Protein: P17787 - metric: balanced_accuracy
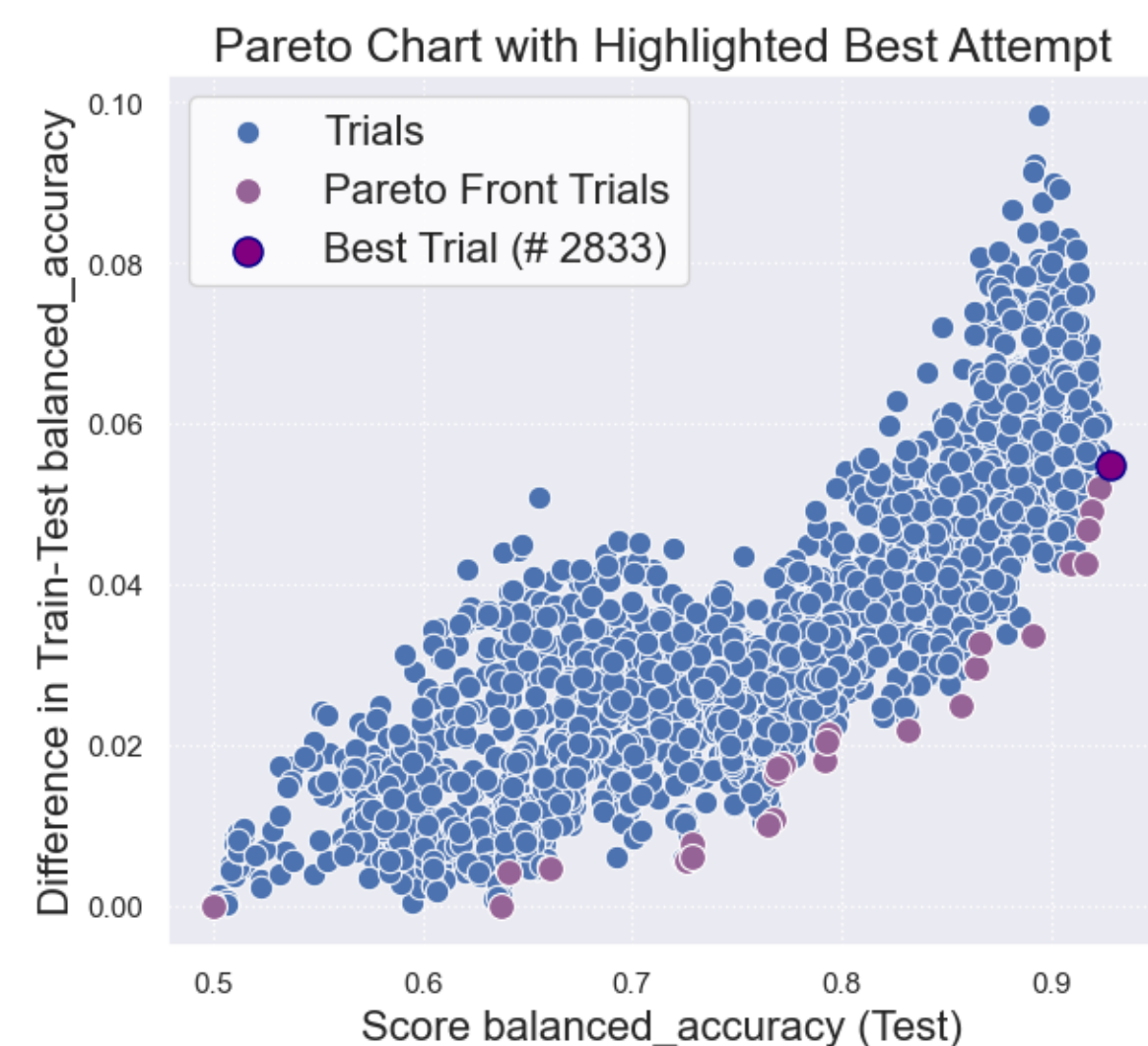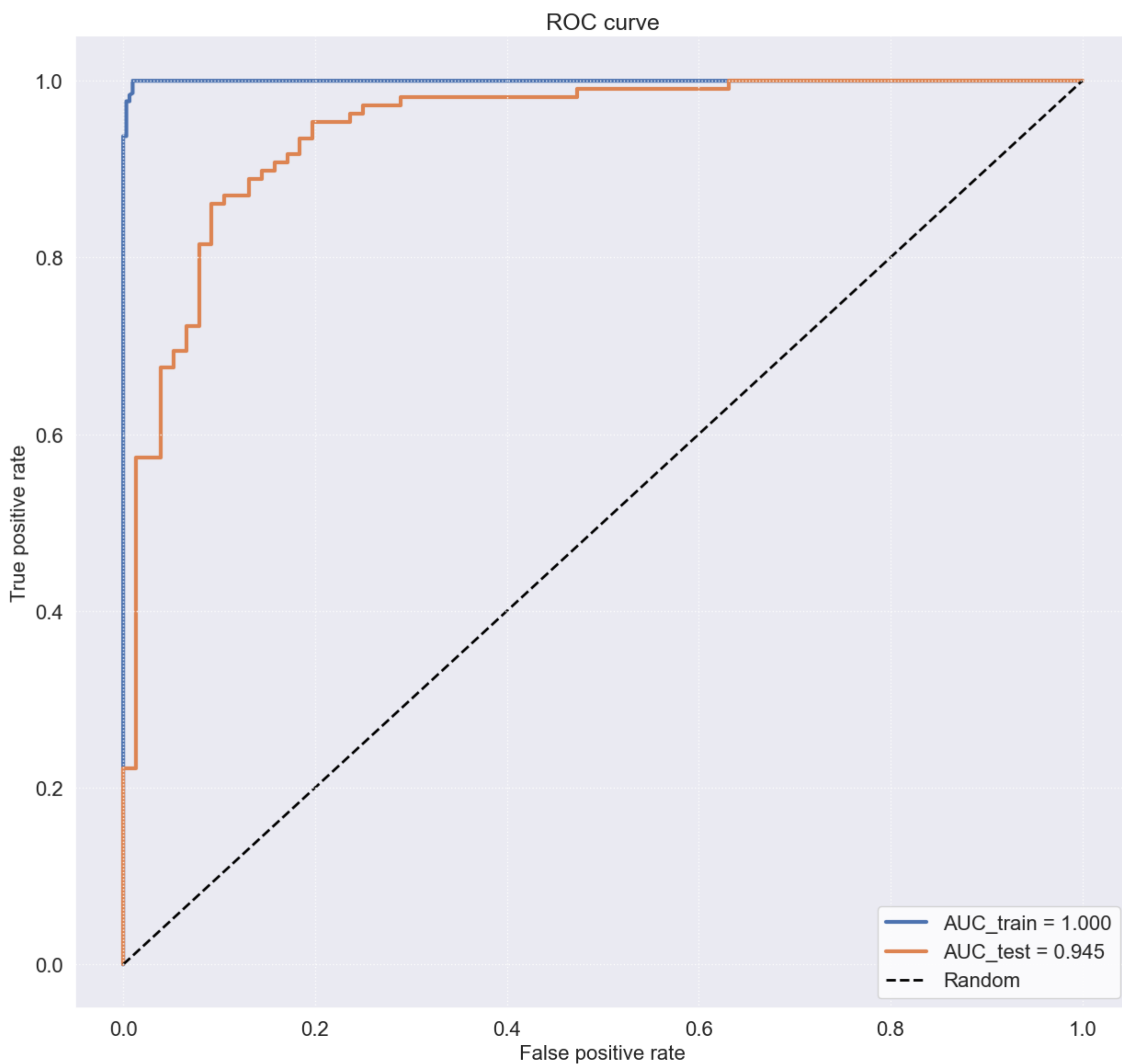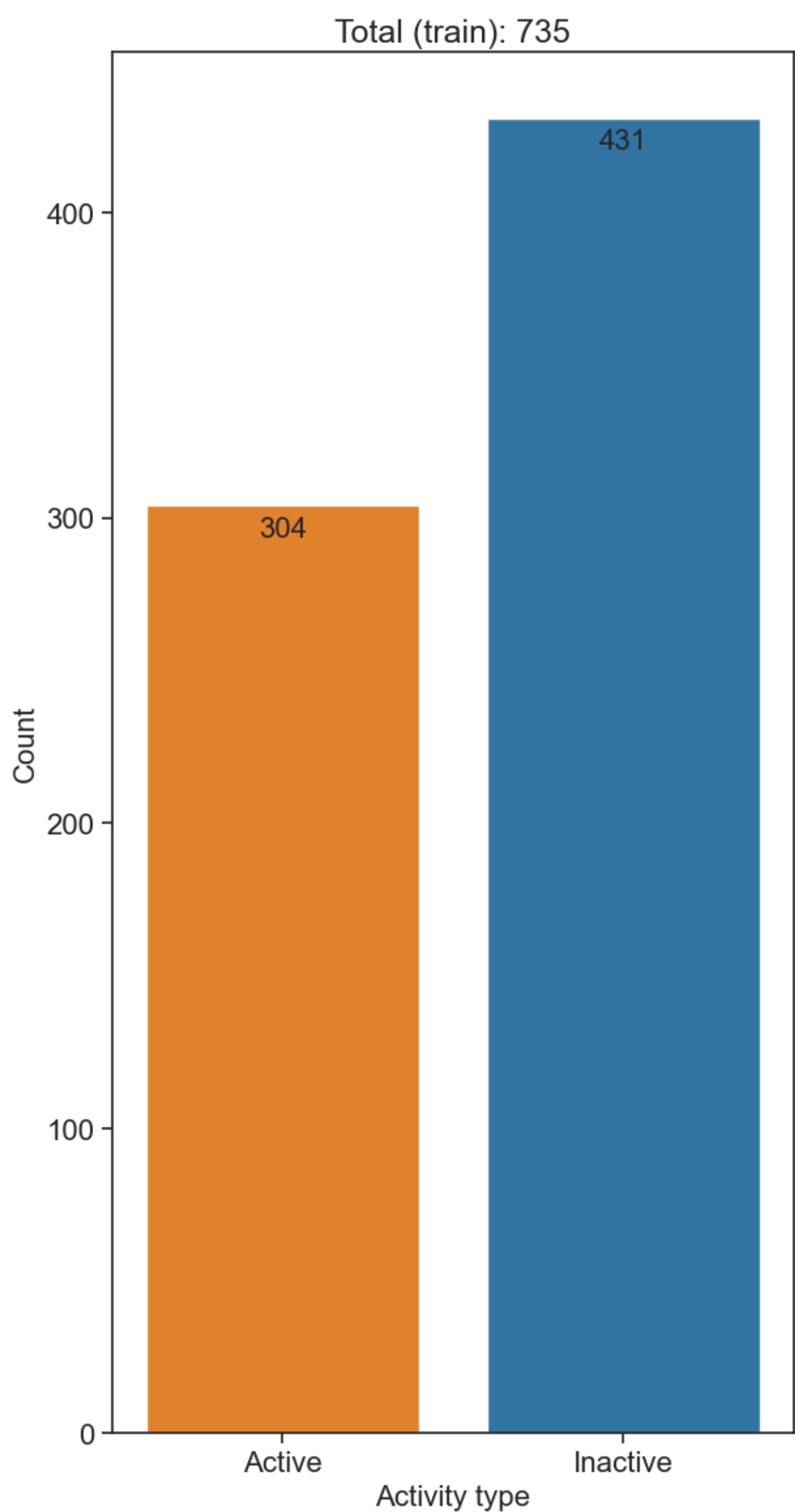
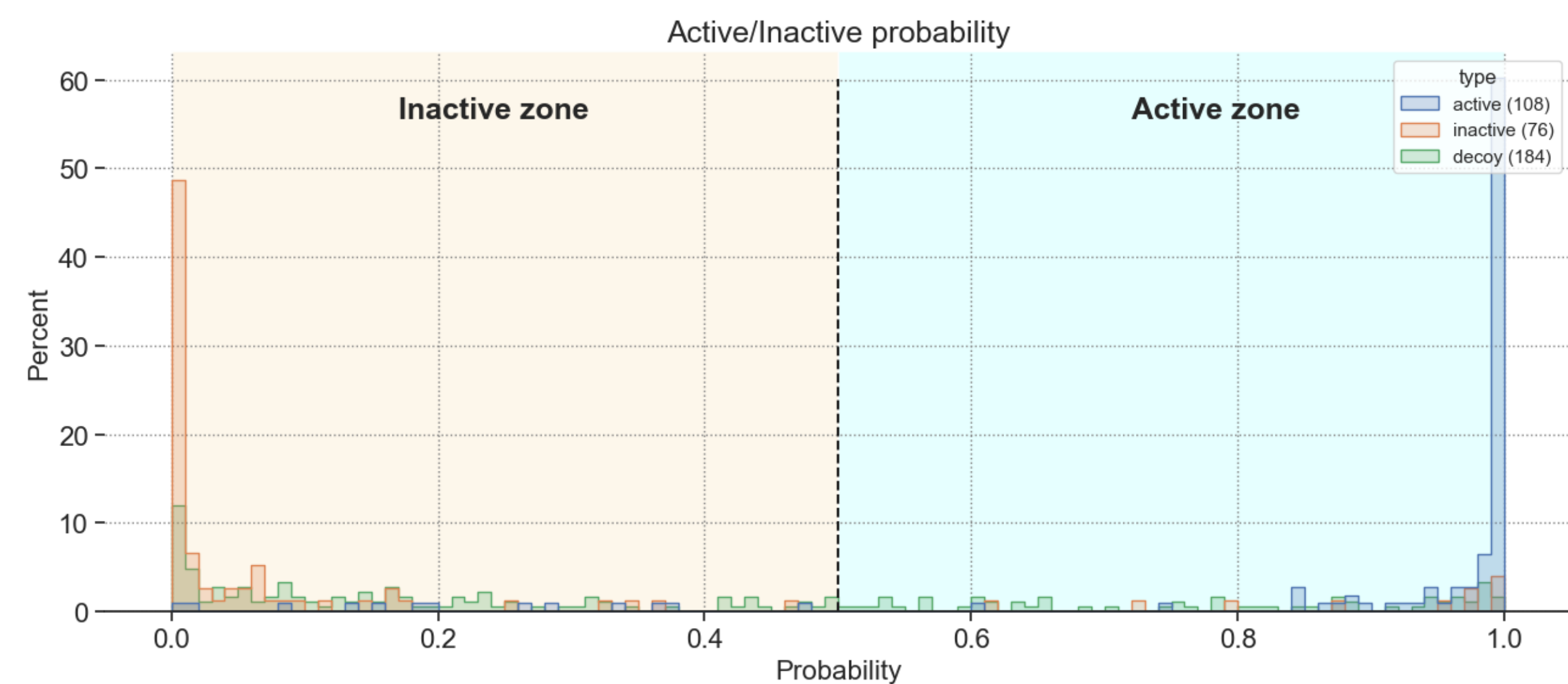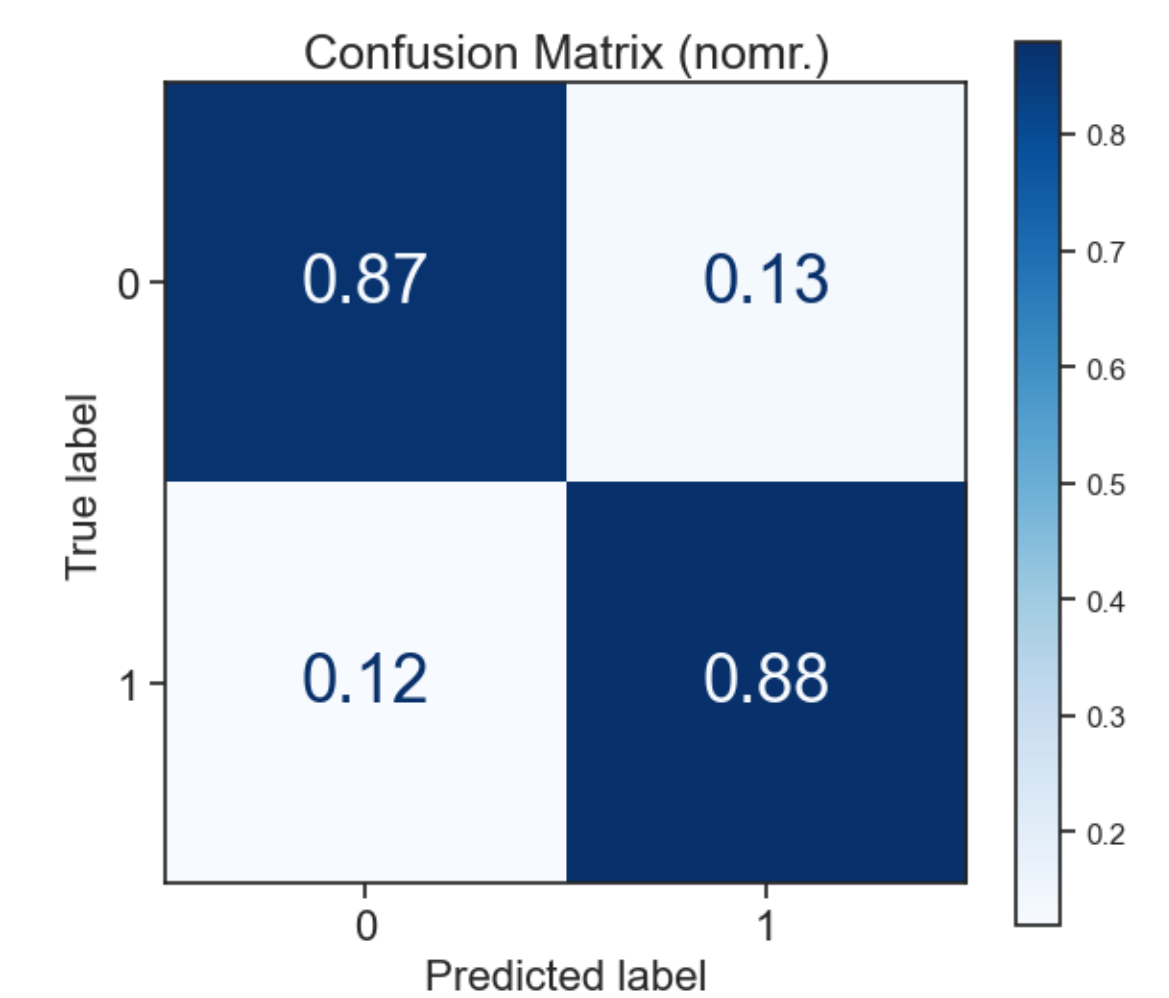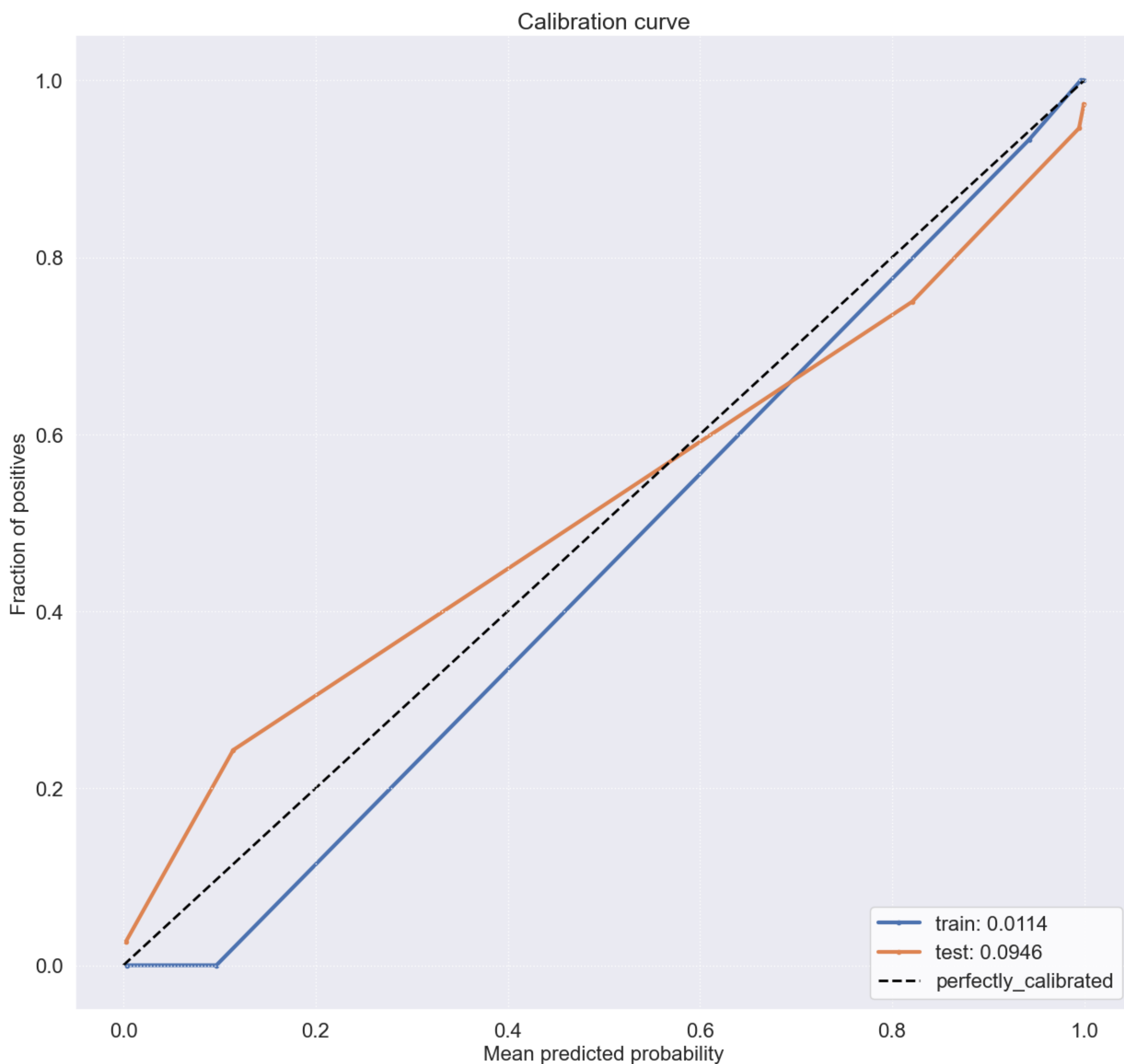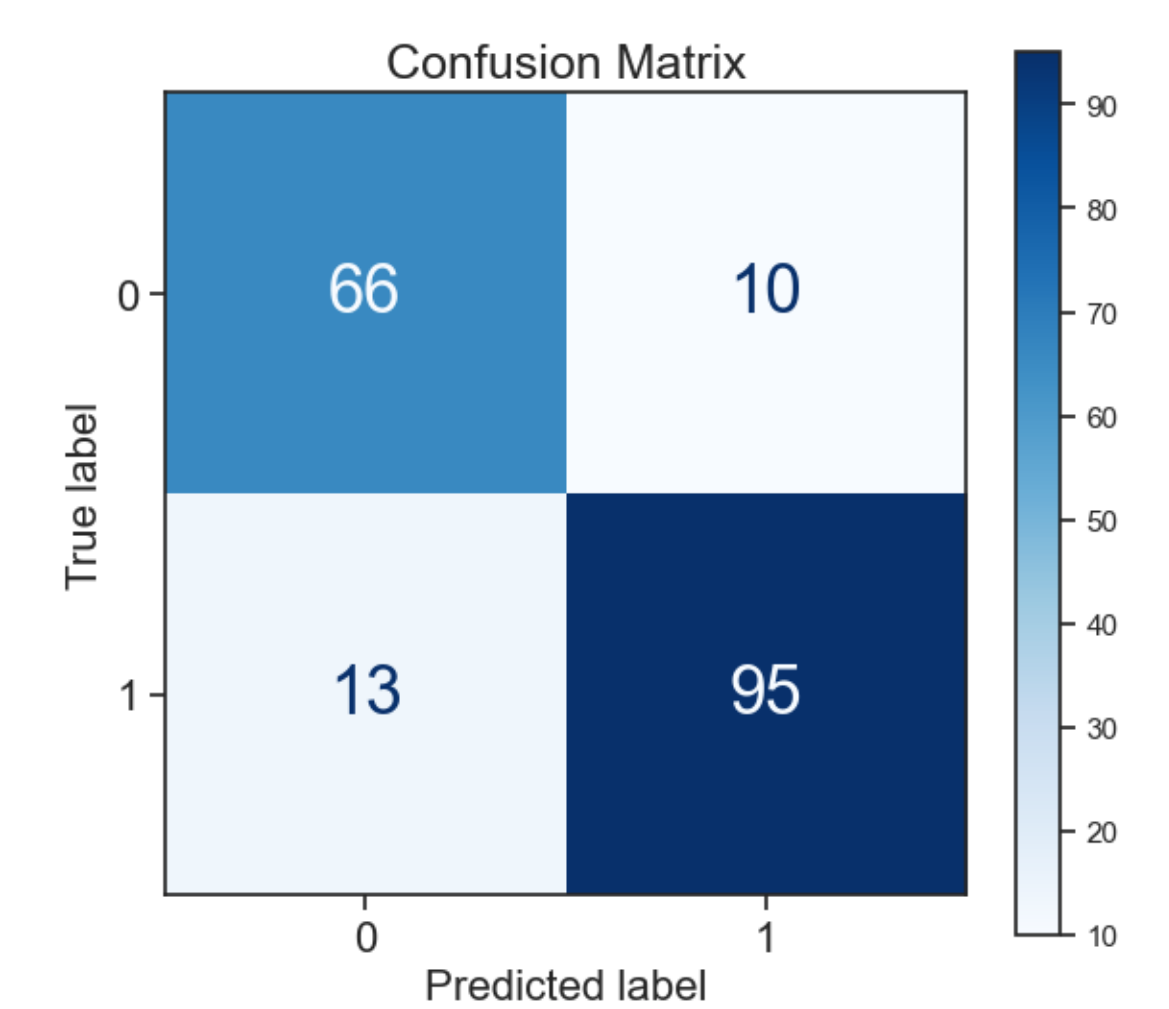| XGBoost Parameters	Value
alpha	1.156
colsample_bytree	0.7552
gamma	0.6145
lambda	4.4763
learning_rate	0.4016
max_depth	7.0
min_child_weight	3.0
n_estimators	570.0
reg_alpha	0.0768
reg_lambda	0.0845
scale_pos_weight	3.7877
subsample	0.8502

Metric	Train	Test
AUC	0.9997	0.9447
accuracy	0.9864	0.875
balanced accuracy	0.9836	0.874
recall_(sens)	1.0	0.8796
specificity	0.9671	0.8684
precision	0.9773	0.9048
f1_score	0.9864	0.8753

Protein: P17787 - metric: f1

Total (train): 735

Active: 304
Inactive: 431

ROC curve

AUC_train = 0.999
AUC_test = 0.952
Random

Pareto Chart with Highlighted Best Attempt

Trials
Pareto Front Trials
Best Trial (# 1268)

XGBoost Parameters
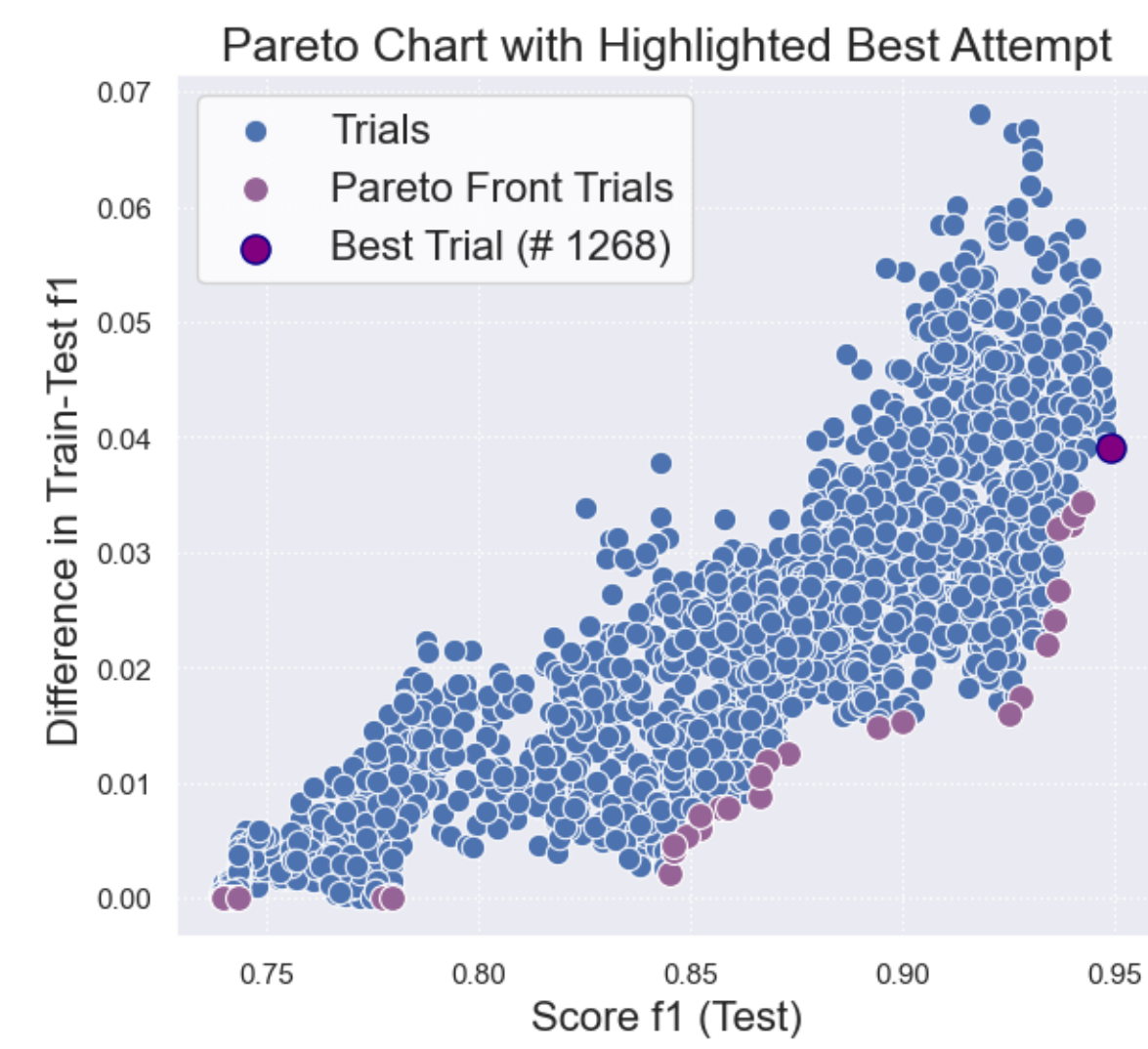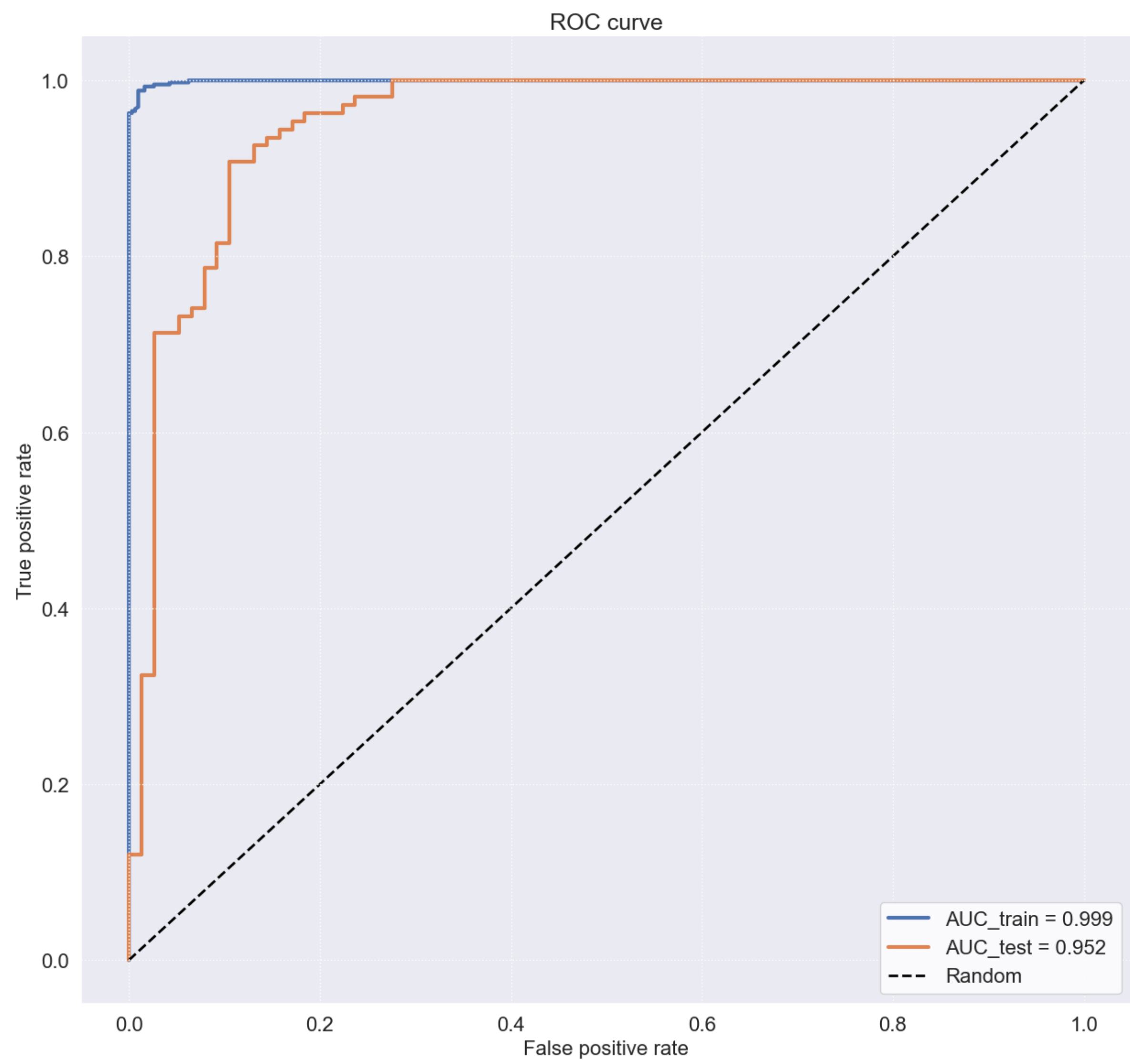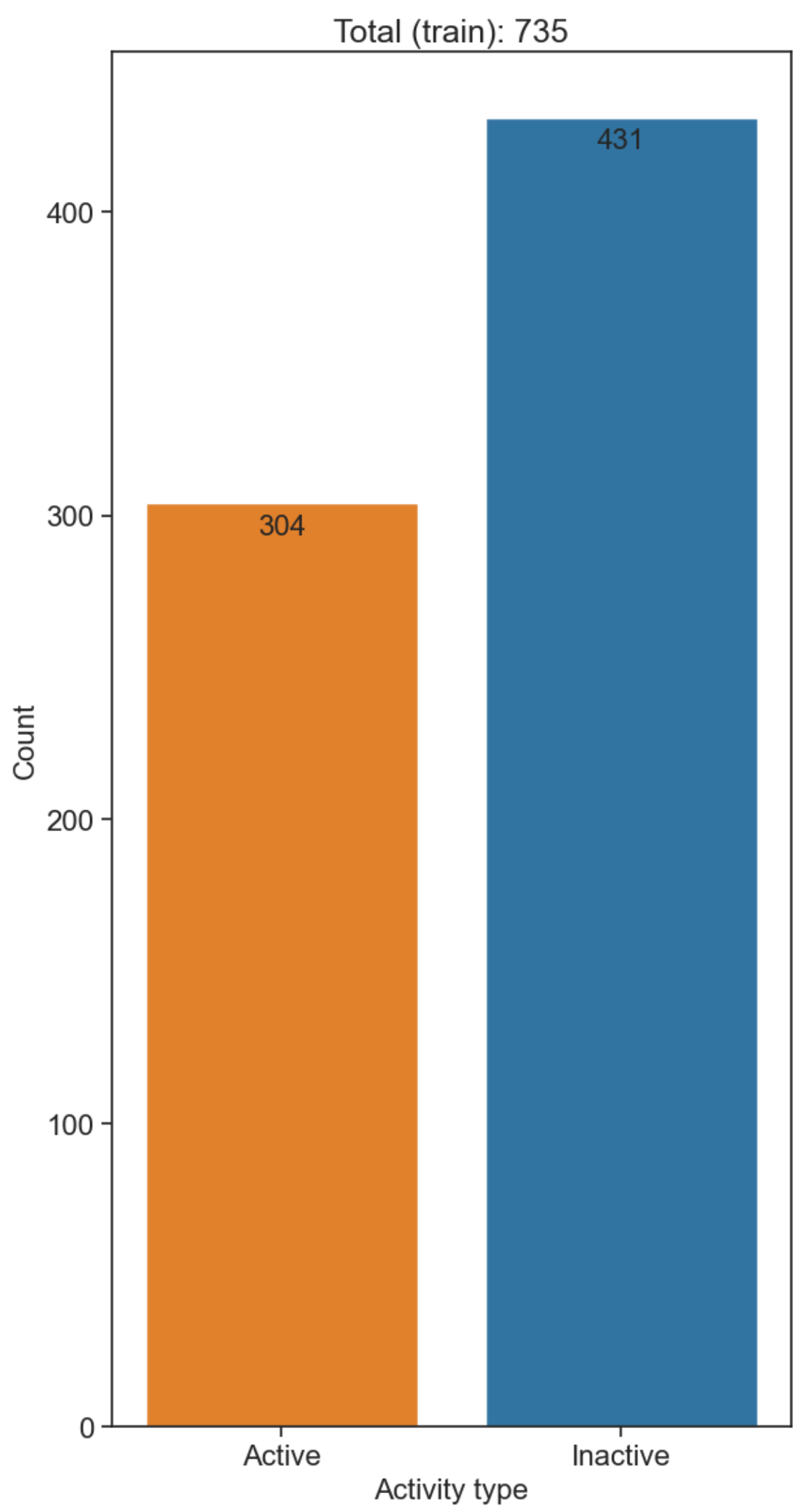
XGBoost Parameters	Value
alpha	4.9843
colsample_bytree	0.7821
gamma	0.0217
lambda	2.419
learning_rate	0.0352
max_depth	10.0
min_child_weight	1.0
n_estimators	266.0
reg_alpha	0.9045
reg_lambda	0.6381
scale_pos_weight	1.2428
subsample	0.7009

Confusion Matrix

	0	1
0	66	10
1	10	98

Confusion Matrix (nomr.)

	0	1
0	0.87	0.13
1	0.093	0.91

Calibration curve

train: 0.0167
test: 0.0811
perfectly_calibrated

Active/Inactive probability

Inactive zone
Active zone

type
active (108)
inactive (76)
decoy (184)
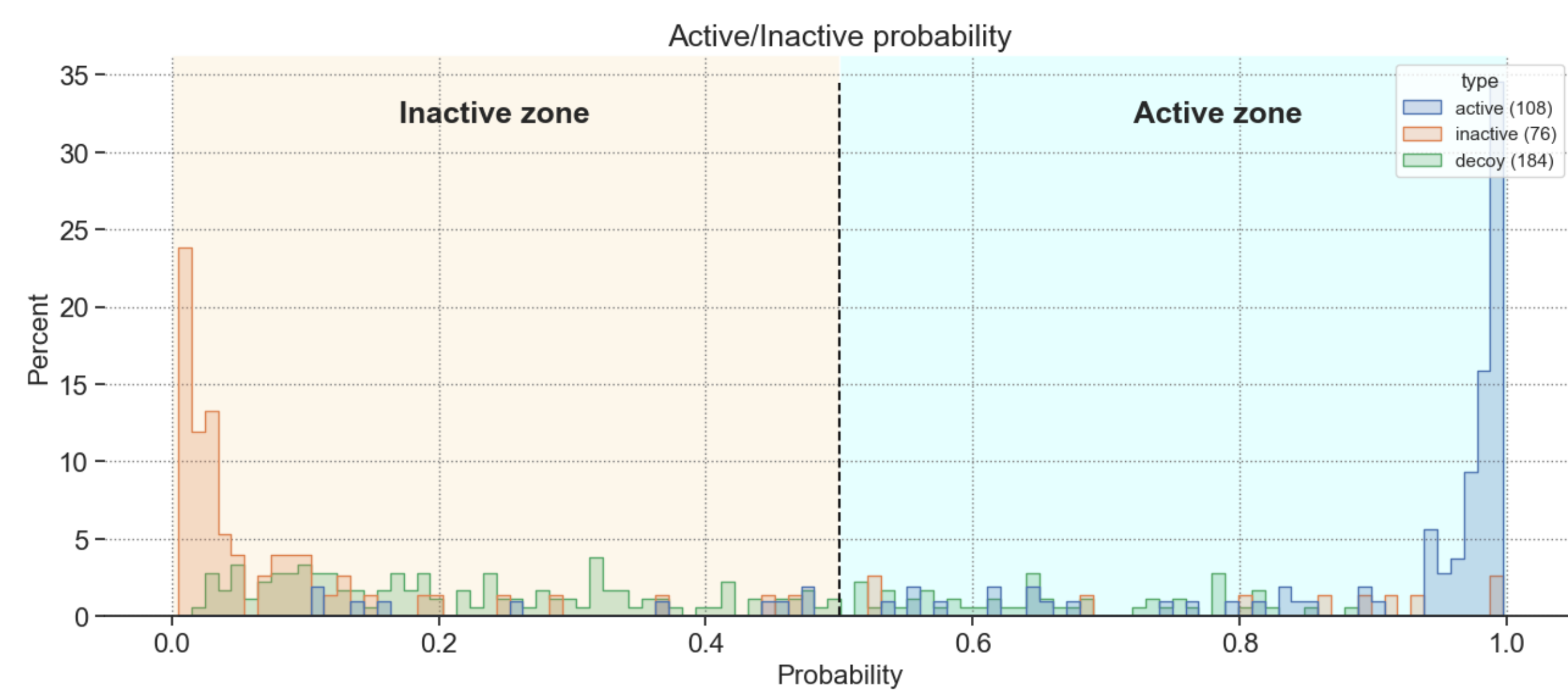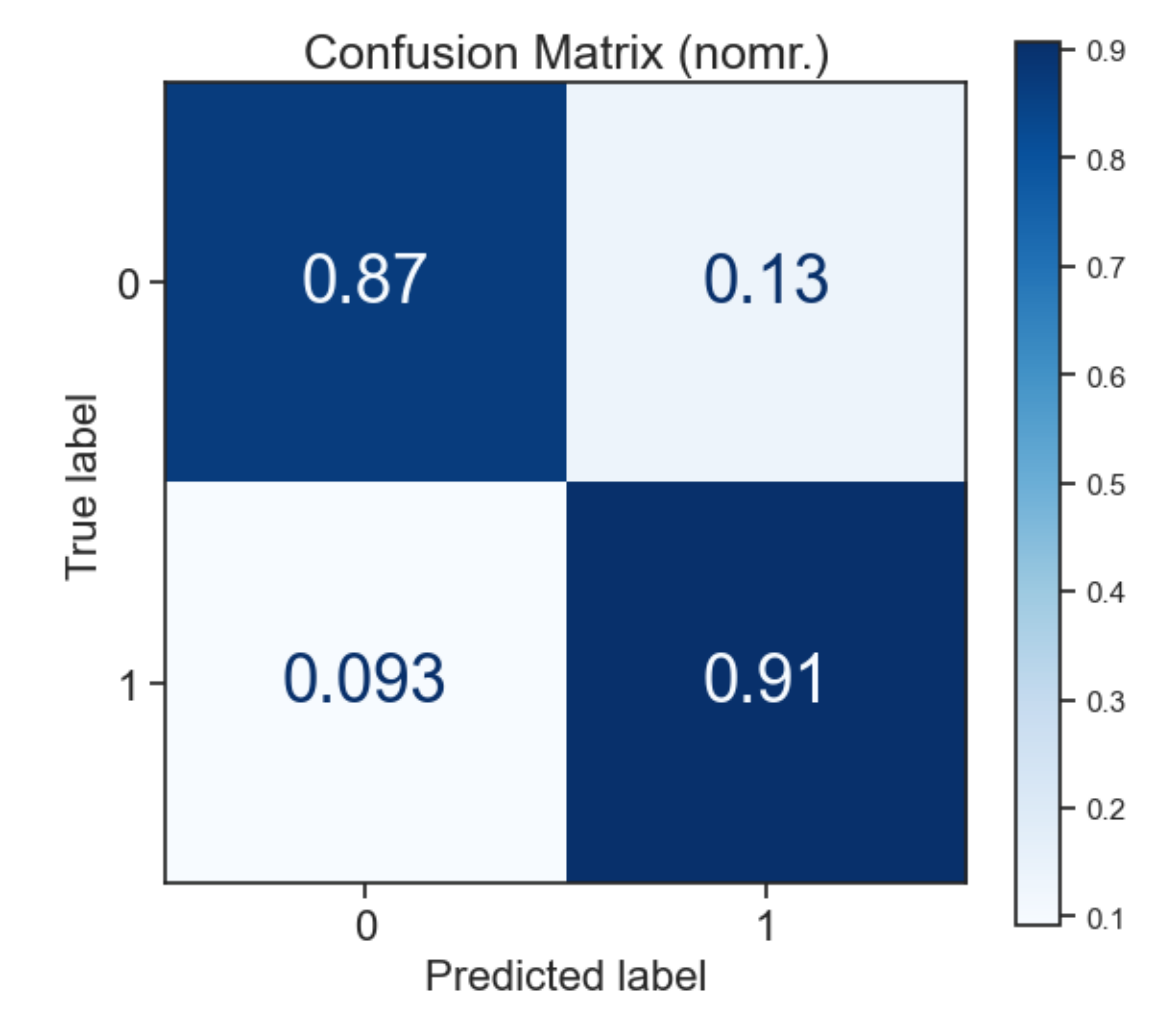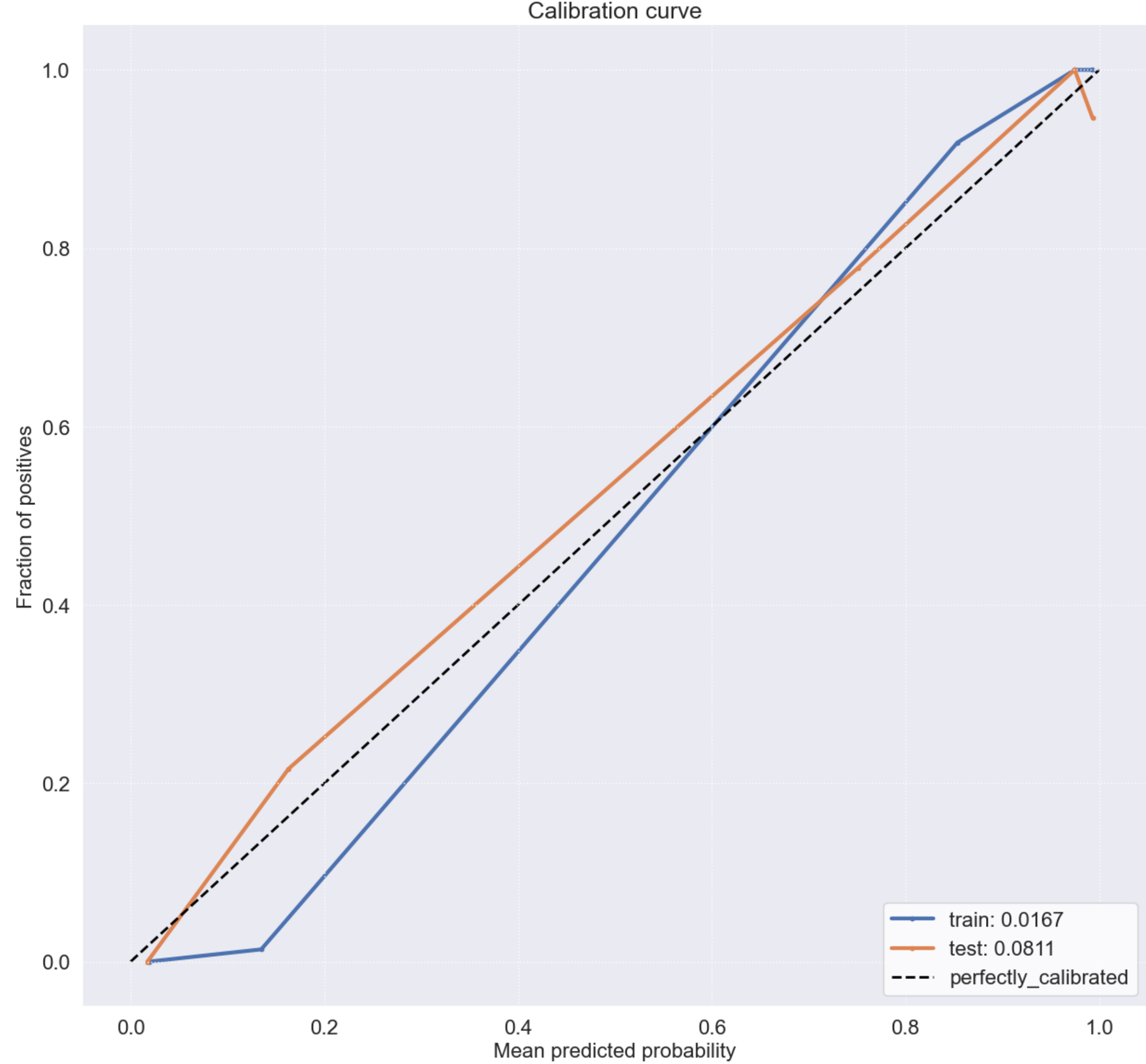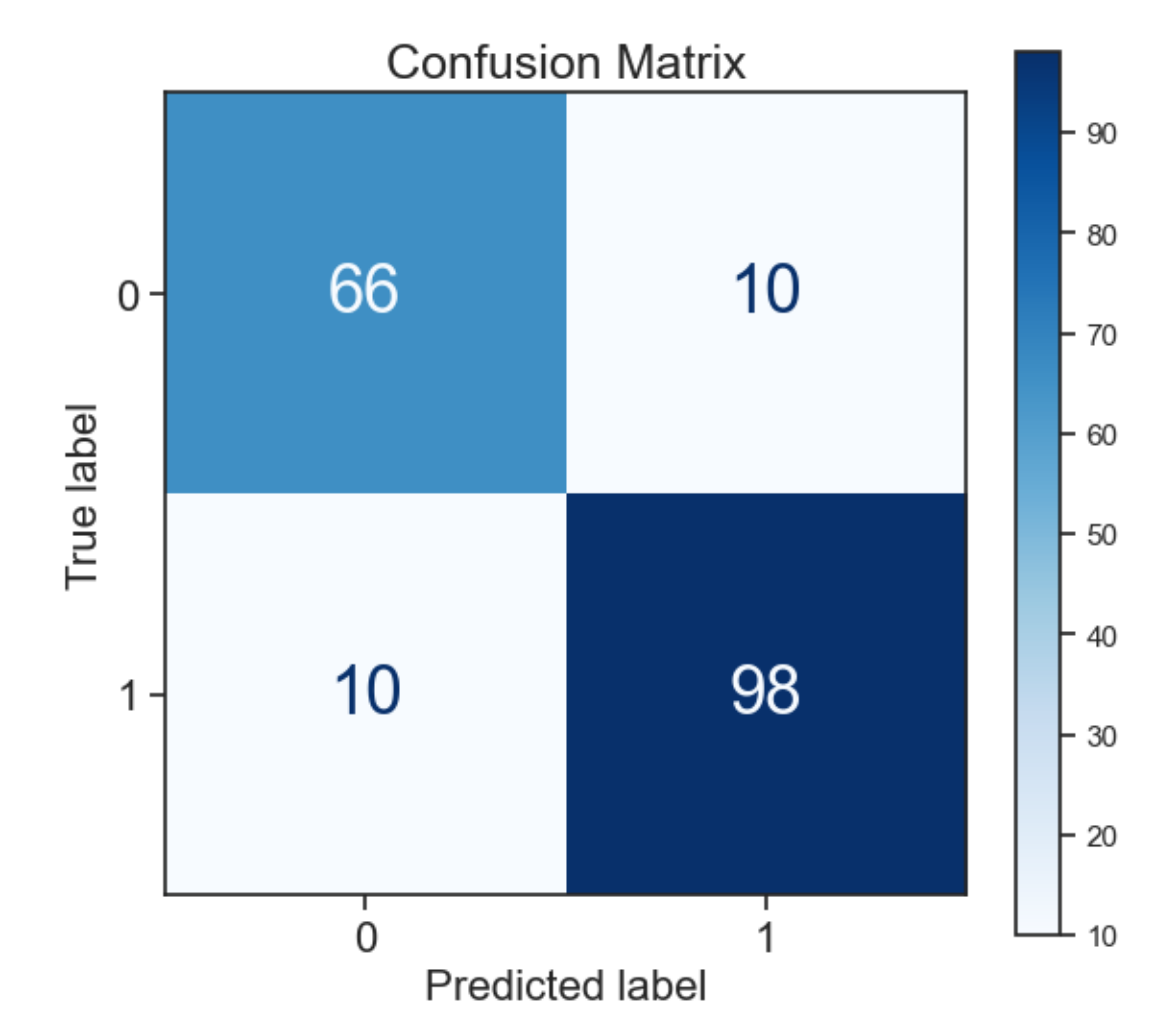
Metric	Train	Test
AUC	0.9994	0.9516
accuracy	0.985	0.8913
balanced accuracy	0.9829	0.8879
recall_(sens)	0.9954	0.9074
specificity	0.9704	0.8684
precision	0.9795	0.9074
f1_score	0.985	0.8913

Protein: P17787 - metric: precision

Total (train): 735

ROC curve

AUC_train = 0.997
AUC_test = 0.953
Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 1630)
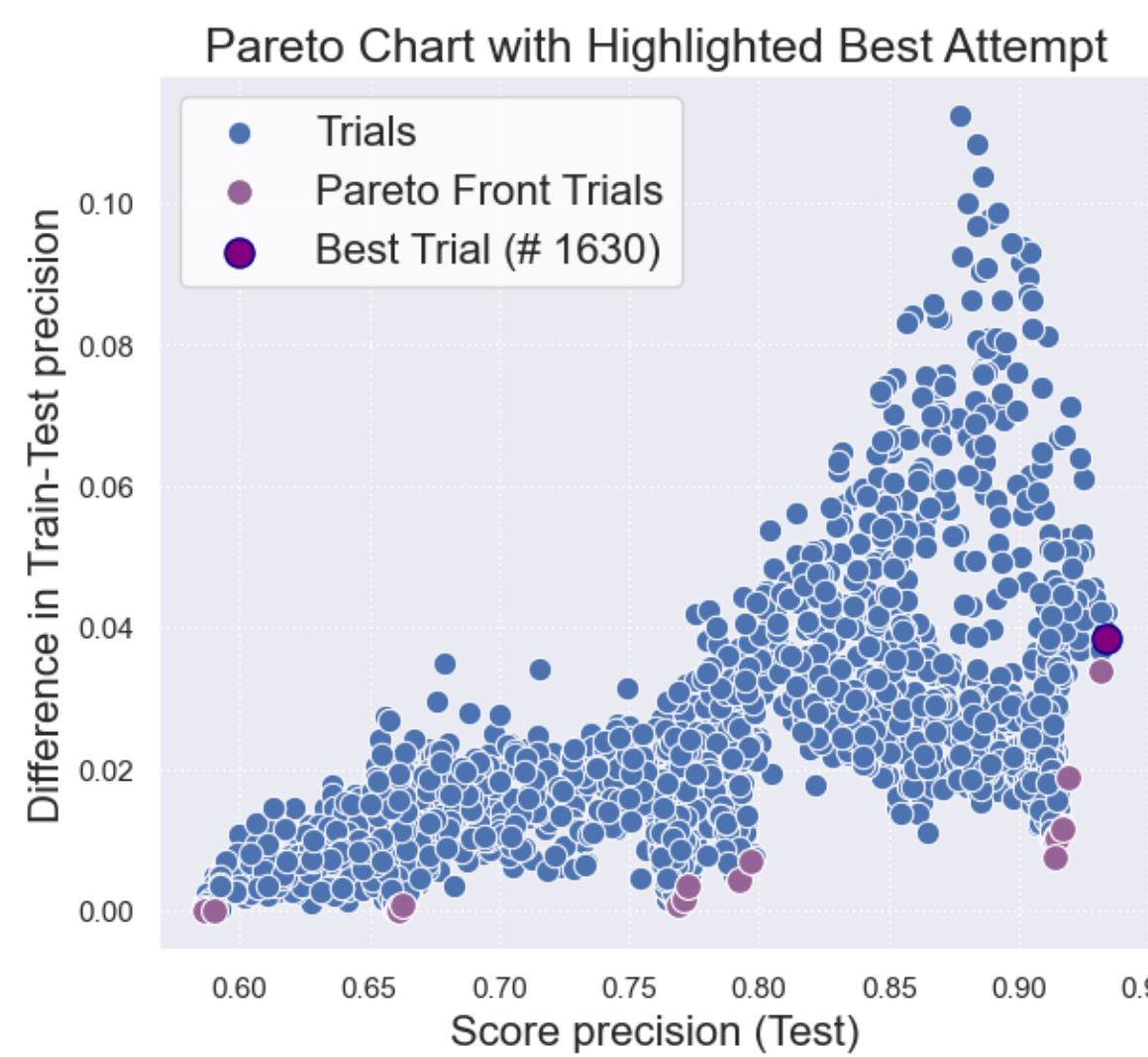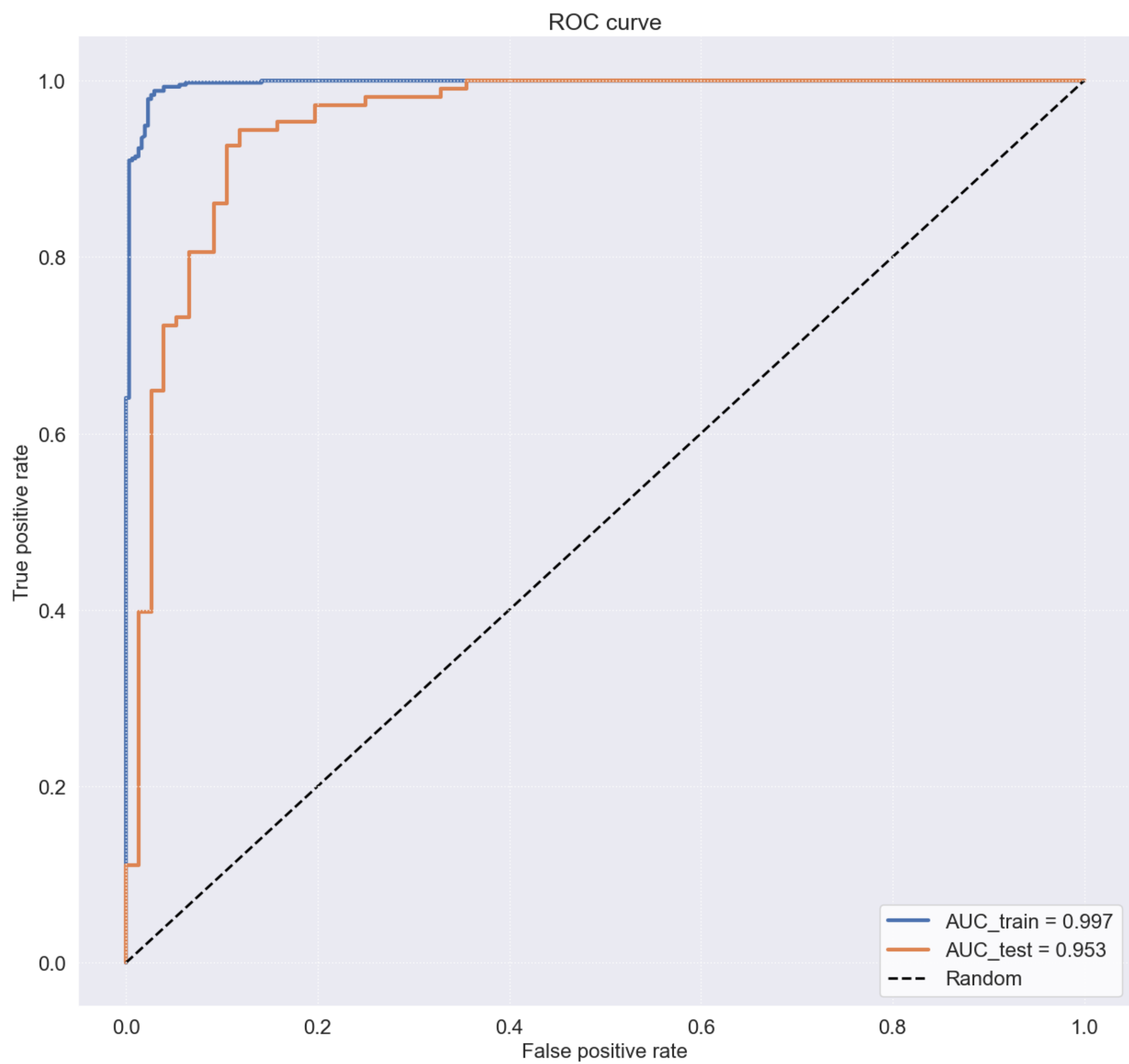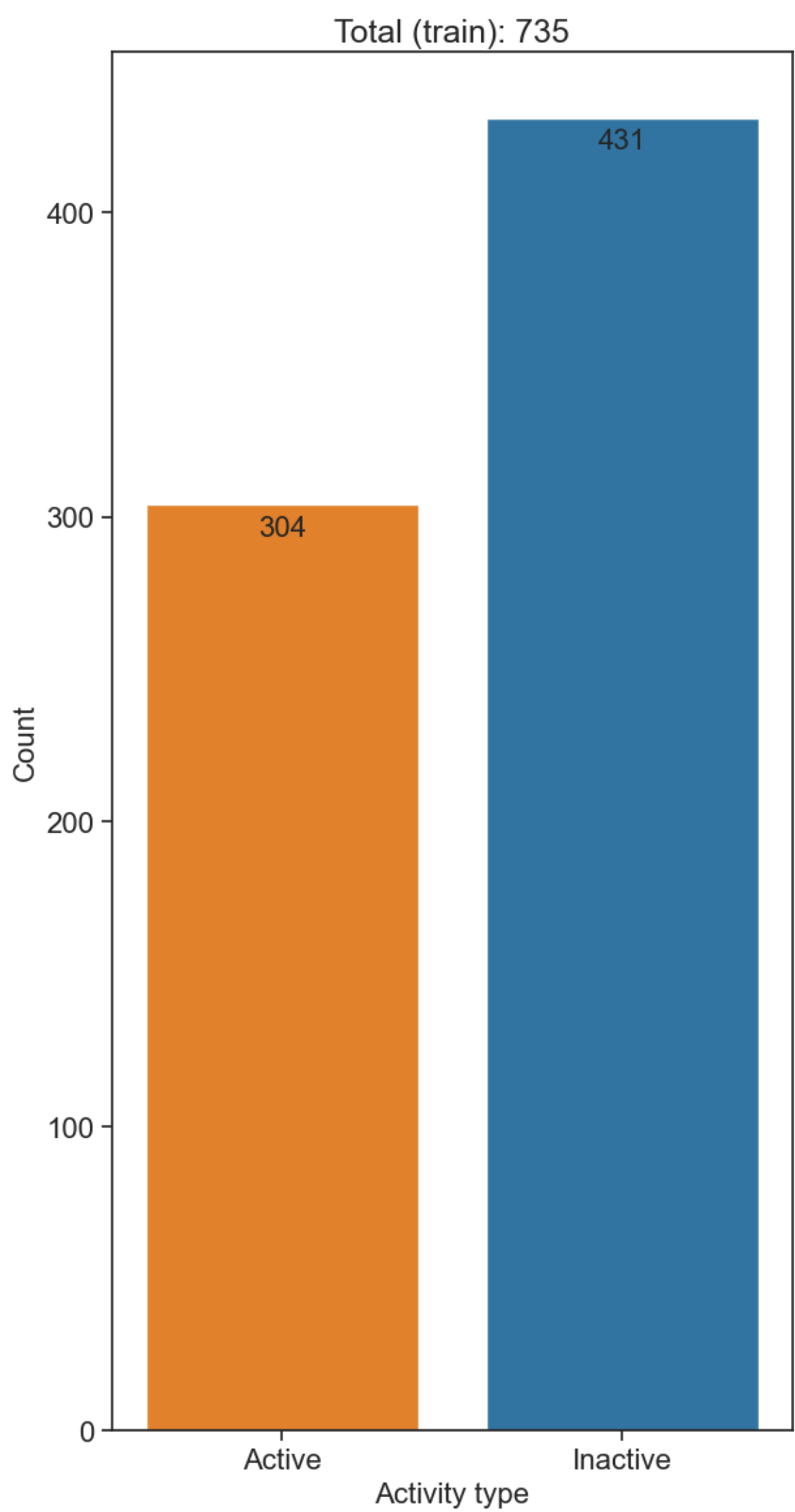
XGBoost Parameters	Value
alpha	4.4523
colsample_bytree	0.6271
gamma	0.0387
lambda	4.4874
learning_rate	0.0552
max_depth	3.0
min_child_weight	2.0
n_estimators	861.0
reg_alpha	0.5181
reg_lambda	0.6903
scale_pos_weight	1.0795
subsample	0.5406

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0232
test: 0.078
perfectly_calibrated

Active/Inactive probability

type
active (108)
inactive (76)
decoy (184)
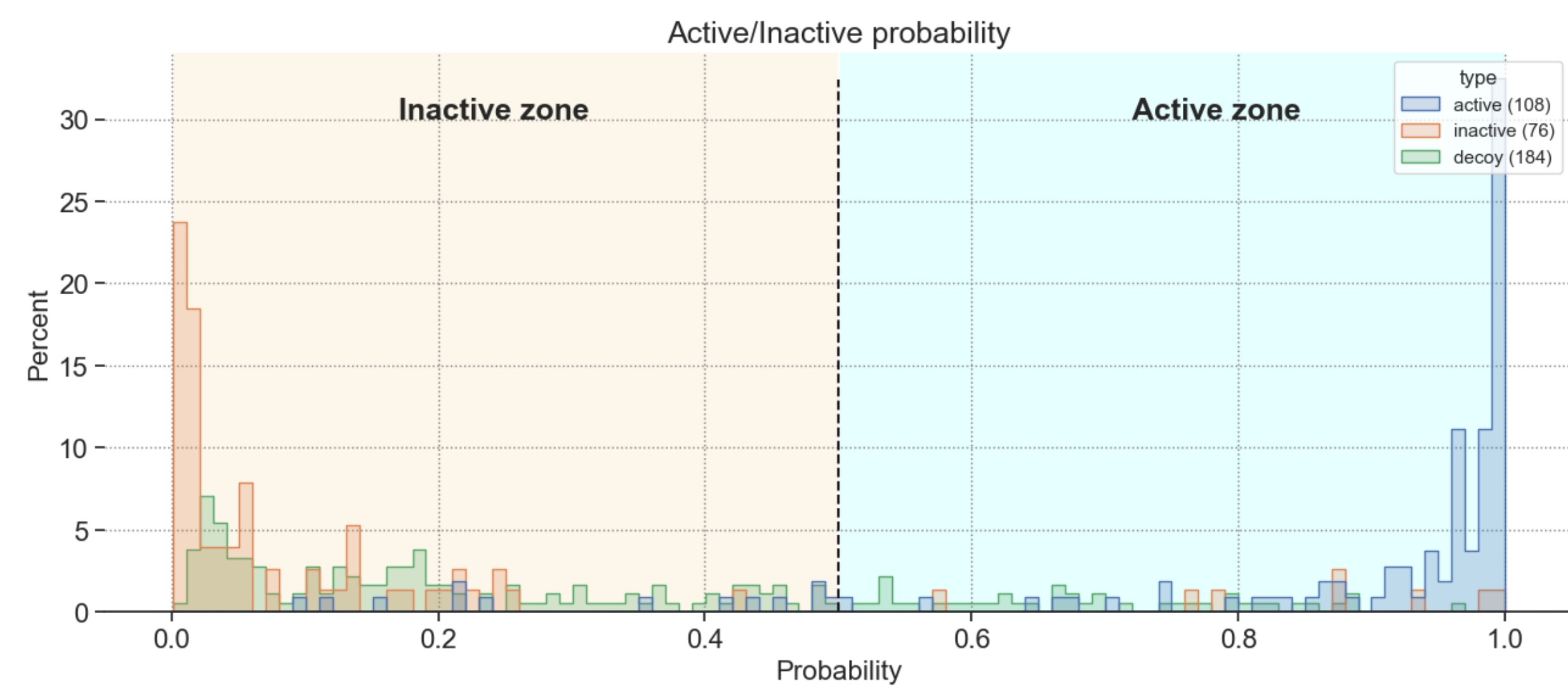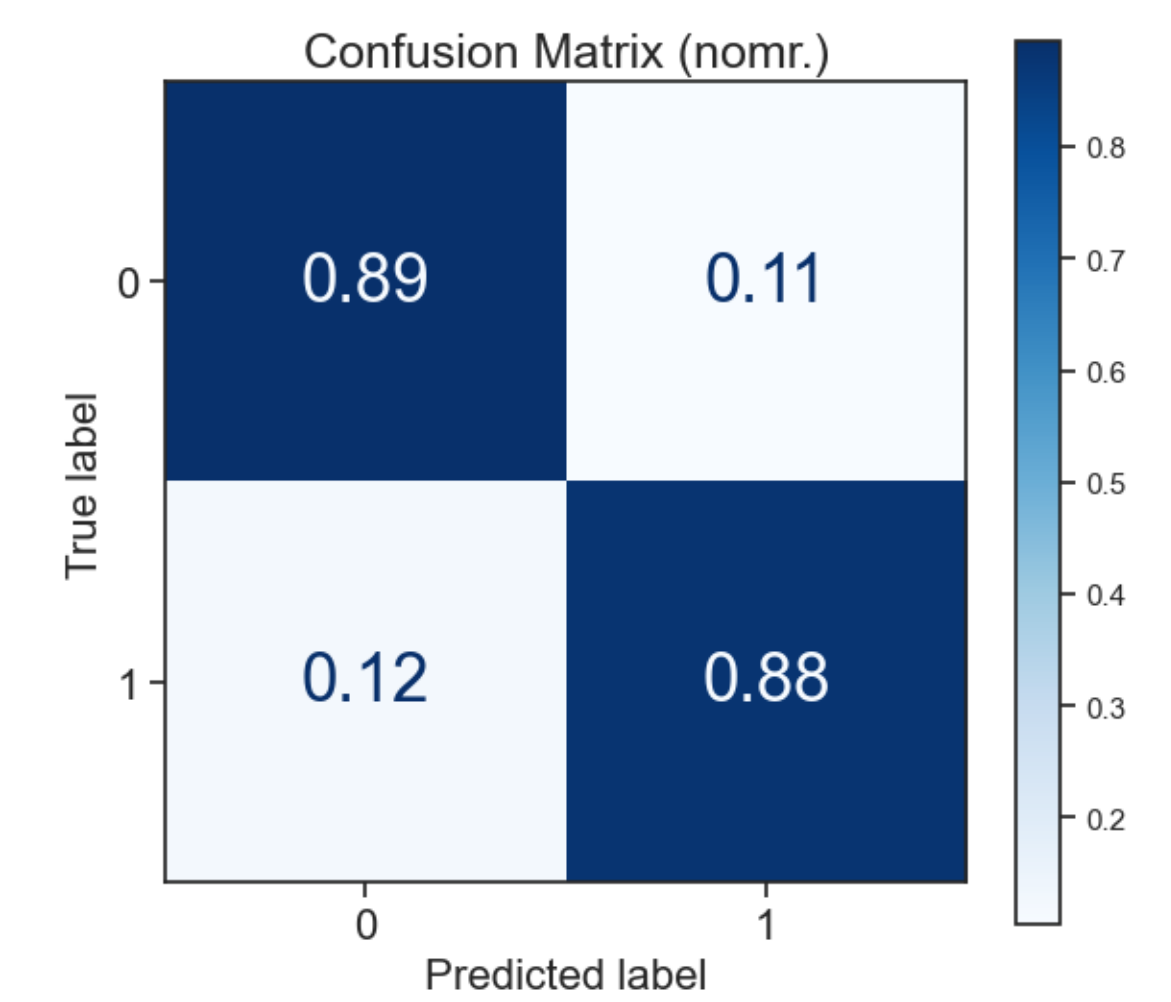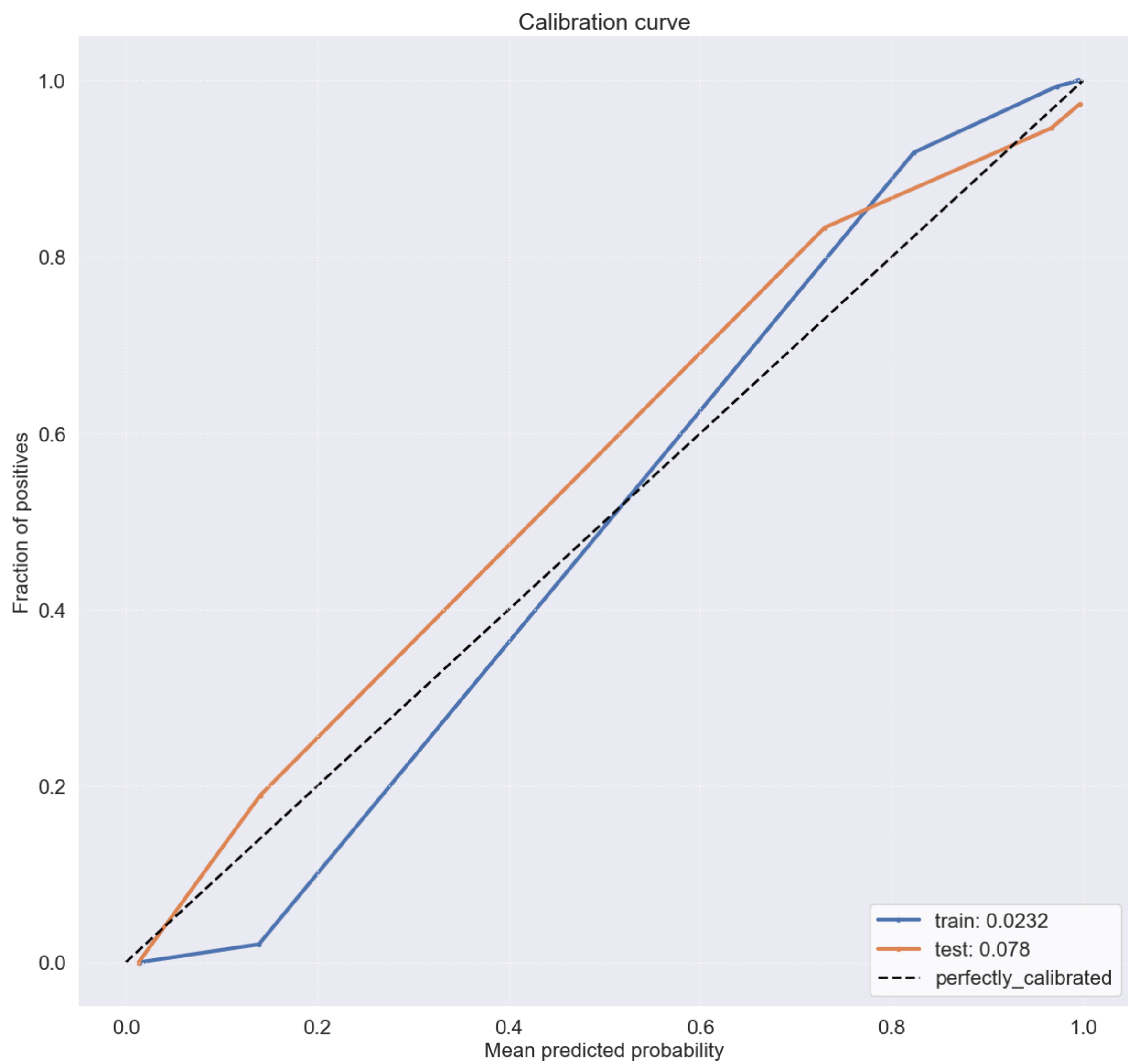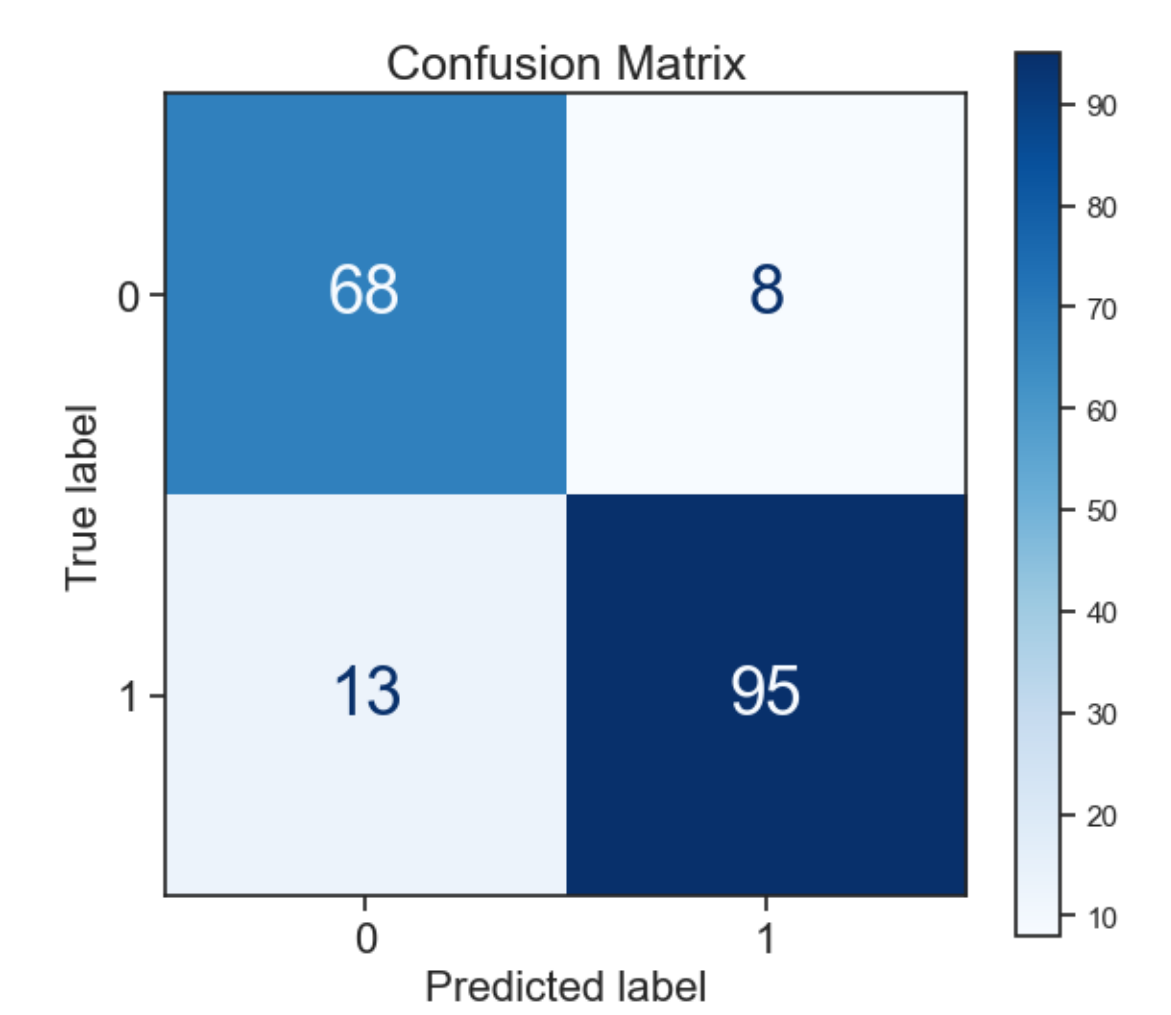
Metric	Train	Test
AUC	0.9968	0.9535
accuracy	0.9796	0.8859
balanced accuracy	0.9778	0.8872
recall_(sens)	0.9884	0.8796
specificity	0.9671	0.8947
precision	0.9771	0.9223
f1_score	0.9796	0.8863

Protein: P17787 - metric: SoftVotingXGBoostEnsemble
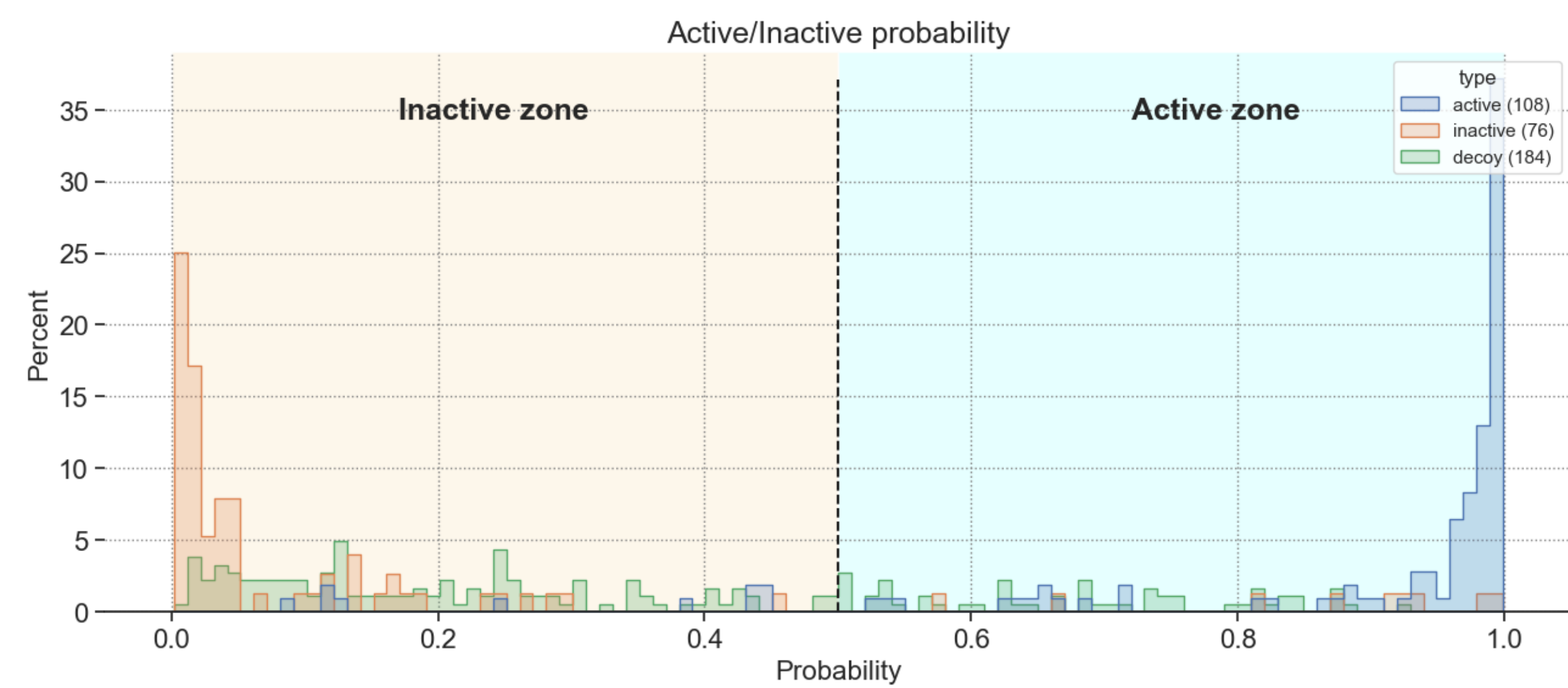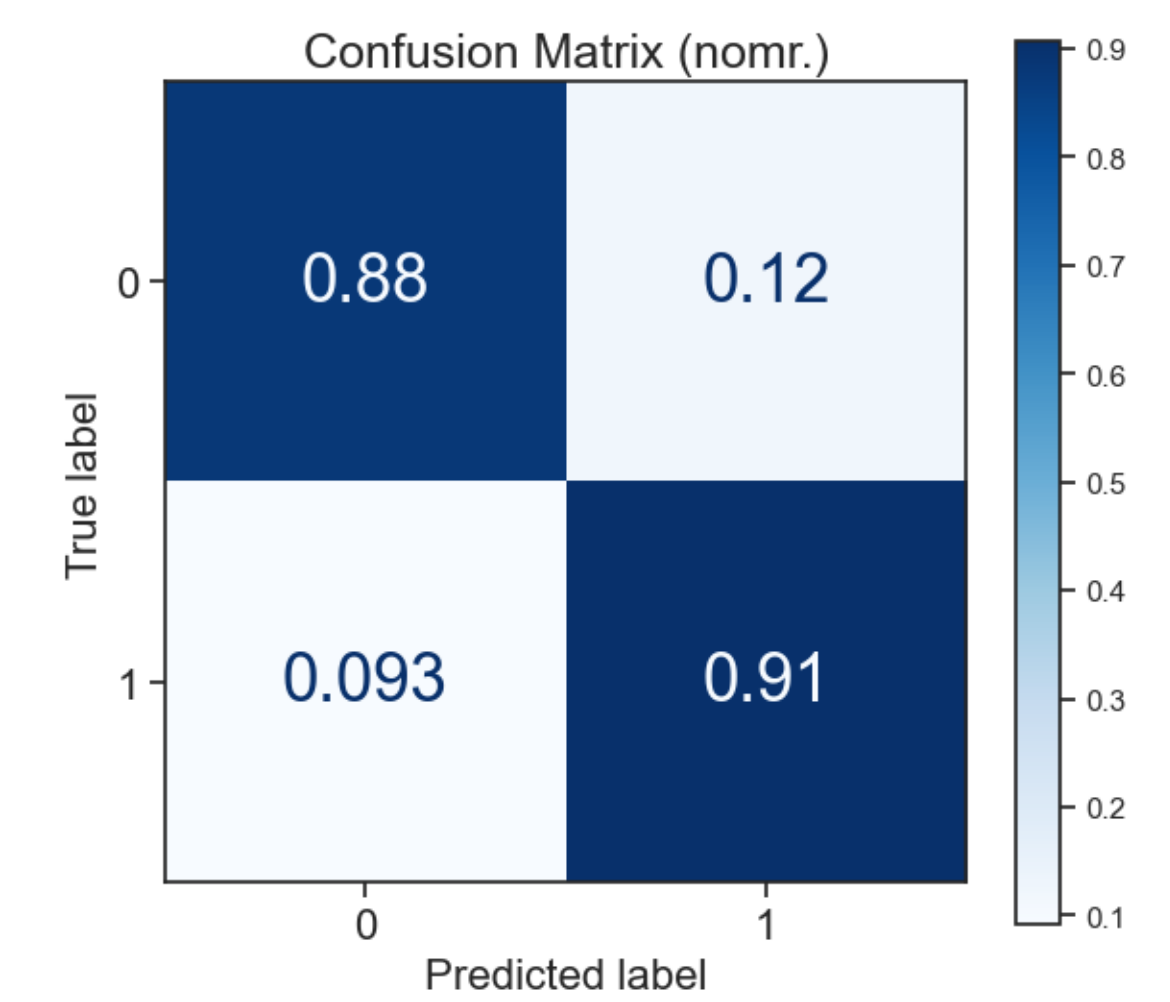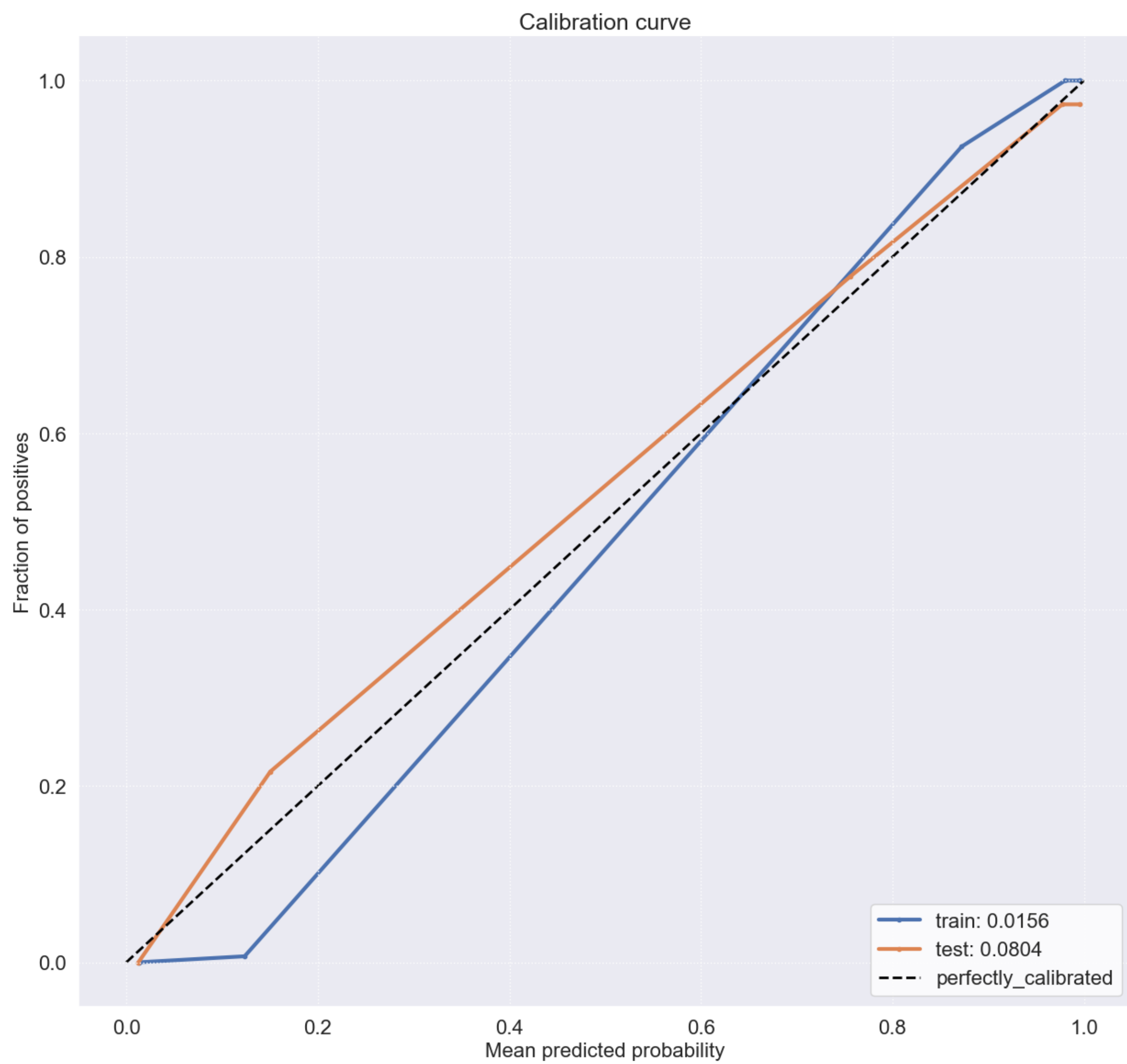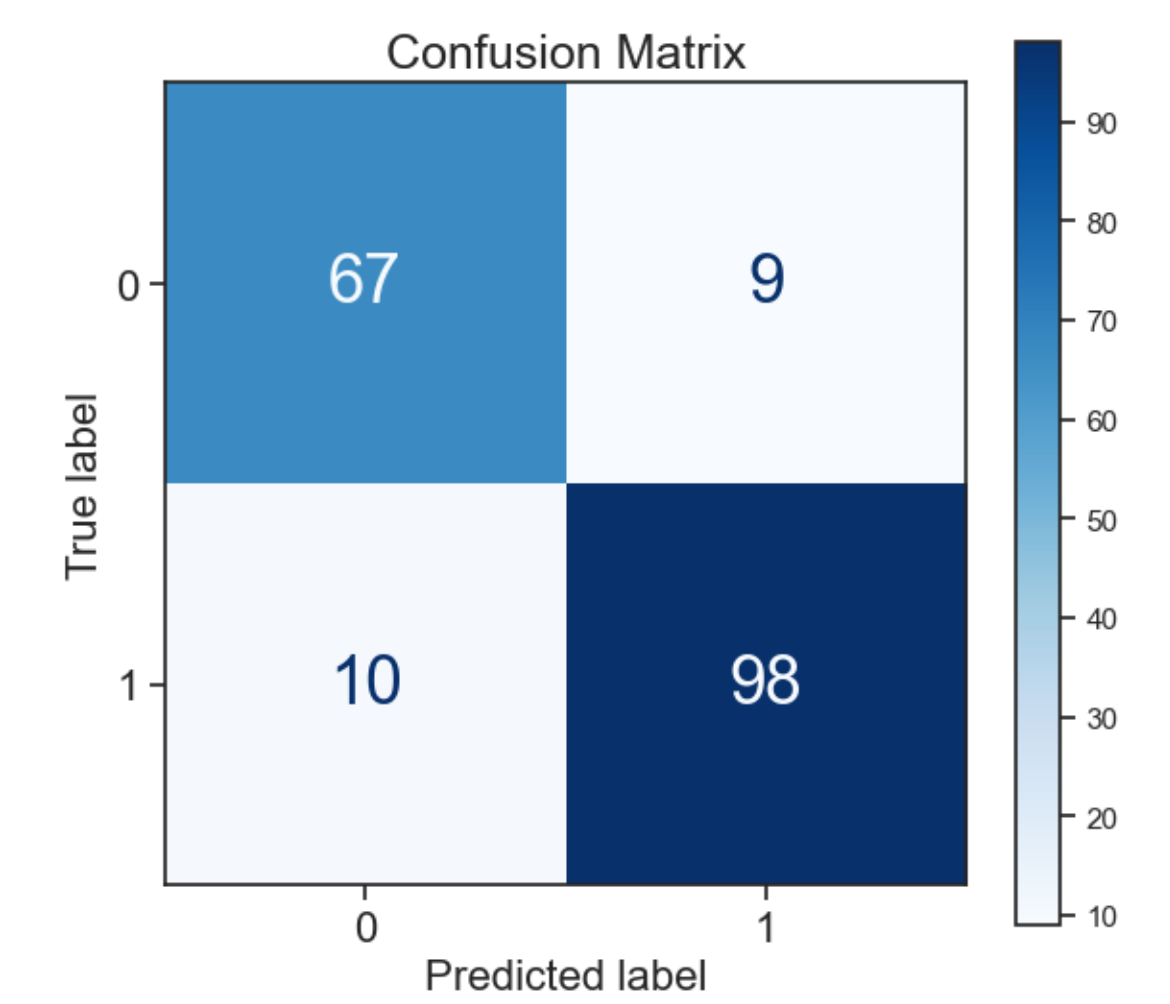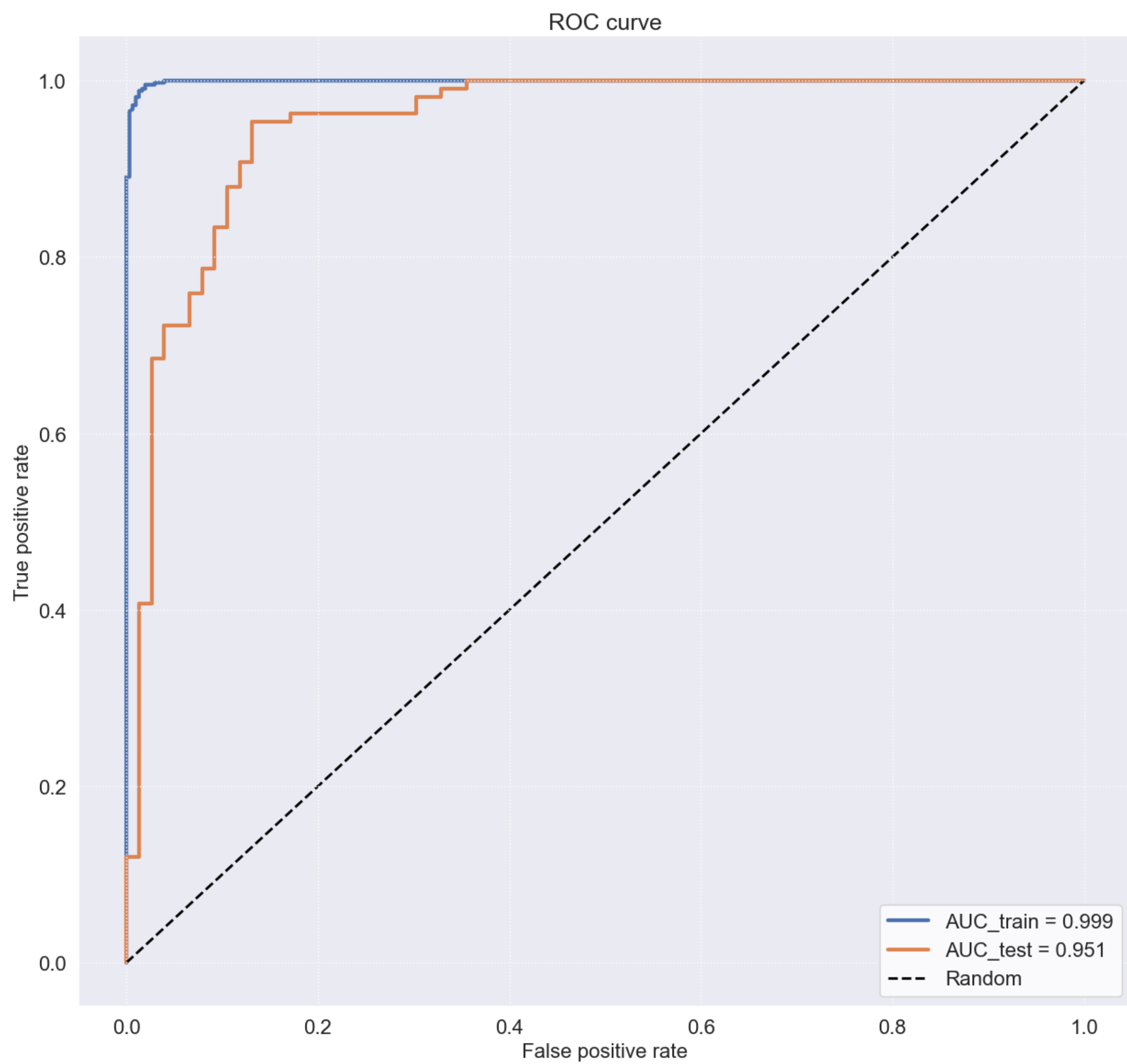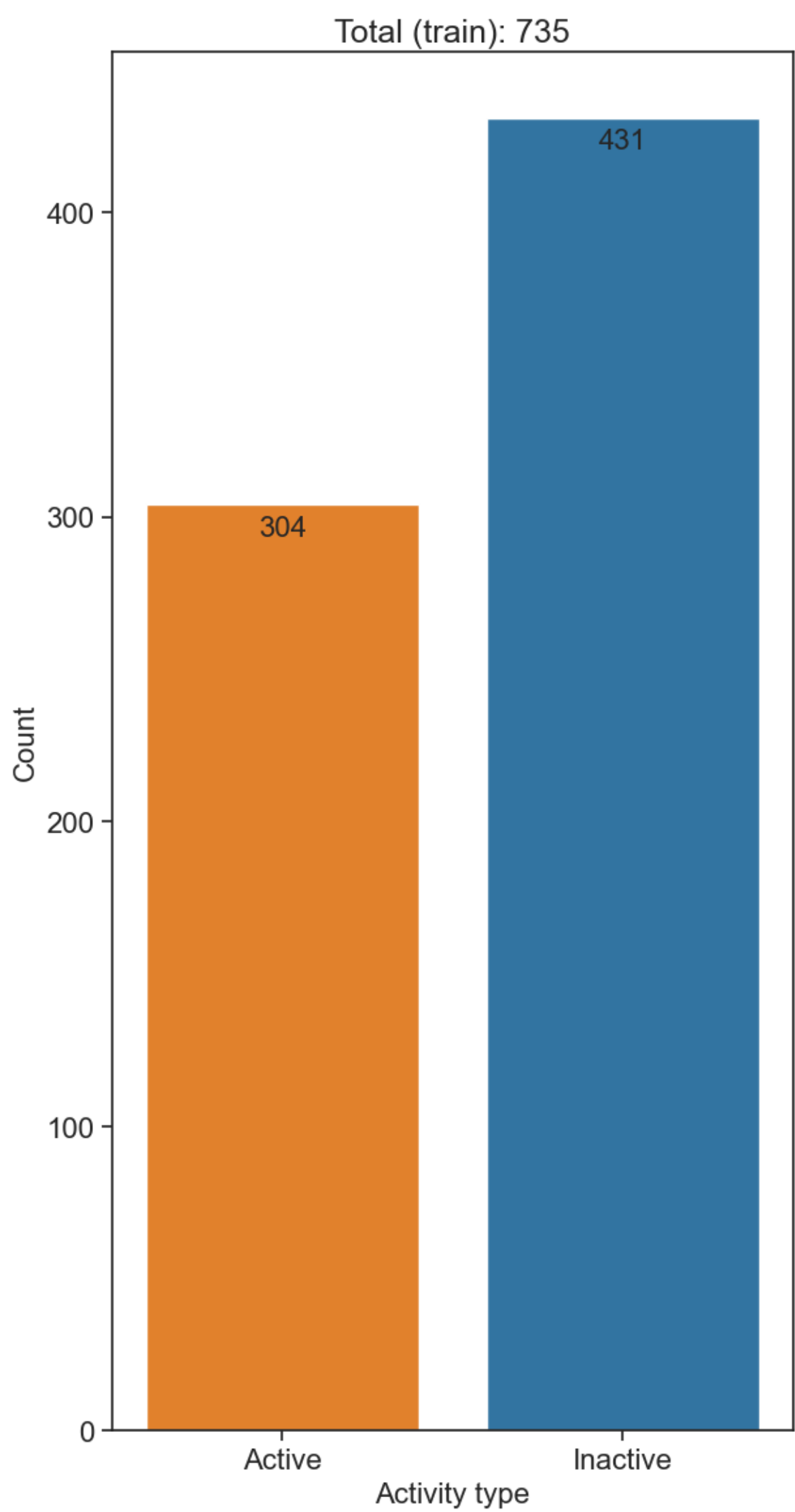
Metric	Train	Test
AUC	0.9992	0.9508
accuracy	0.985	0.8967
balanced accuracy	0.9824	0.8945
recall_(sens)	0.9977	0.9074
specificity	0.9671	0.8816
precision	0.9773	0.9159
f1_score	0.985	0.8968