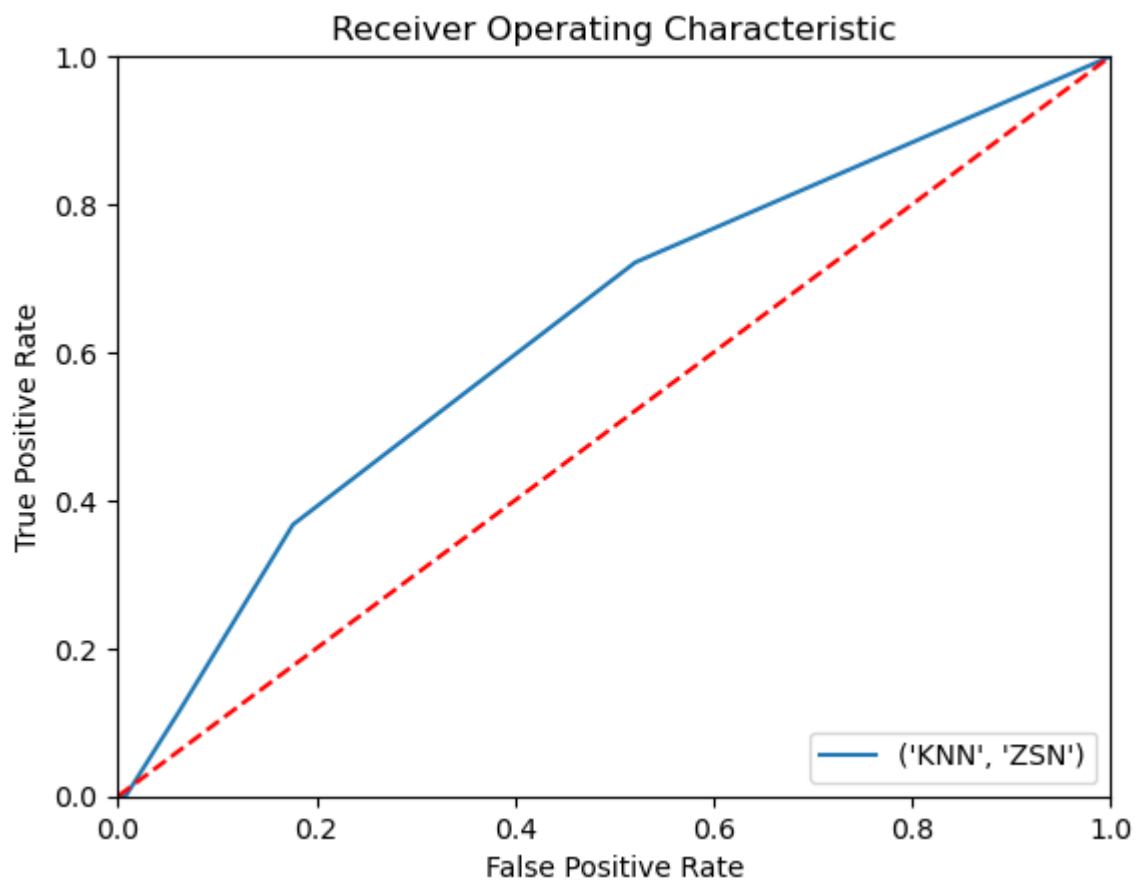
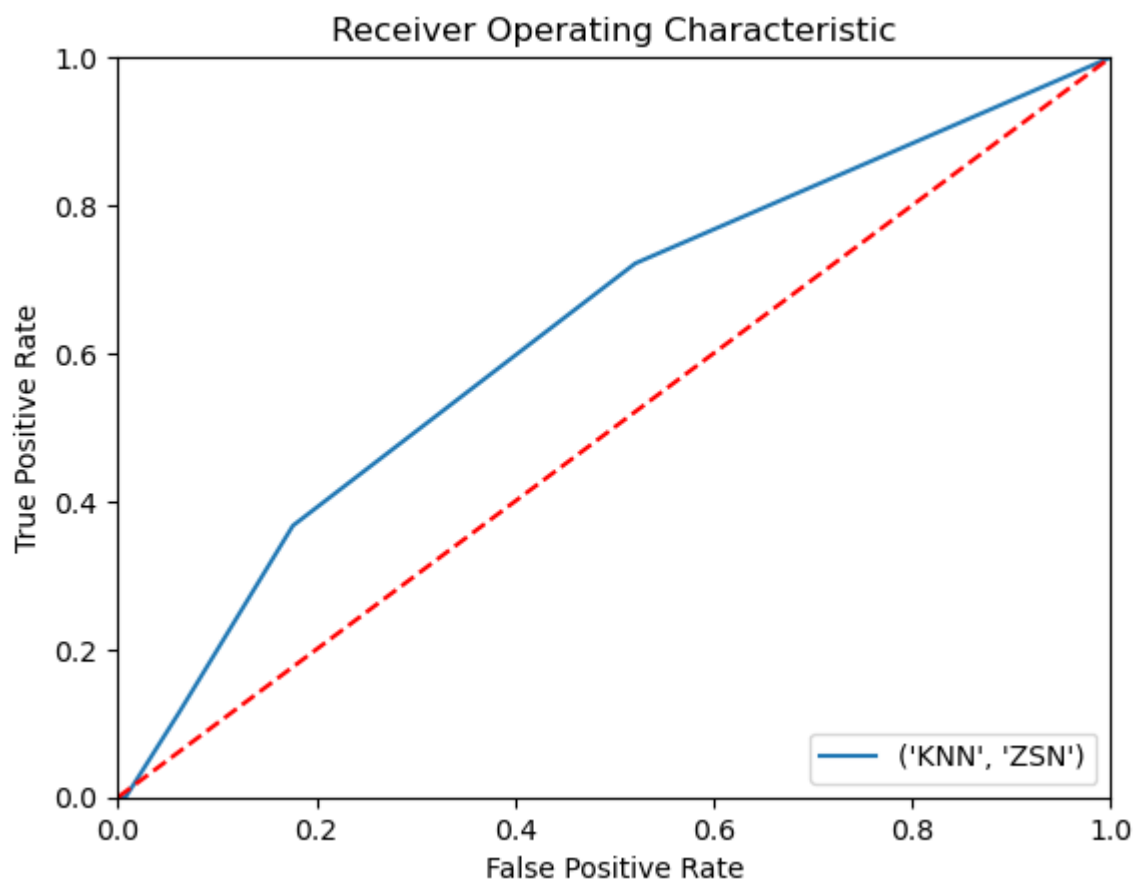


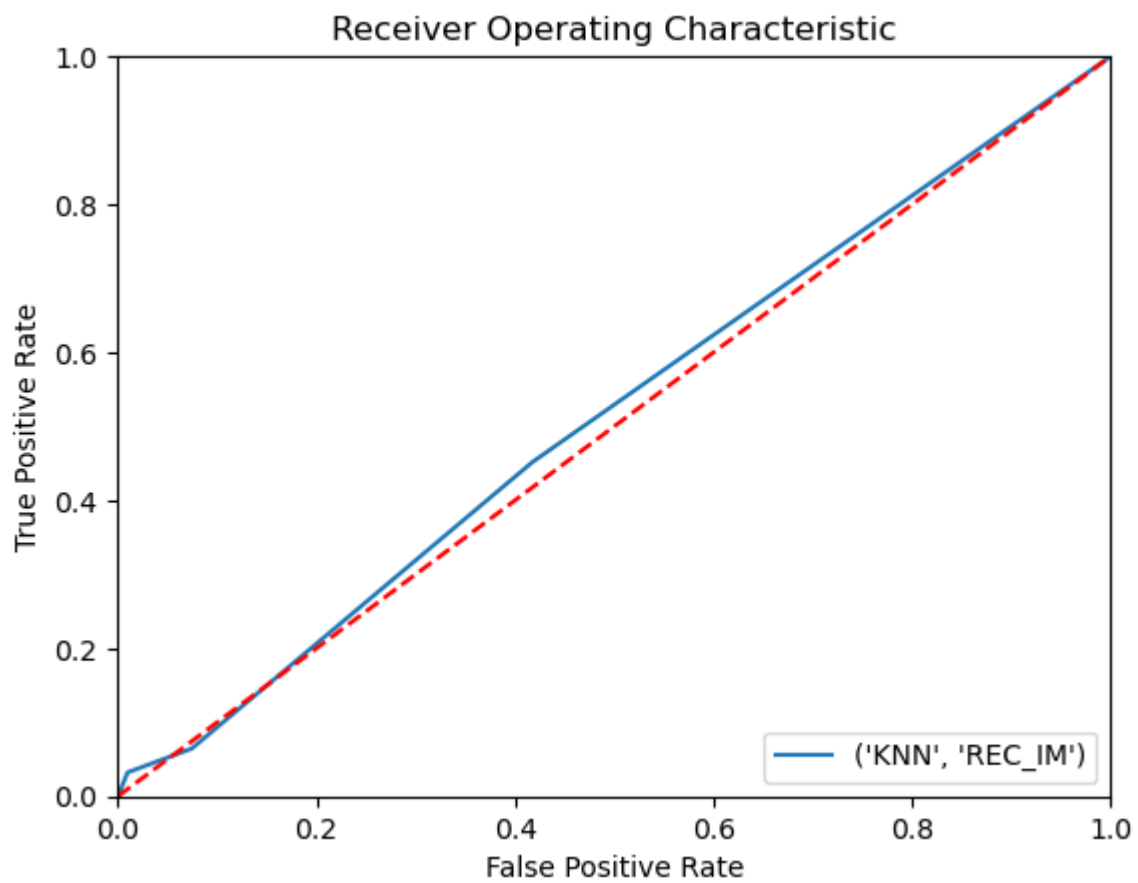
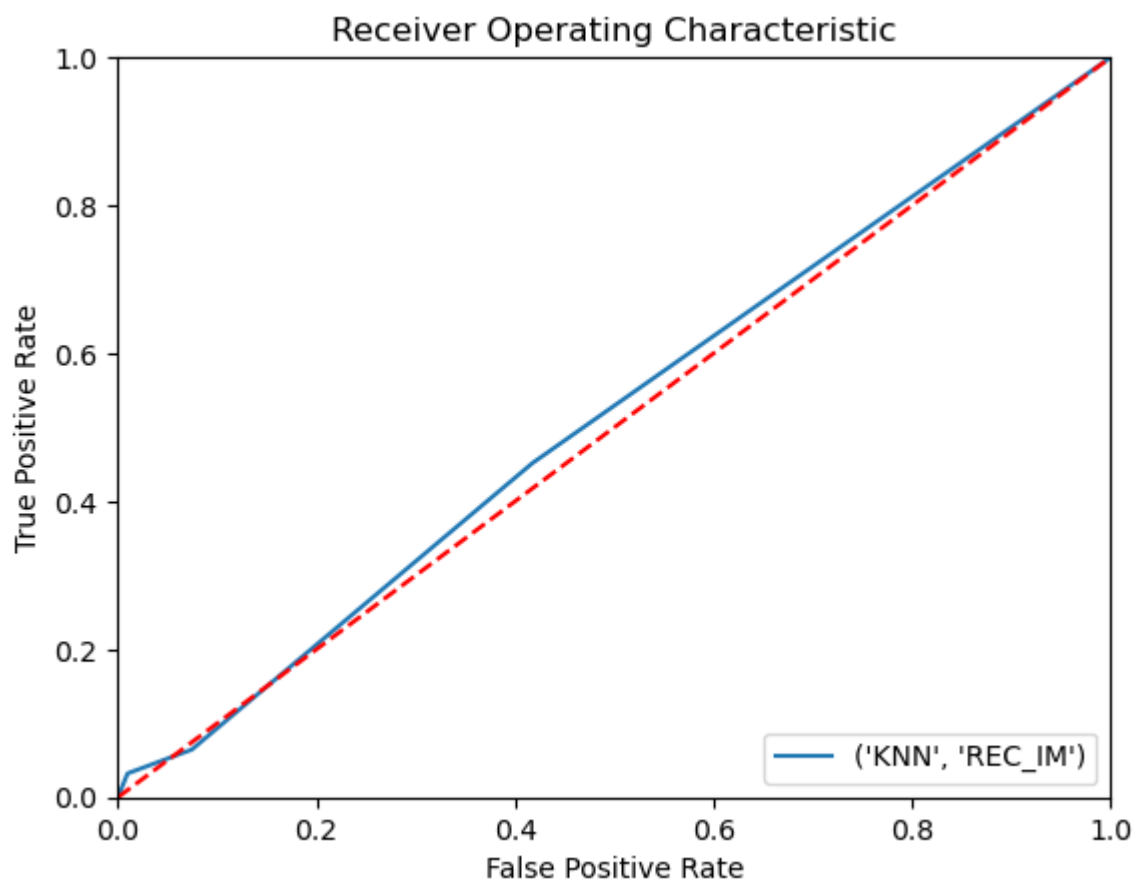


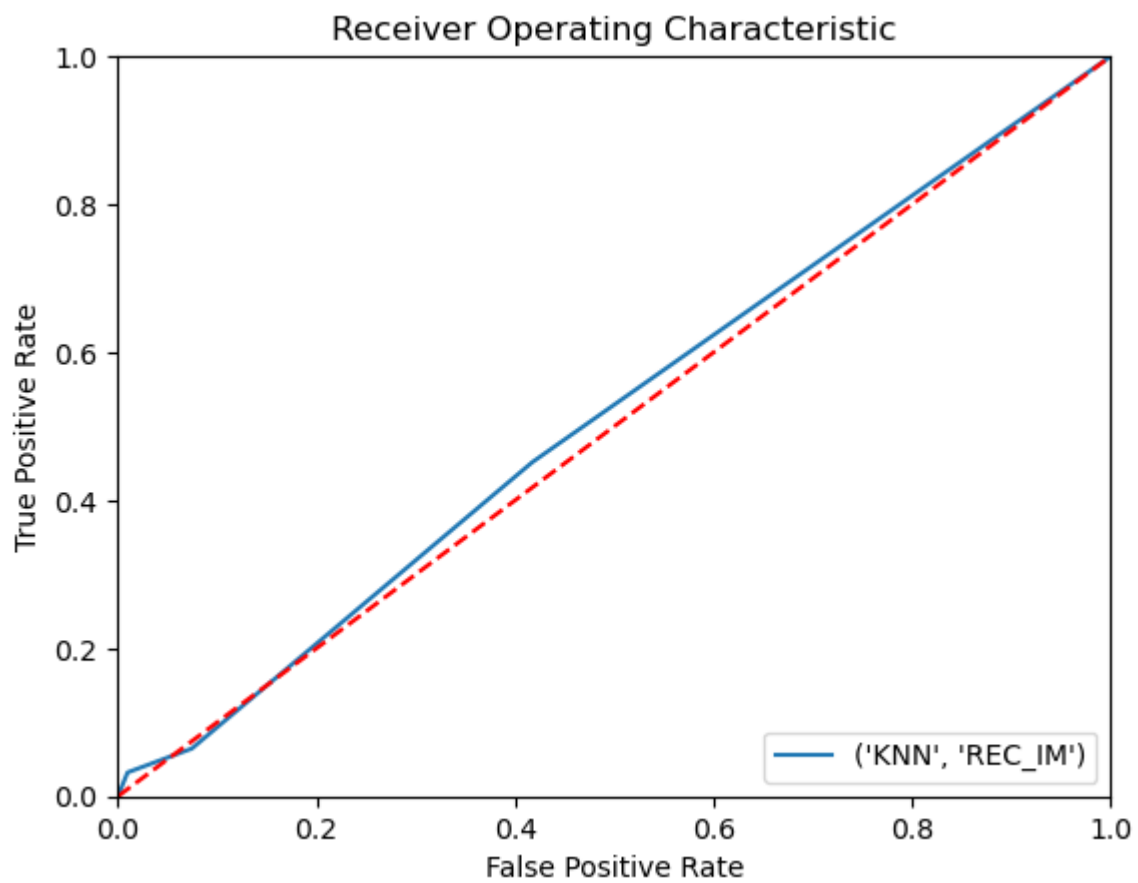
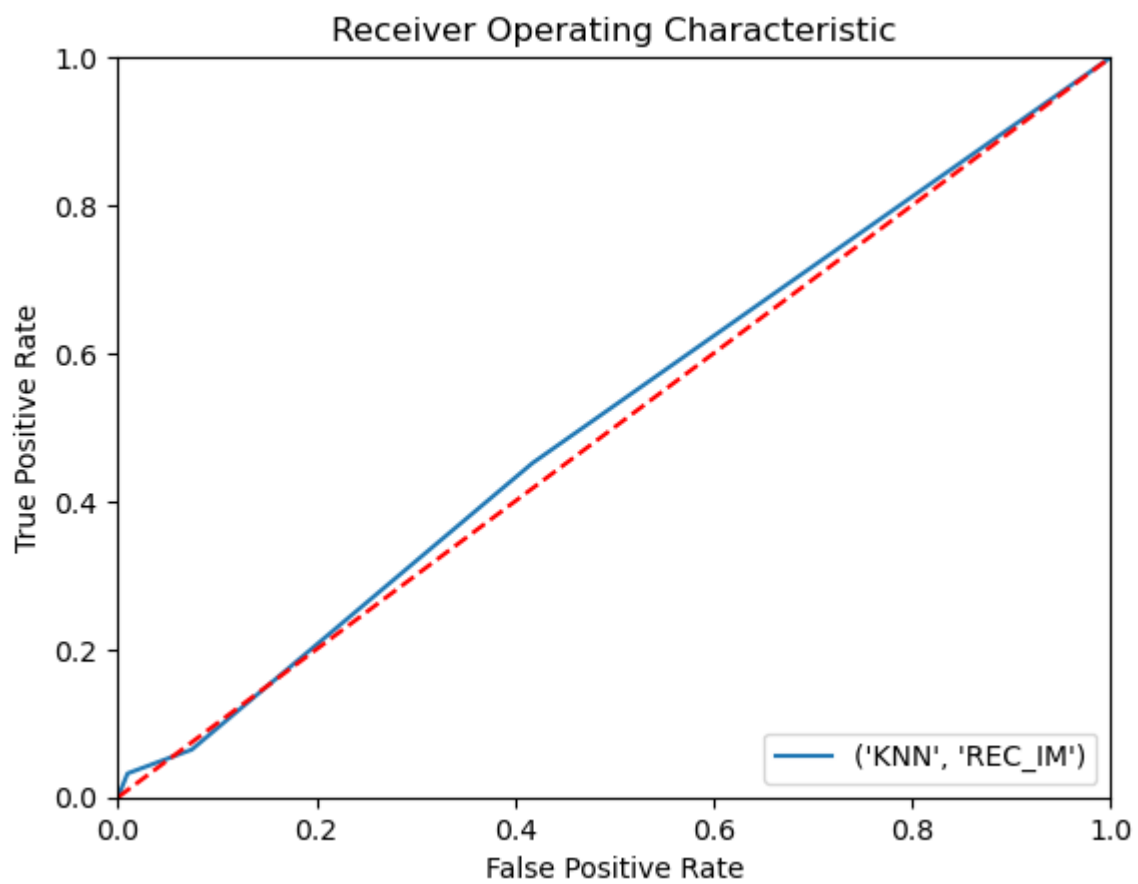


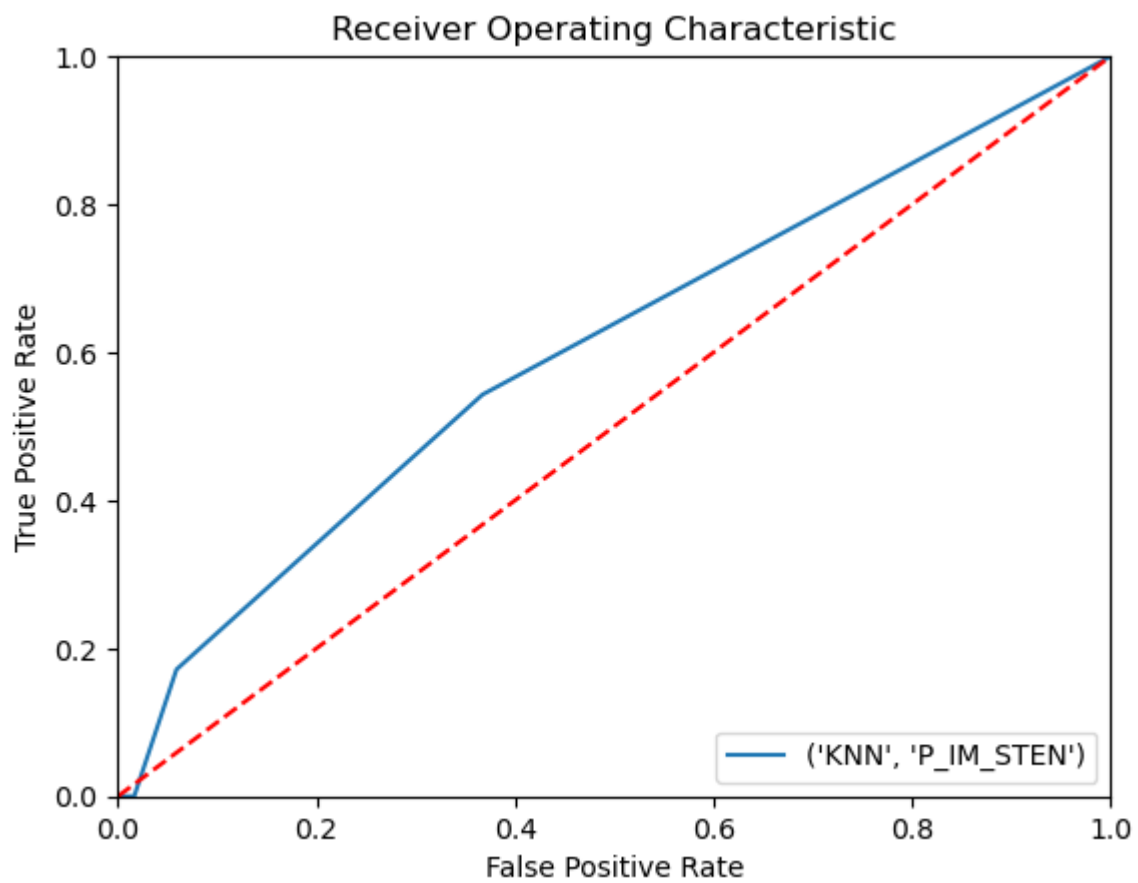
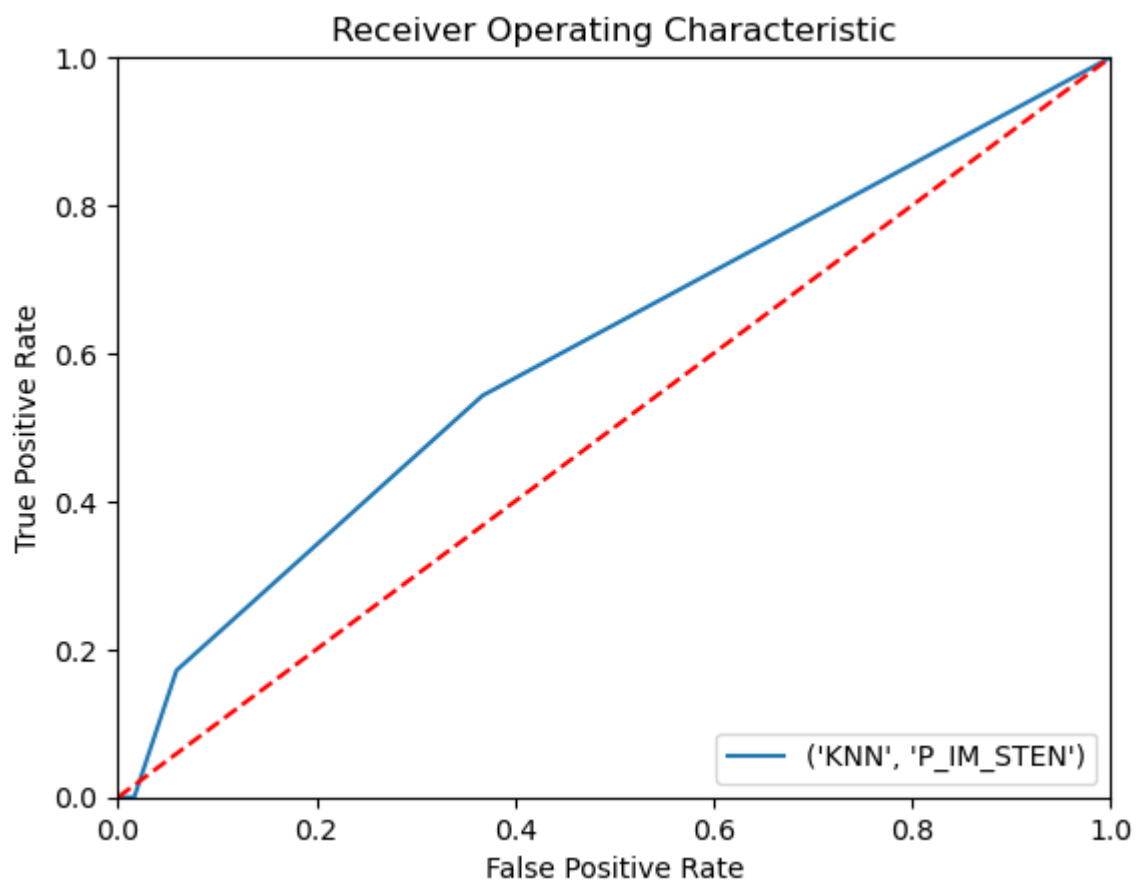


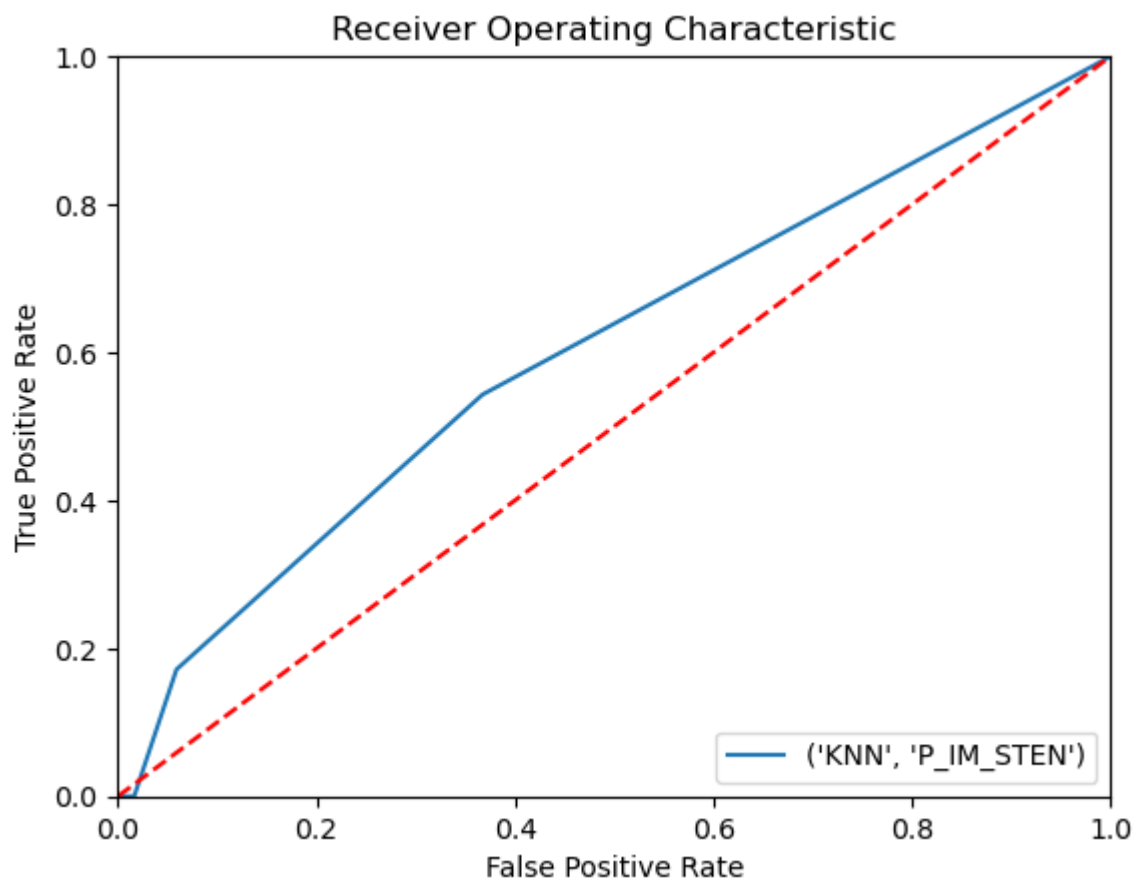
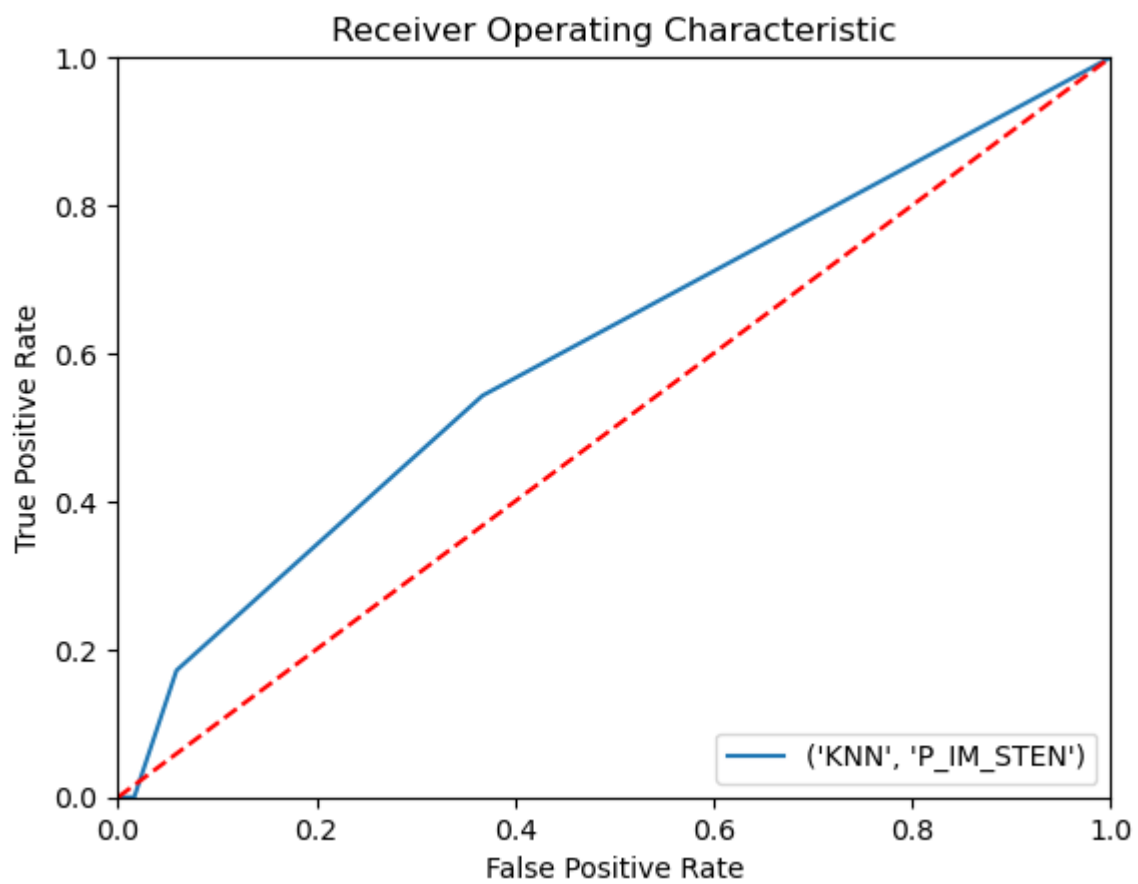


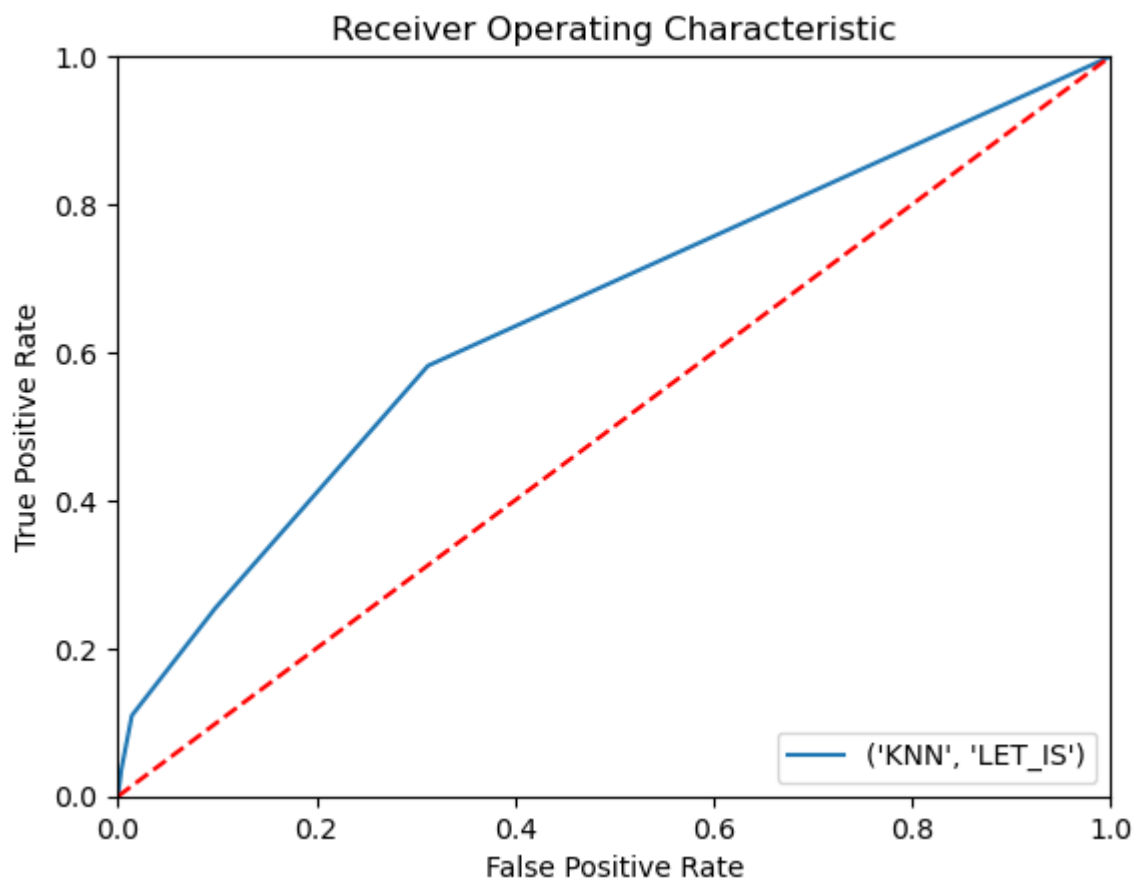
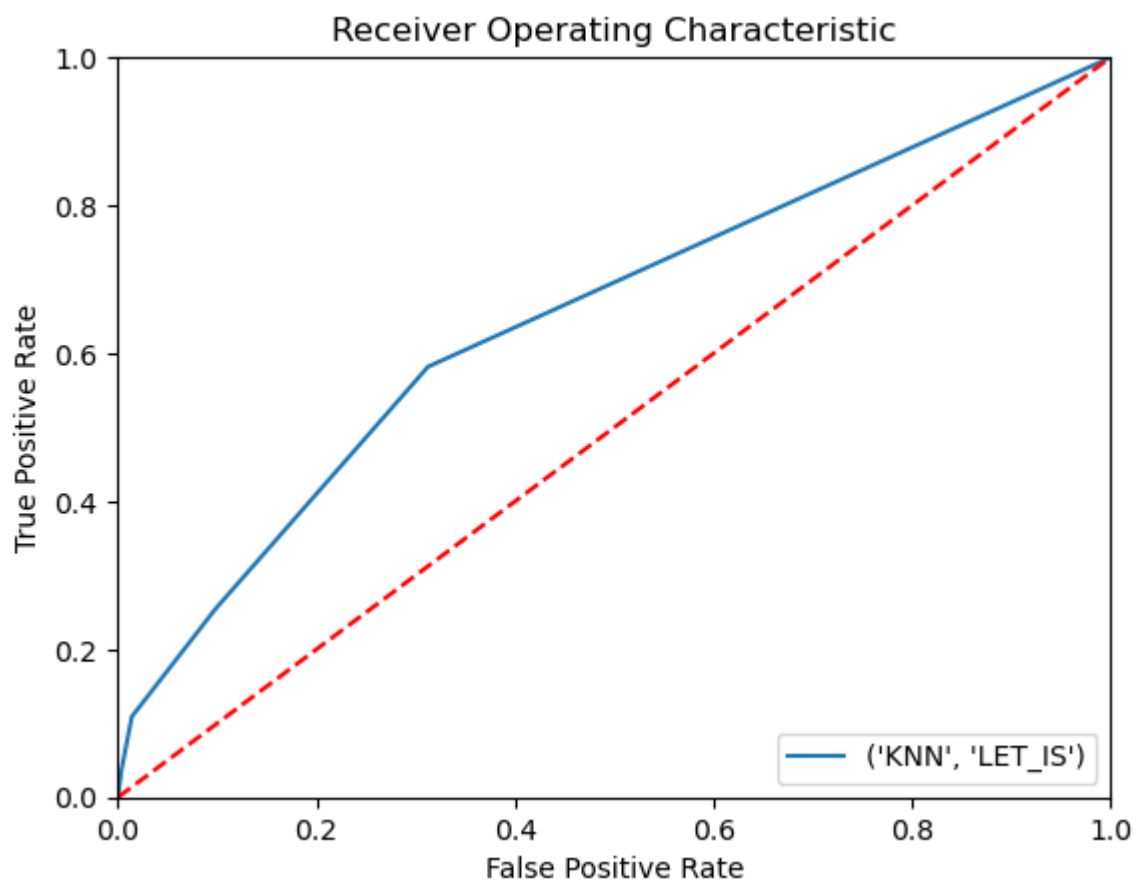


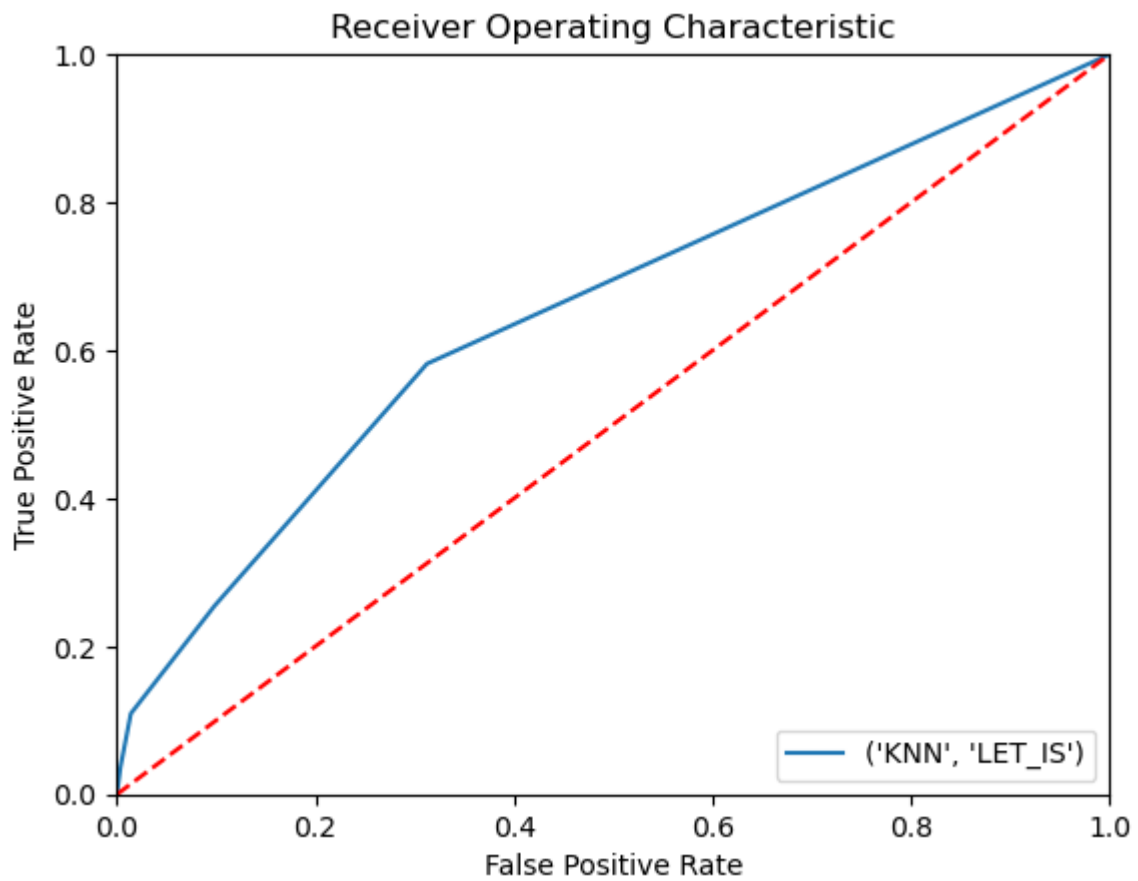
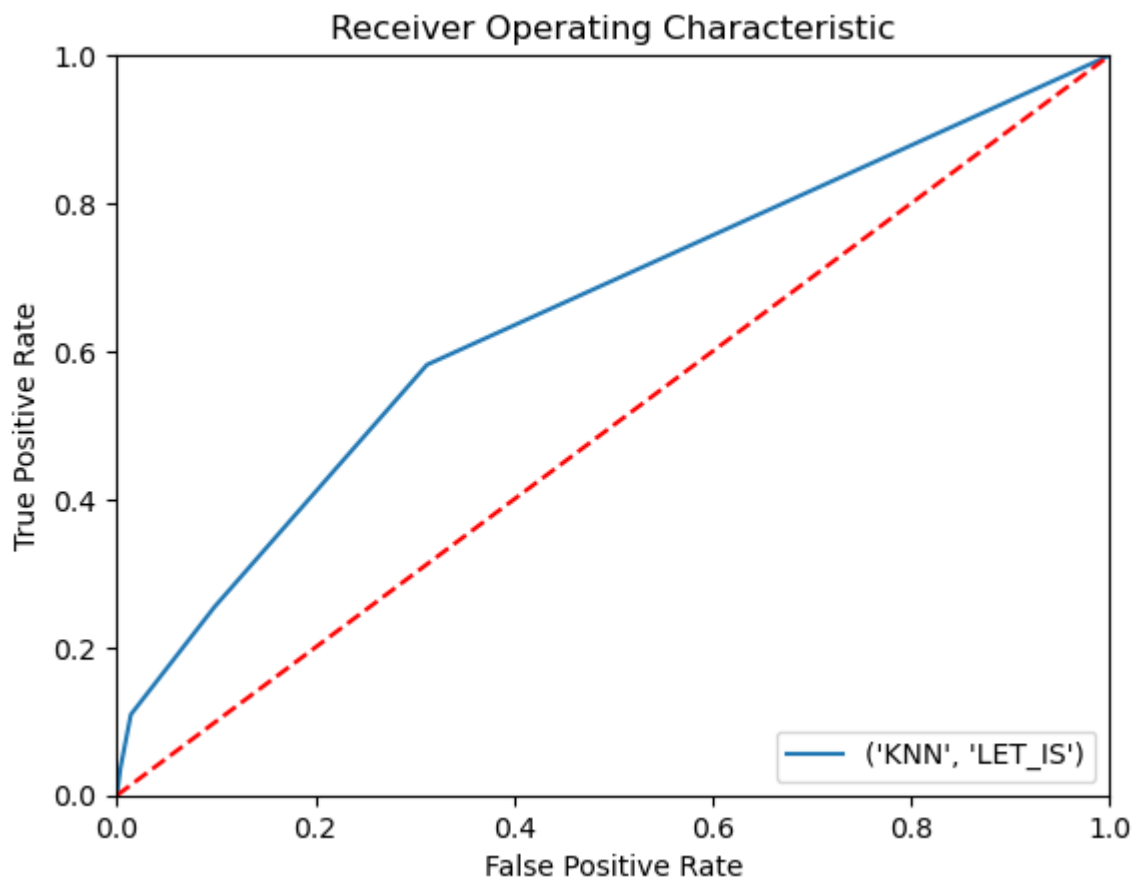


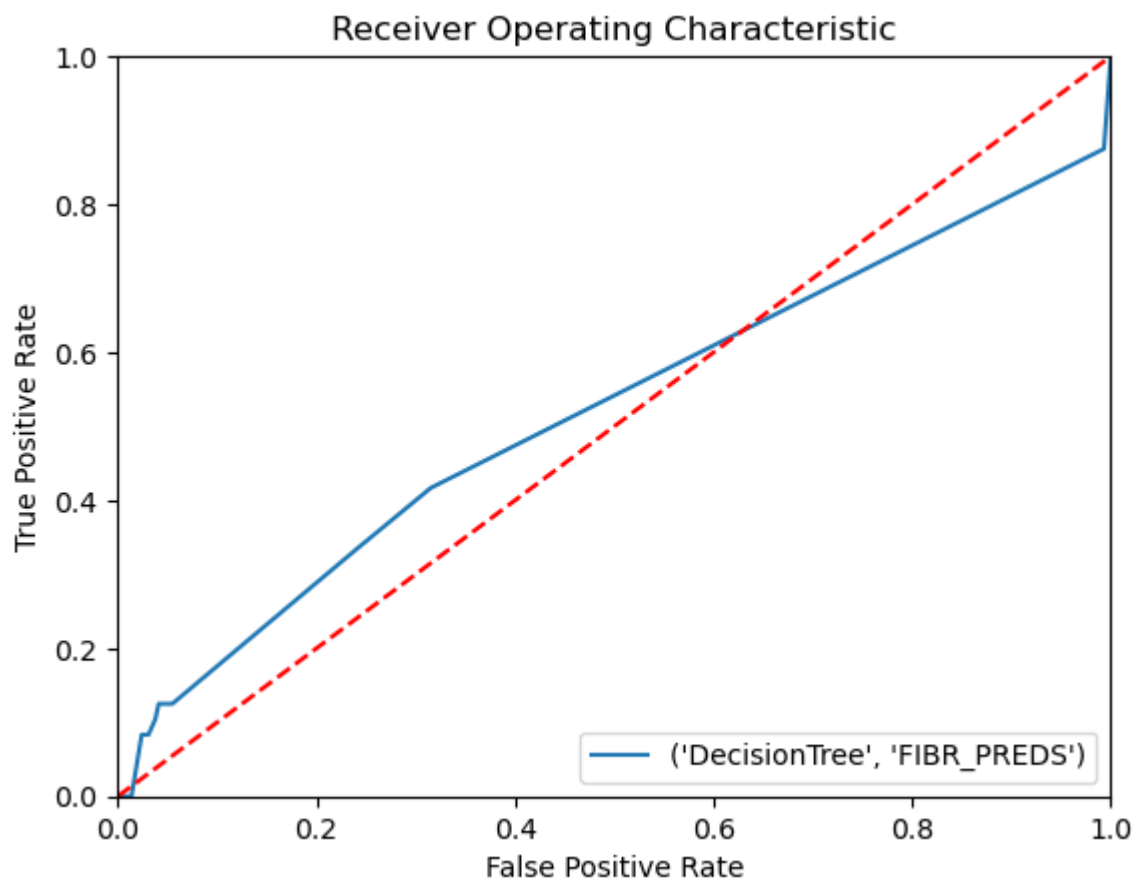
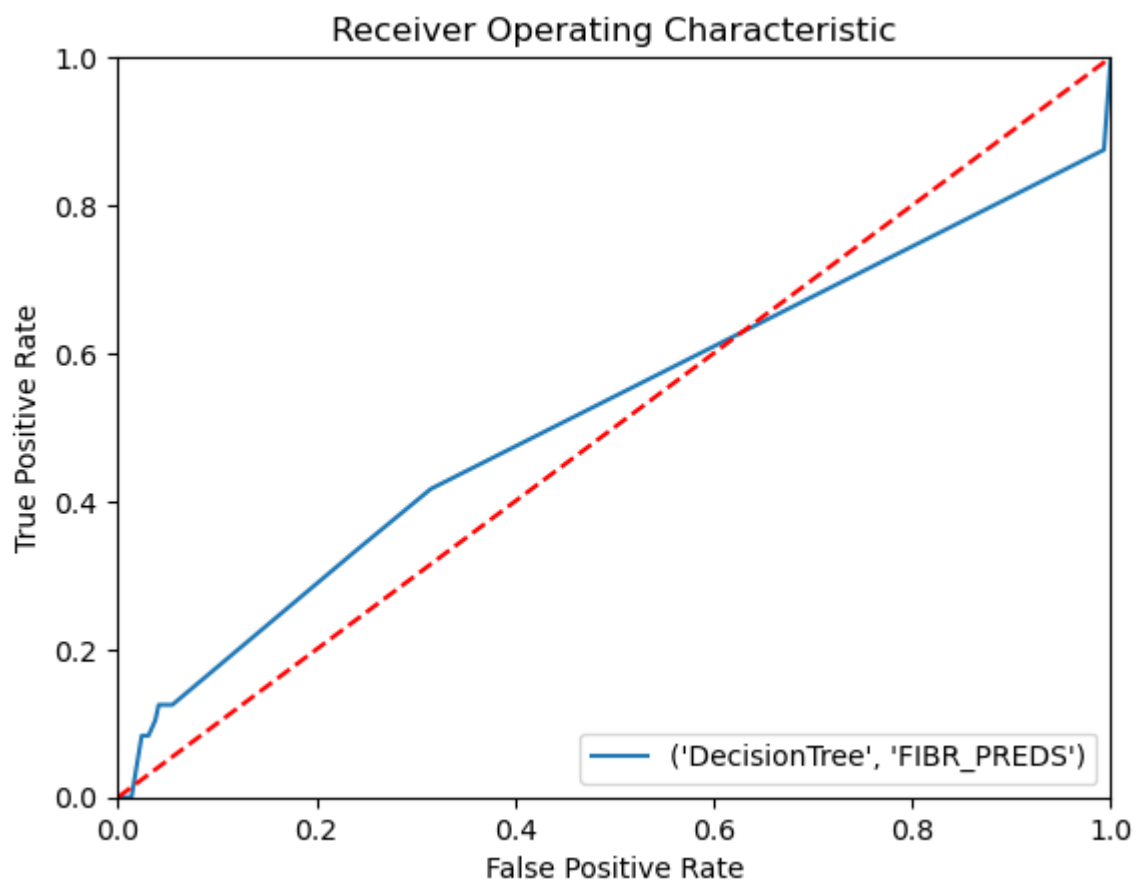


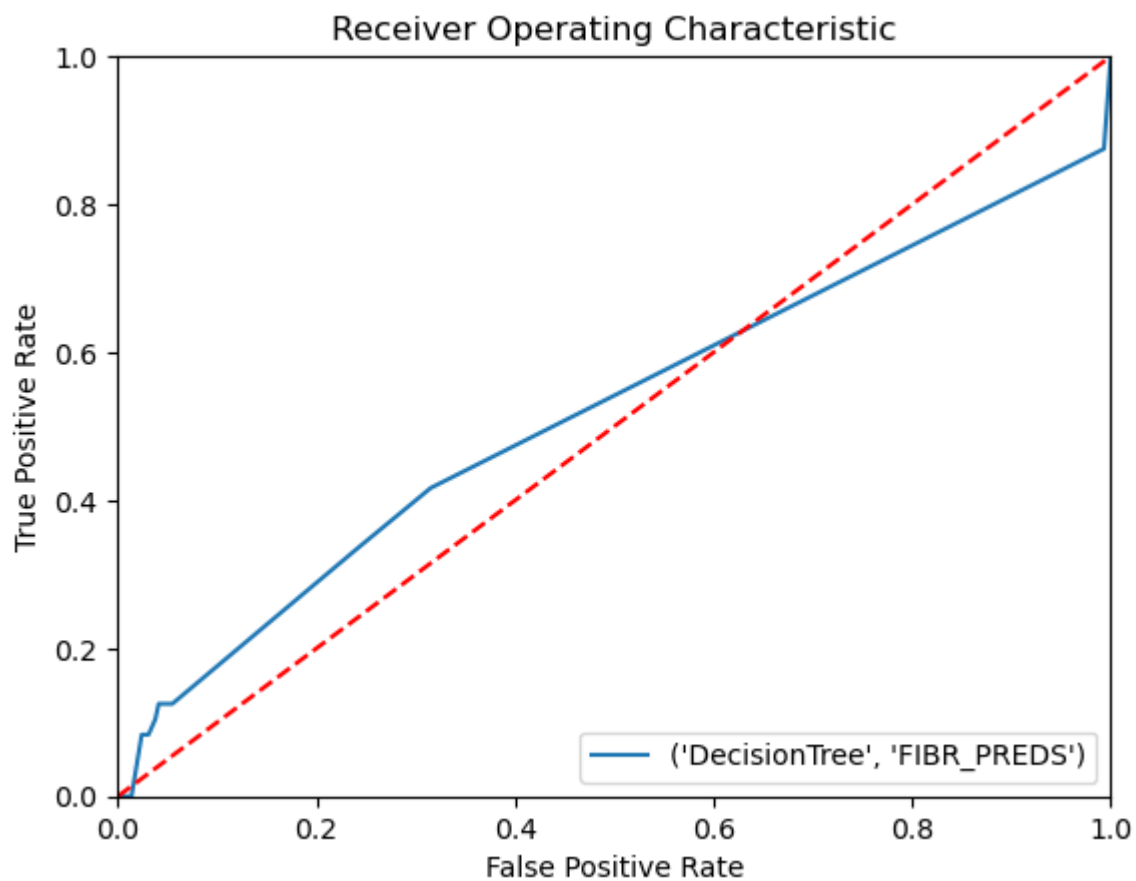
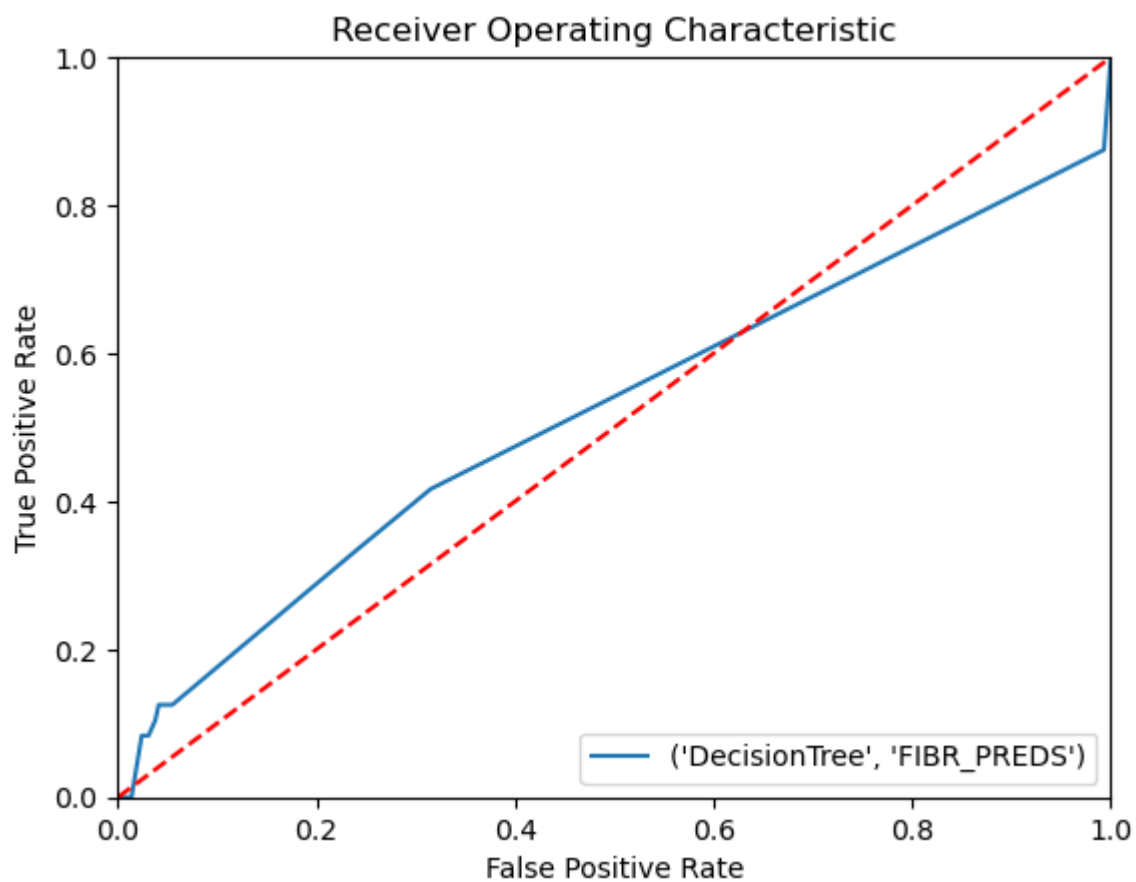


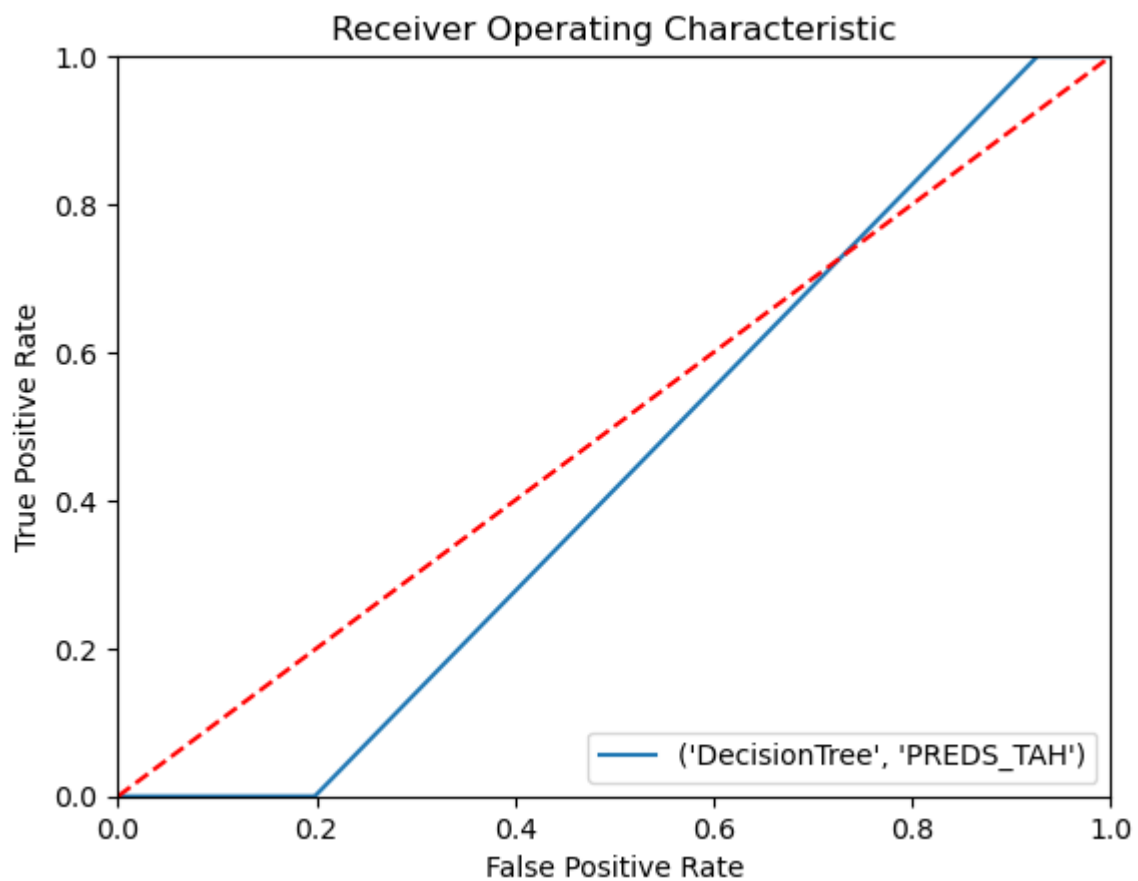
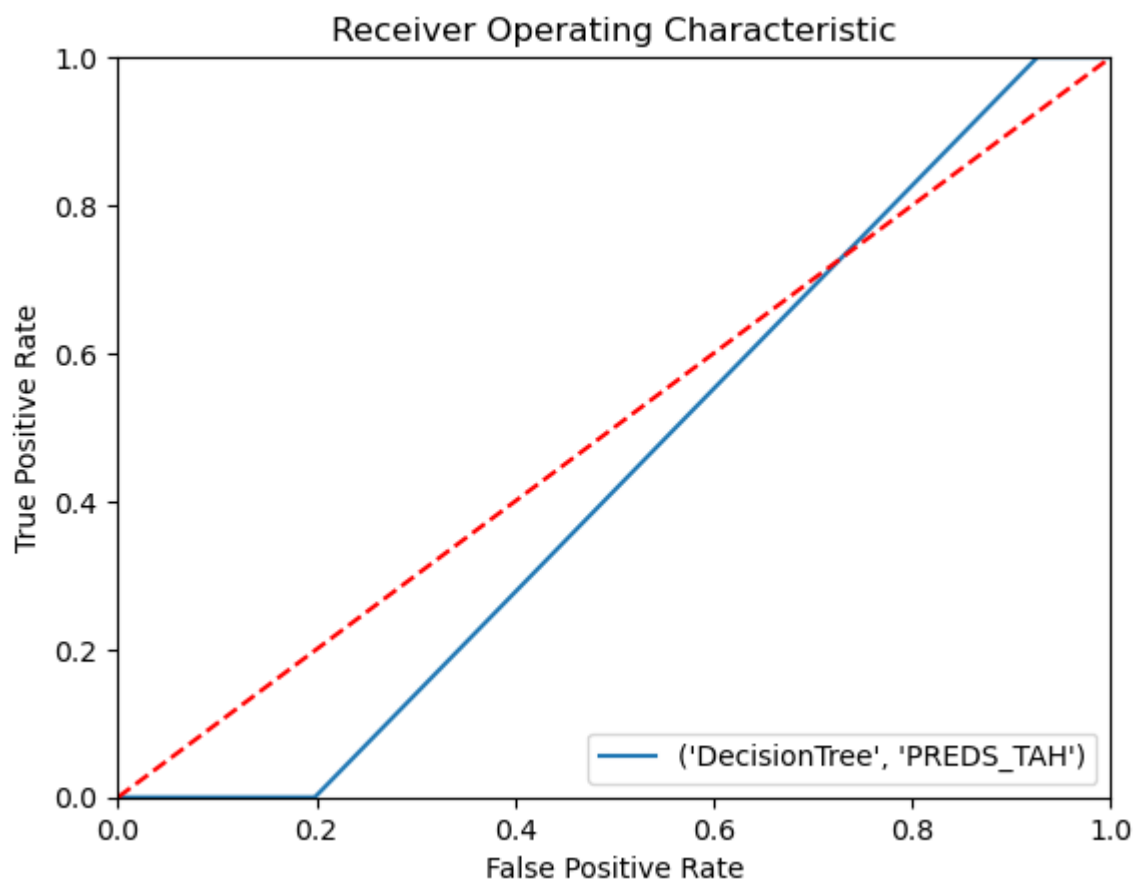


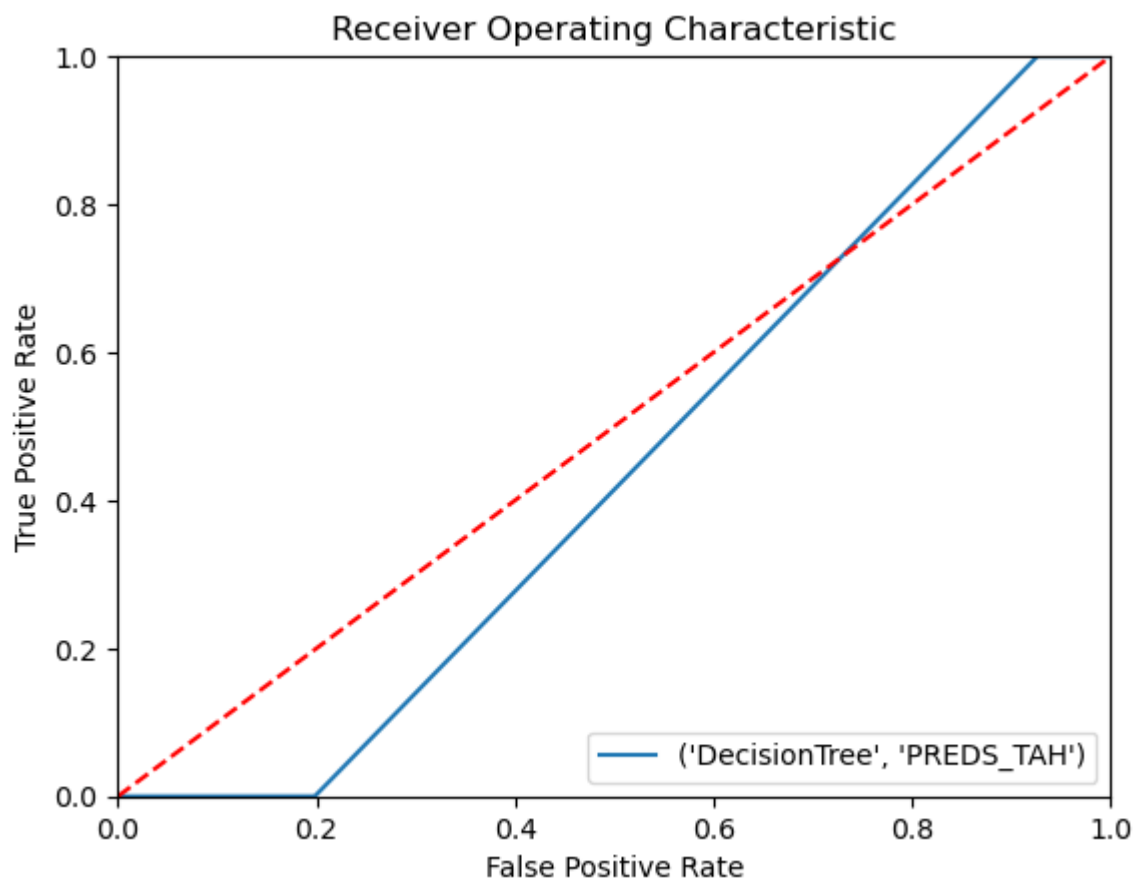
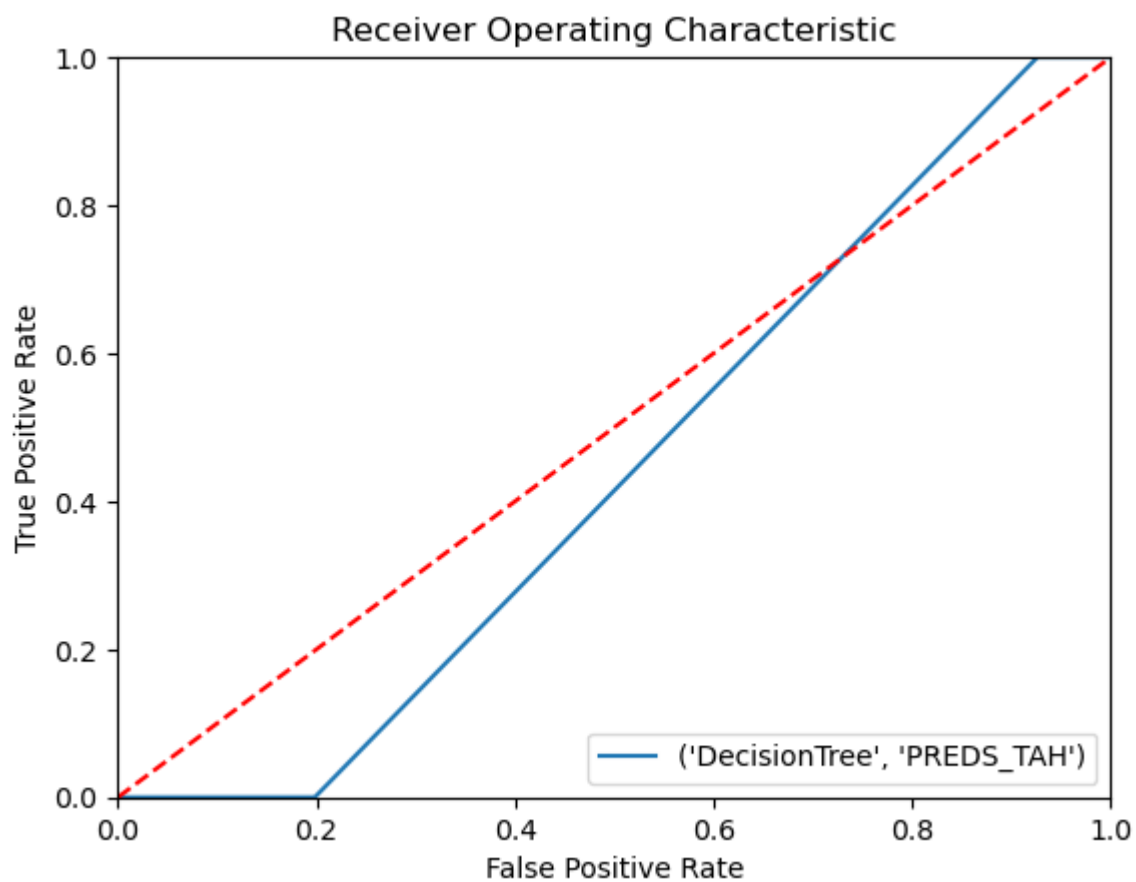


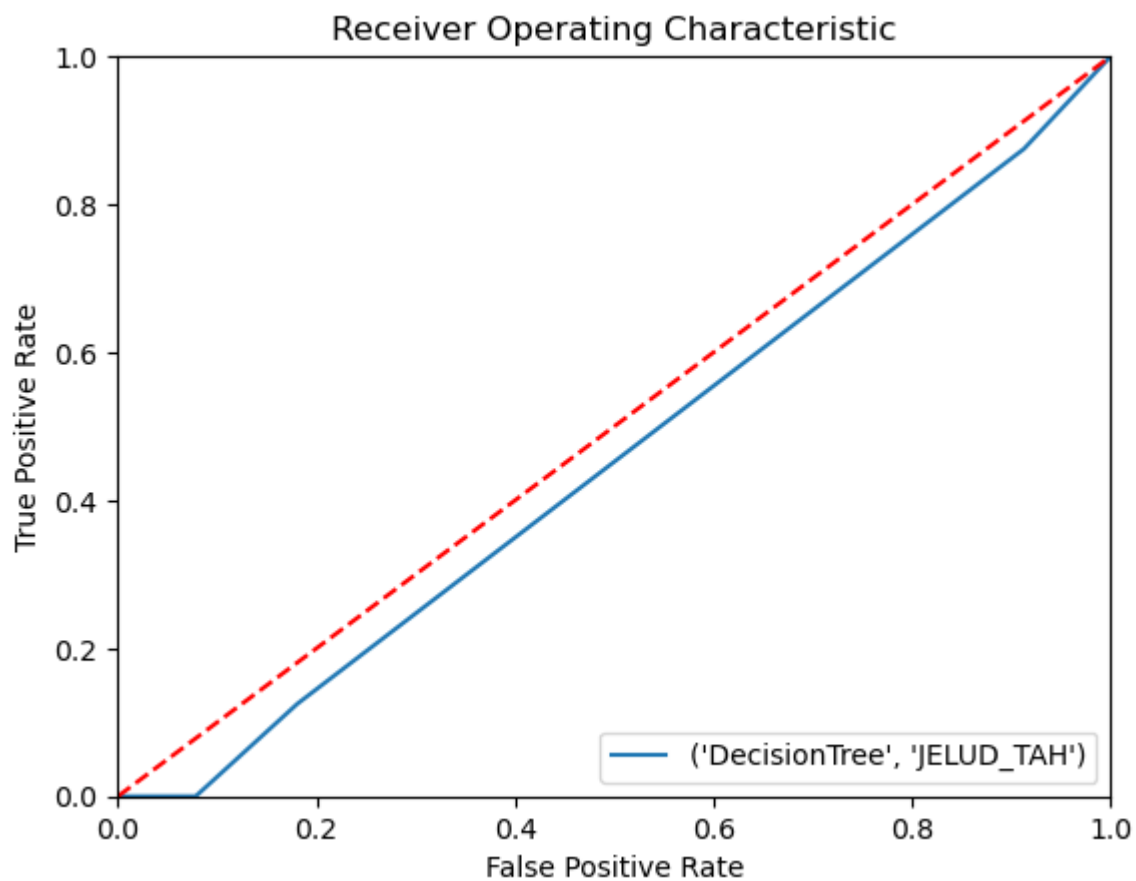
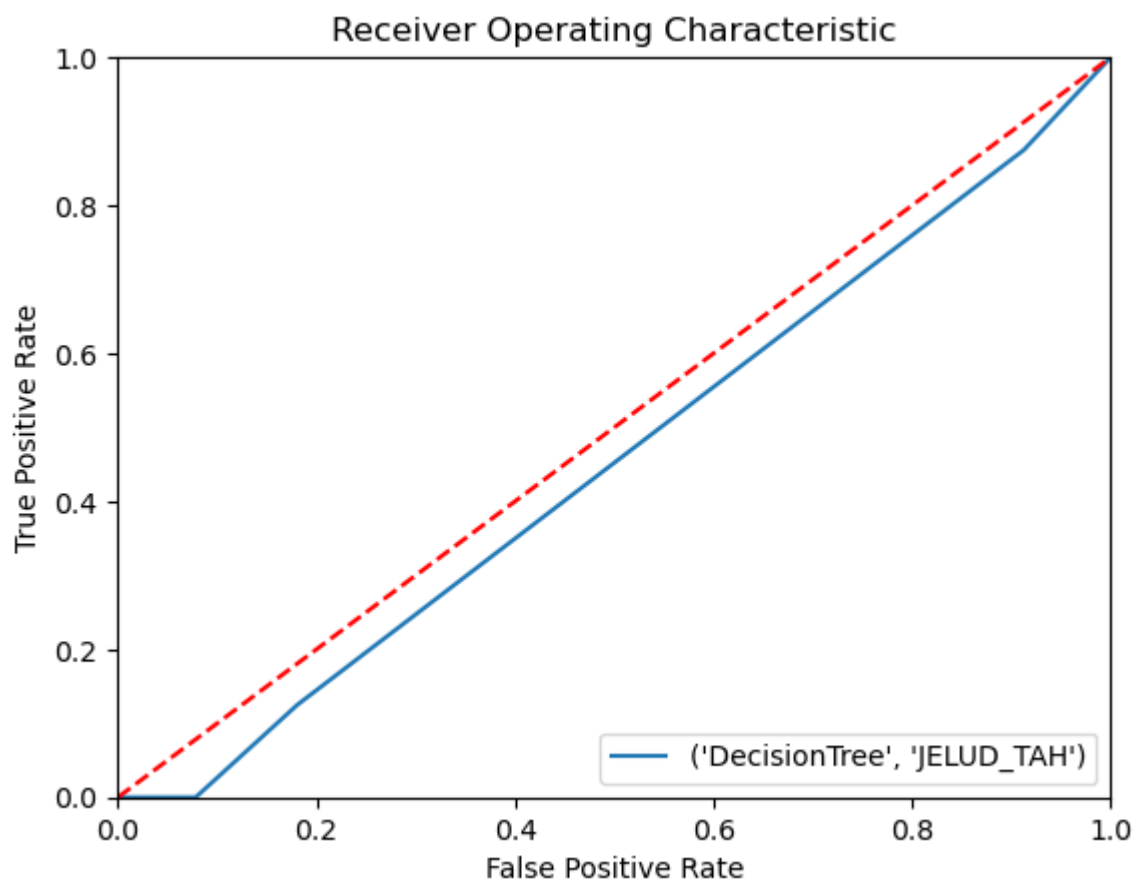


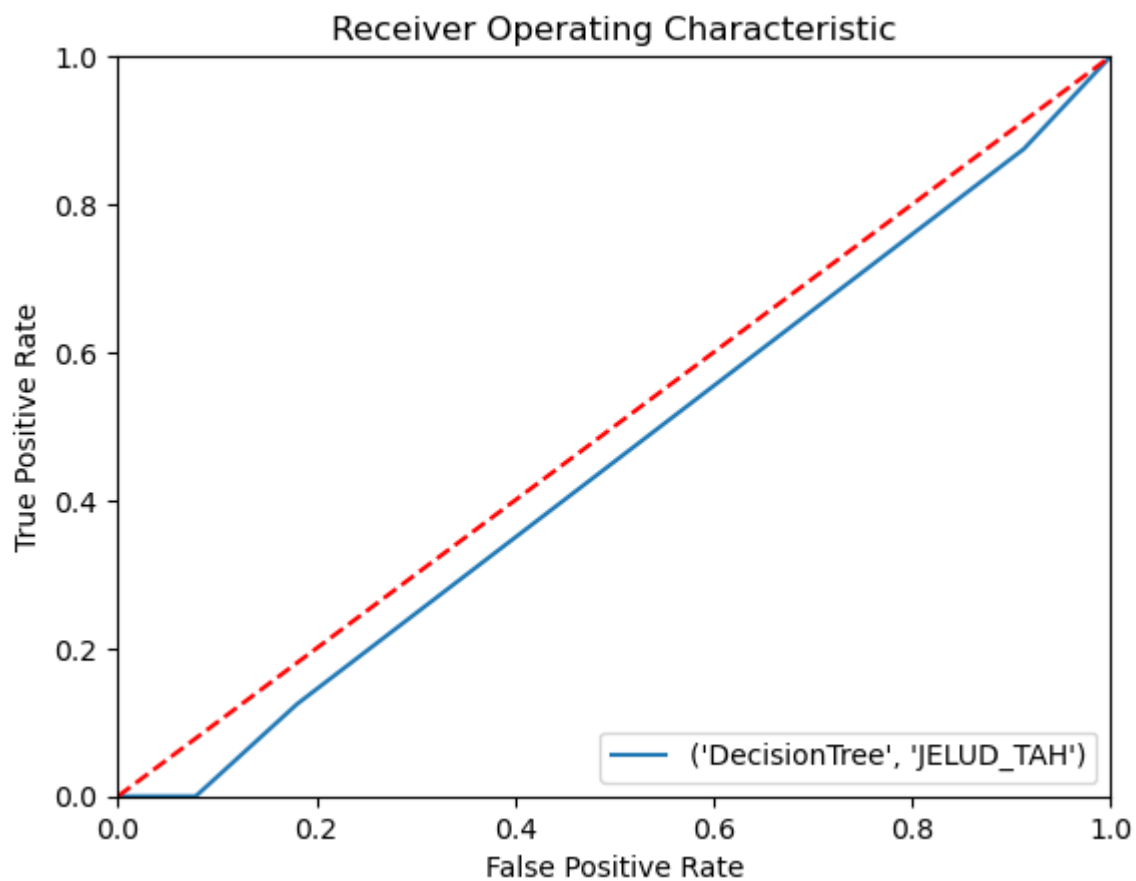
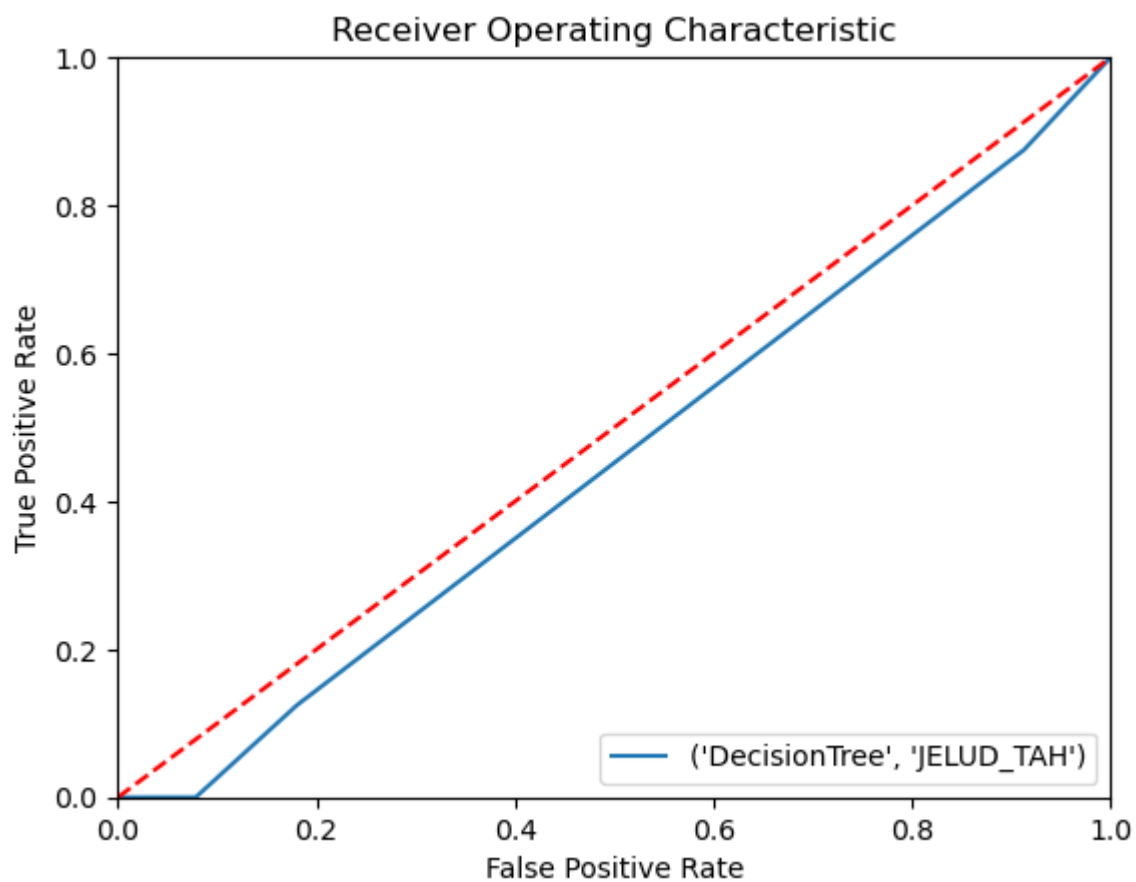


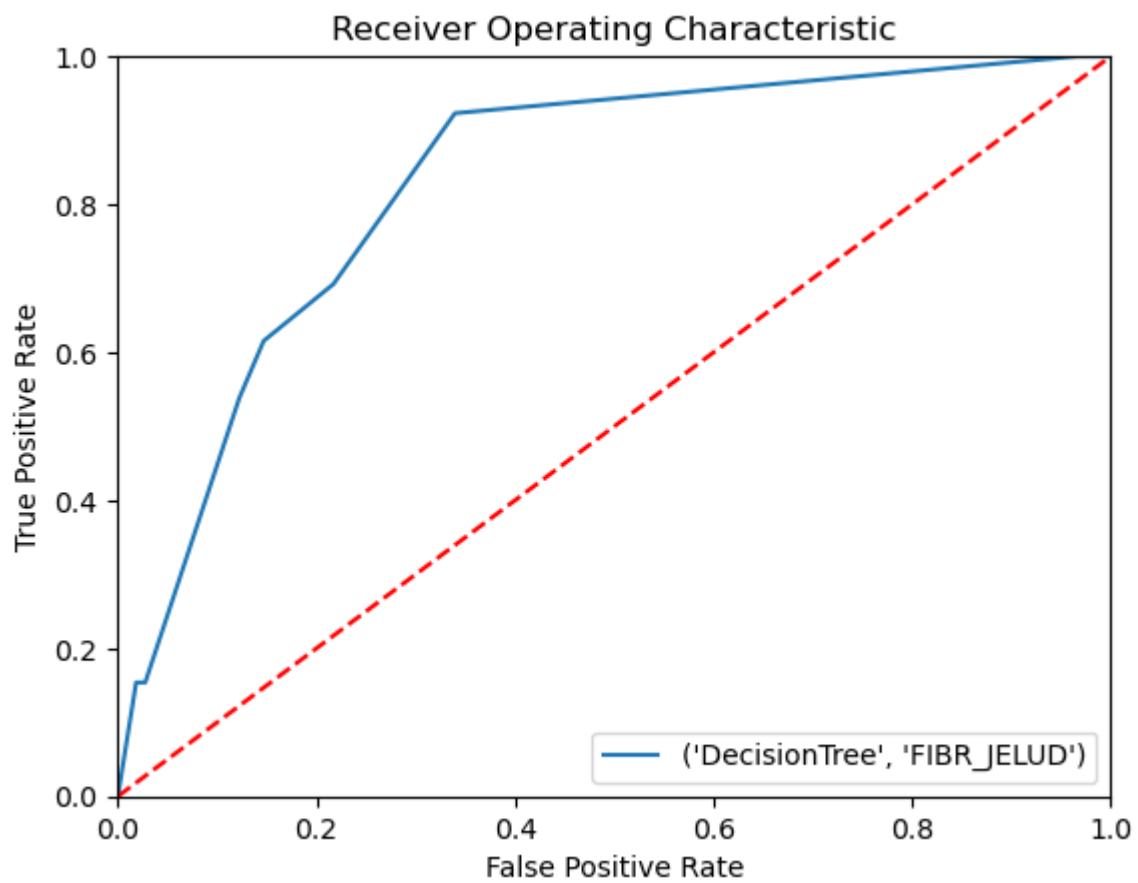
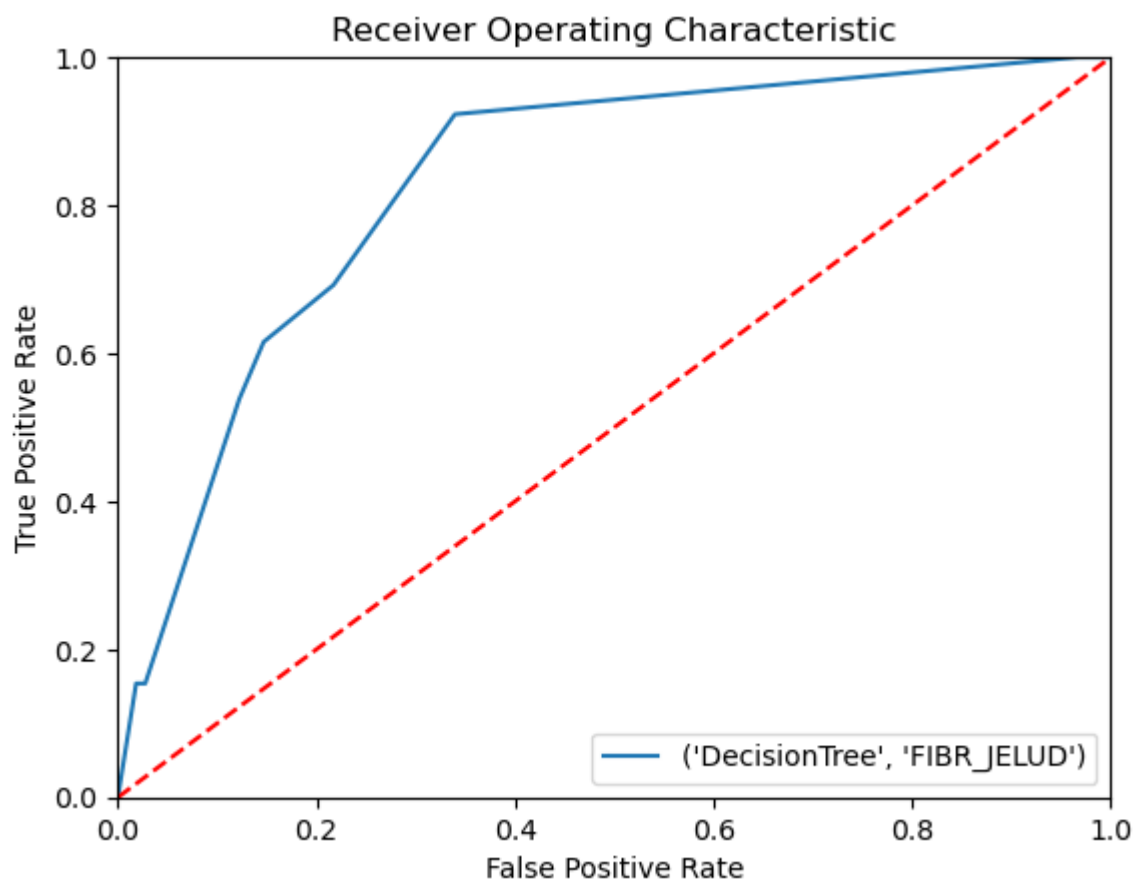


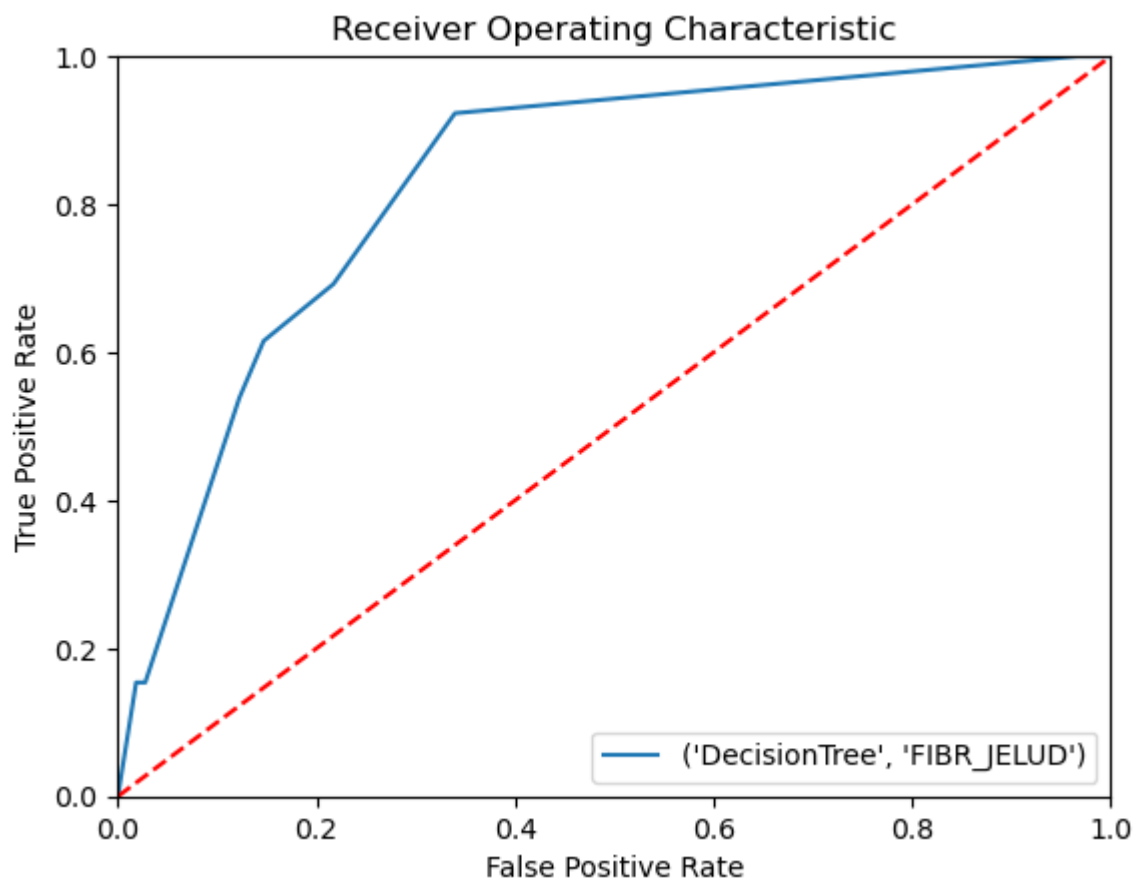
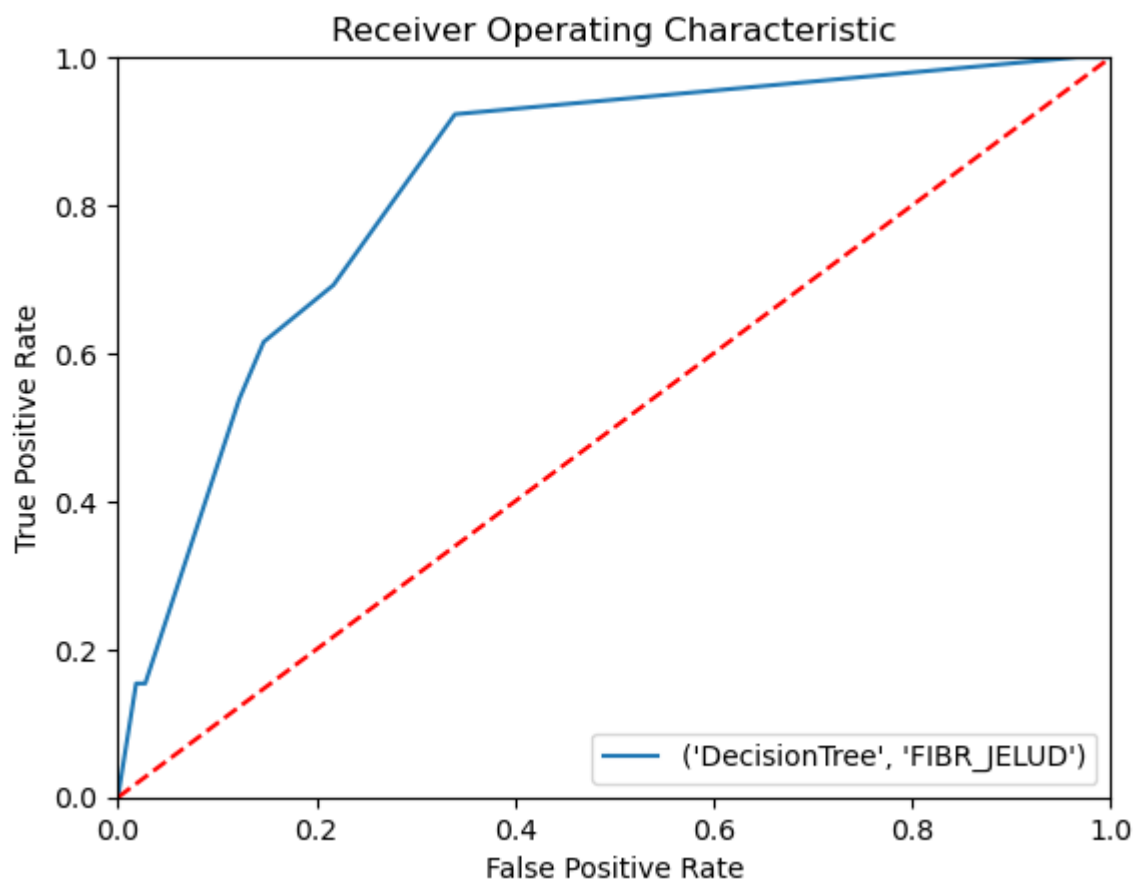


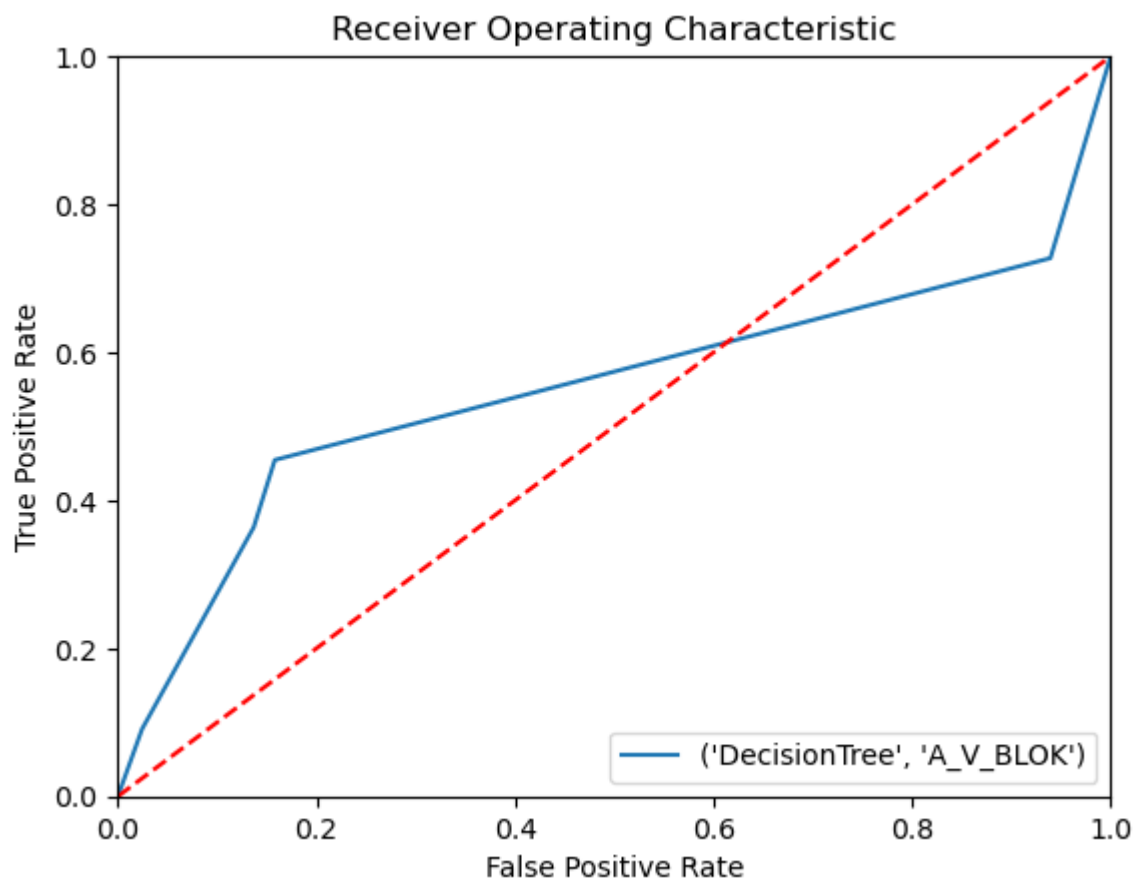
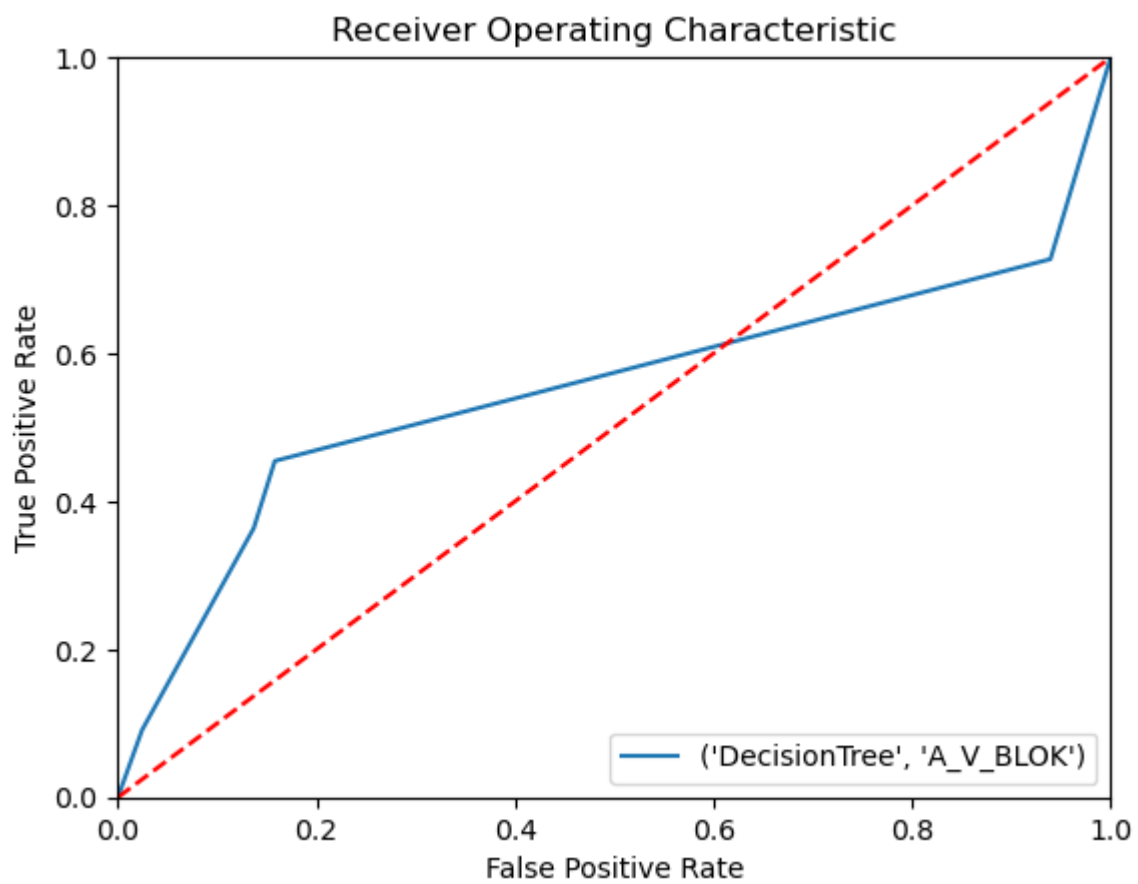


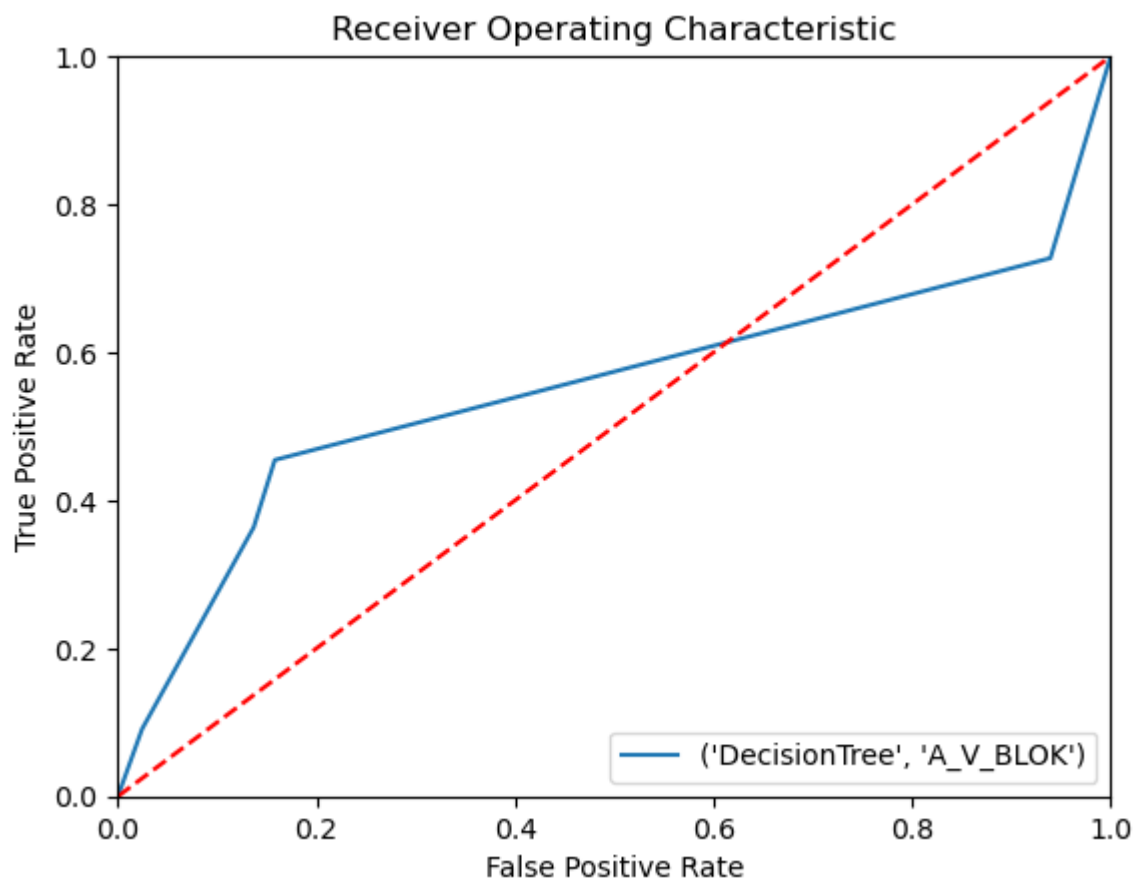
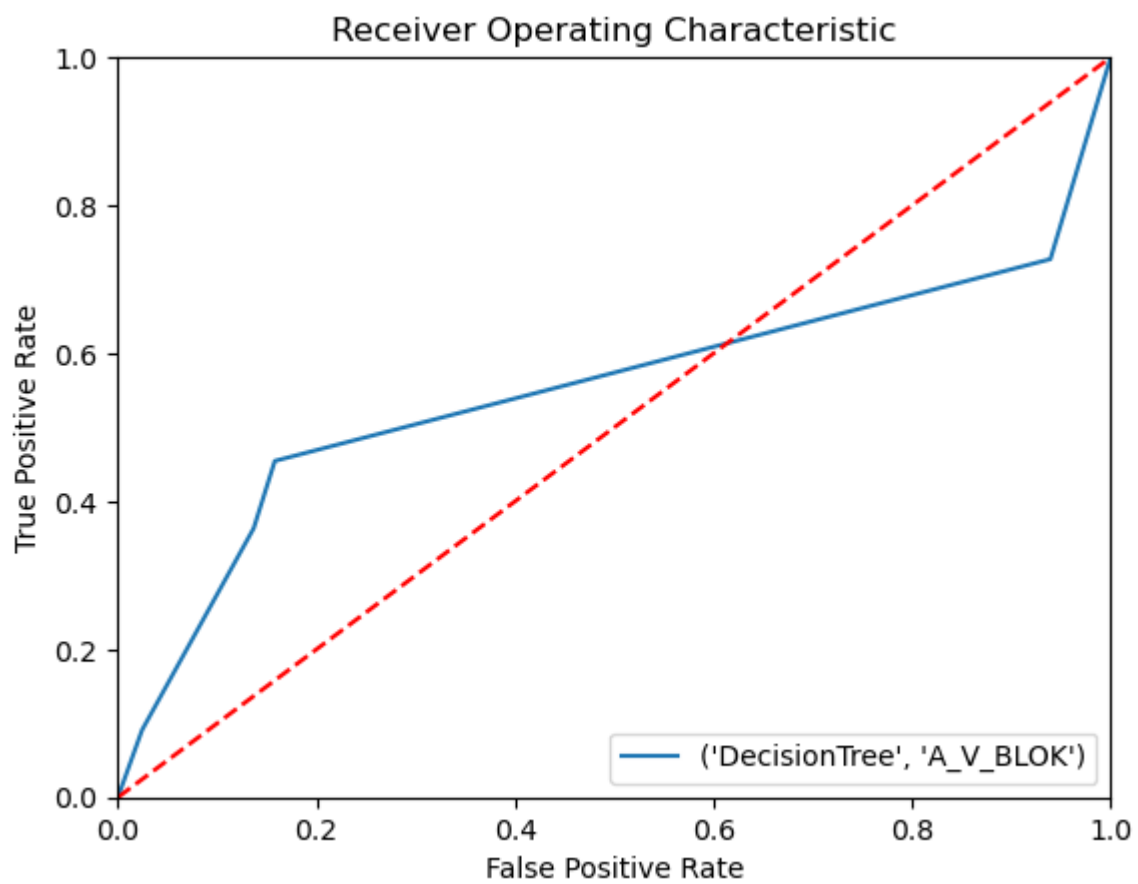


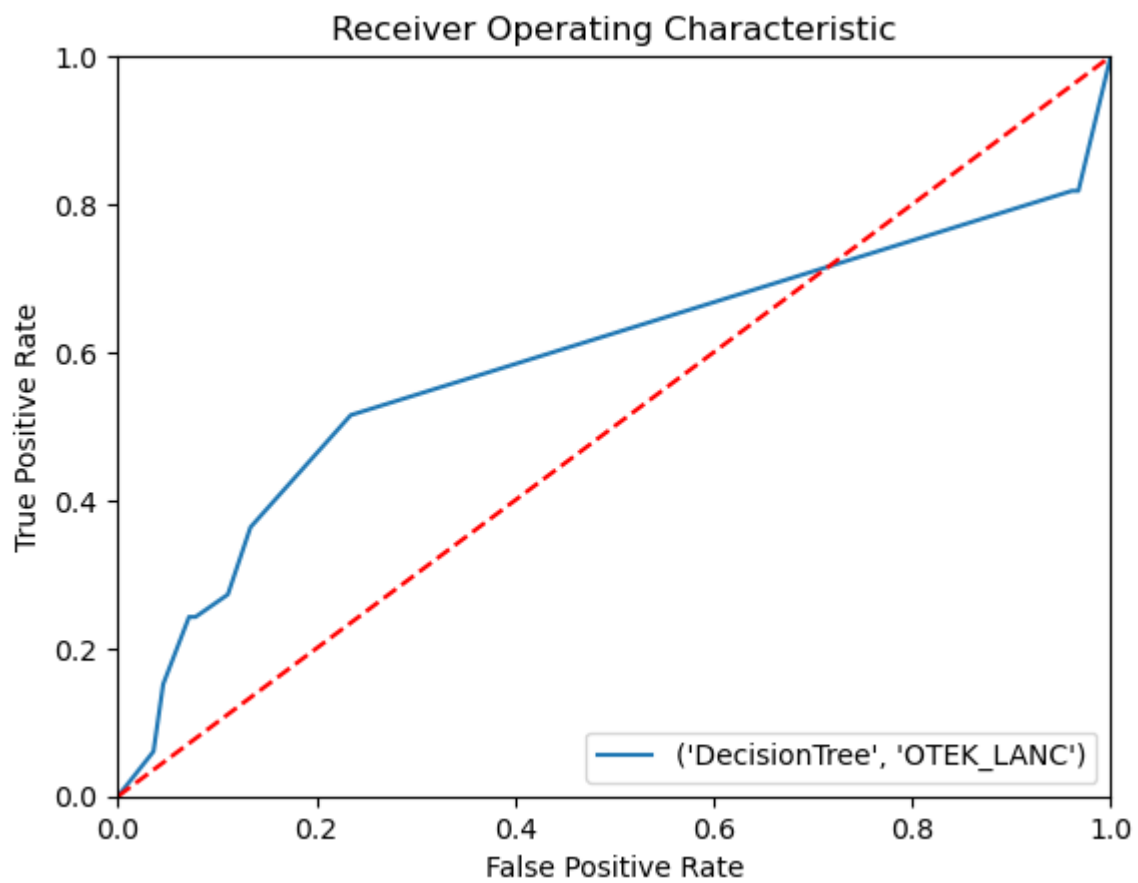
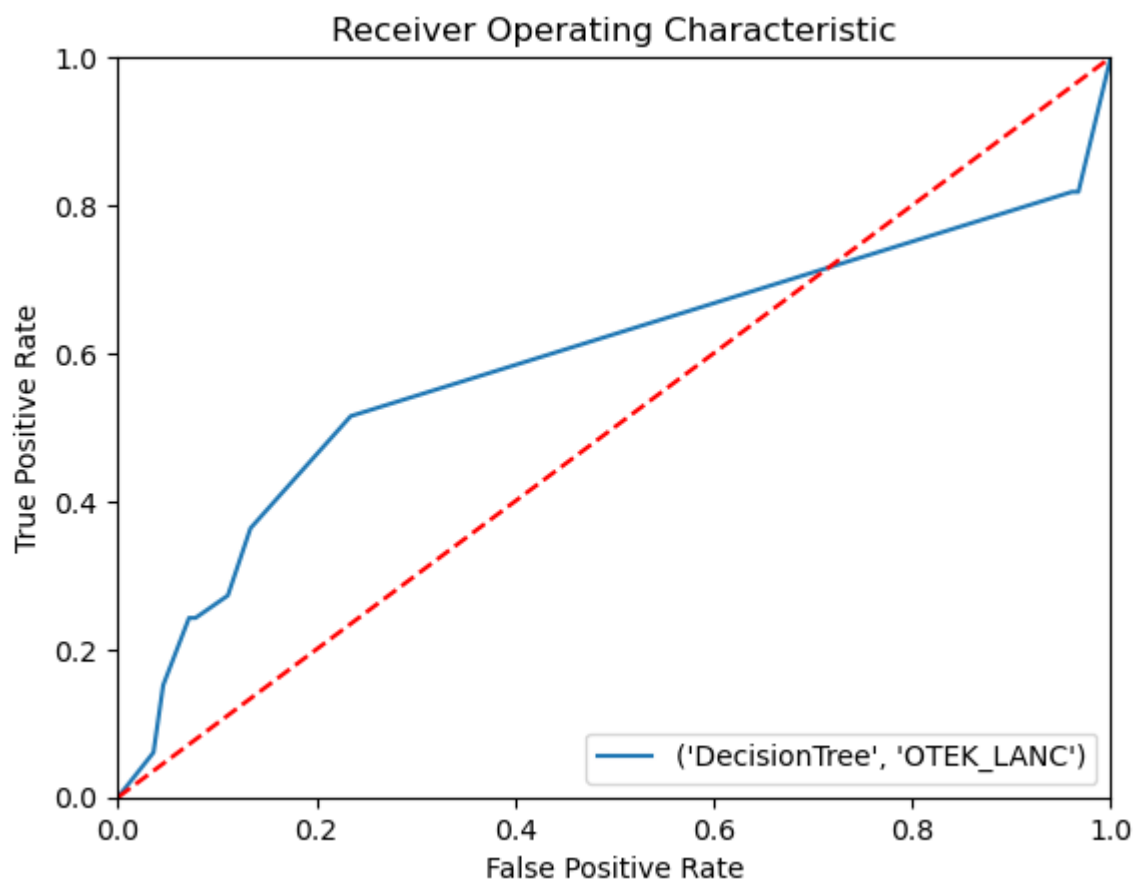


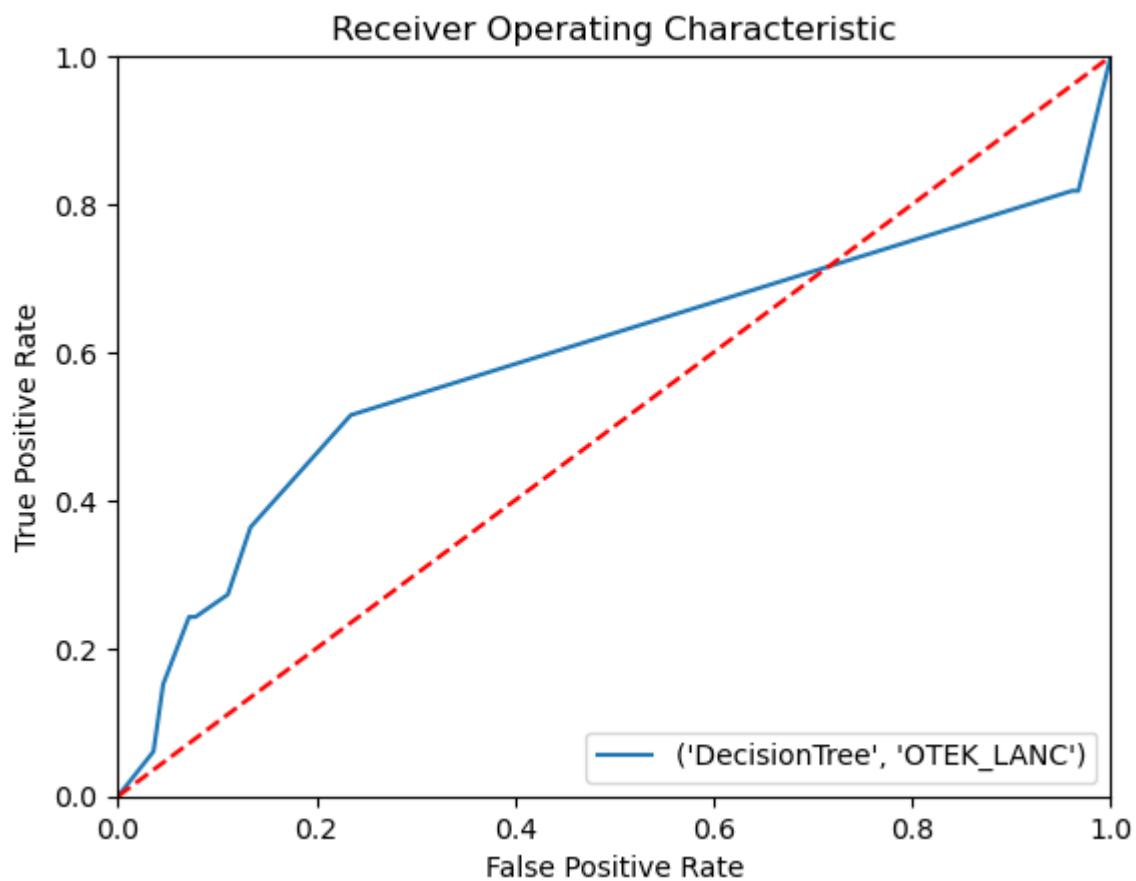
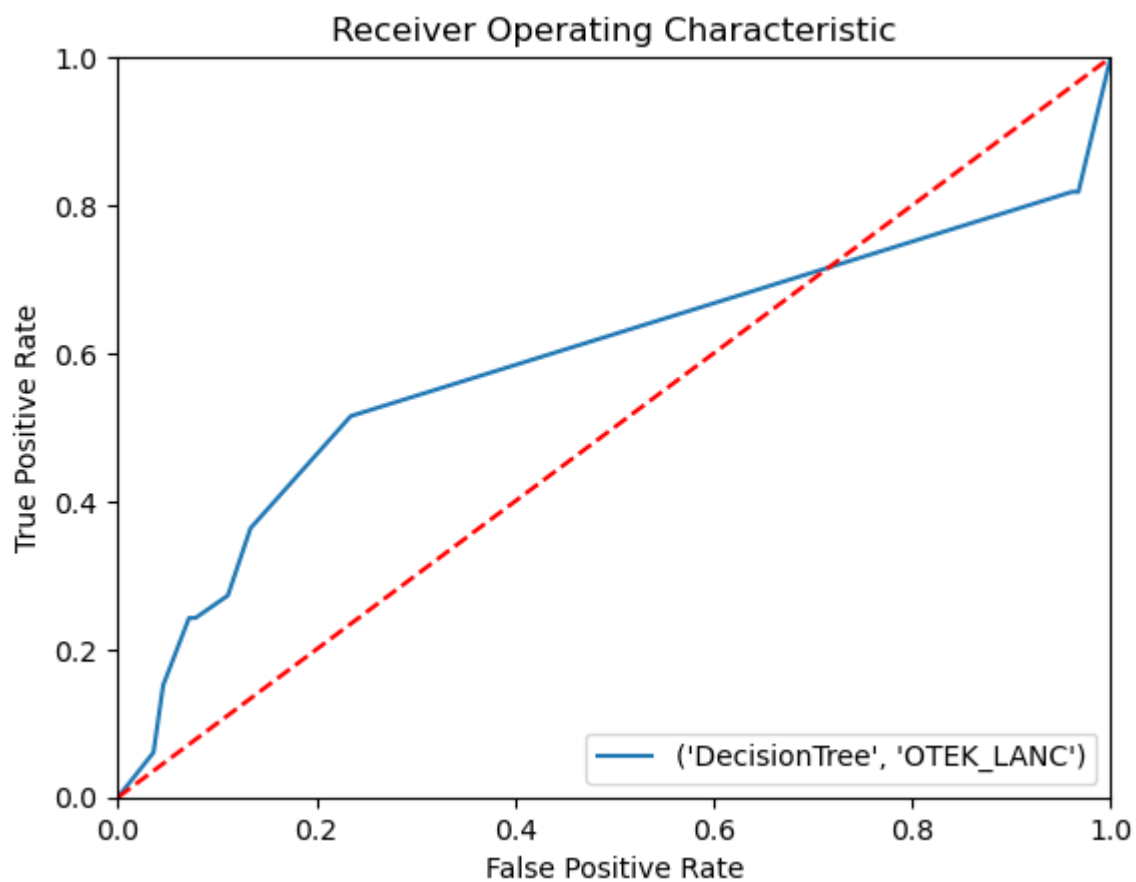


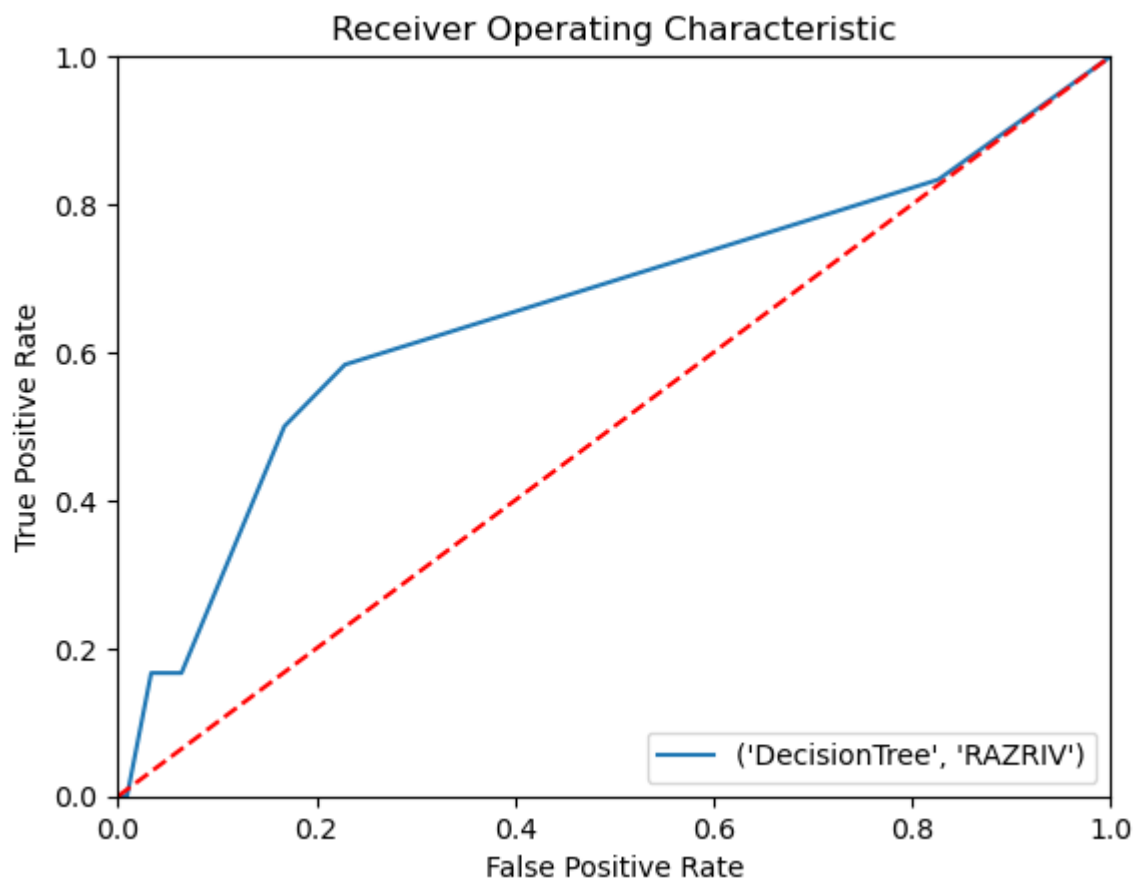
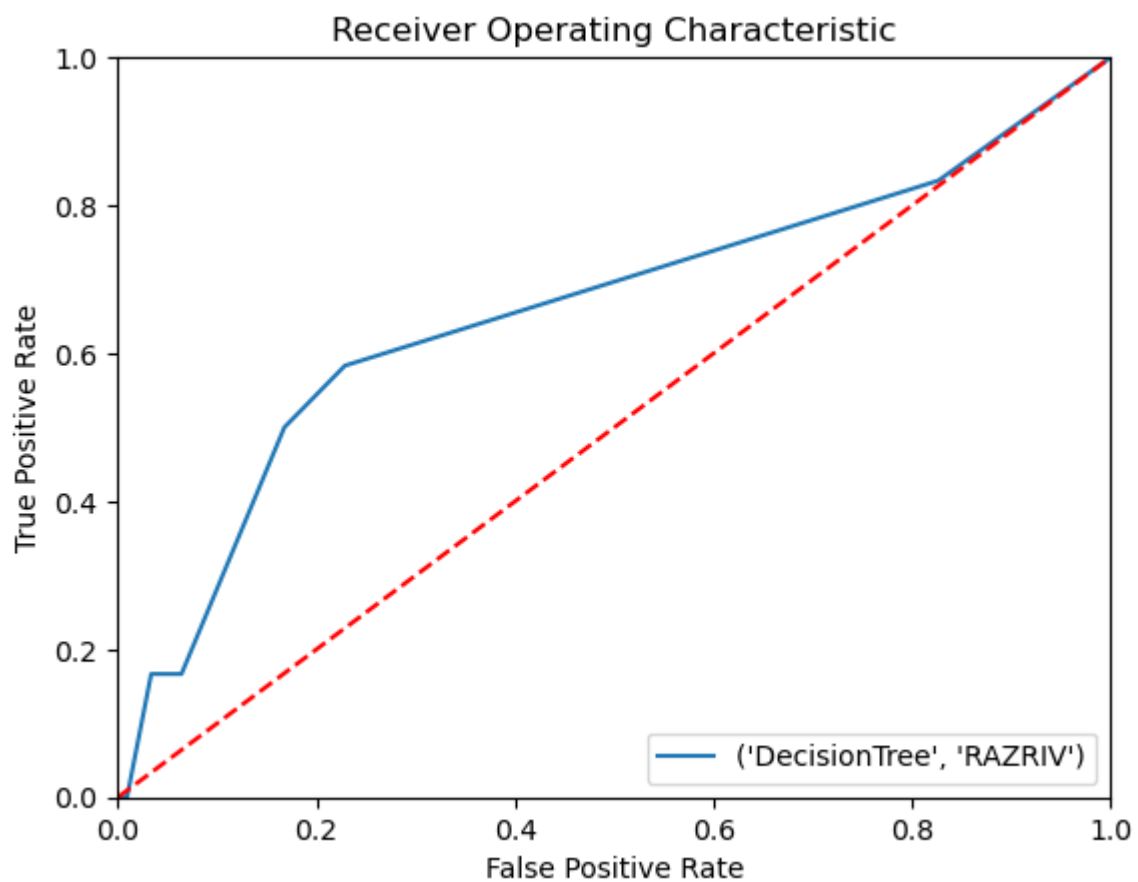


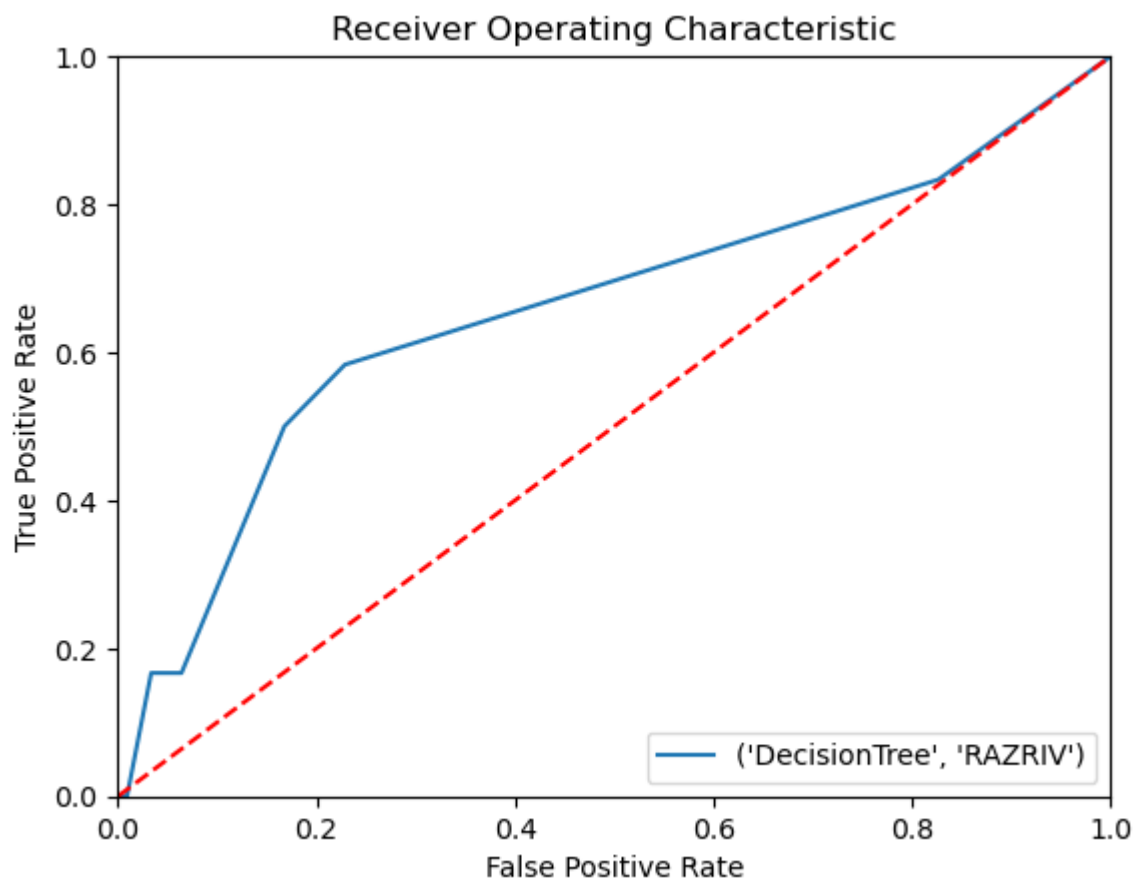
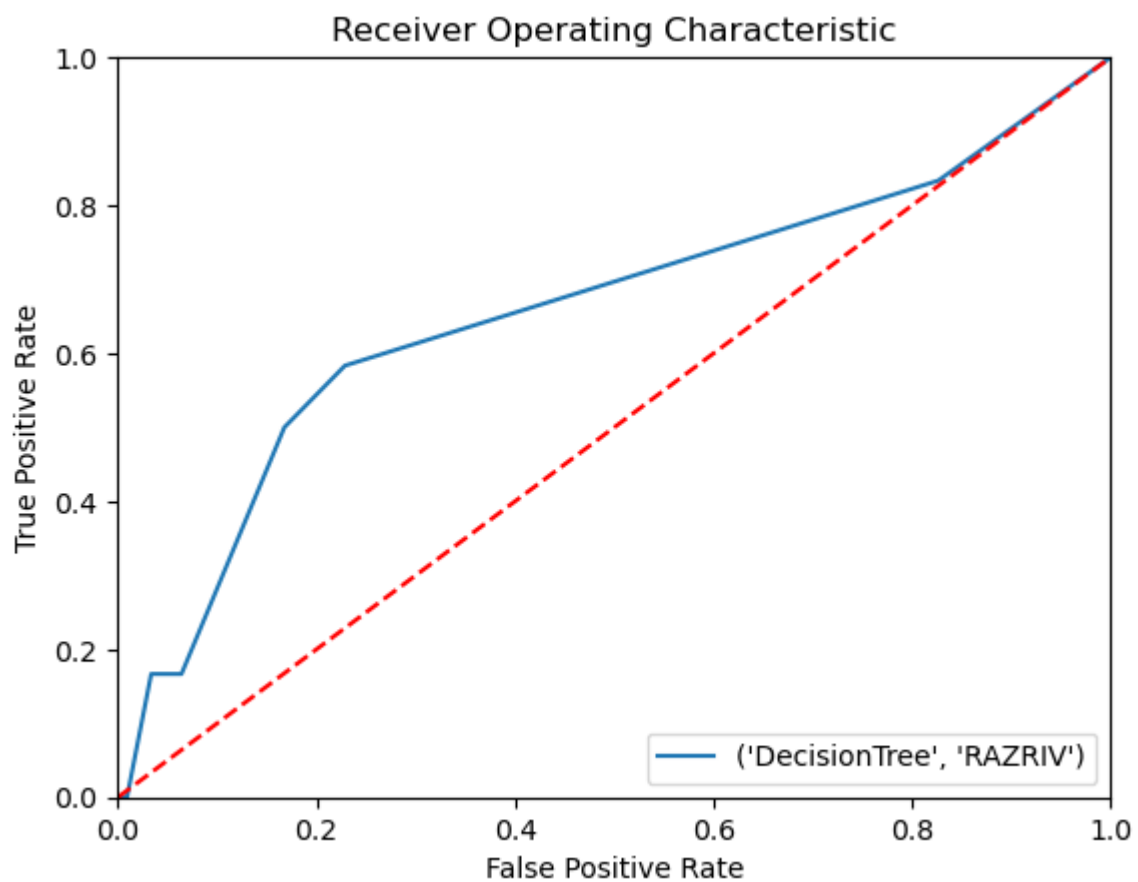


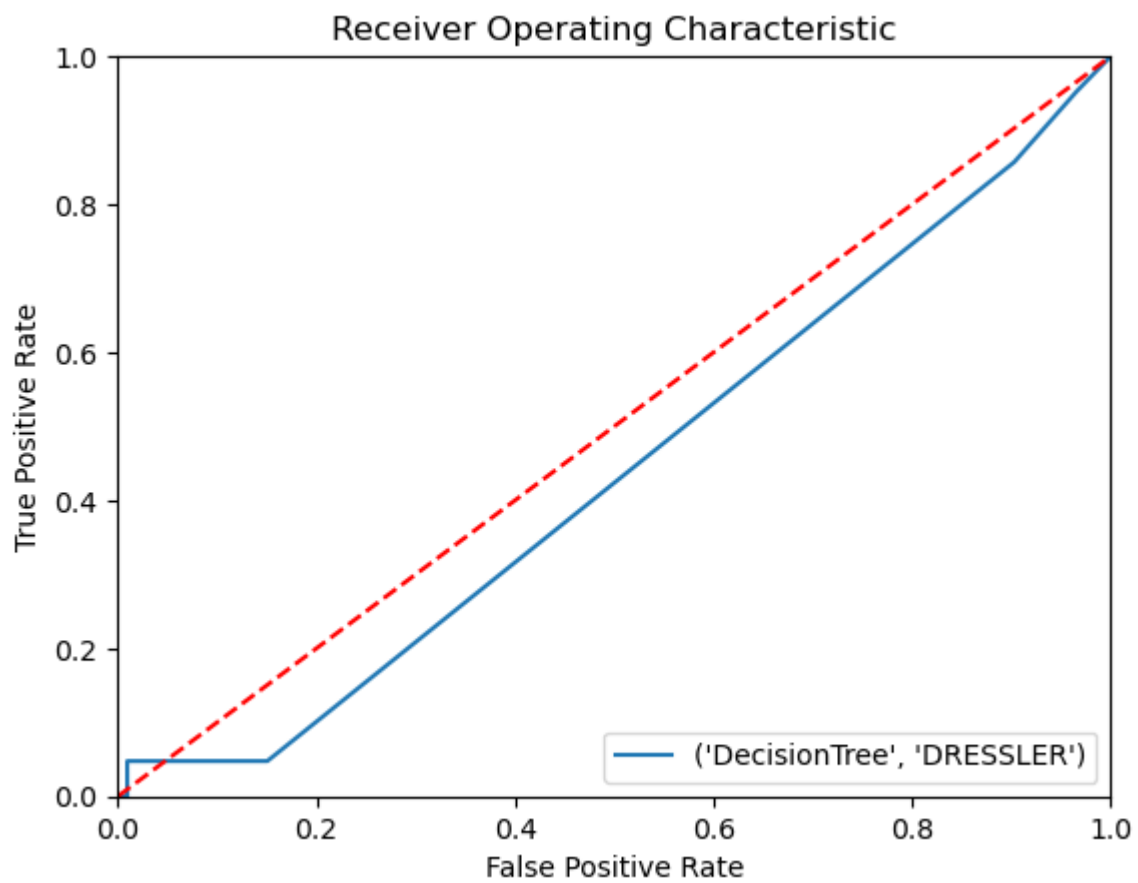
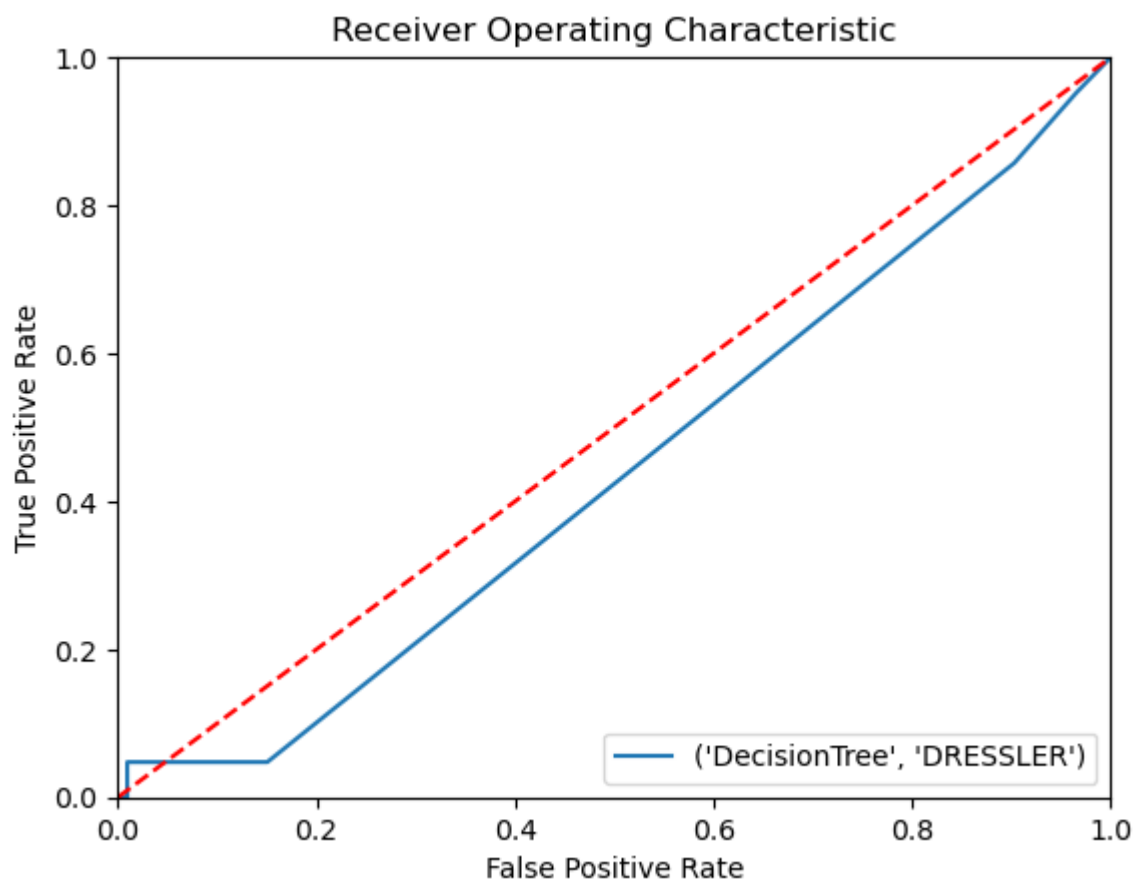


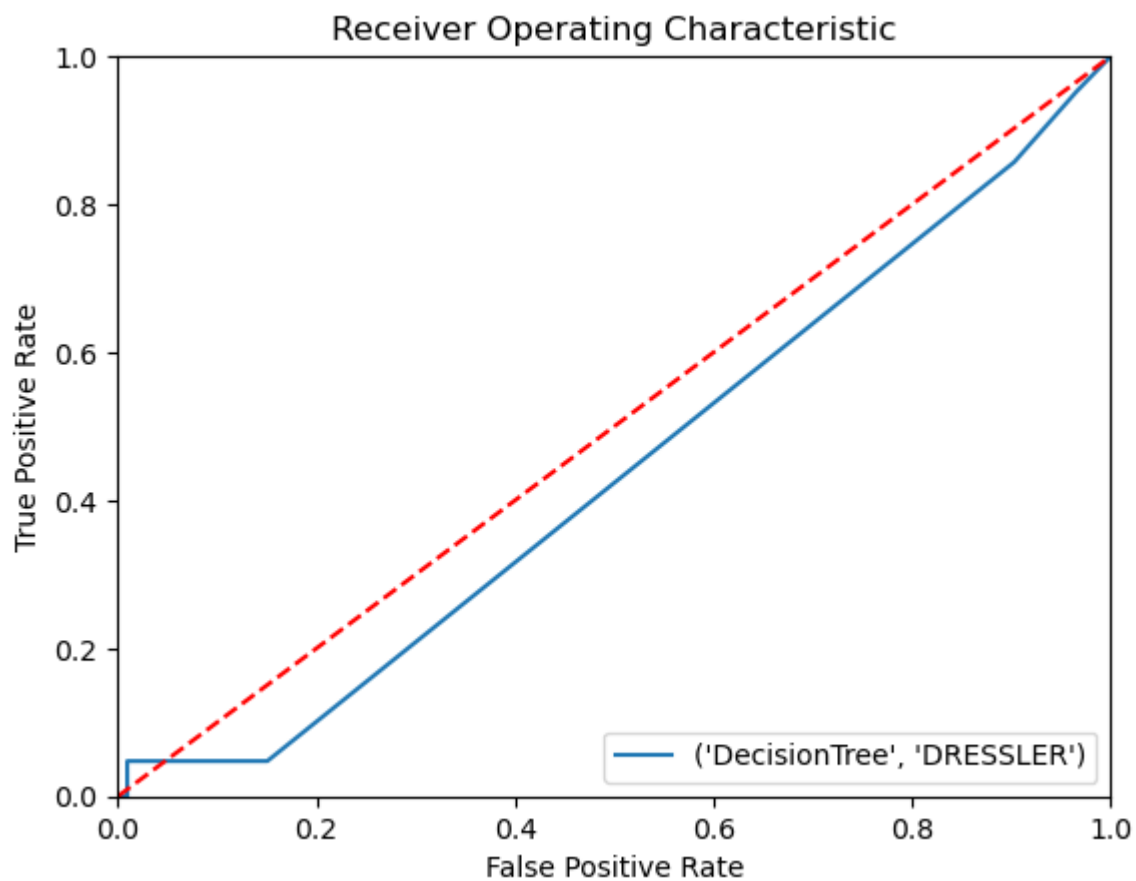
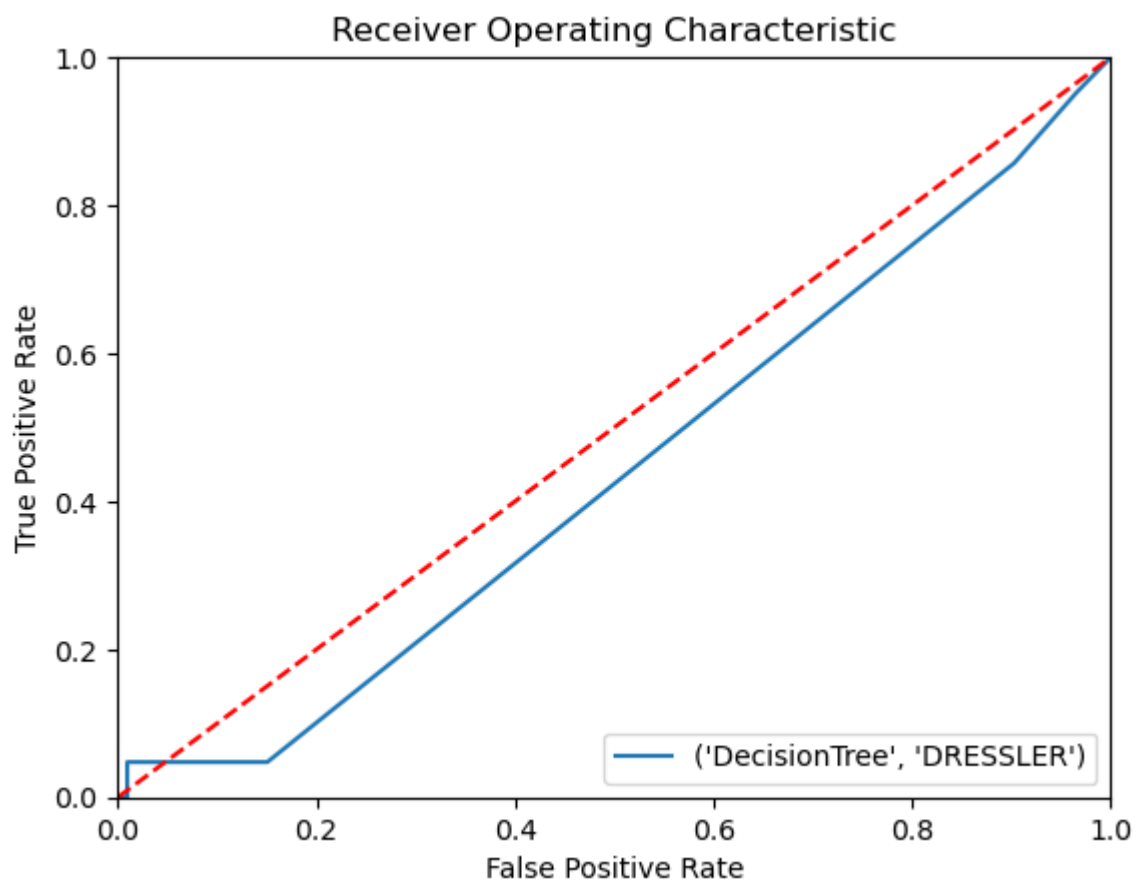


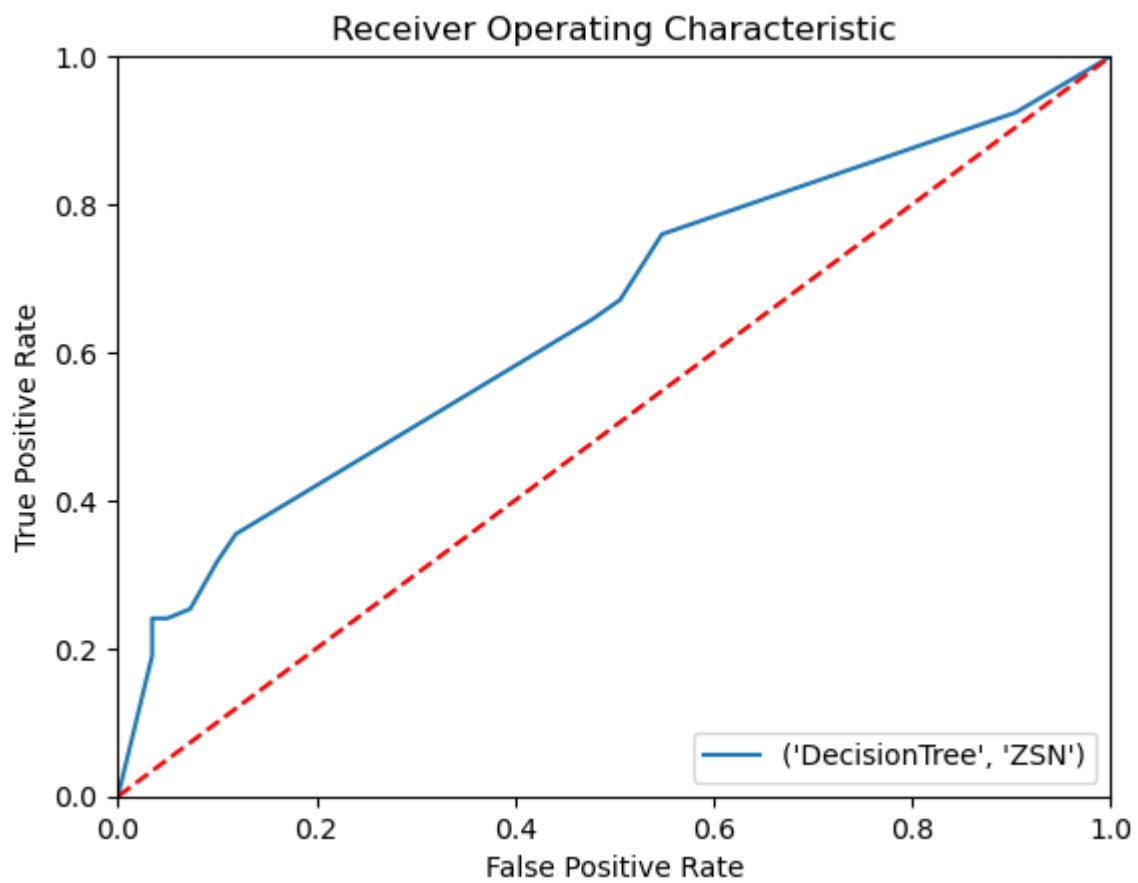
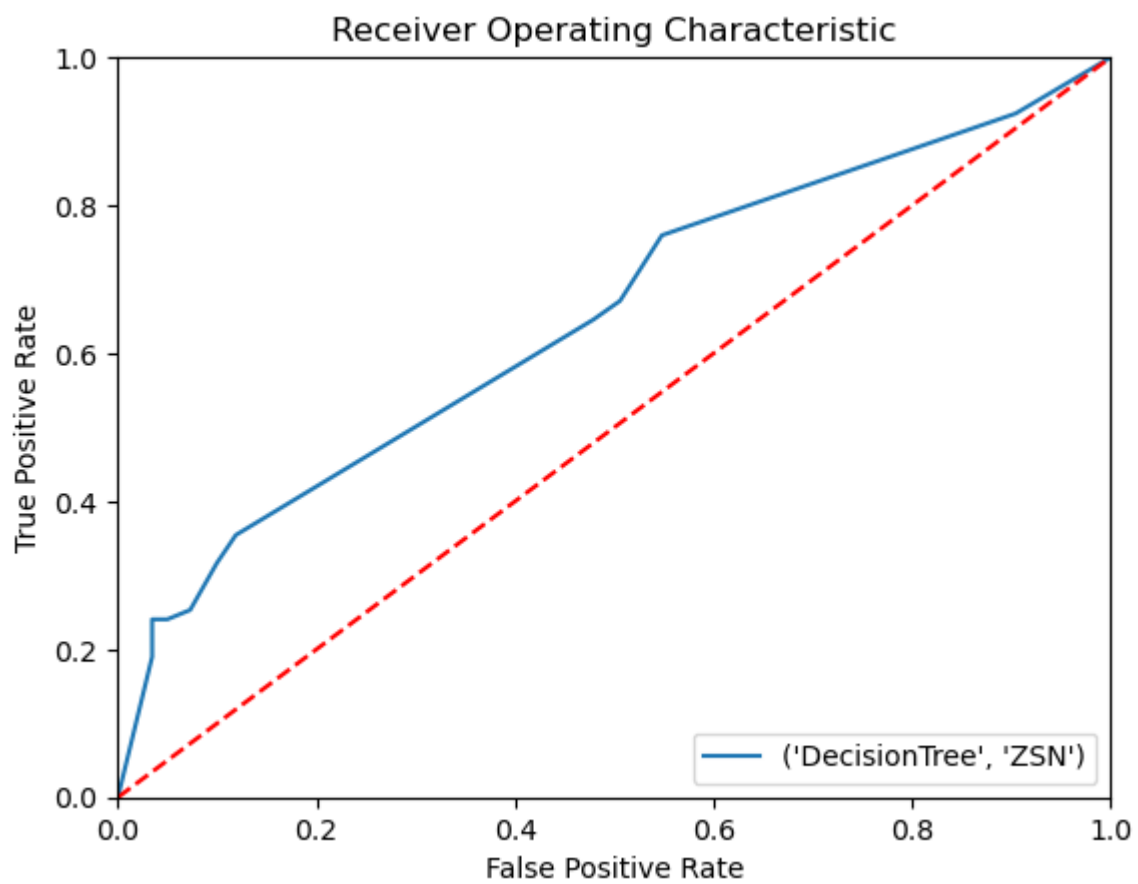


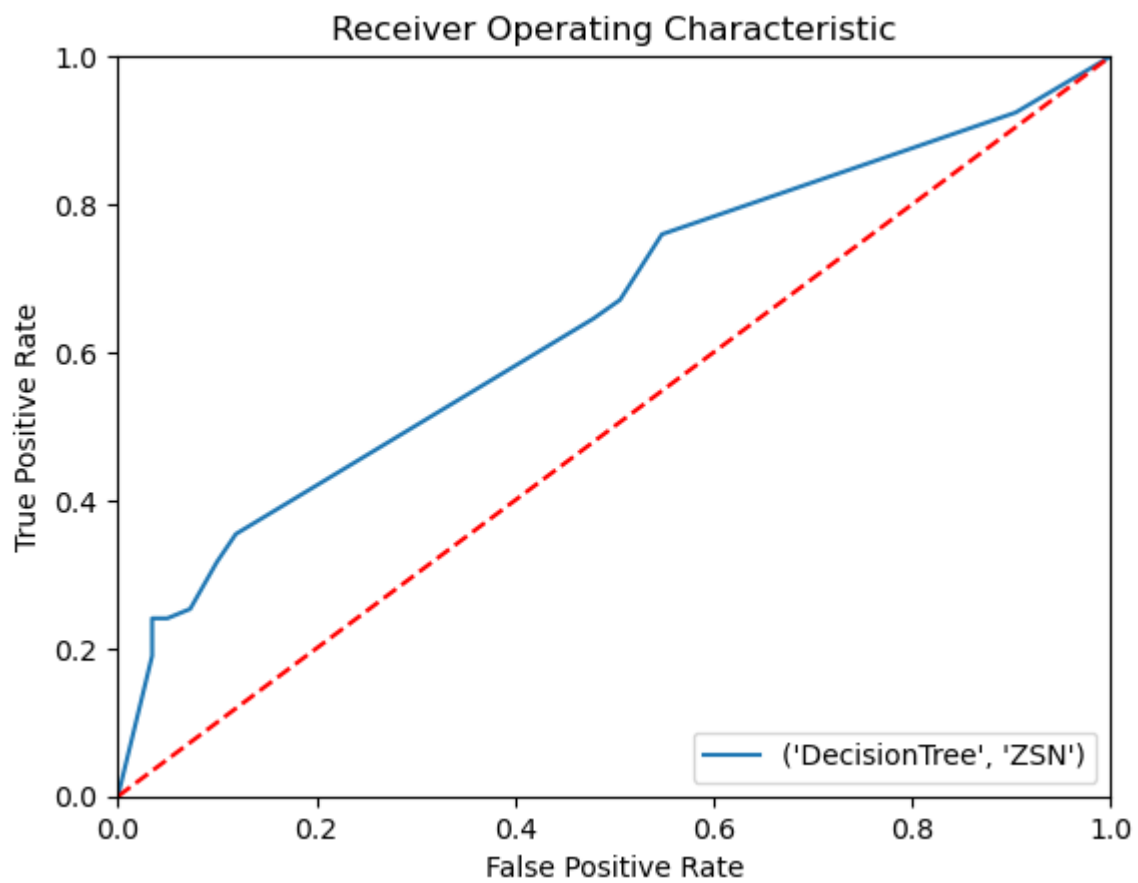
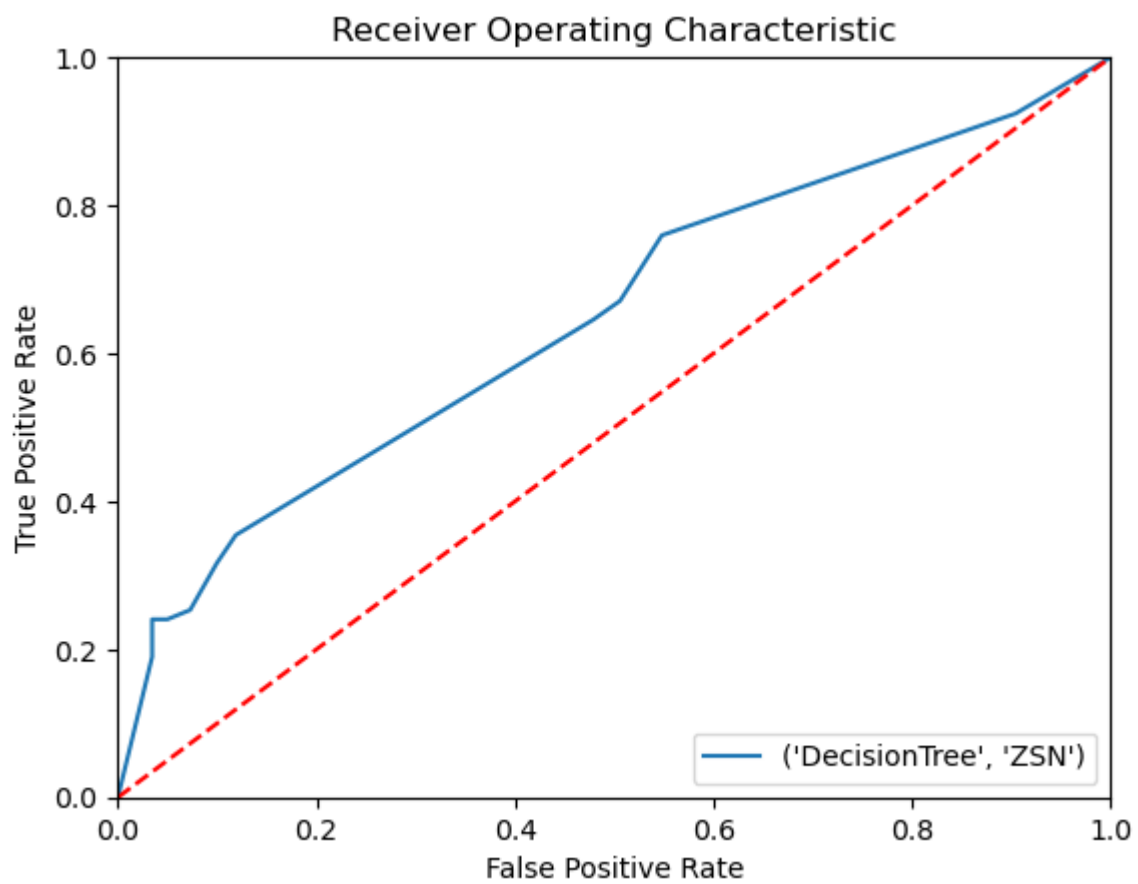


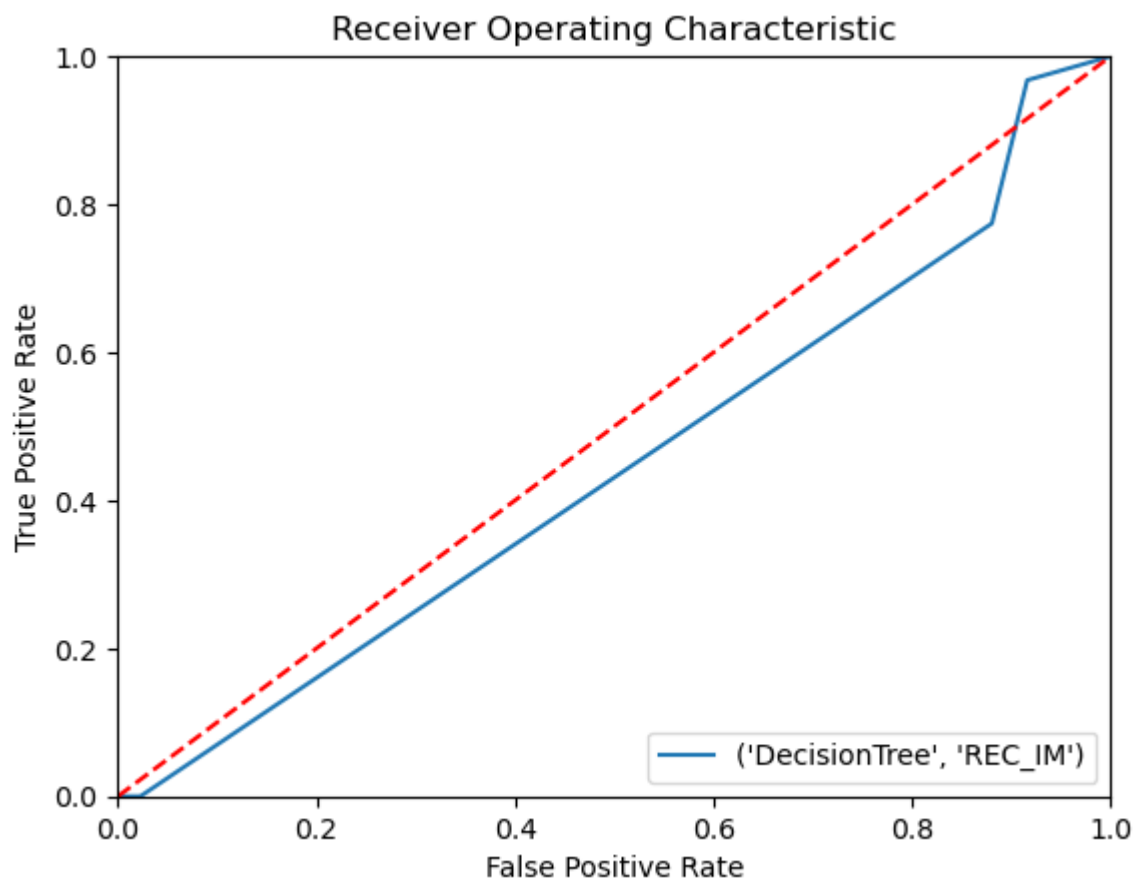
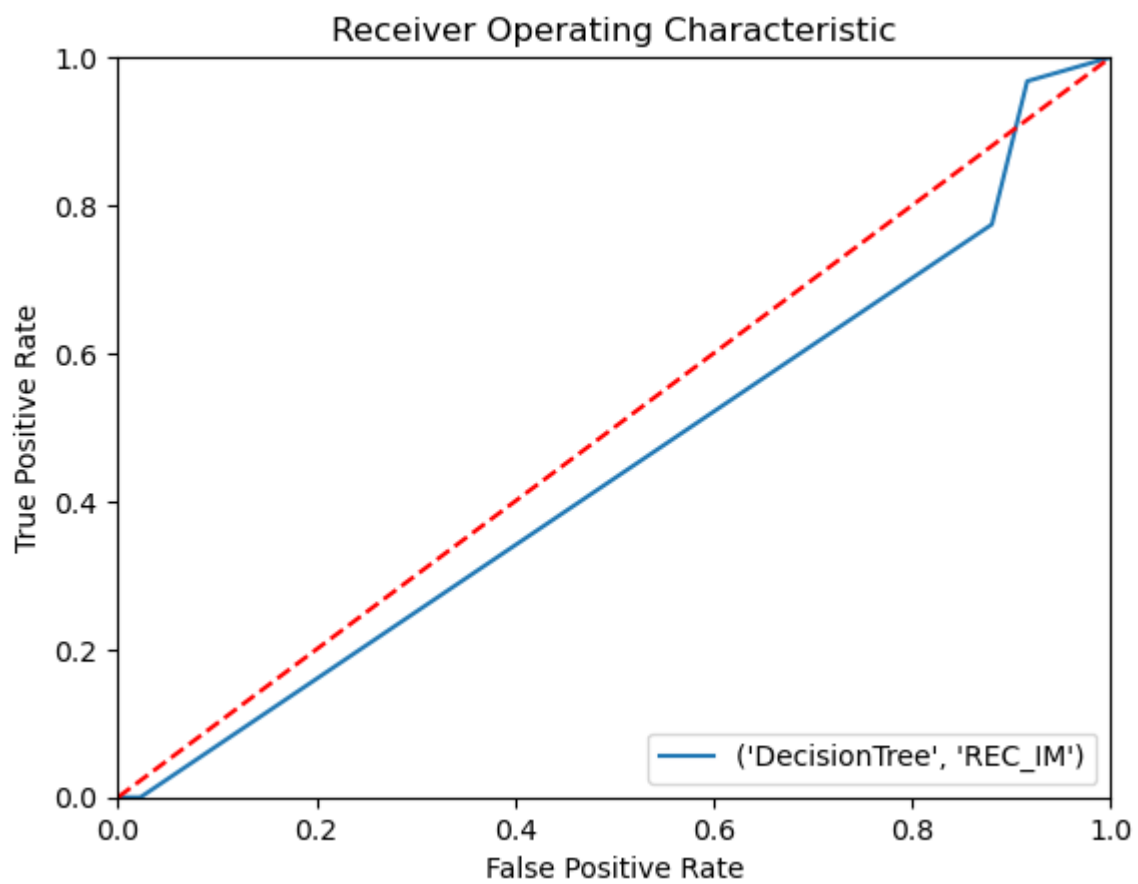


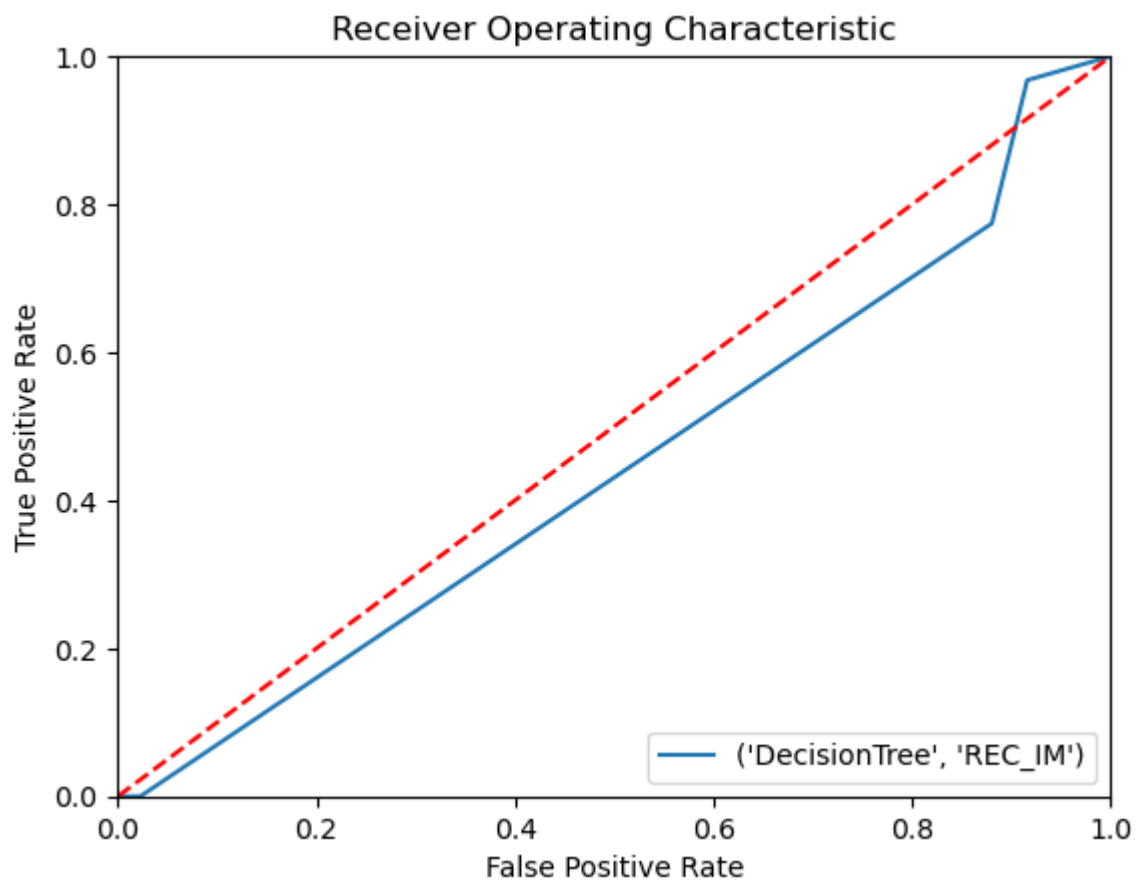
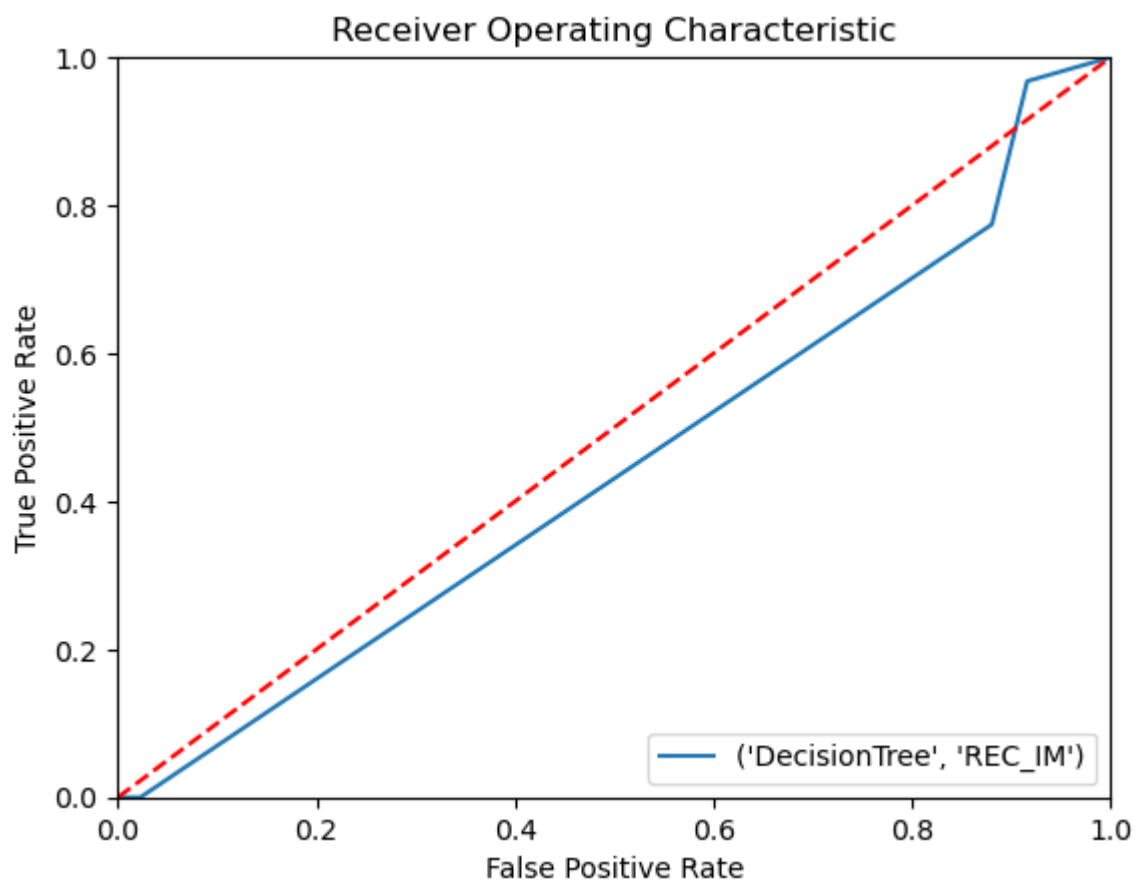


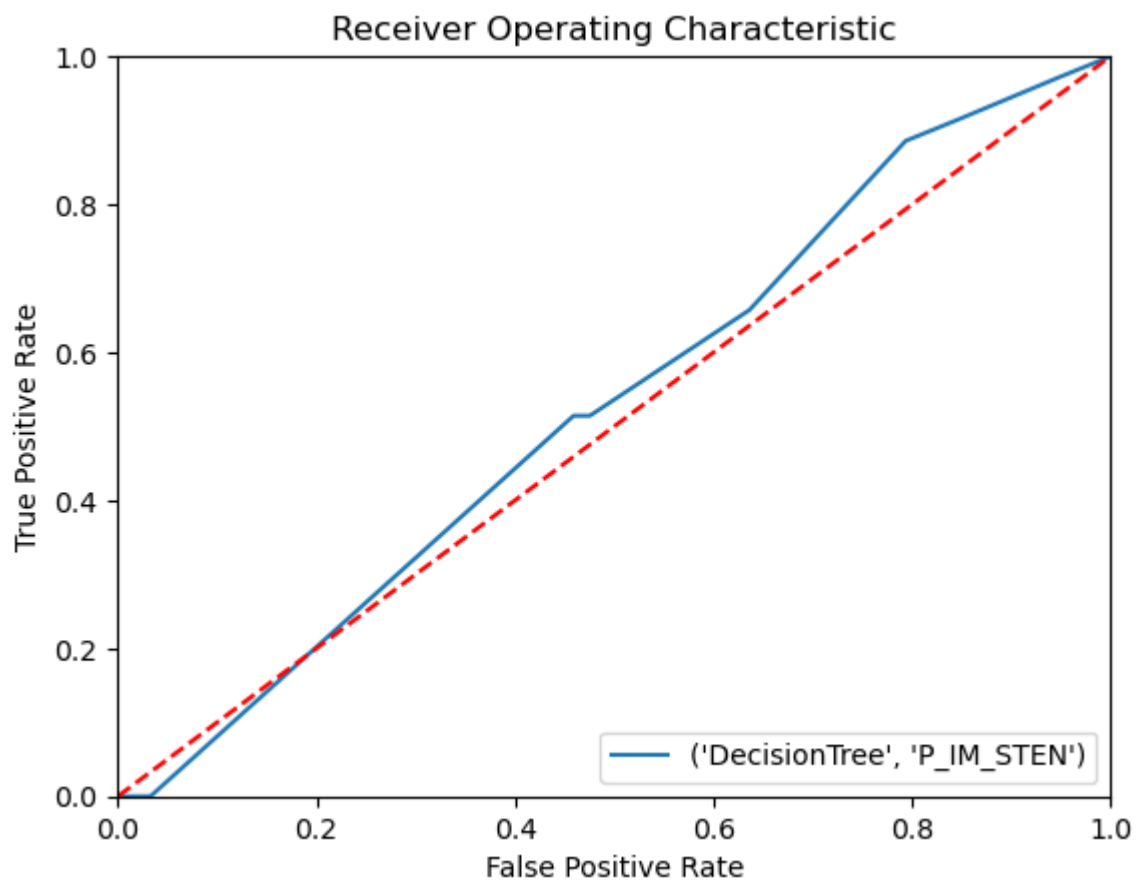
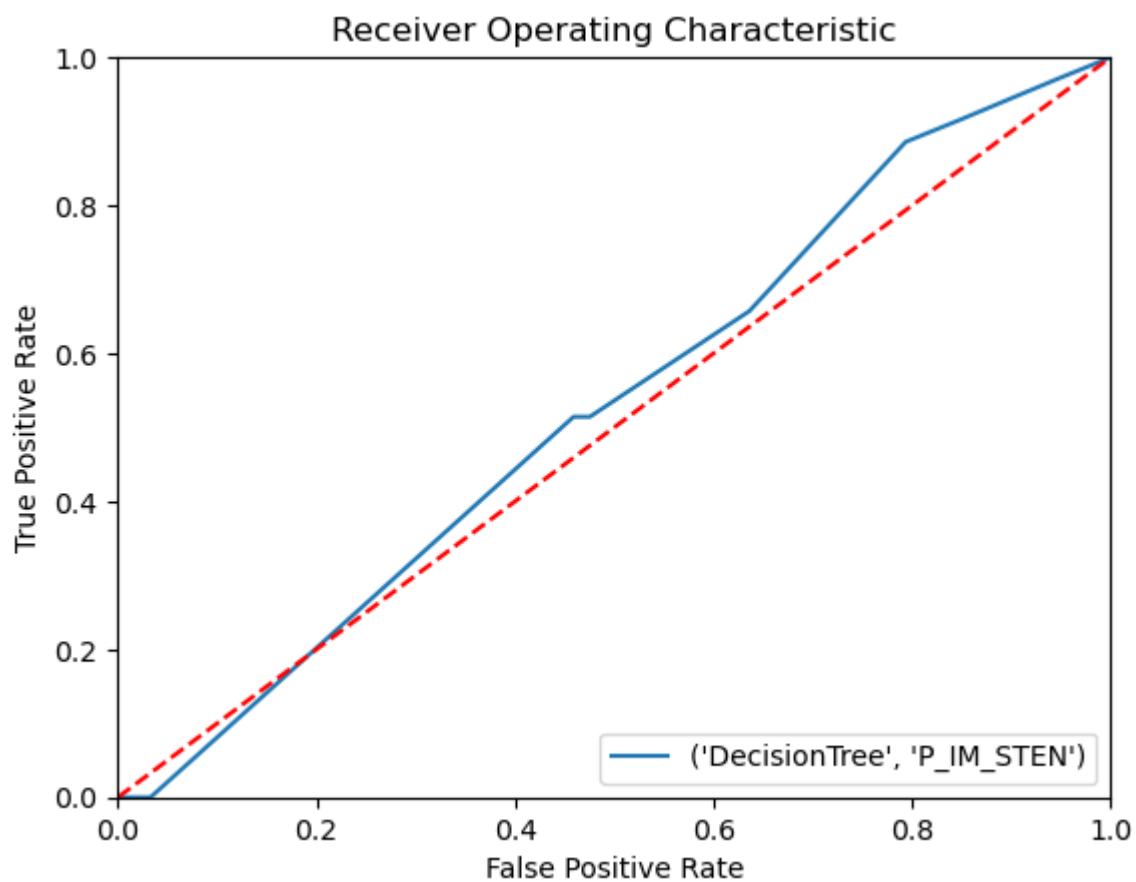


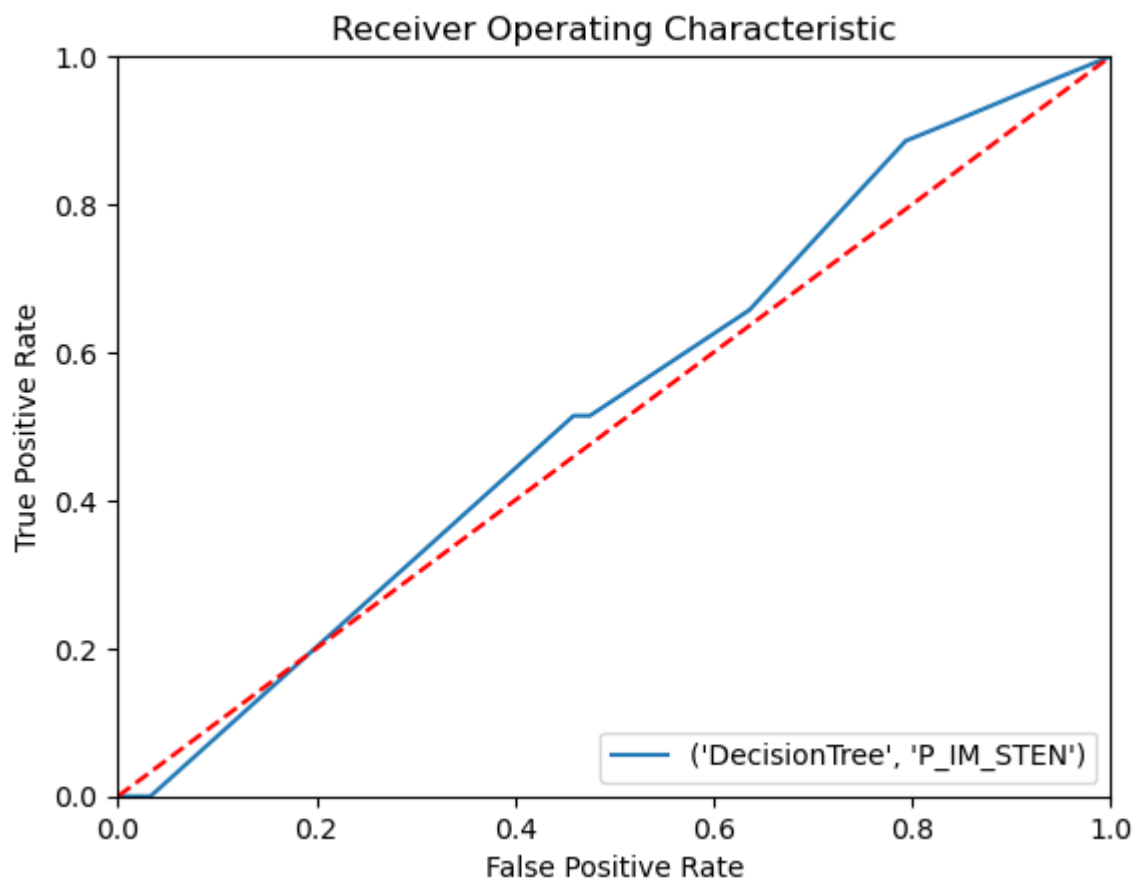
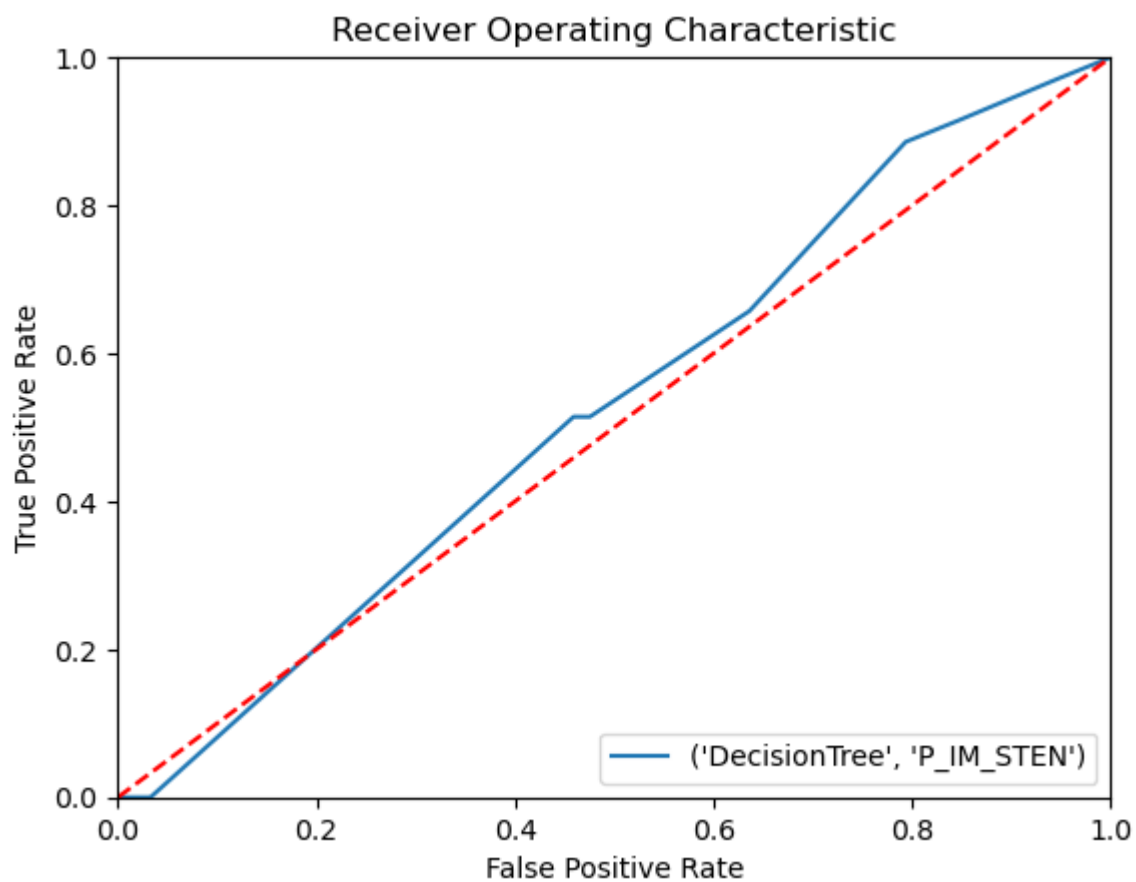


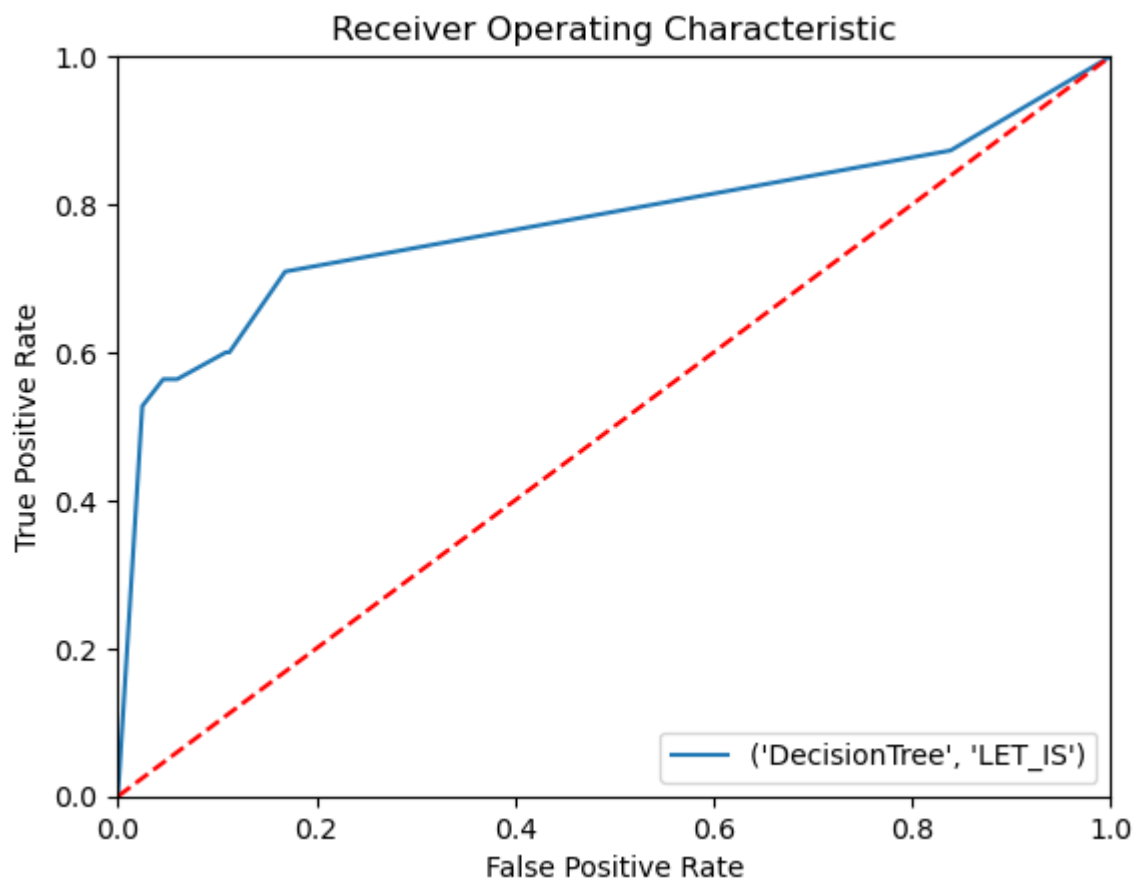
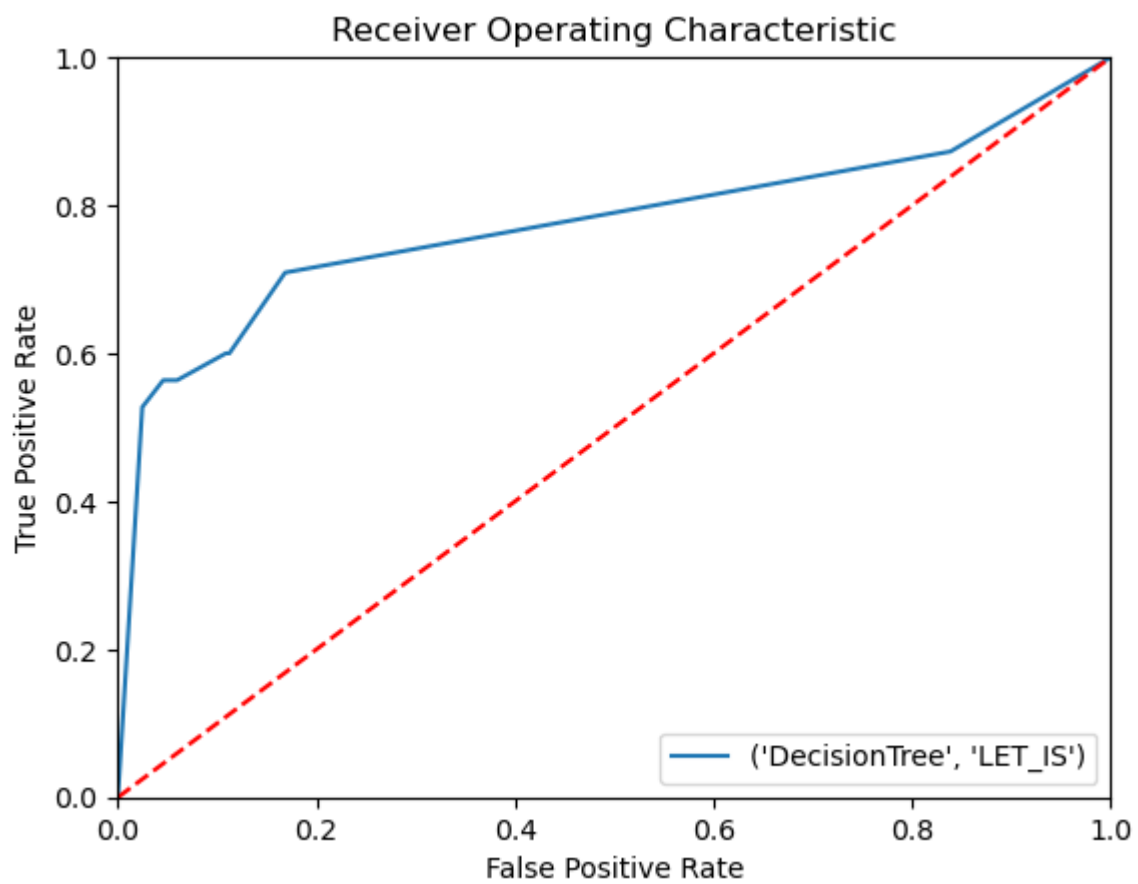


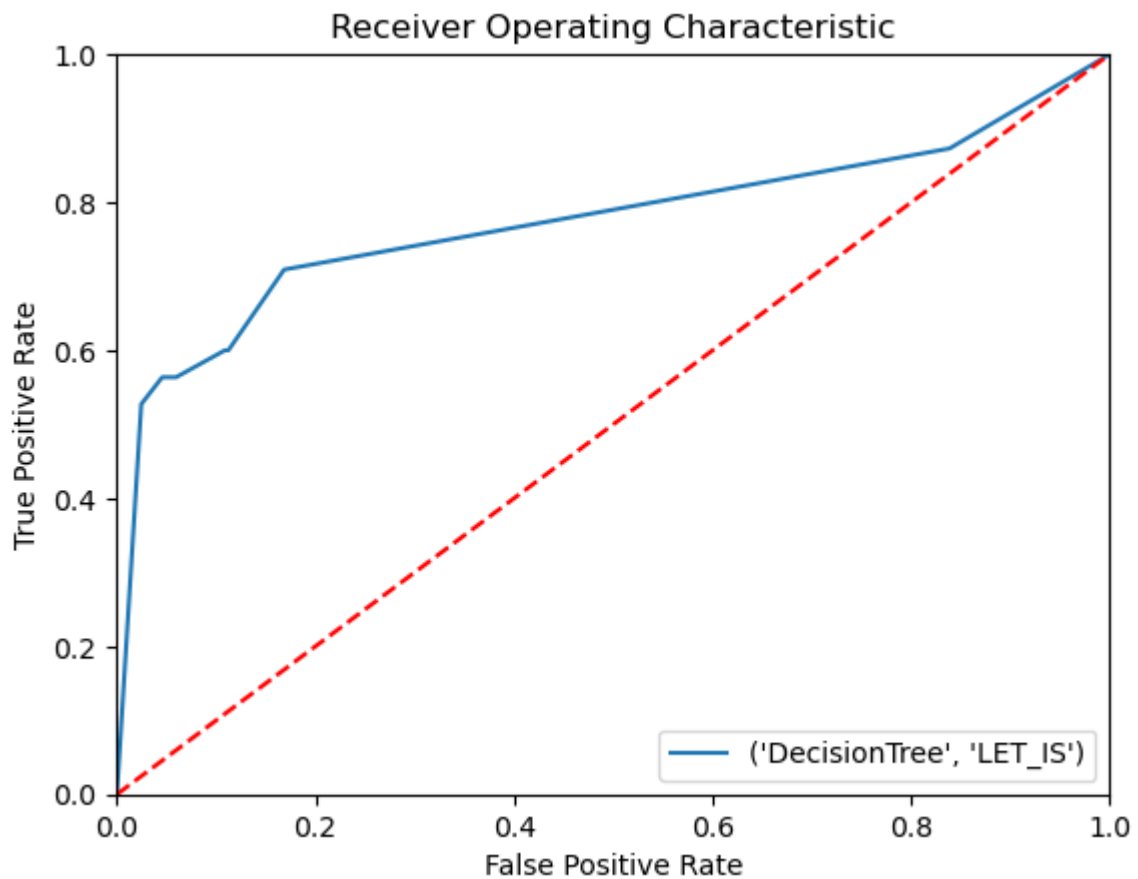
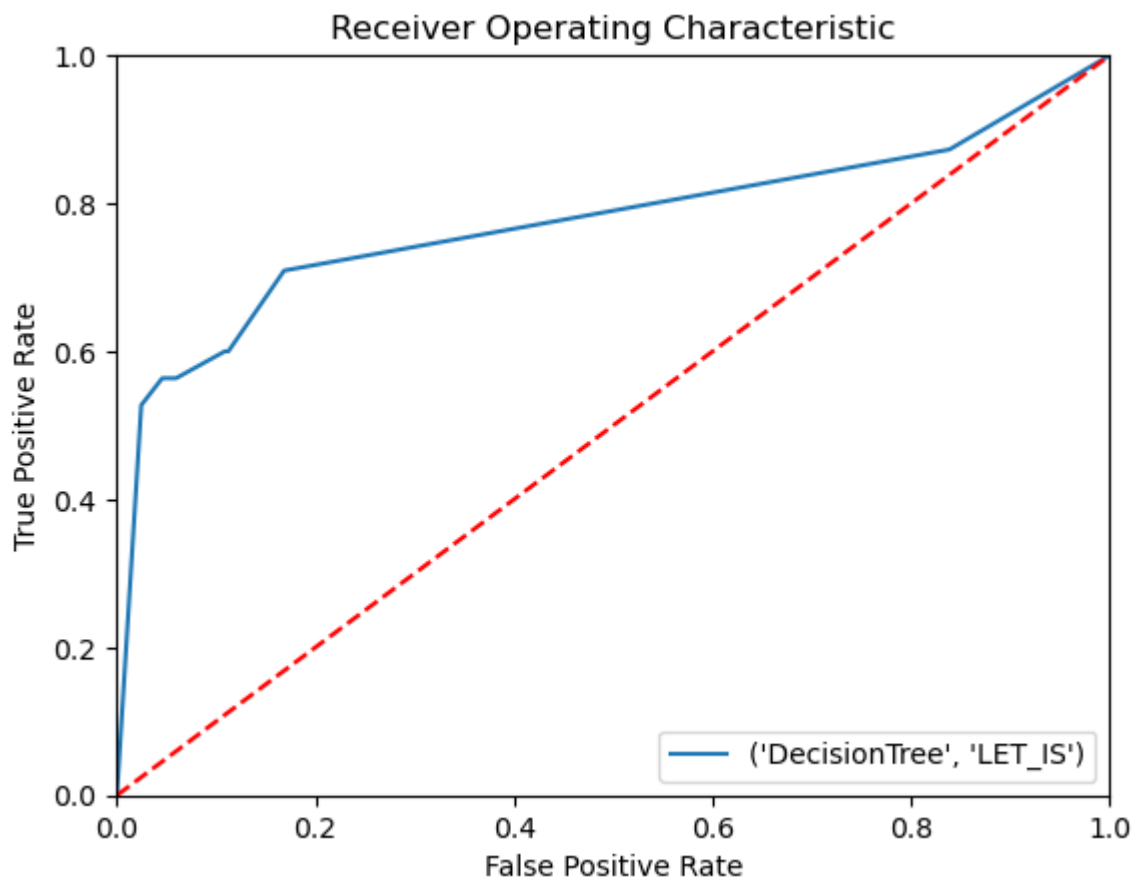












```
In [8]: acc_df=pd.DataFrame(columns=y.columns, index=models.keys())  
for i in range(0, len(acc)):
```

```
acc_df.iloc[i] = acc[i]
acc_df
```

Out[8]:

	FIBR_PREDS	PREDS_TAH	JELUD_TAH	FIBR_JELUD	A_V_BLOK	OTEK_LANC	RAZRIV
LogisticRegression	85.29412	98.82353	96.76471	94.70588	95.29412	88.52941	96.17647
XGBoostClassifier	85.88235	99.41176	97.64706	95.58824	96.17647	89.41176	96.47059
KNN	84.70588	99.11765	97.64706	96.17647	96.76471	90.0	96.47059
DecisionTree	84.41176	98.82353	96.76471	95.58824	94.70588	86.17647	93.82353

In [9]:

```
best_model = pd.DataFrame(columns=['model', 'acc'], index=[acc_df.columns])
for i in range(acc_df.shape[1]):
    best_model.iloc[i] = [acc_df.index[np.argmax(acc_df.iloc[:, i])], round(np.max(acc_df.iloc[:, i]), 5)]
best_model
```

Out[9]:

	model	acc
FIBR_PREDS	XGBoostClassifier	85.88235
PREDS_TAH	XGBoostClassifier	99.41176
JELUD_TAH	XGBoostClassifier	97.64706
FIBR_JELUD	KNN	96.17647
A_V_BLOK	KNN	96.76471
OTEK_LANC	KNN	90.0
RAZRIV	XGBoostClassifier	96.47059
DRESSLER	LogisticRegression	94.11765
ZSN	XGBoostClassifier	80.29412
REC_IM	KNN	90.29412
P_IM_STEN	XGBoostClassifier	89.11765
LET_IS	XGBoostClassifier	93.52941

Here we are going to try with RBF and Linear SVM models. This was neglected above because of for-loop issues and predict values using sklearn. There are issues with outputting SVM predictions as a binary probability in a data frame of multiple responses even though each response (columns) is binary.

In [10]:

```
acc2=[]
for model_name, model in models2.items():
    clf2 = model
    acc_model2=[]
    for i in range(0, 12):
        clf2.fit(X_train, y_train.iloc[:, i])
        acc_feature2 = round(clf2.score(X_test, y_test.iloc[:, i]) * 100, 5)
        acc_model2.append(acc_feature2)
    acc2.append(acc_model2)
```

```
In [11]: acc_df2=pd.DataFrame(columns=y.columns, index=models2.keys())
for i in range(0, len(acc2)):
    acc_df2.iloc[i] = acc2[i]
acc_df2
```

```
Out[11]:
```

	FIBR_PREDS	PREDS_TAH	JELUD_TAH	FIBR_JELUD	A_V_BLOK	OTEK_LANC	RAZRIV
LogisticRegression	85.29412	98.82353	96.76471	94.70588	95.29412	88.52941	96.17647
XGBoostClassifier	85.88235	99.41176	97.64706	95.58824	96.17647	89.41176	96.47059
KNN	84.70588	99.11765	97.64706	96.17647	96.76471	90.0	96.47059
Linear SVM	85.58824	99.11765	97.64706	96.17647	96.76471	89.70588	96.47059
RBF SVM	85.88235	99.41176	97.64706	96.17647	96.76471	90.29412	96.47059
DecisionTree	84.70588	98.52941	97.05882	95.29412	94.70588	85.88235	93.52941

```
In [12]: best_model2 = pd.DataFrame(columns=['model', 'acc'], index=[acc_df2.columns])
for i in range(acc_df2.shape[1]):
    best_model2.iloc[i] = [acc_df2.index[np.argmax(acc_df2.iloc[:, i])], round(np.max(
best_model2
```

```
Out[12]:
```

	model	acc
FIBR_PREDS	XGBoostClassifier	85.88235
PREDS_TAH	XGBoostClassifier	99.41176
JELUD_TAH	XGBoostClassifier	97.64706
FIBR_JELUD	KNN	96.17647
A_V_BLOK	KNN	96.76471
OTEK_LANC	RBF SVM	90.29412
RAZRIV	XGBoostClassifier	96.47059
DRESSLER	LogisticRegression	94.11765
ZSN	XGBoostClassifier	80.29412
REC_IM	RBF SVM	90.88235
P_IM_STEN	Linear SVM	89.70588
LET_IS	XGBoostClassifier	93.52941