

Model Evaluation Report

Crypto Volatility Detection

Generated: 2025-11-13 02:47

Milestone 3: Model Training & Evaluation

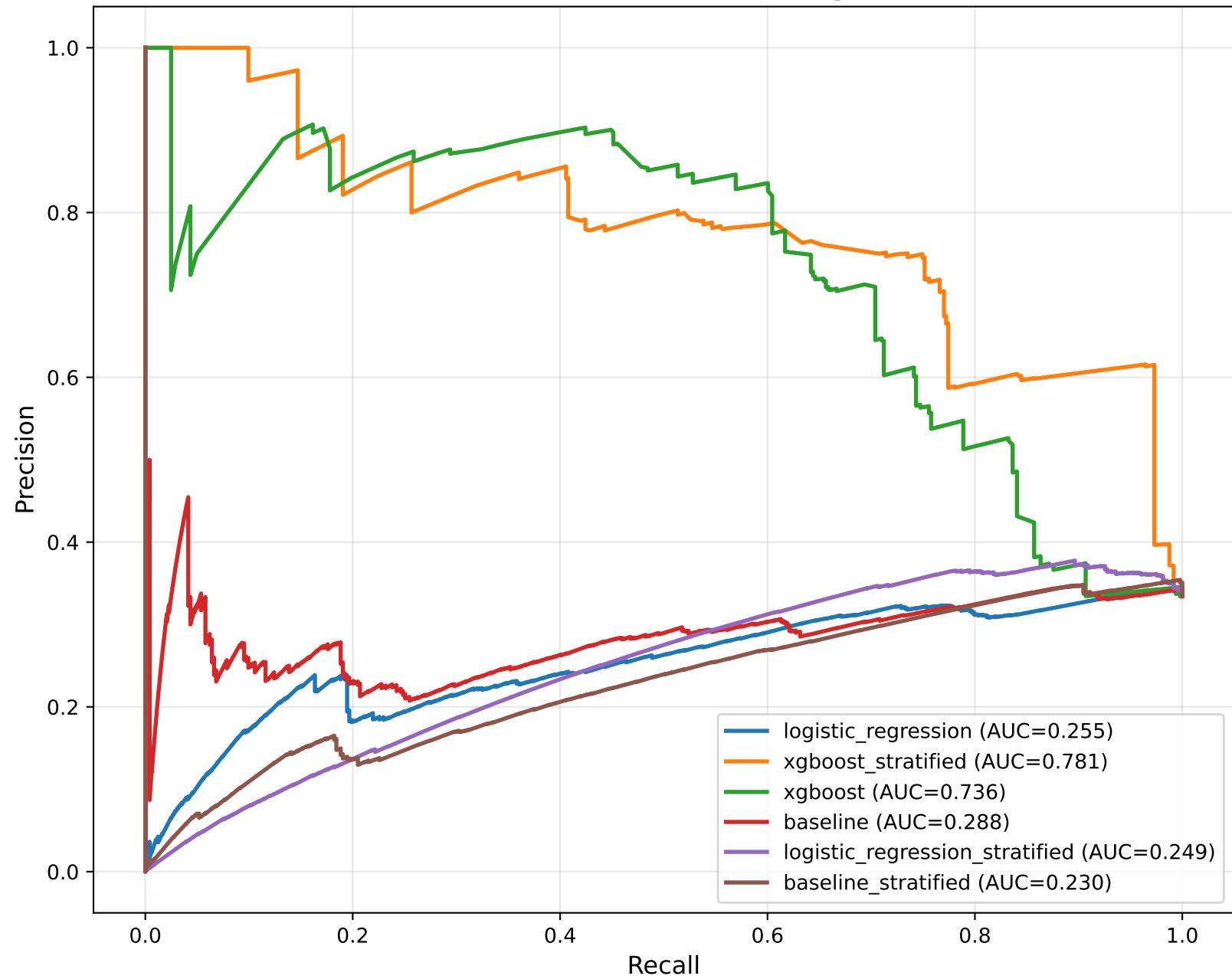
Model Performance Comparison

Model	PR-AUC	F1-Score	Precision	Recall	ROC-AUC
logistic_regression	0.2549	0.4241	0.3100	0.6708	0.3611
xgboost_stratified	0.7815	0.6851	0.5287	0.9731	0.8802
xgboost	0.7359	0.3994	0.8741	0.2588	0.7968
baseline	0.2881	0.0000	0.0000	0.0000	0.4110
logistic_regression_stratified	0.2491	0.5013	0.3635	0.8075	0.3298
baseline_stratified	0.2295	0.1694	0.1356	0.2257	0.2711

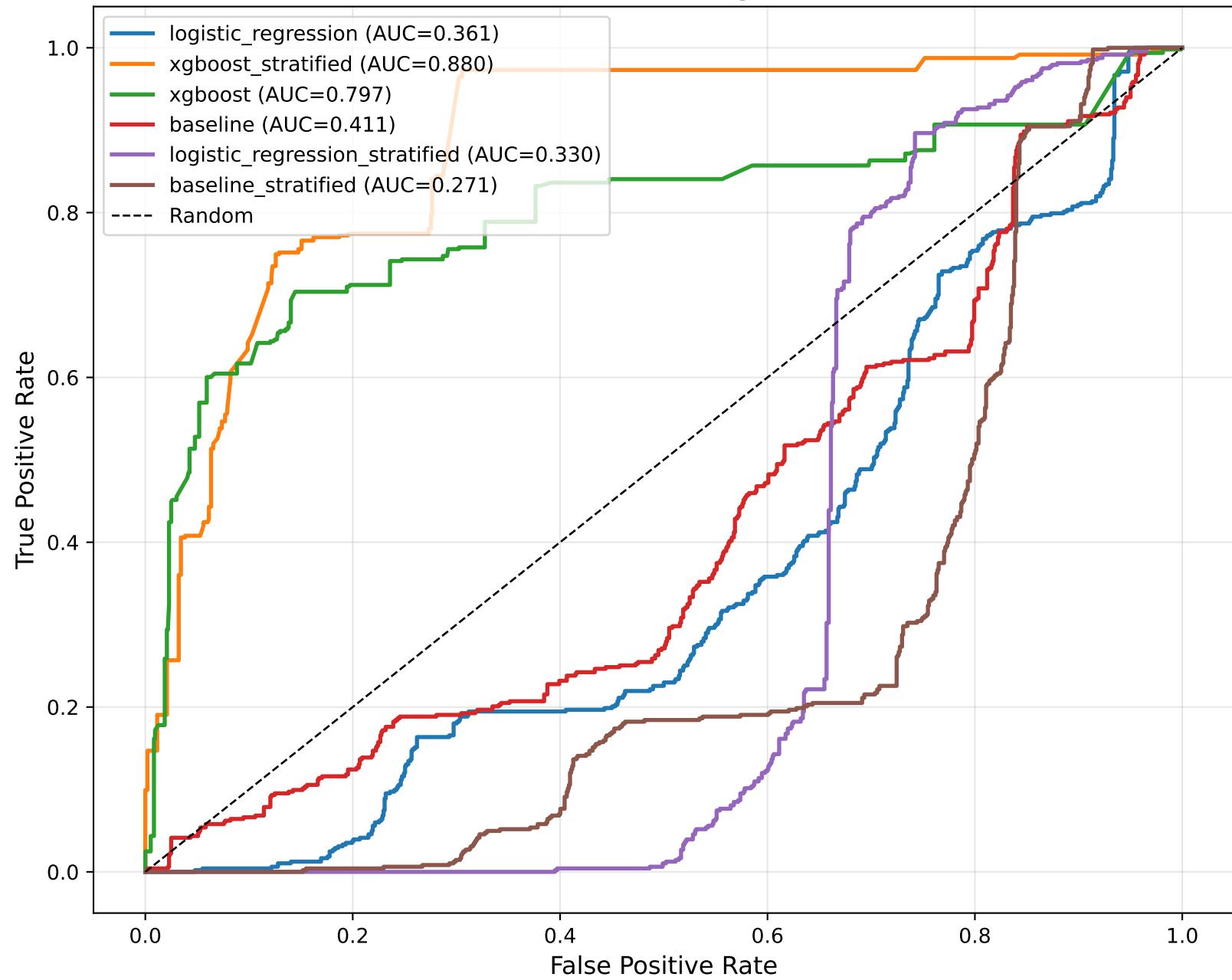
Primary Metric: PR-AUC (best: 0.7815)

PR-AUC is preferred due to class imbalance (spike events are rare)

Precision-Recall Curves Comparison

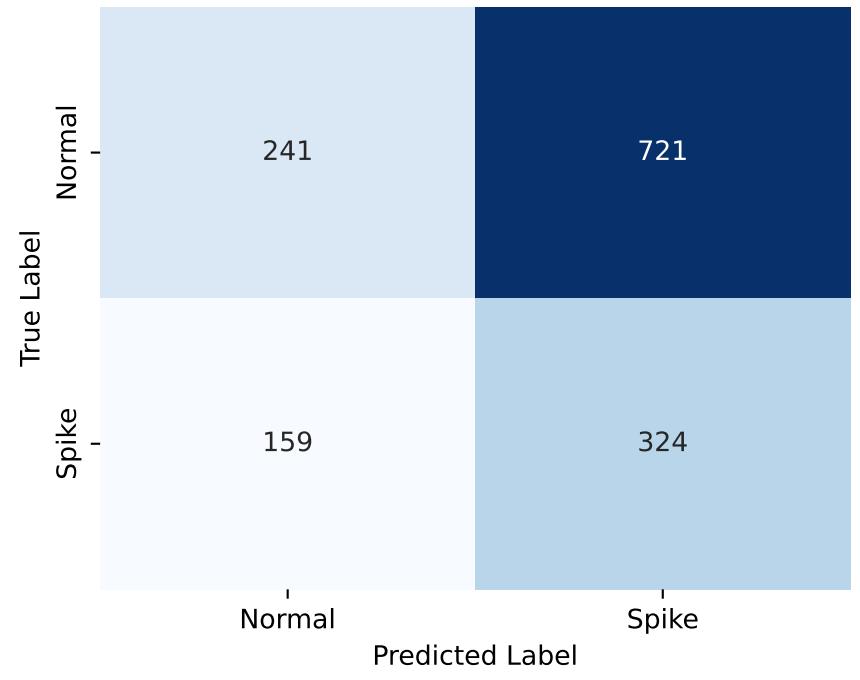


ROC Curves Comparison

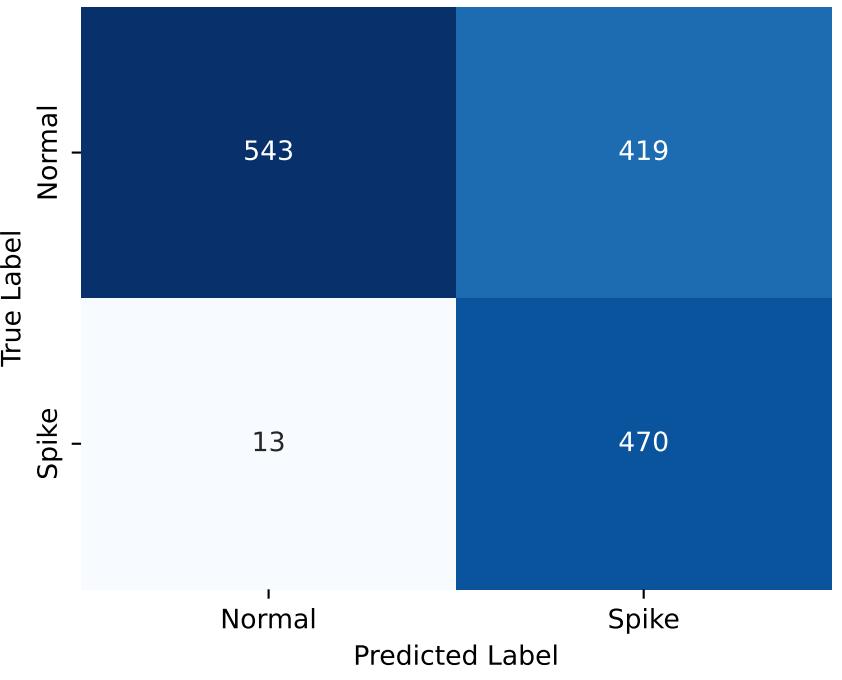


Confusion Matrices (Test Set)

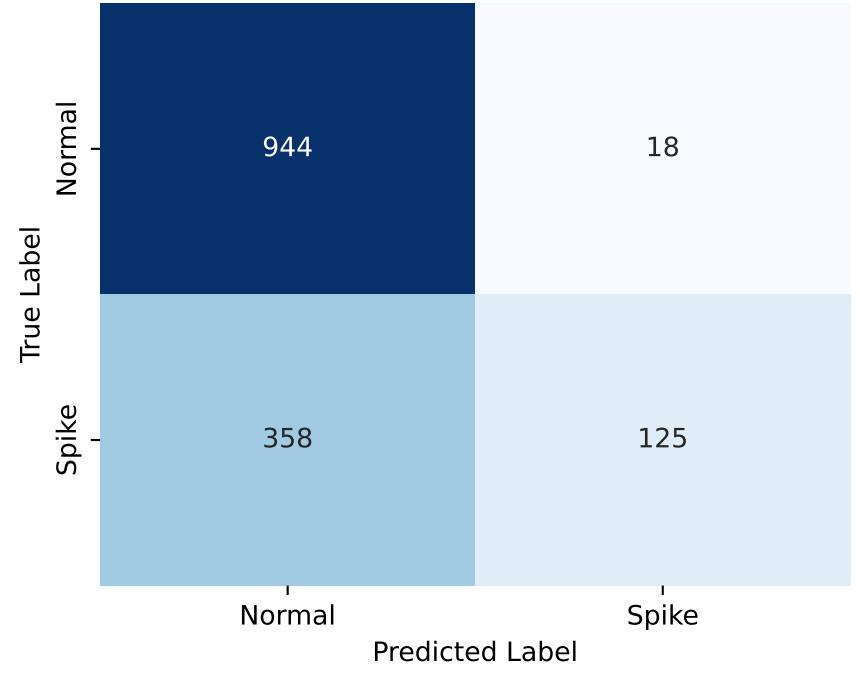
logistic_regression



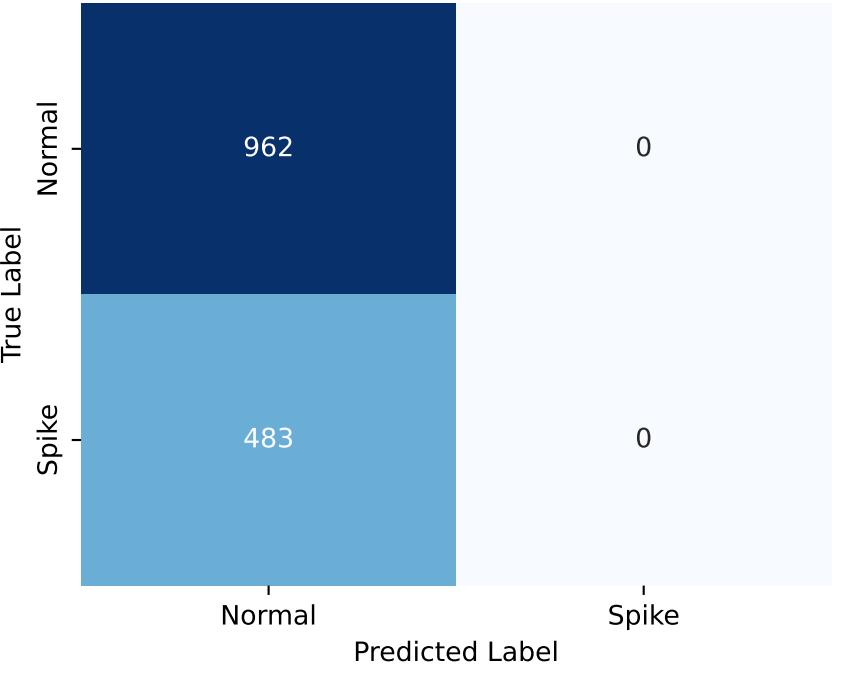
xgboost_stratified



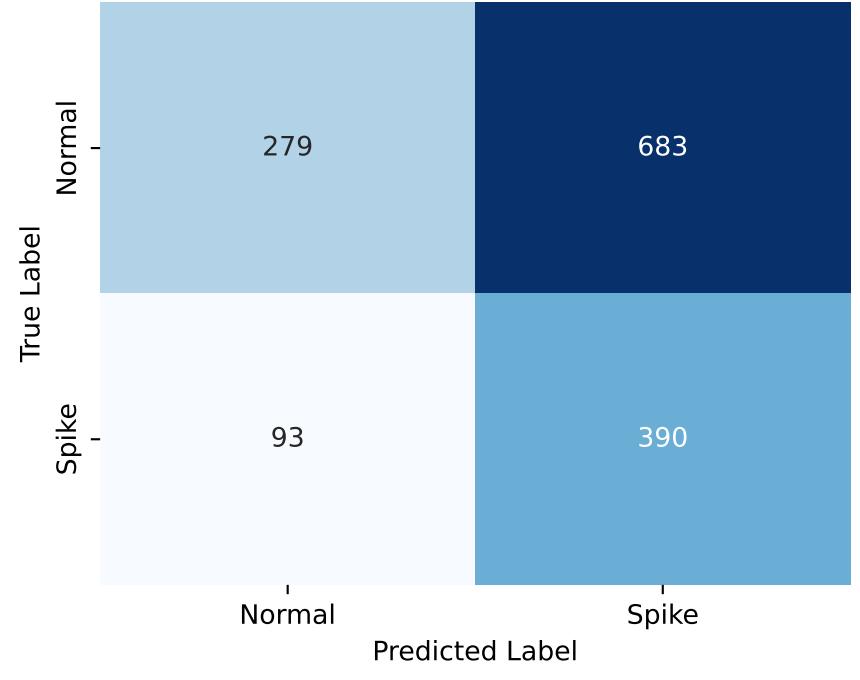
xgboost



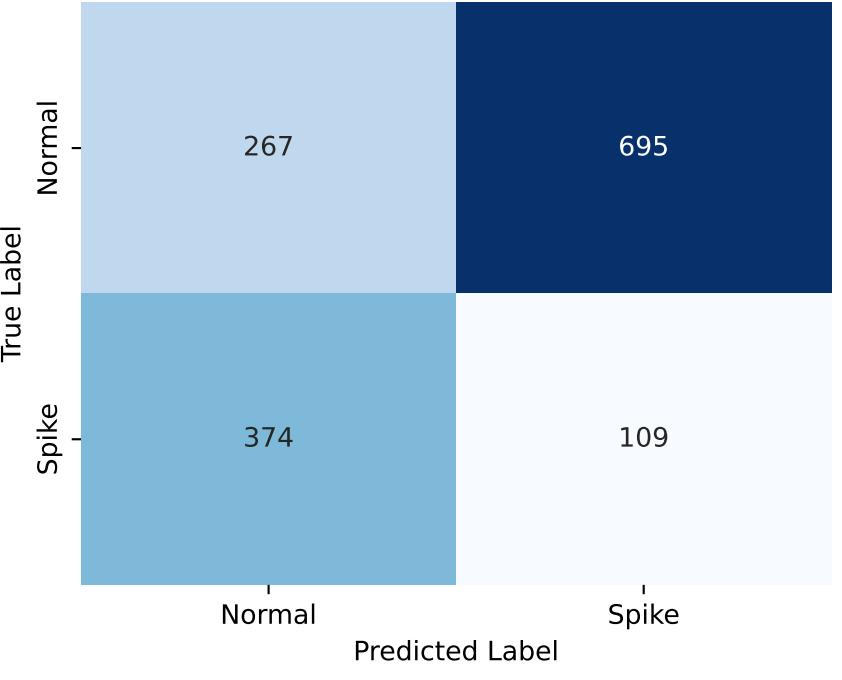
baseline



logistic_regression_stratified

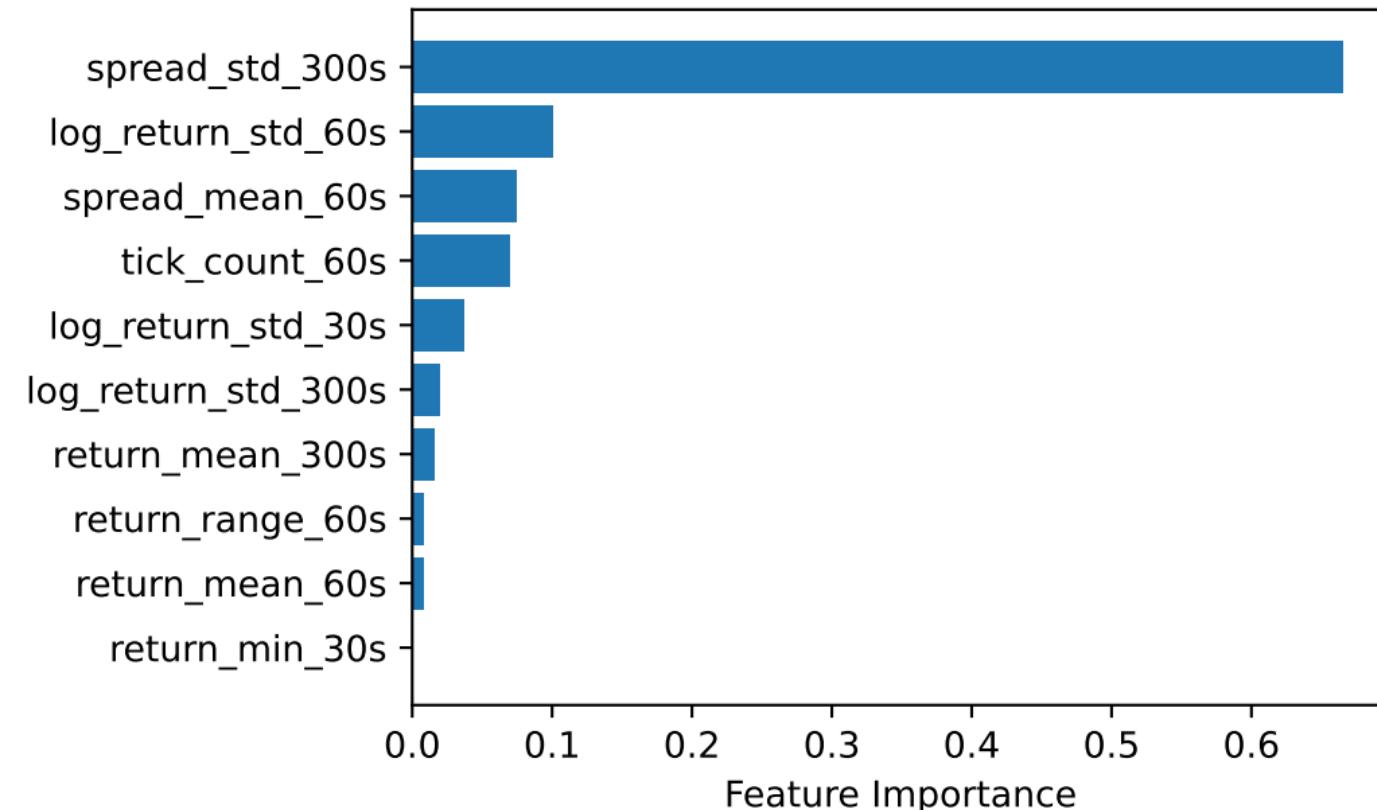


baseline_stratified



Feature Importance

xgboost_stratified (Importance)



xgboost (Importance)

