Protein: O15379 - metric: balanced_accuracy

Total (train): 180

ROC curve

AUC_train = 1.000
AUC_test = 0.916
Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
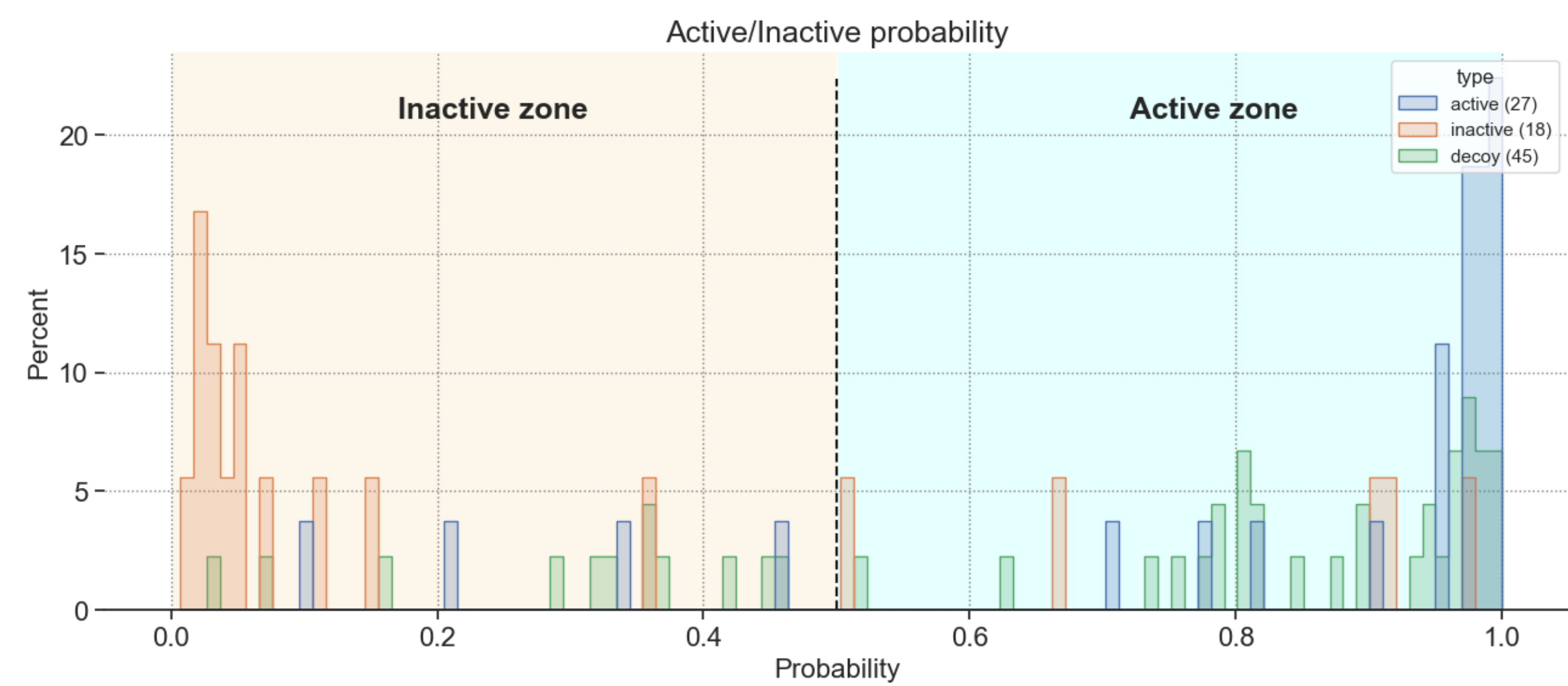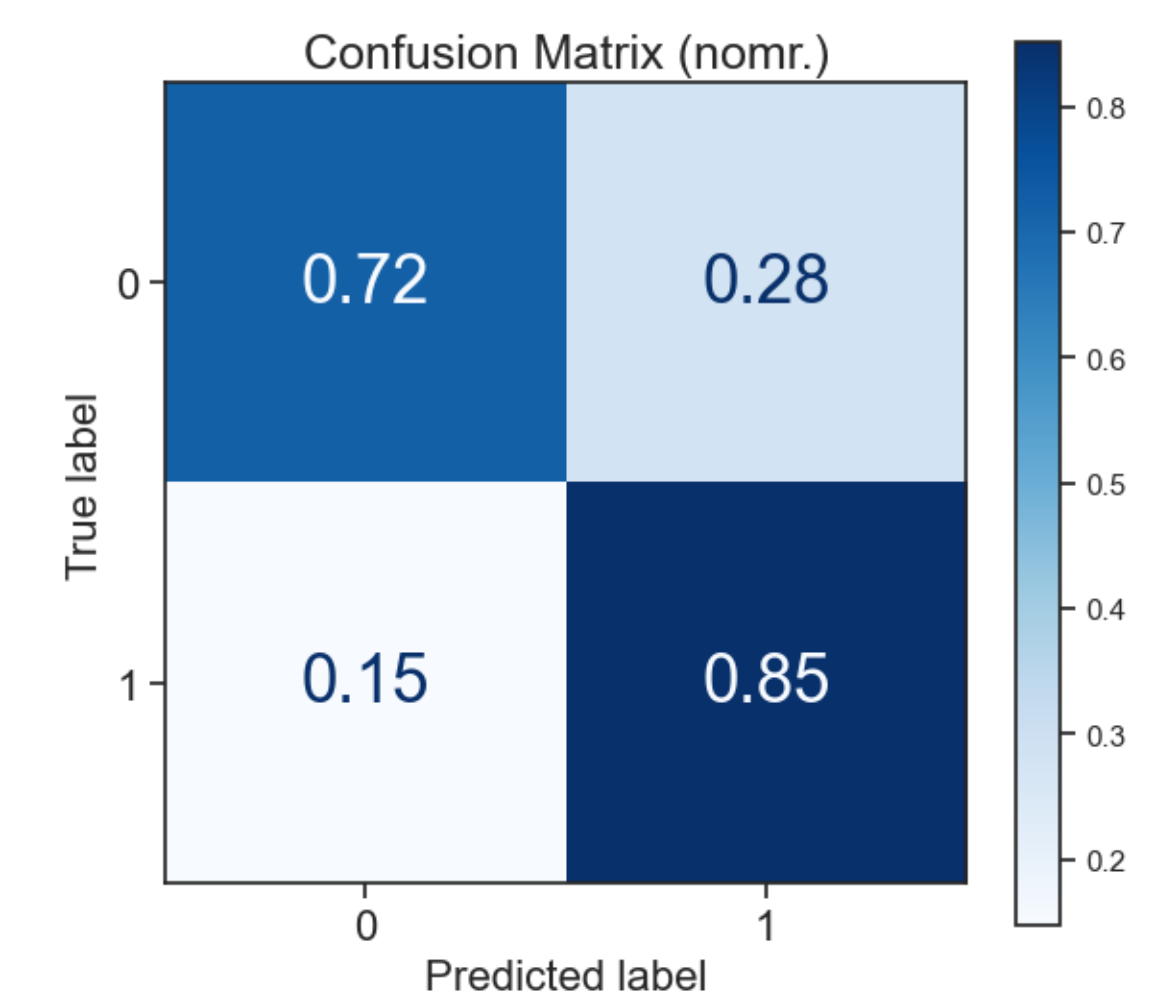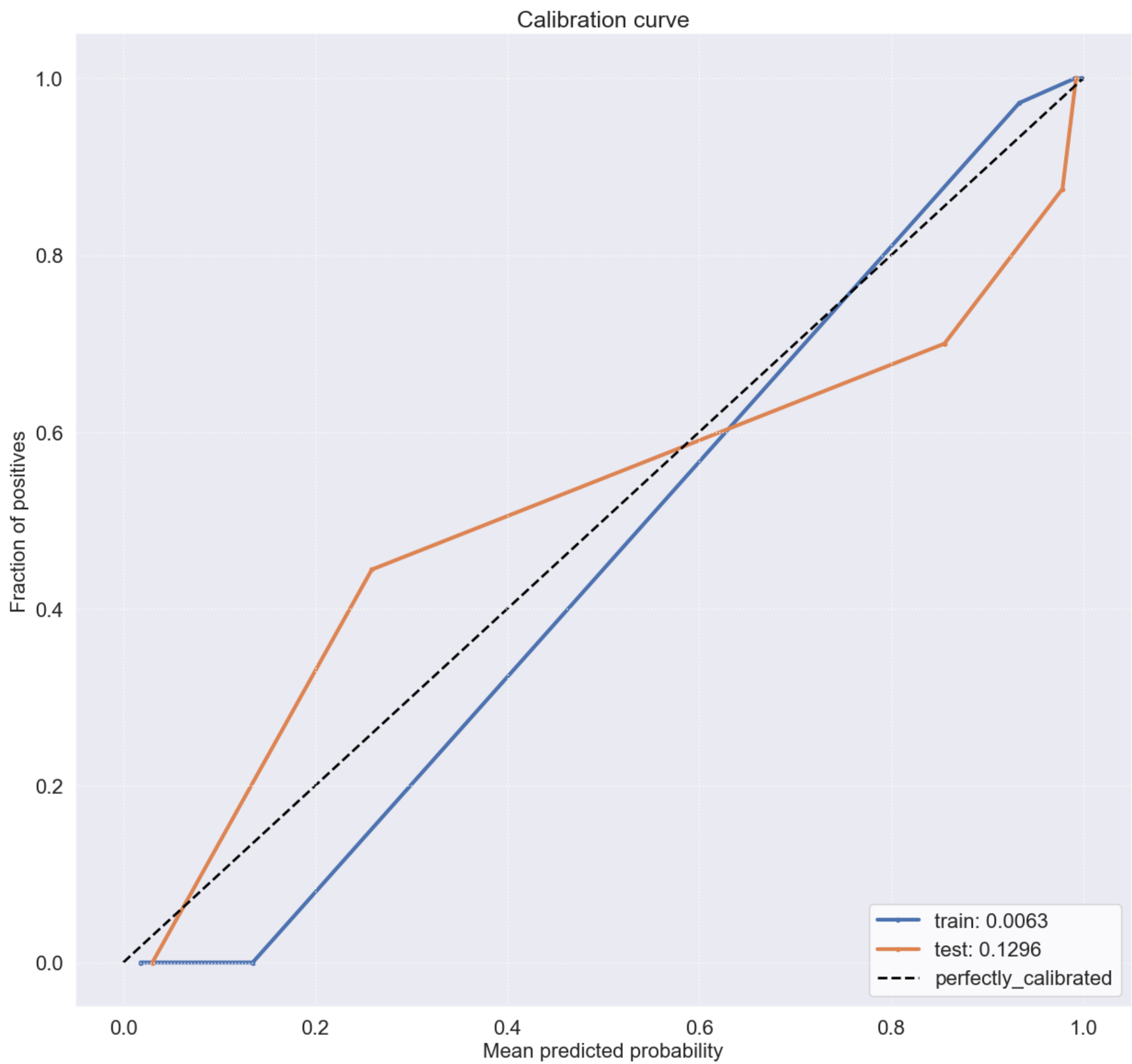- Best Trial (# 2549)

Confusion Matrix

	Predicted 0	Predicted 1
True 0	13	5
True 1	4	23

Calibration curve

train: 0.0063
test: 0.1296
perfectly_calibrated
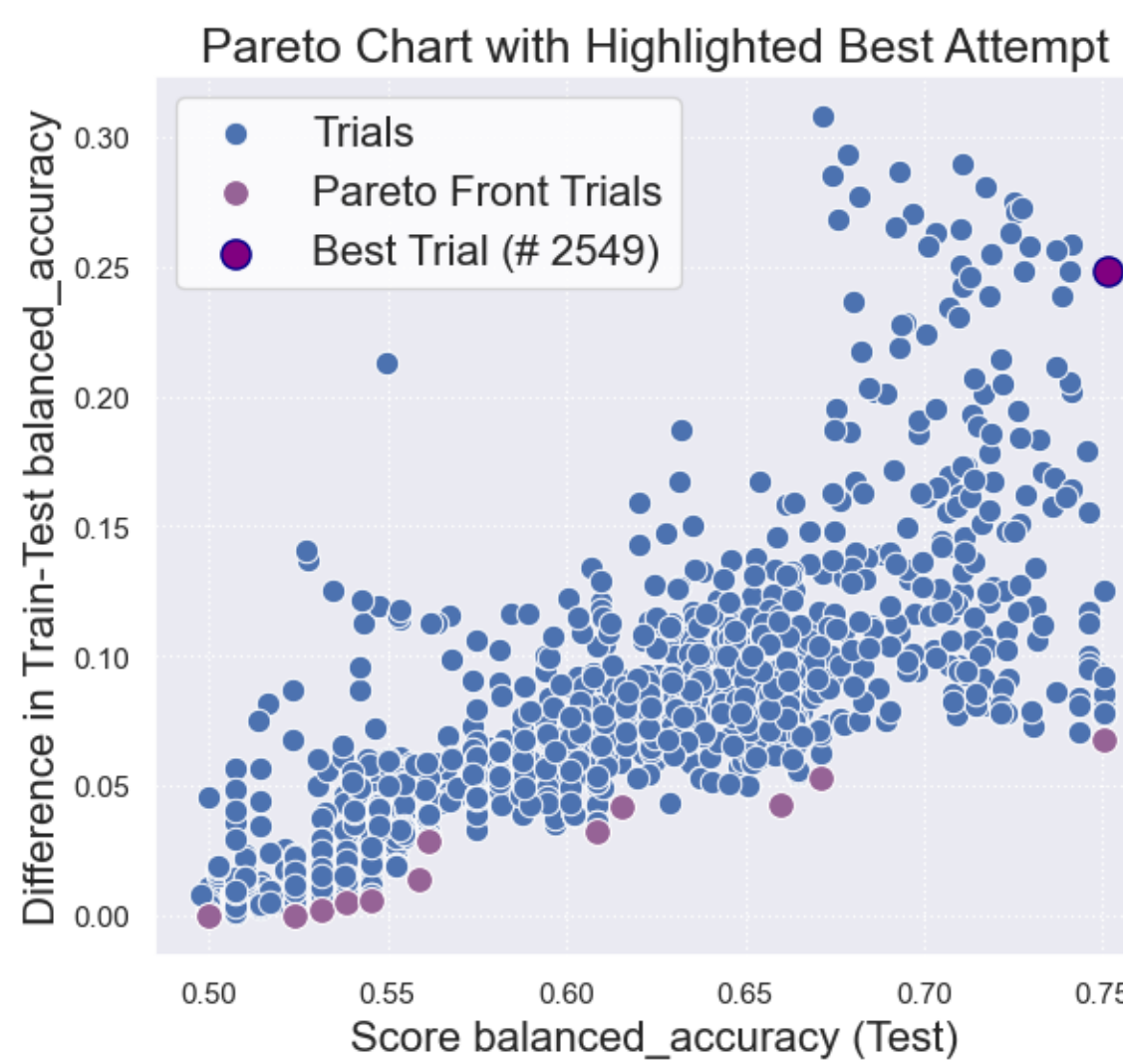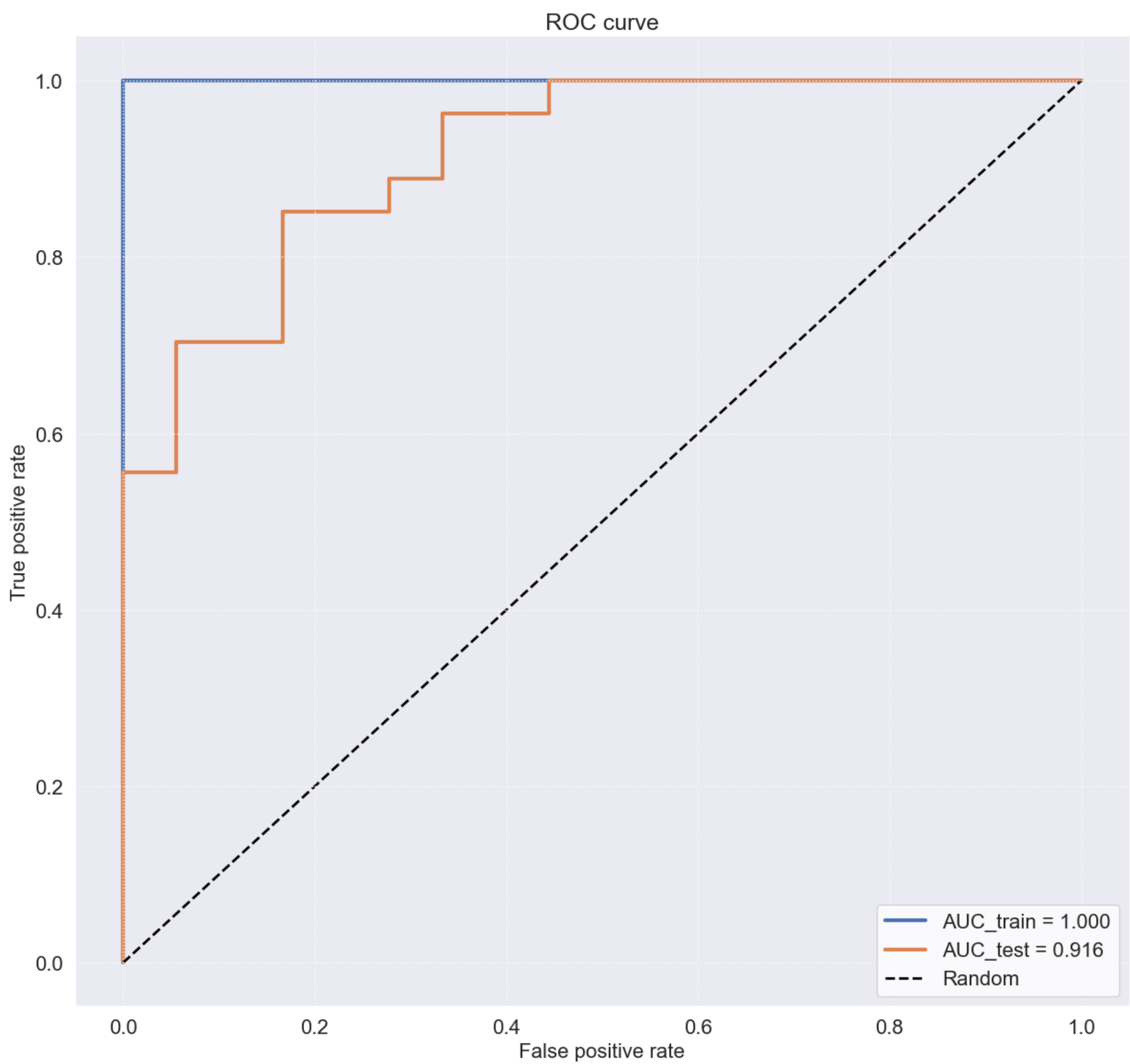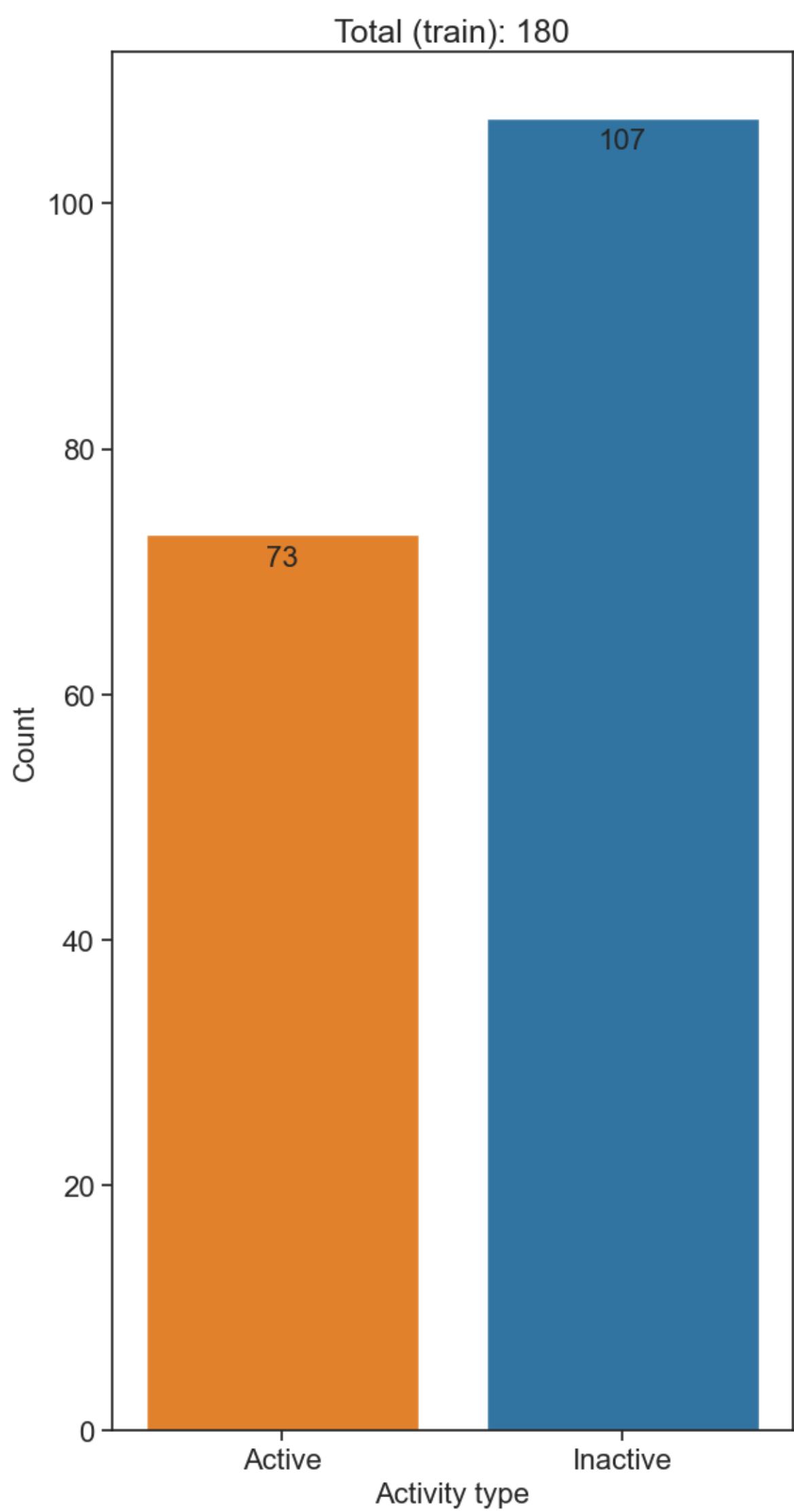
XGBoost Parameters	Value
alpha	1.8044
colsample_bytree	0.729
gamma	0.2105
lambda	4.0283
learning_rate	0.1086
max_depth	3.0
min_child_weight	1.0
n_estimators	643.0
reg_alpha	0.0888
reg_lambda	0.2601
scale_pos_weight	2.8085
subsample	0.8232

Confusion Matrix (nomr.)

	Predicted 0	Predicted 1
True 0	0.72	0.28
True 1	0.15	0.85

Active/Inactive probability

type
- active (27)
- inactive (18)
- decoy (45)
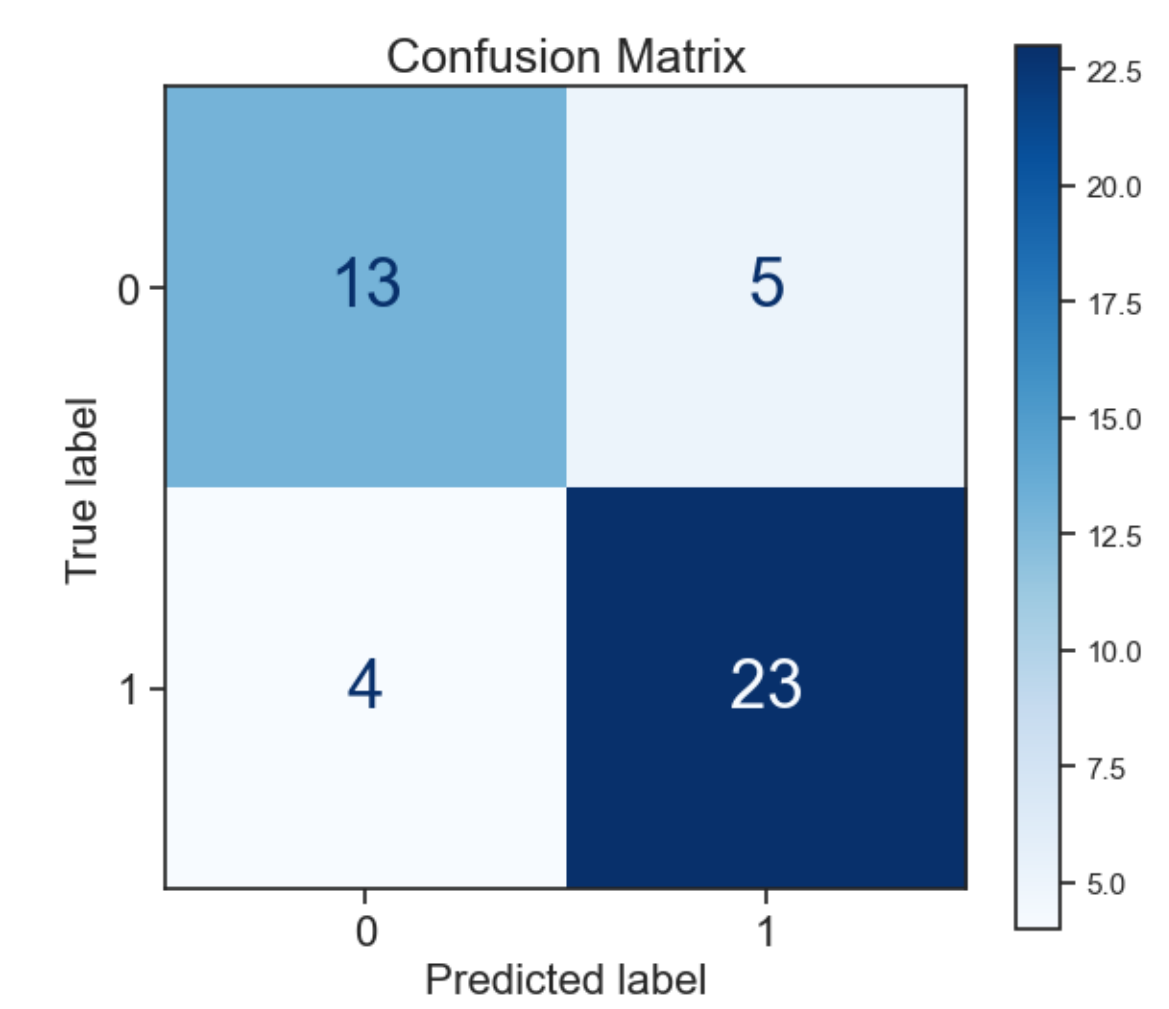
Metric	Train	Test
AUC	1.0	0.9156
accuracy	1.0	0.8
balanced accuracy	1.0	0.787
recall_(sens)	1.0	0.8519
specificity	1.0	0.7222
precision	1.0	0.8214
f1_score	1.0	0.799

Protein: O15379 - metric: f1

Total (train): 180

Bar chart showing Active: 73, Inactive: 107

ROC curve

AUC_train = 0.999
AUC_test = 0.938
Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 2470)

Confusion Matrix

	Predicted 0	Predicted 1
True 0	10	8
True 1	2	25

Confusion Matrix (nomr.)

	Predicted 0	Predicted 1
True 0	0.56	0.44
True 1	0.074	0.93

Calibration curve

train: 0.07
test: 0.131
perfectly_calibrated

Active/Inactive probability

Inactive zone / Active zone

type:
- active (27)
- inactive (18)
- decoy (45)

XGBoost Parameters	Value
alpha	2.0949
colsample_bytree	0.7808
gamma	0.1443
lambda	3.5053
learning_rate	0.0135
max_depth	6.0
min_child_weight	1.0
n_estimators	363.0
reg_alpha	0.8775
reg_lambda	0.9293
scale_pos_weight	4.3336
subsample	0.6411
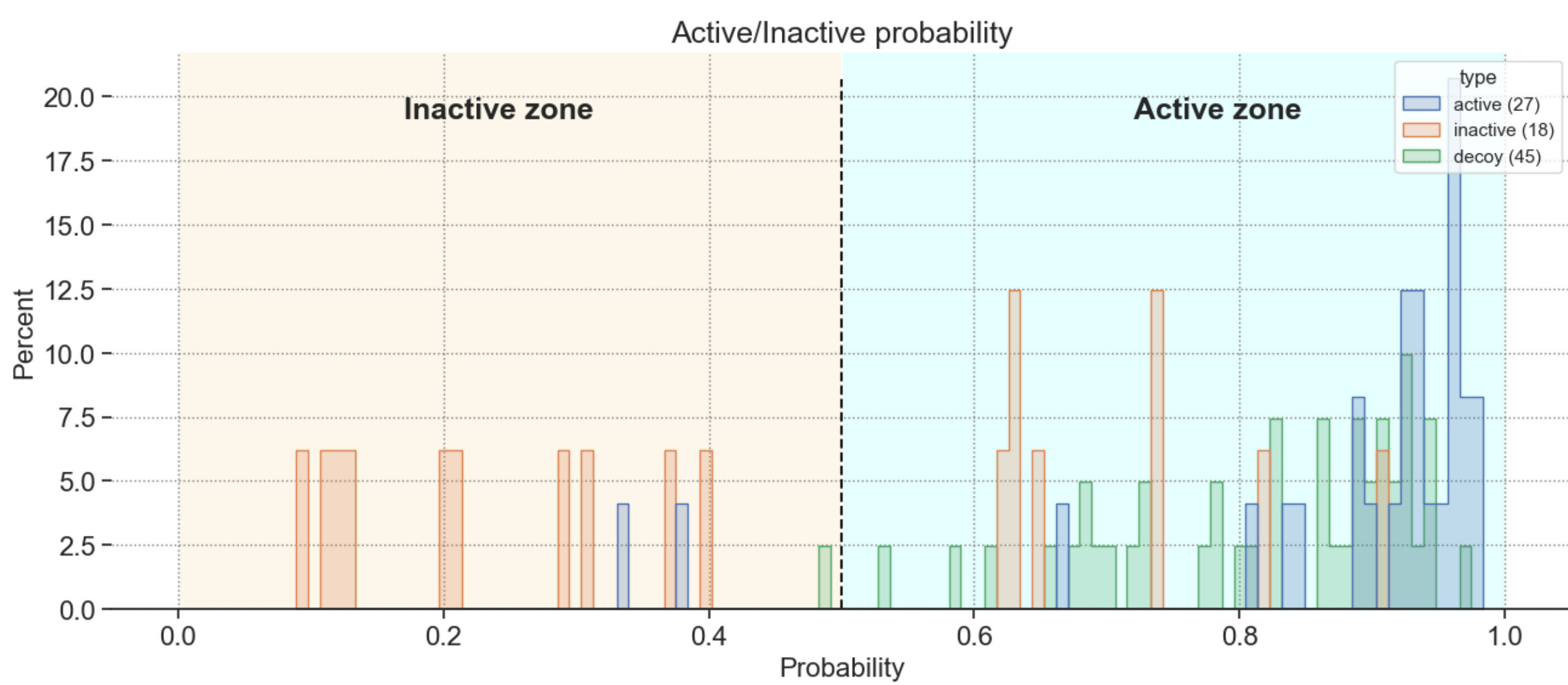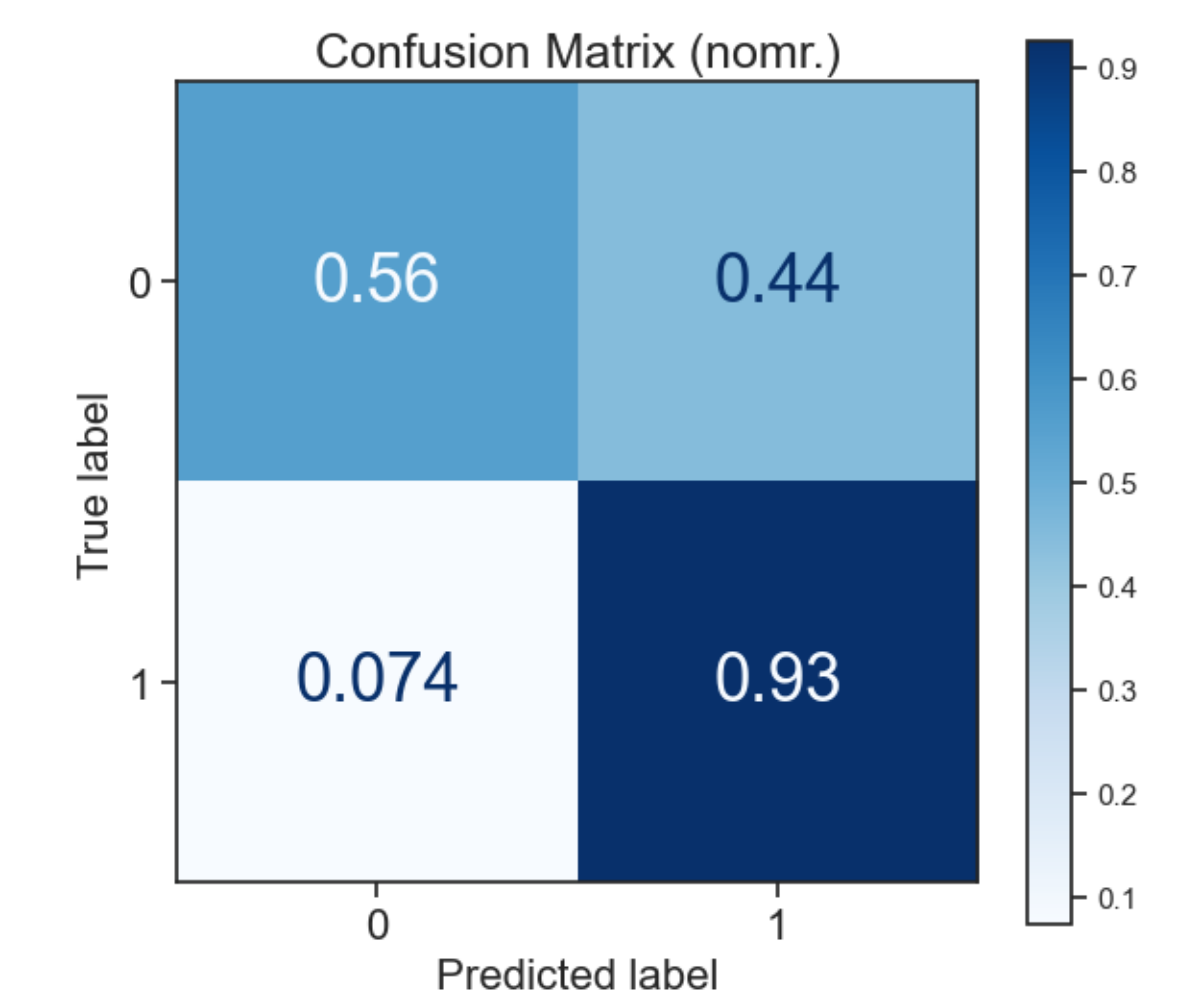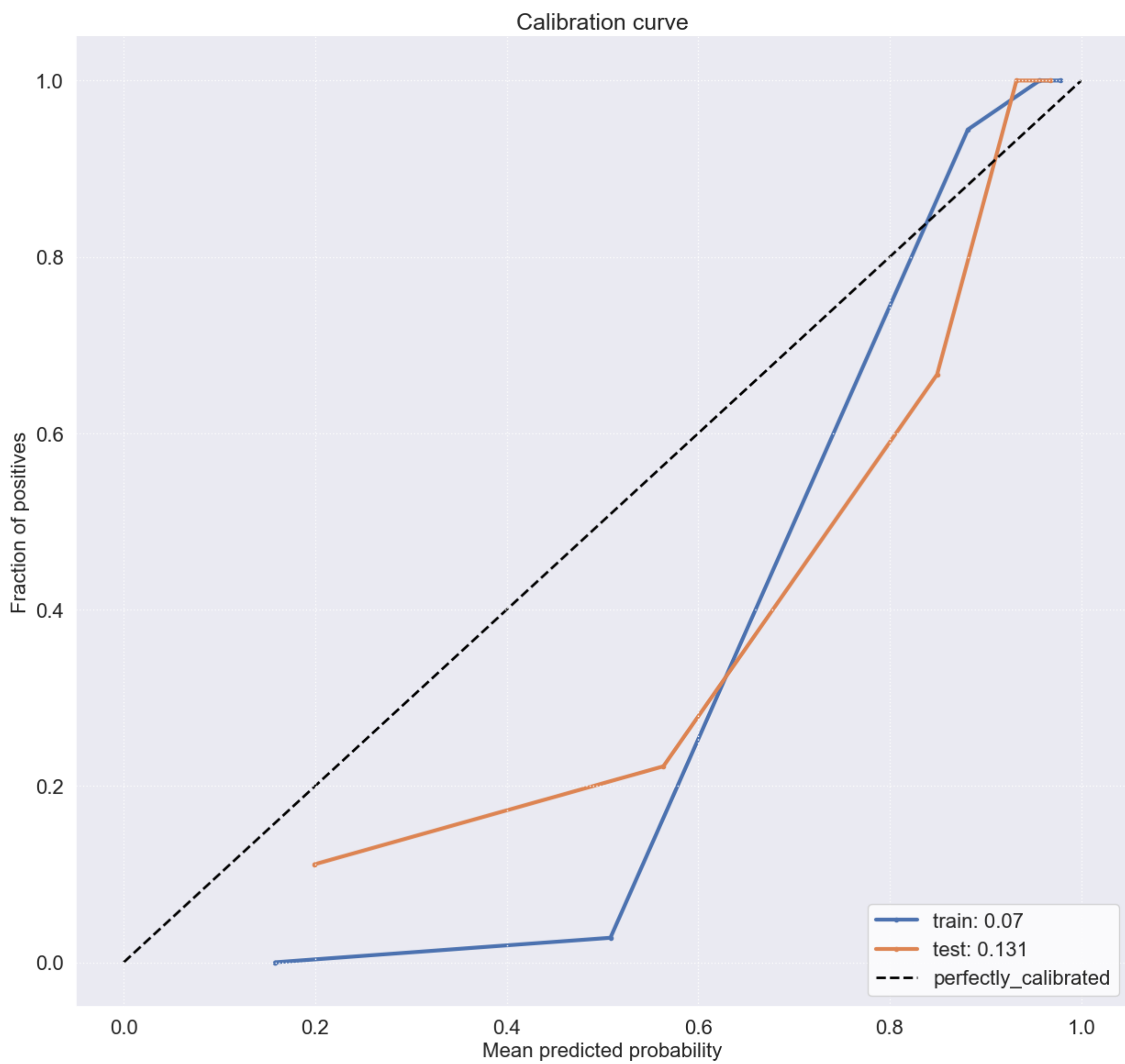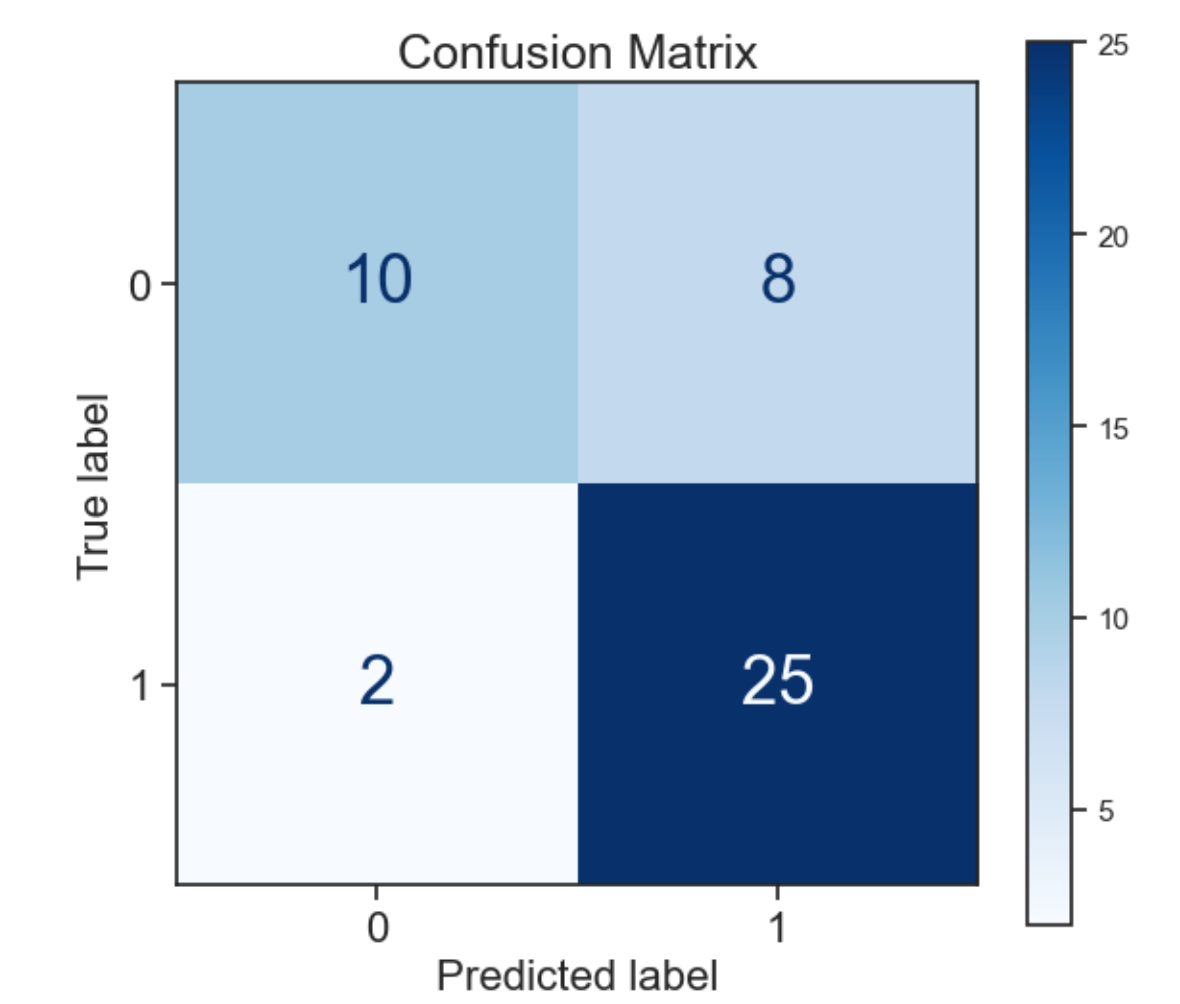
Metric	Train	Test
AUC	0.999	0.9383
accuracy	0.8889	0.7778
balanced accuracy	0.863	0.7407
recall_(sens)	1.0	0.9259
specificity	0.726	0.5556
precision	0.8425	0.7576
f1_score	0.8848	0.7667

Protein: O15379 - metric: precision

Total (train): 180

ROC curve

AUC_train = 1.000
AUC_test = 0.879
Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 2300)
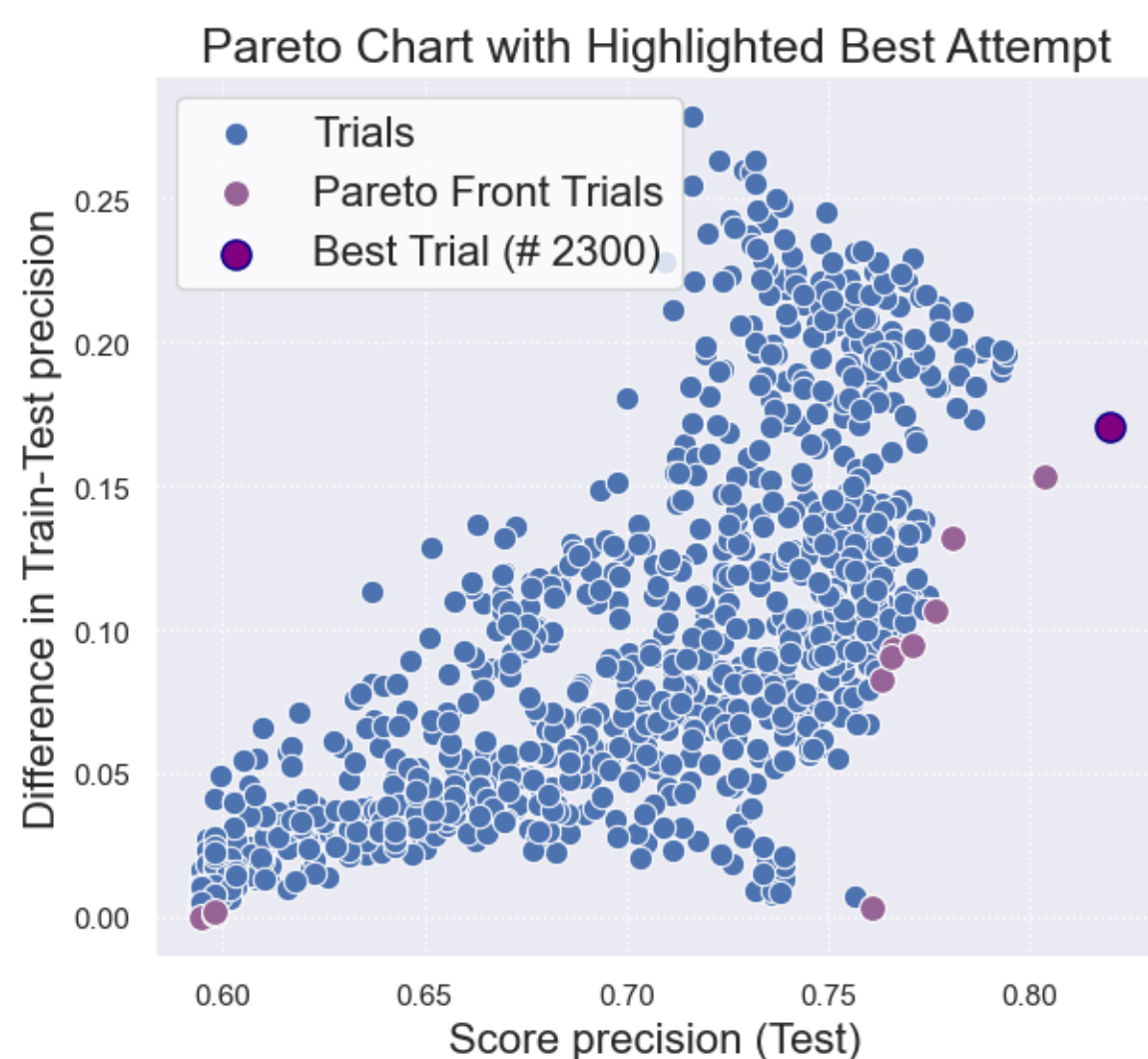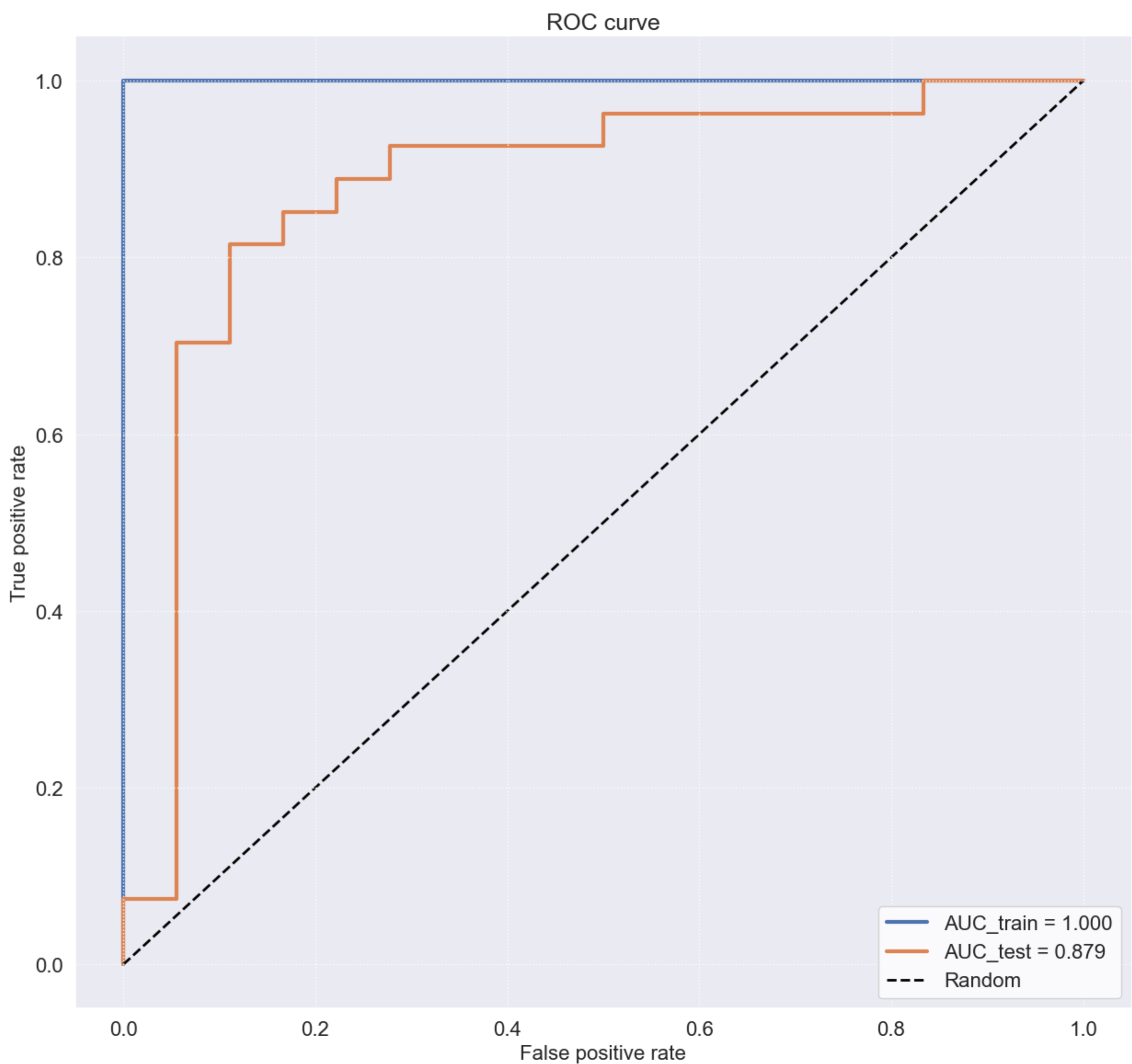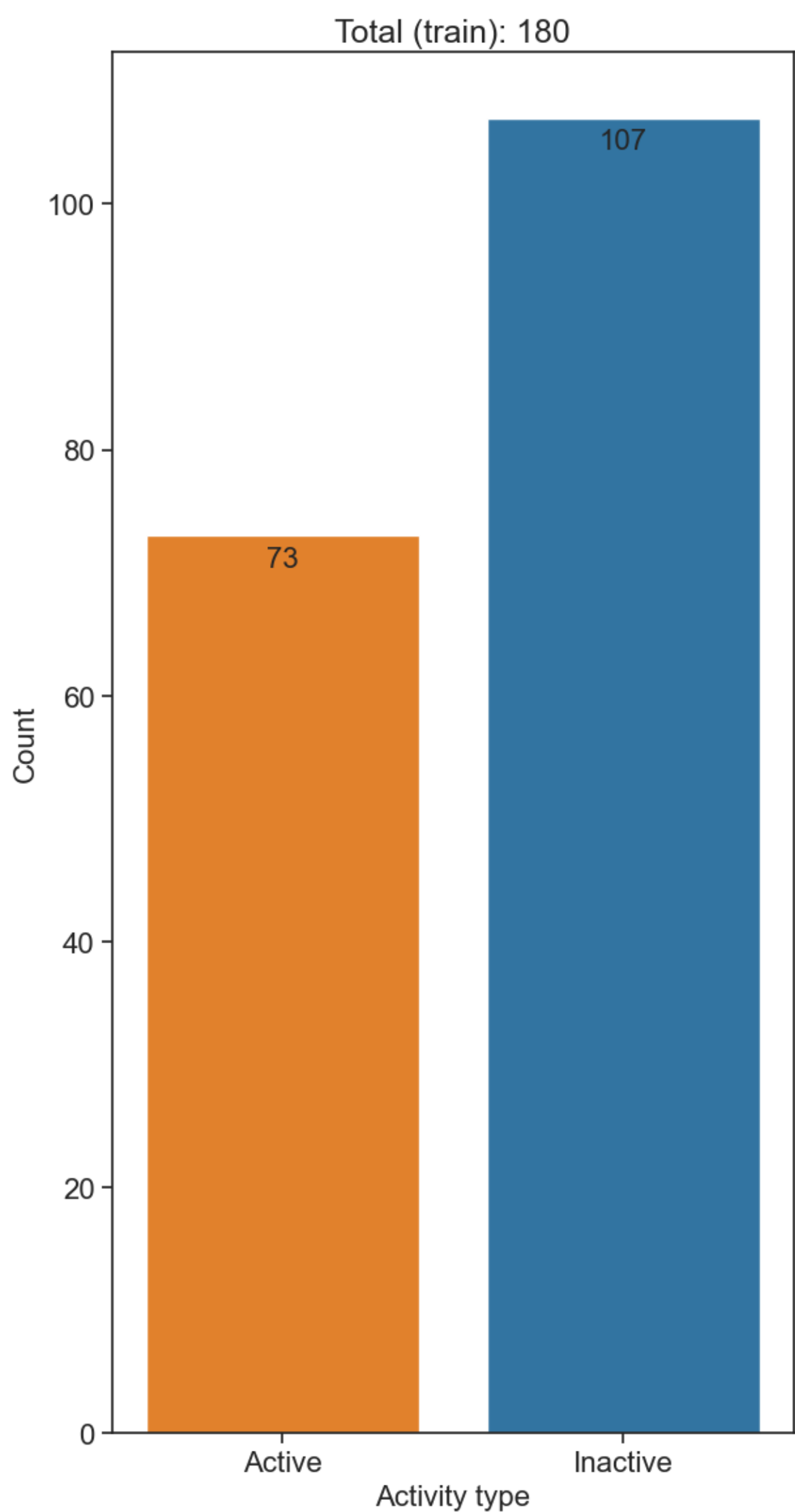
XGBoost Parameters	Value
alpha	0.7536
colsample_bytree	0.597
gamma	0.1661
lambda	0.0661
learning_rate	0.7647
max_depth	10.0
min_child_weight	2.0
n_estimators	787.0
reg_alpha	0.2419
reg_lambda	0.4424
scale_pos_weight	4.0921
subsample	0.8286

Confusion Matrix

	0	1
0	13	5
1	3	24

Confusion Matrix (nomr.)

	0	1
0	0.72	0.28
1	0.11	0.89

Calibration curve

train: 0.0079
test: 0.1328
perfectly_calibrated

Active/Inactive probability

Inactive zone | Active zone

type
- active (27)
- inactive (18)
- decoy (45)
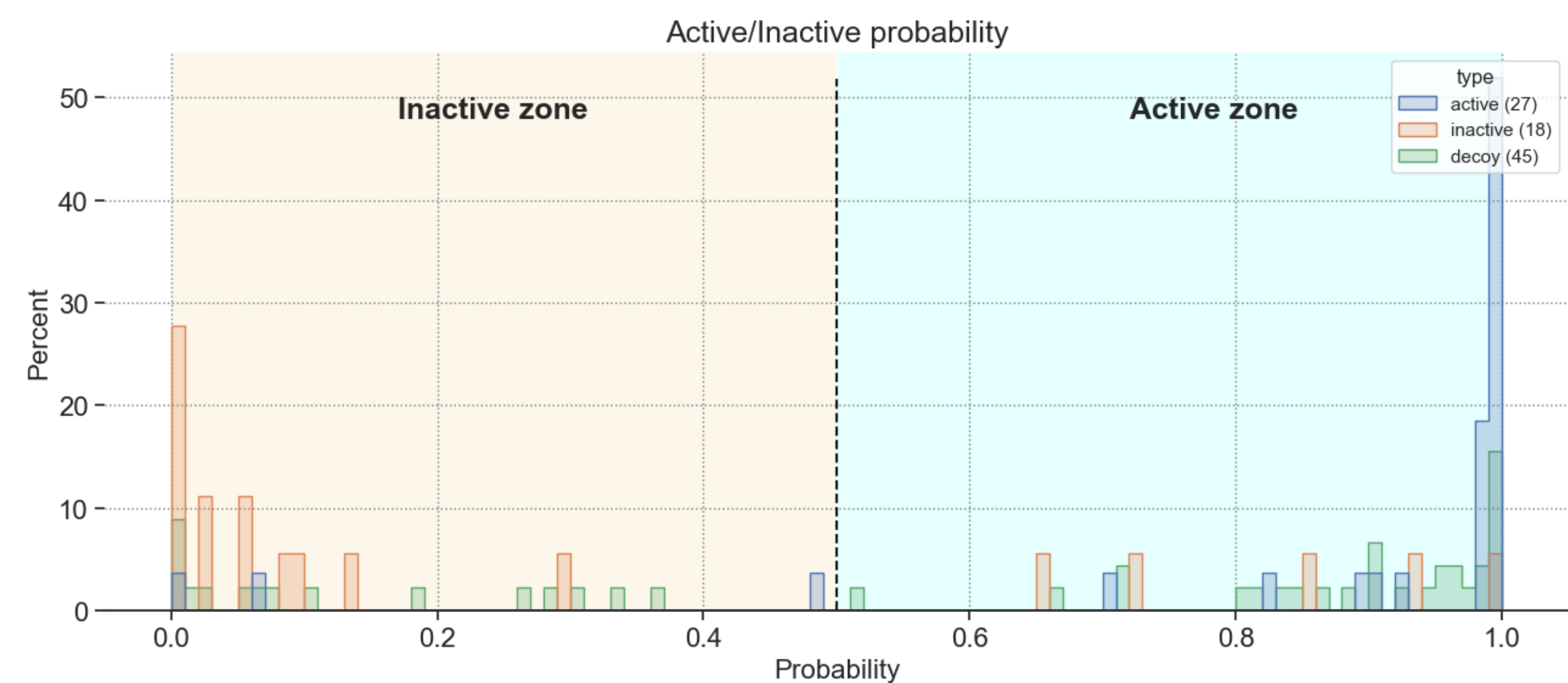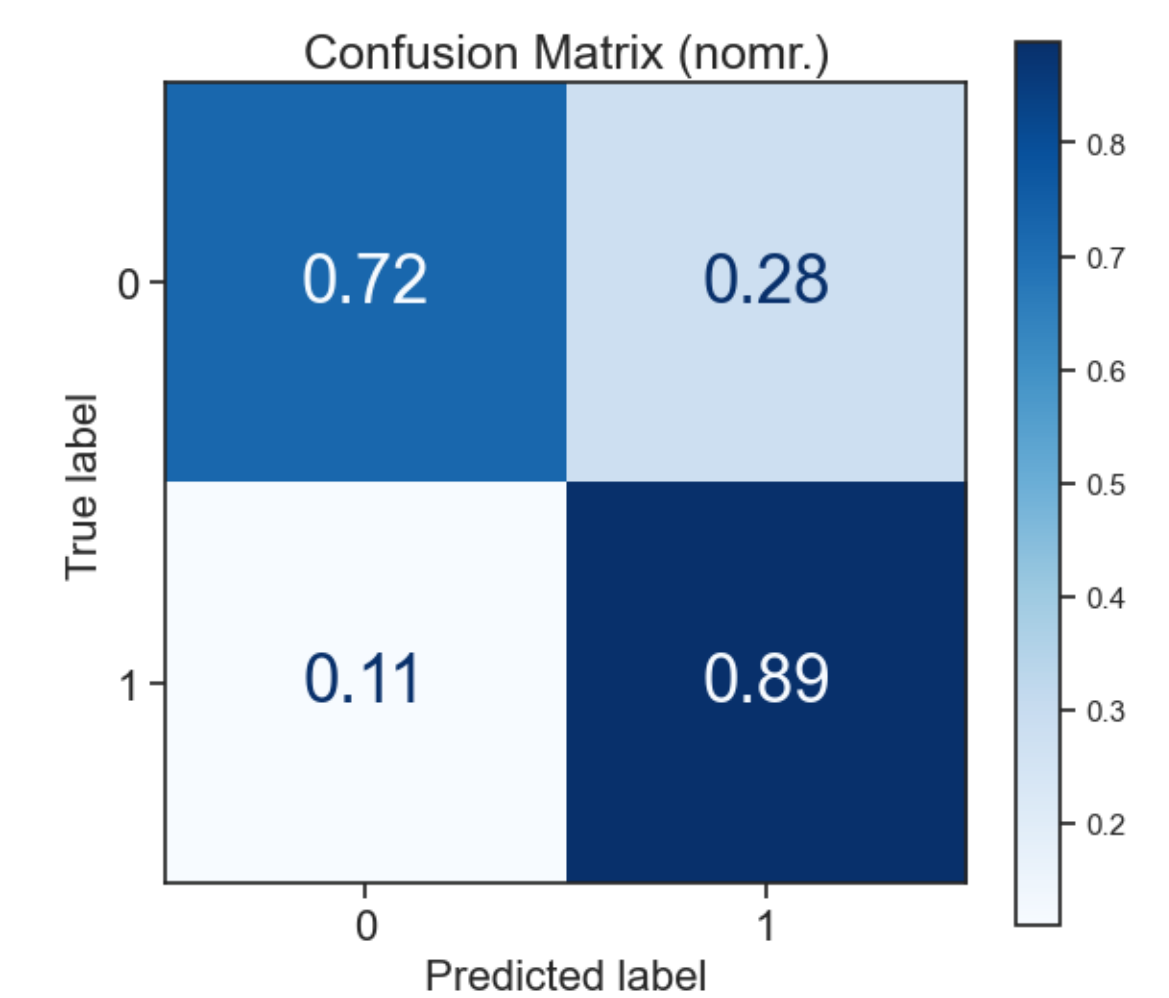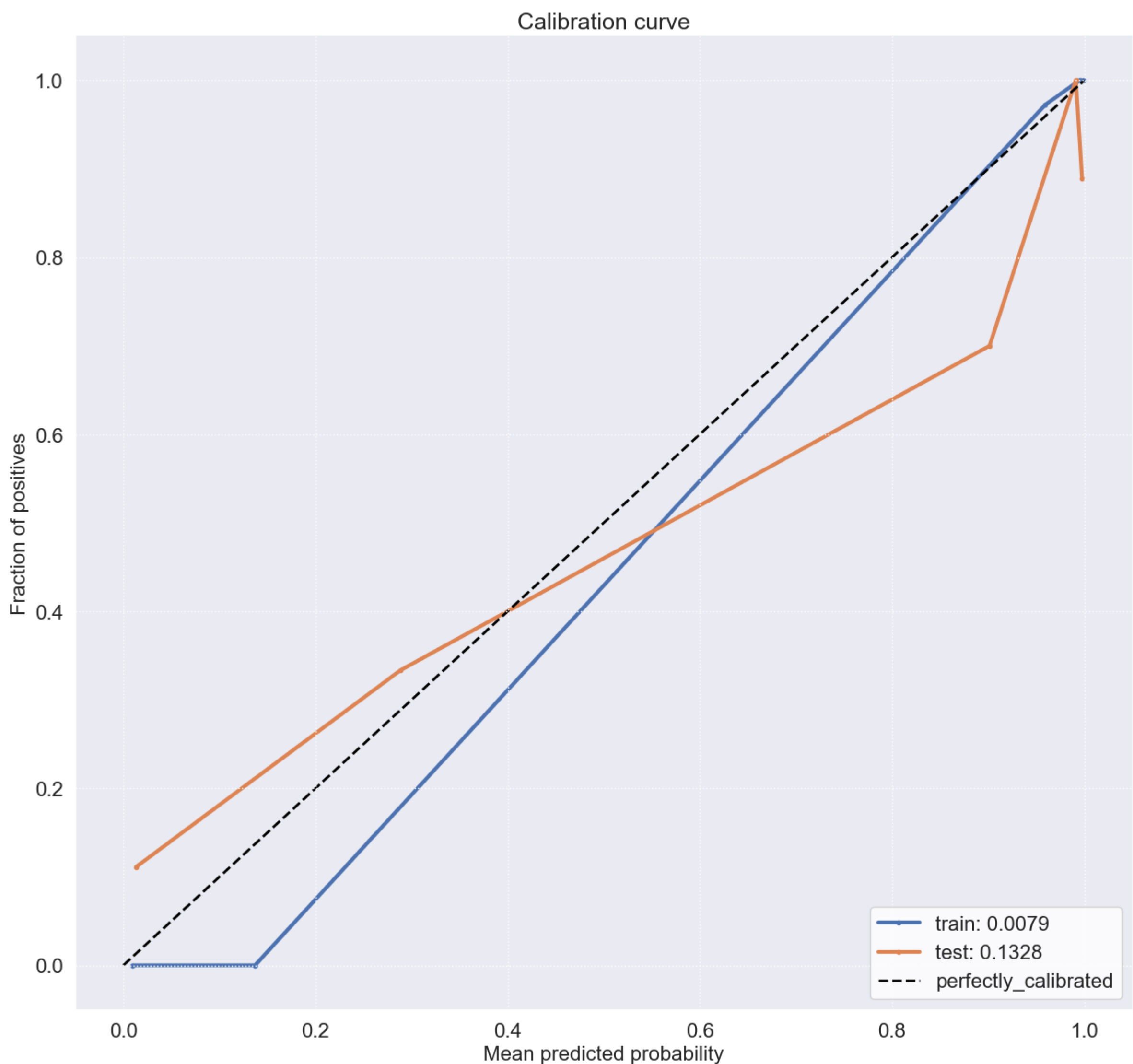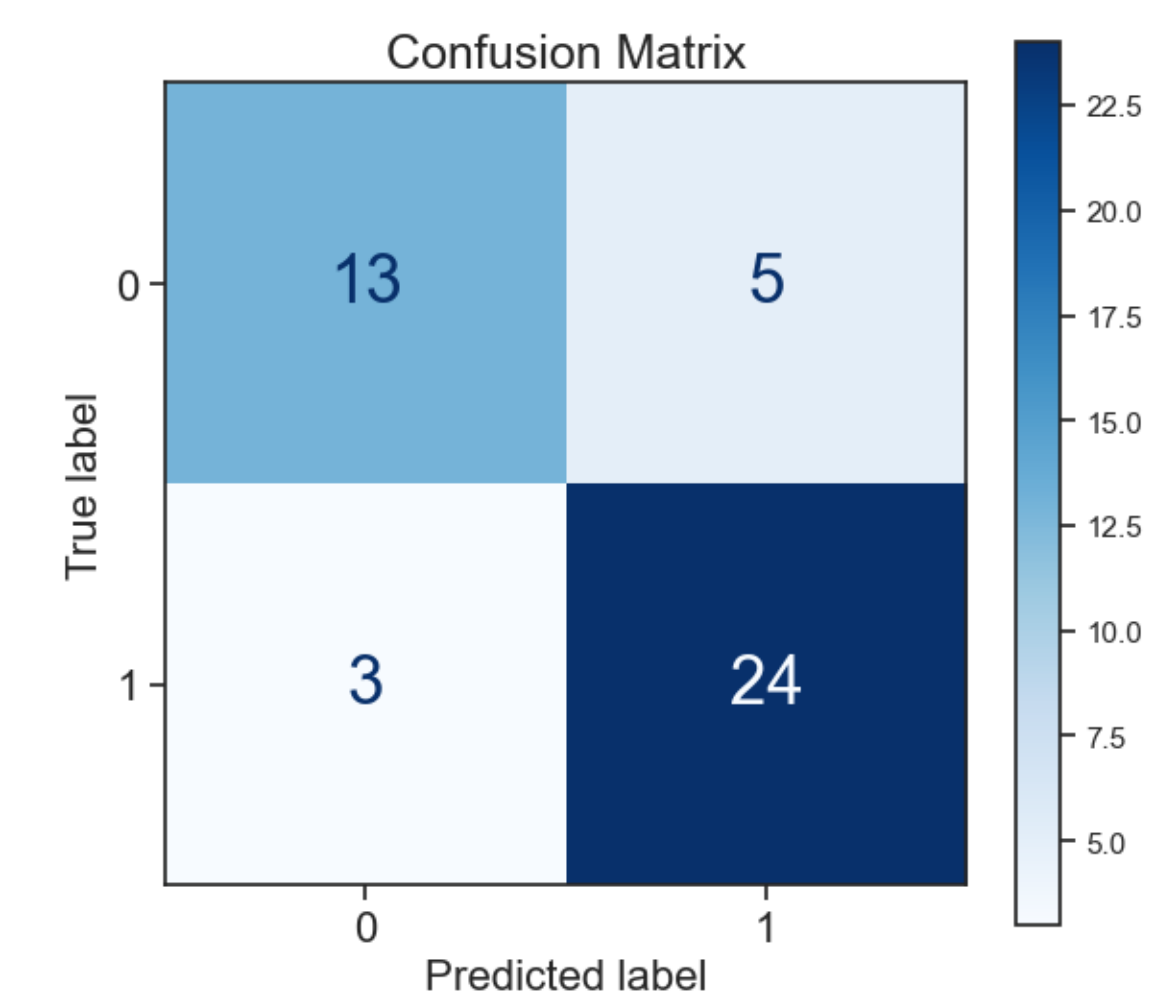
Metric	Train	Test
AUC	1.0	0.8786
accuracy	0.9944	0.8222
balanced accuracy	0.9932	0.8056
recall_(sens)	1.0	0.8889
specificity	0.9863	0.7222
precision	0.9907	0.8276
f1_score	0.9944	0.8202

Protein: O15379 - metric: SoftVotingXGBoostEnsemble
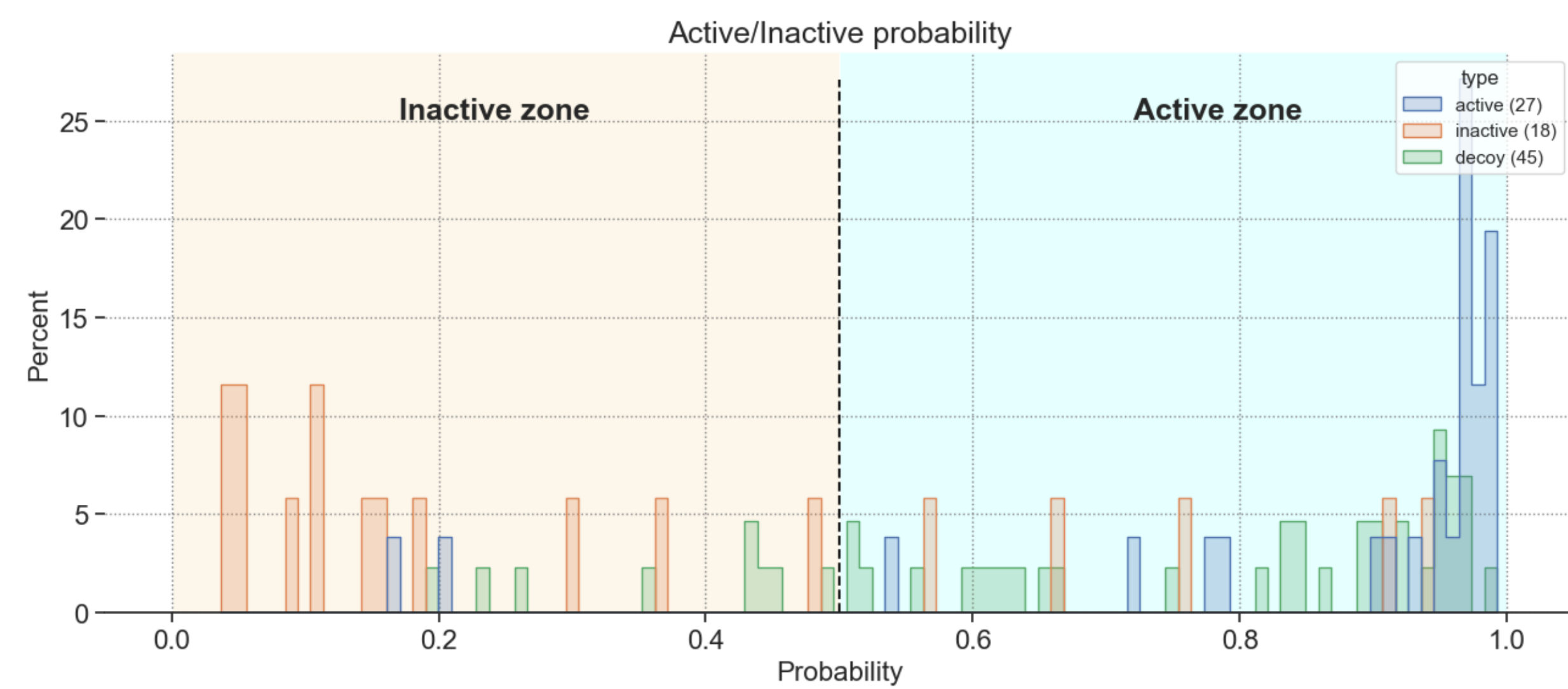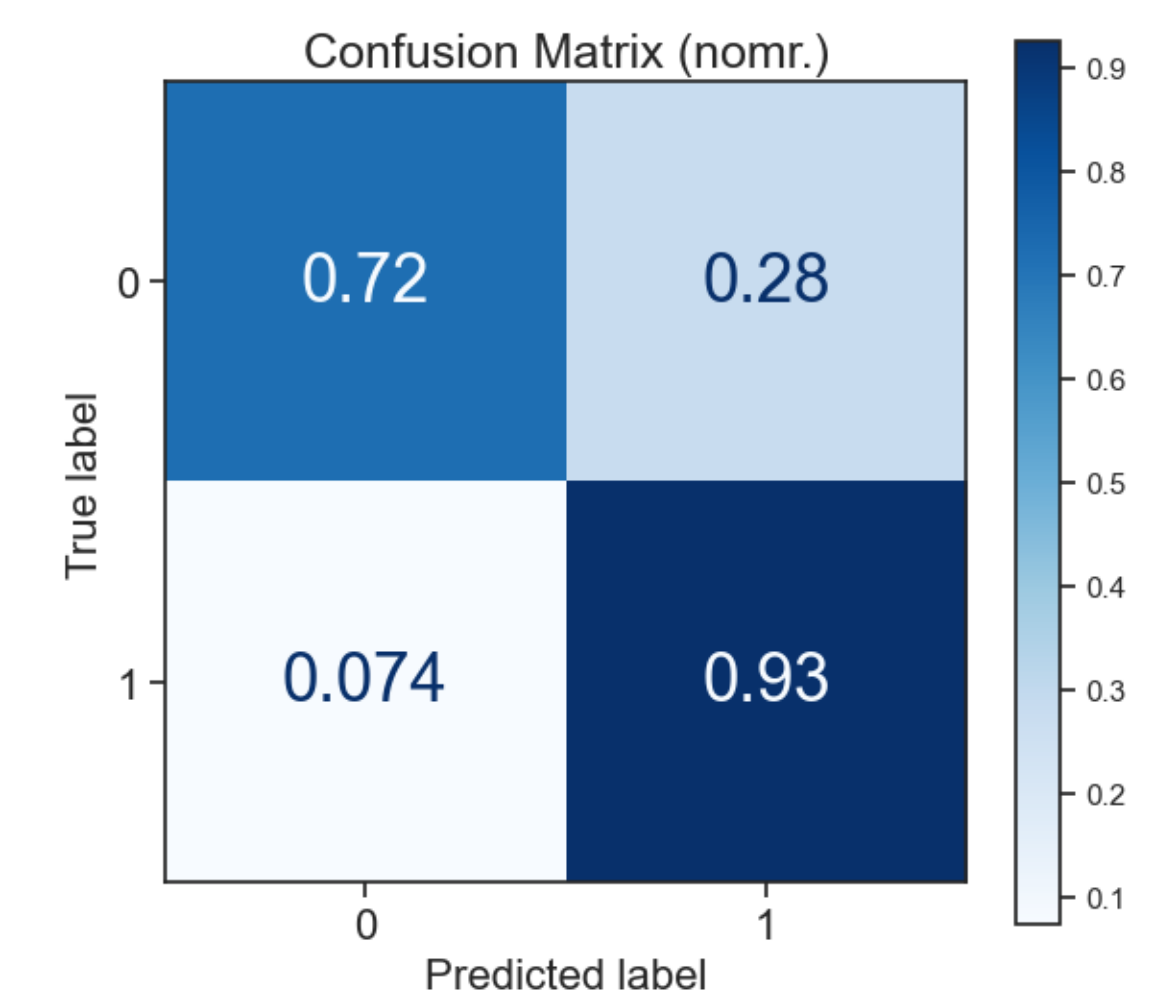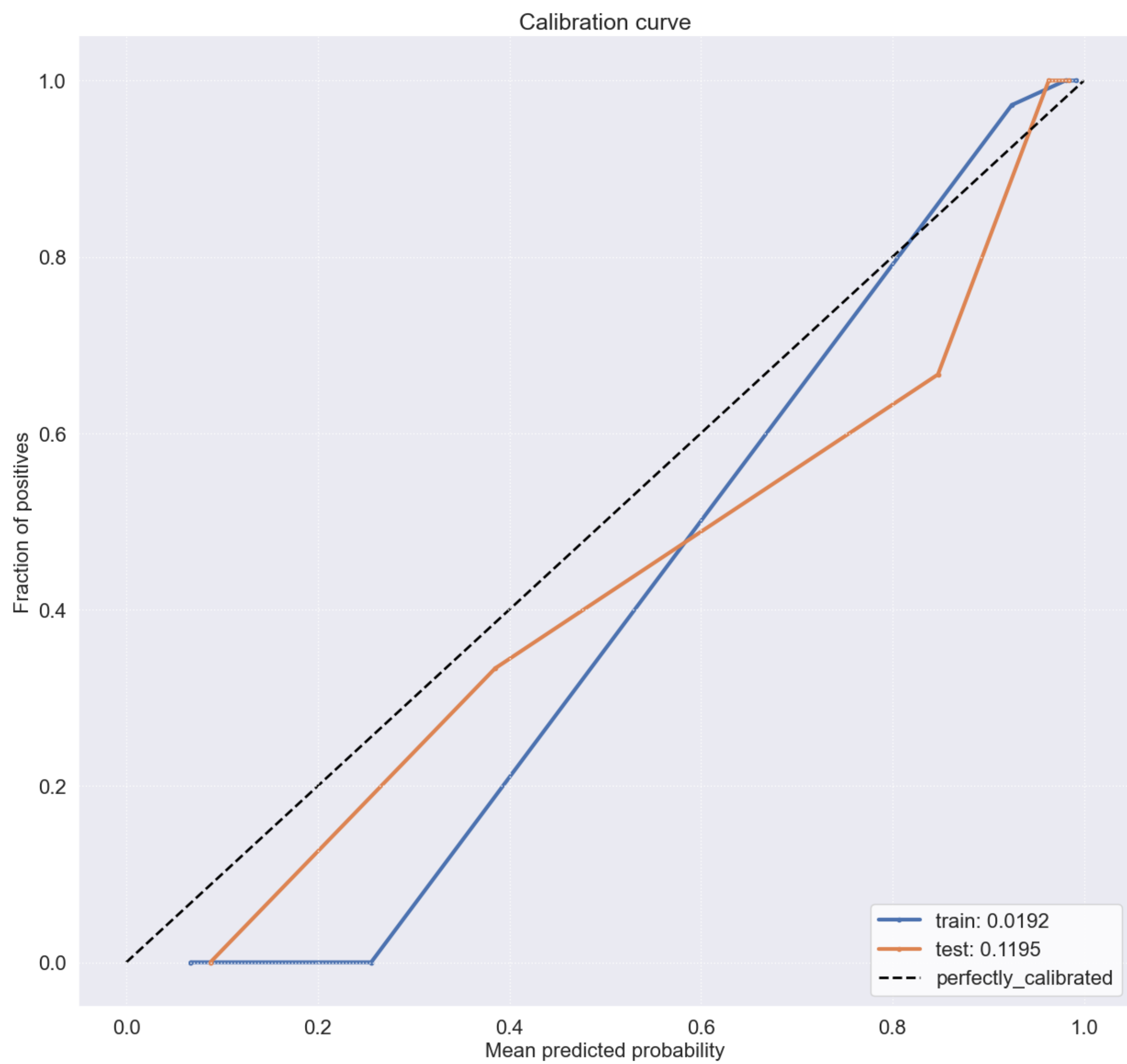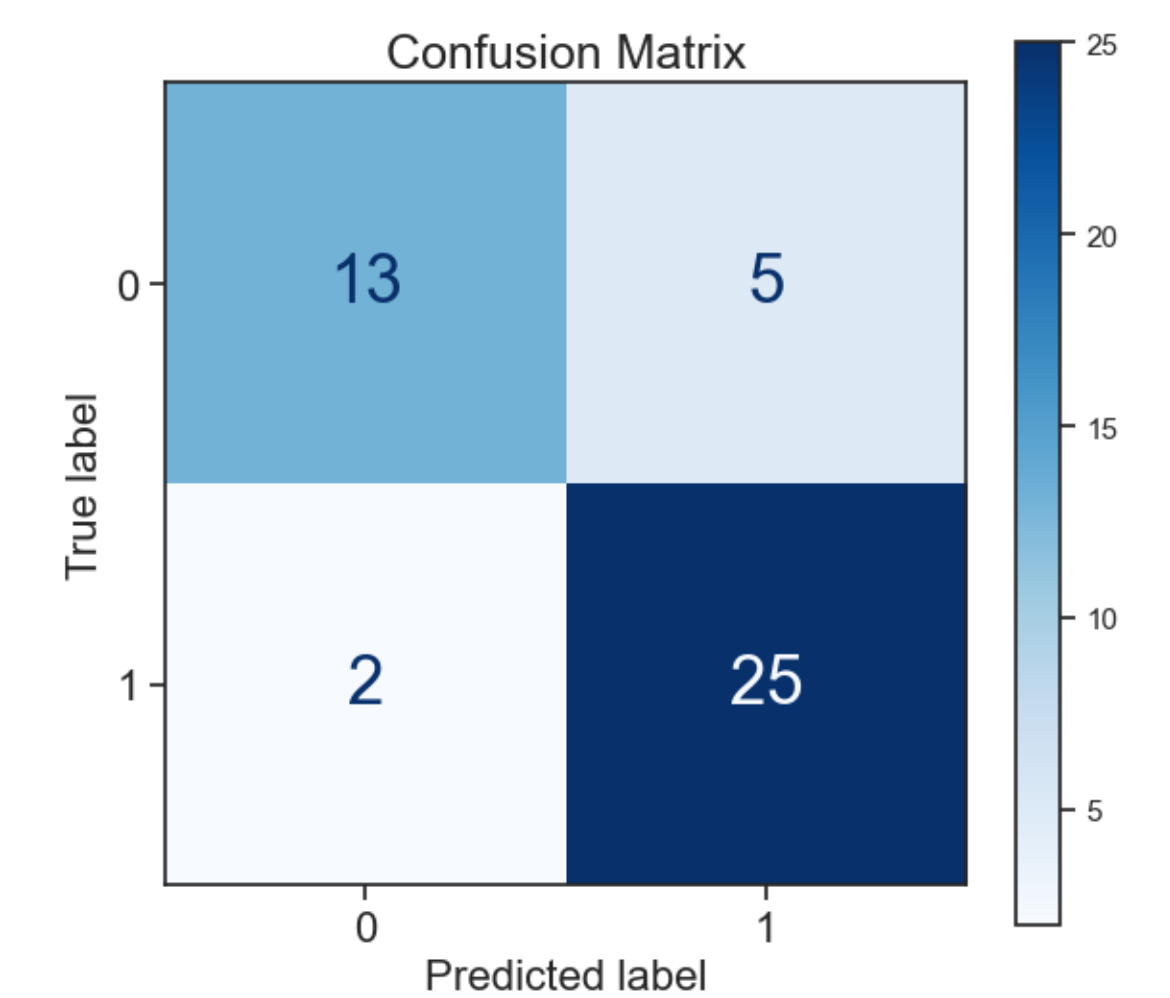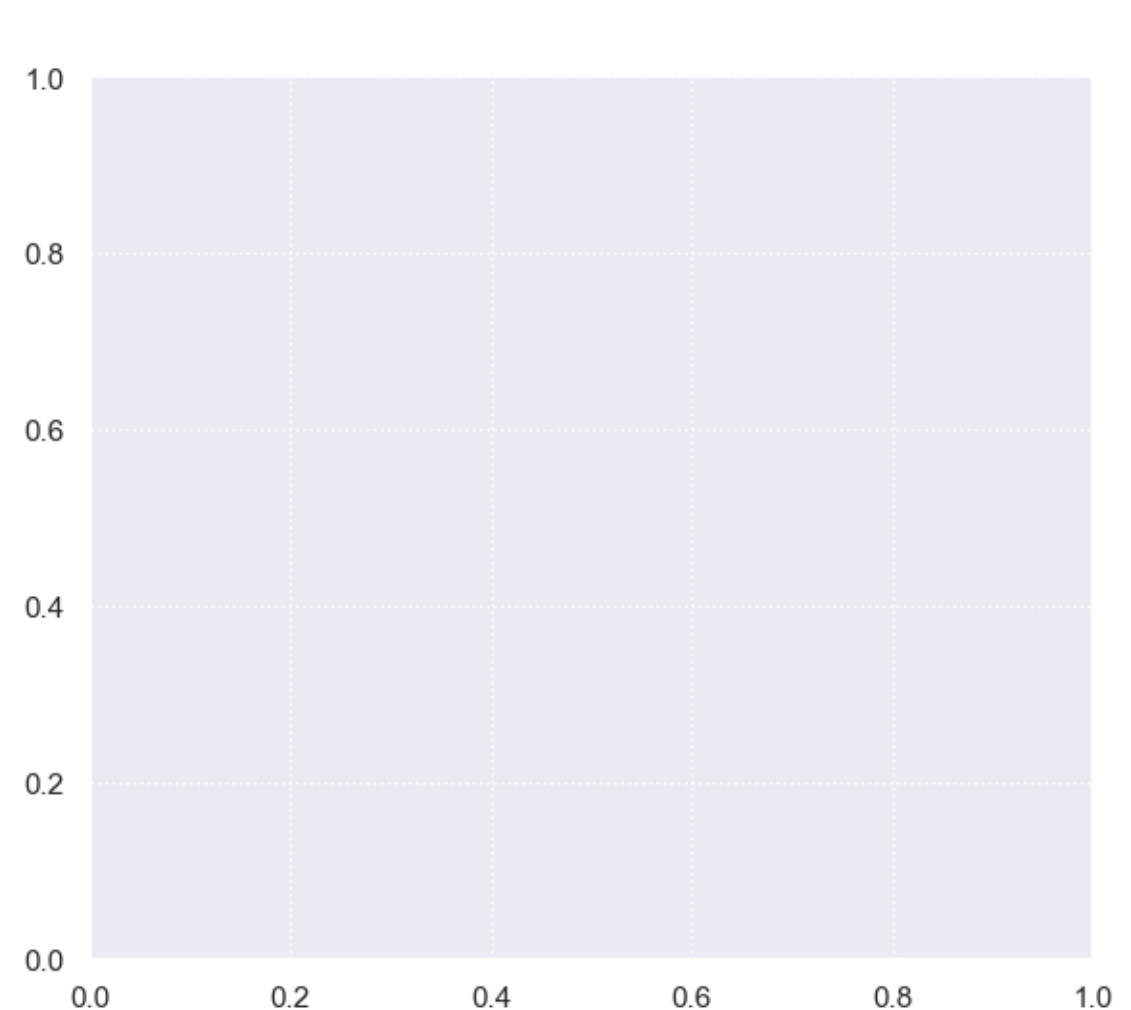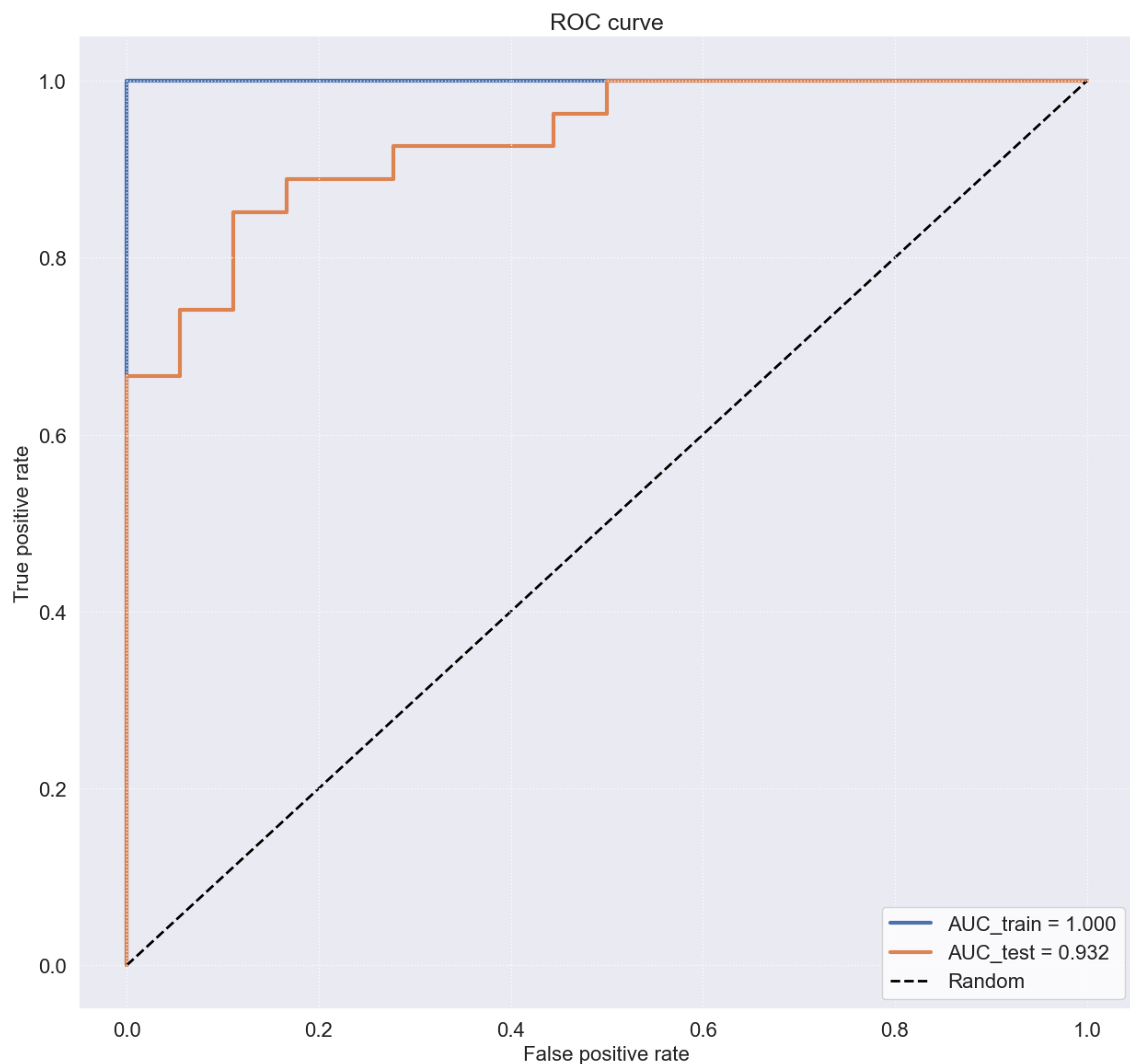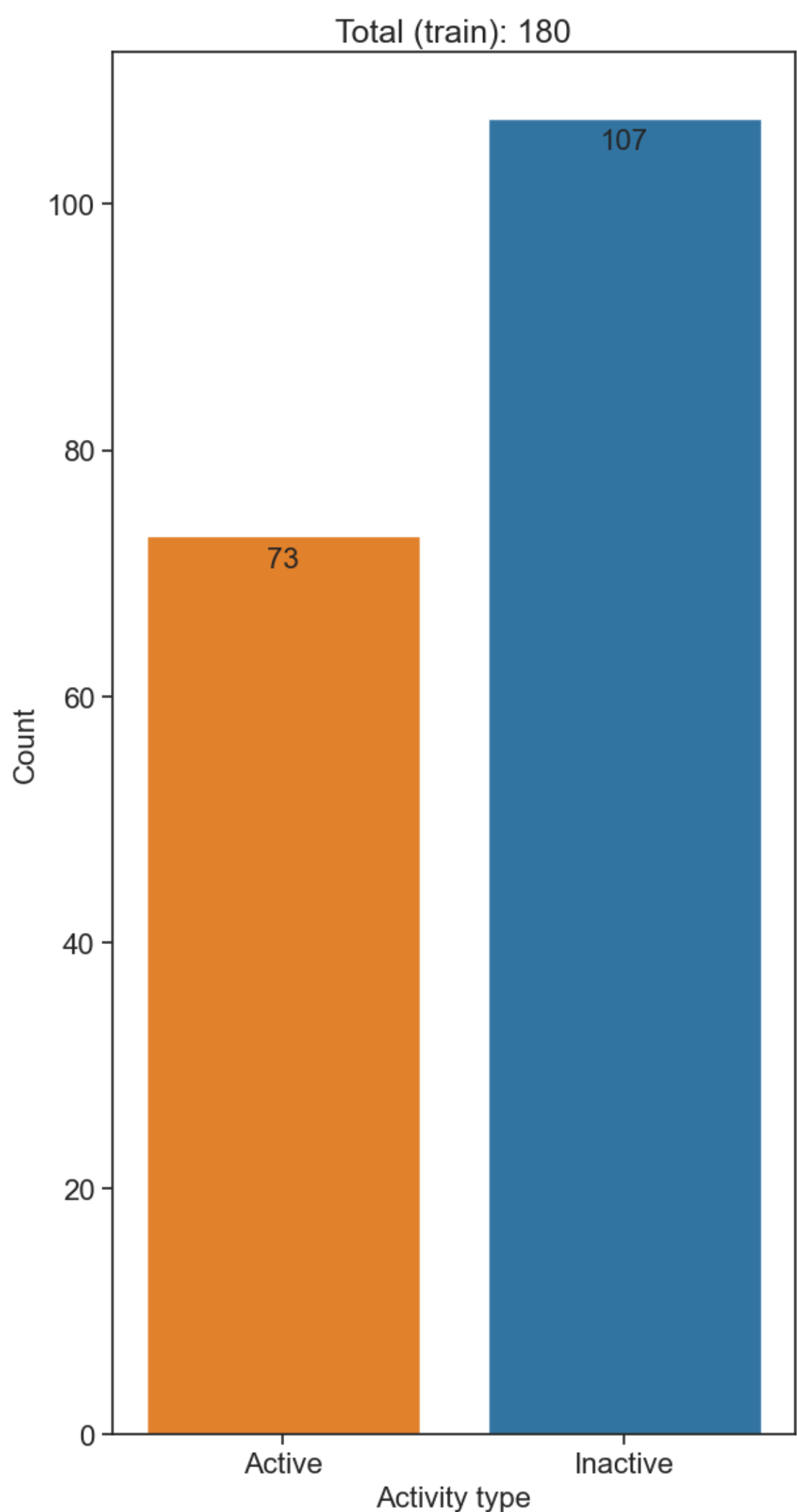
Metric	Train	Test
AUC	1.0	0.9321
accuracy	0.9889	0.8444
balanced accuracy	0.9863	0.8241
recall_(sens)	1.0	0.9259
specificity	0.9726	0.7222
precision	0.9817	0.8333
f1_score	0.9889	0.8415