Protein: P04150 - metric: balanced_accuracy

Total (train): 1420

ROC curve

AUC_train = 1.000
AUC_test = 0.942
Random

Pareto Chart with Highlighted Best Attempt

- Trials
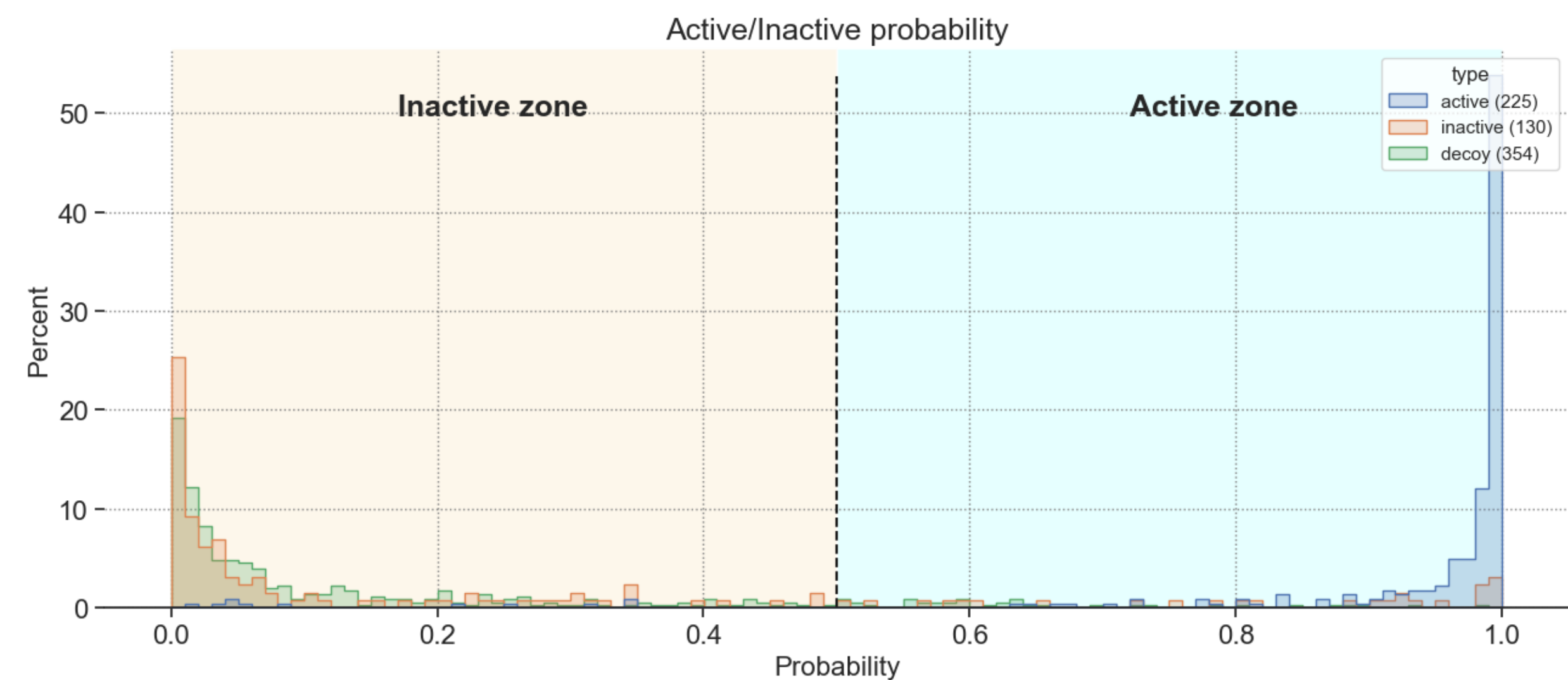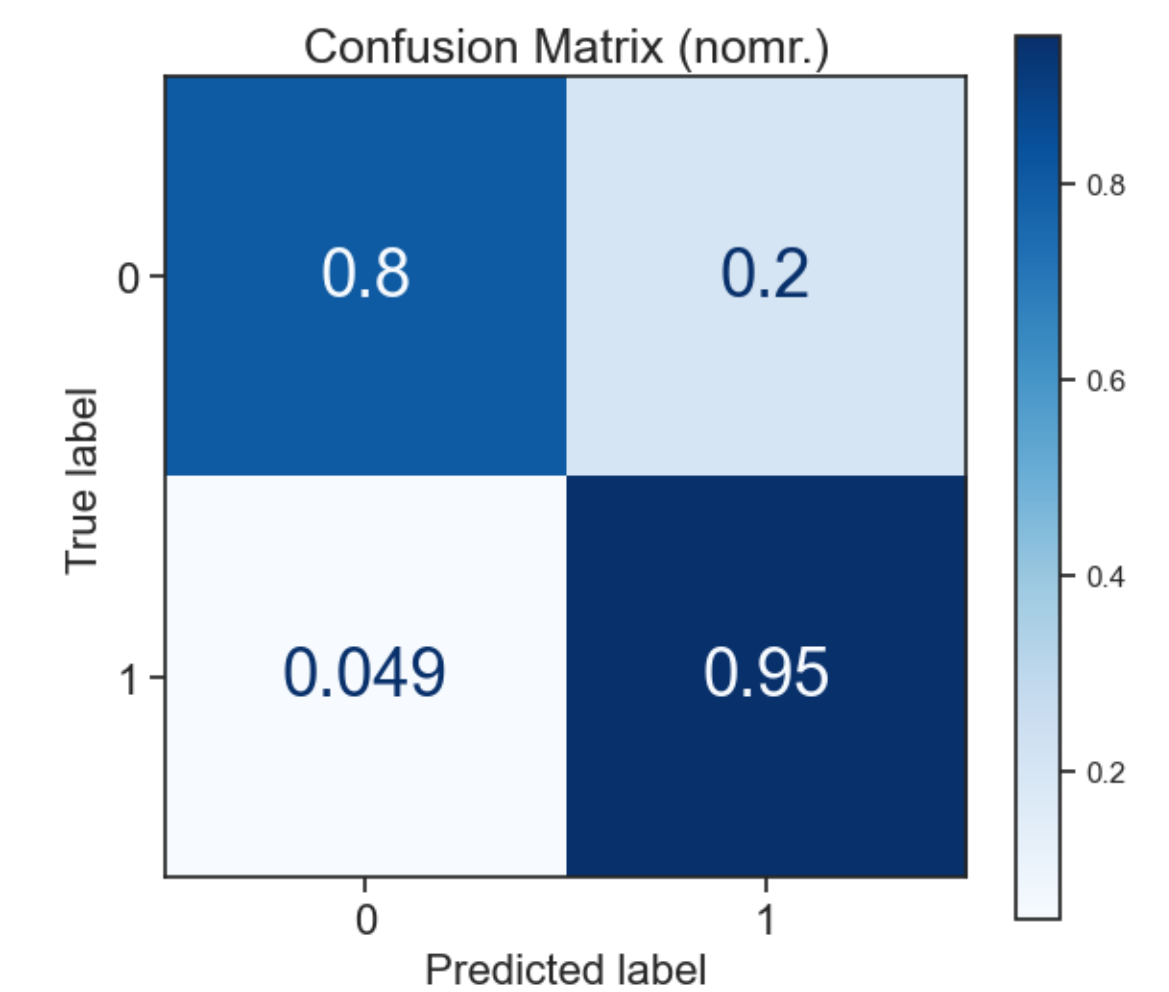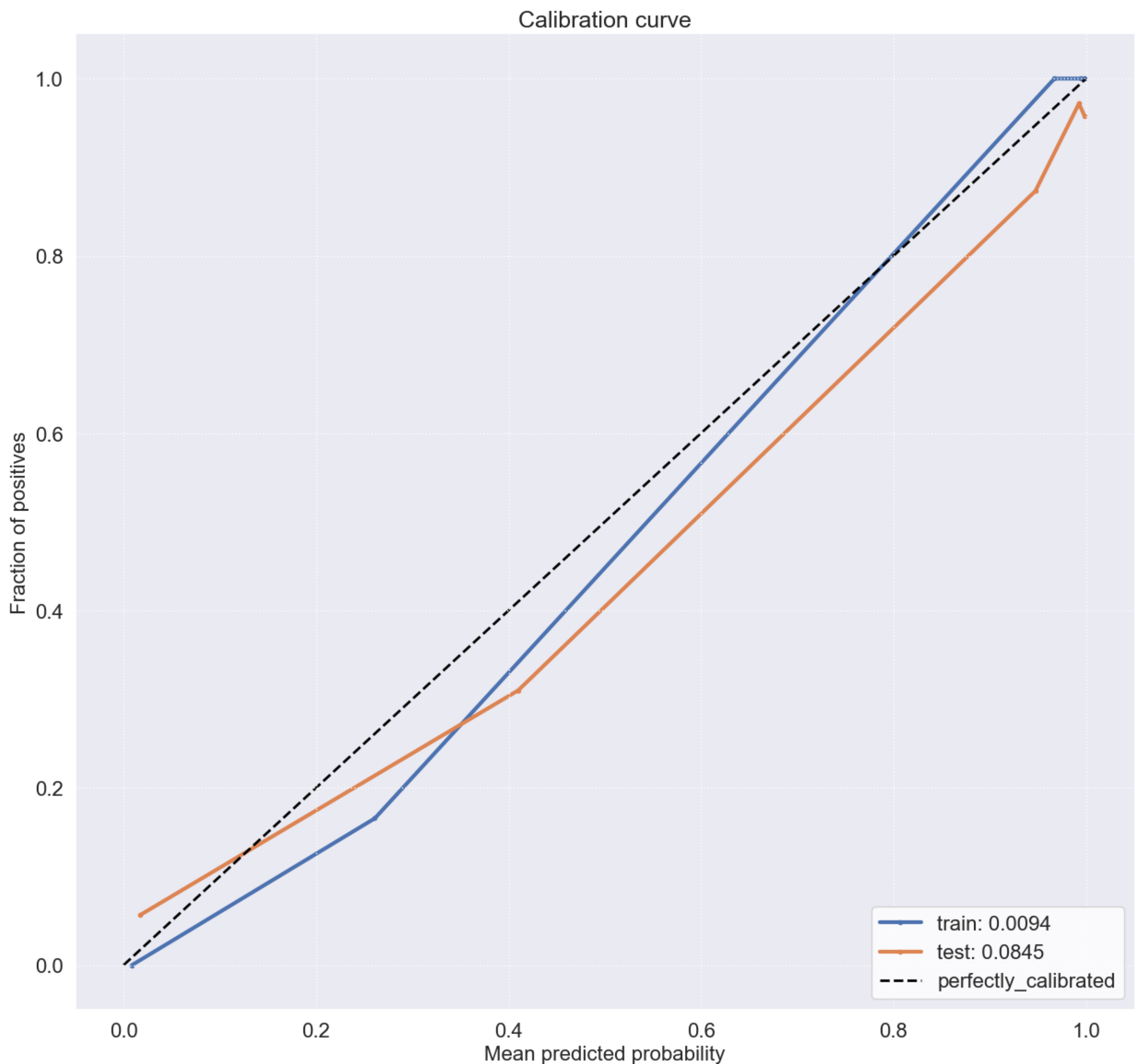- Pareto Front Trials
- Best Trial (# 2194)
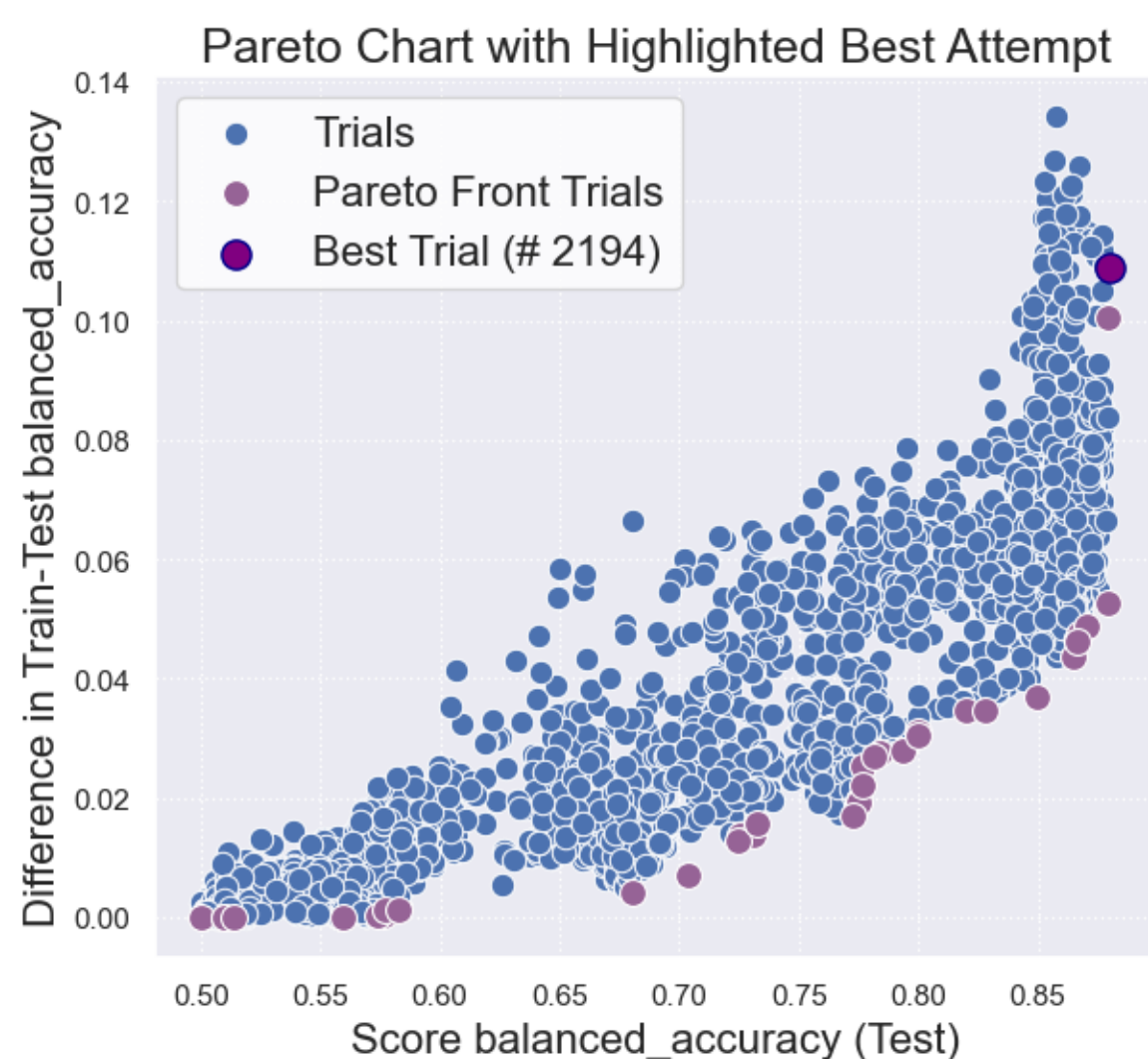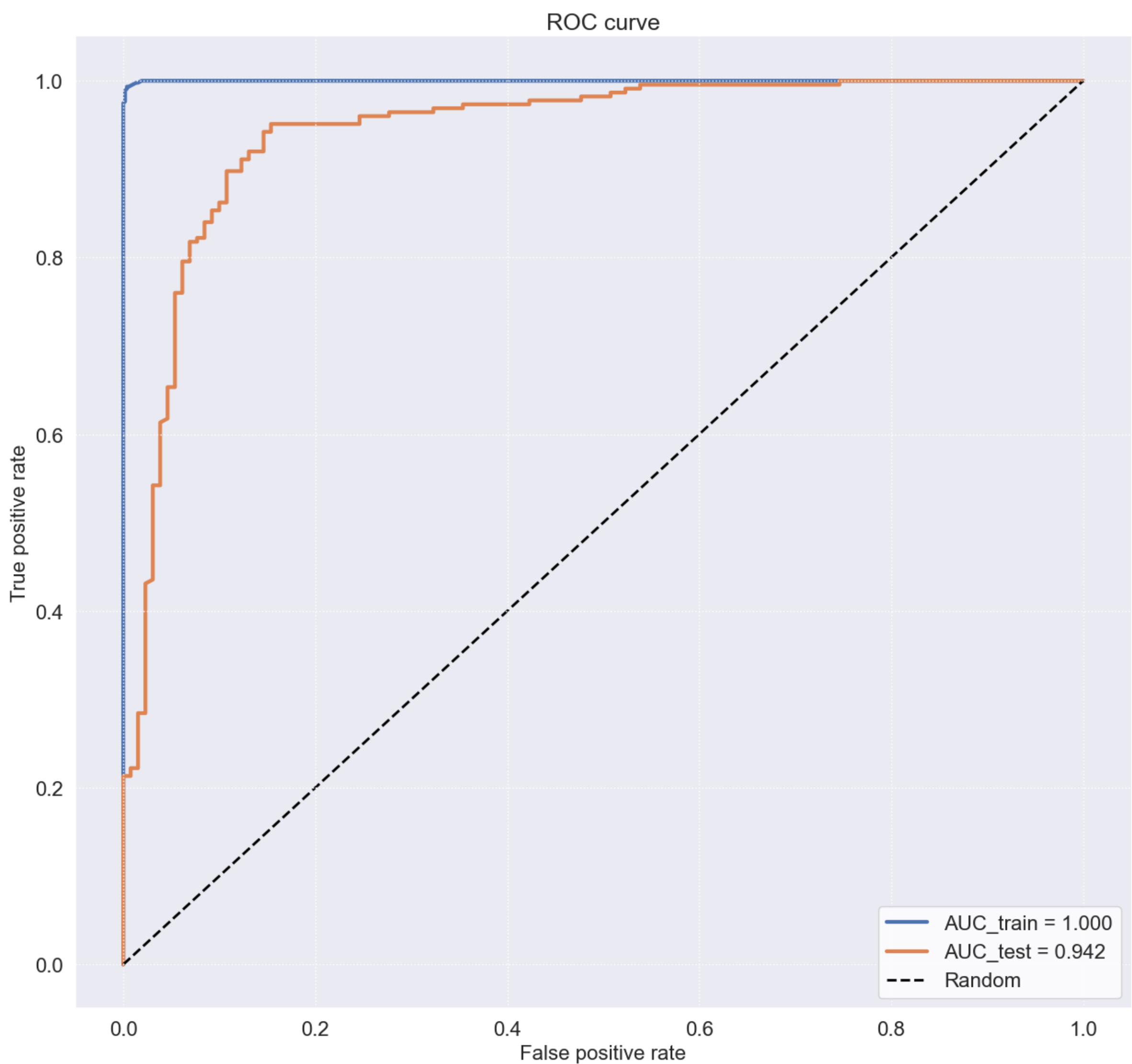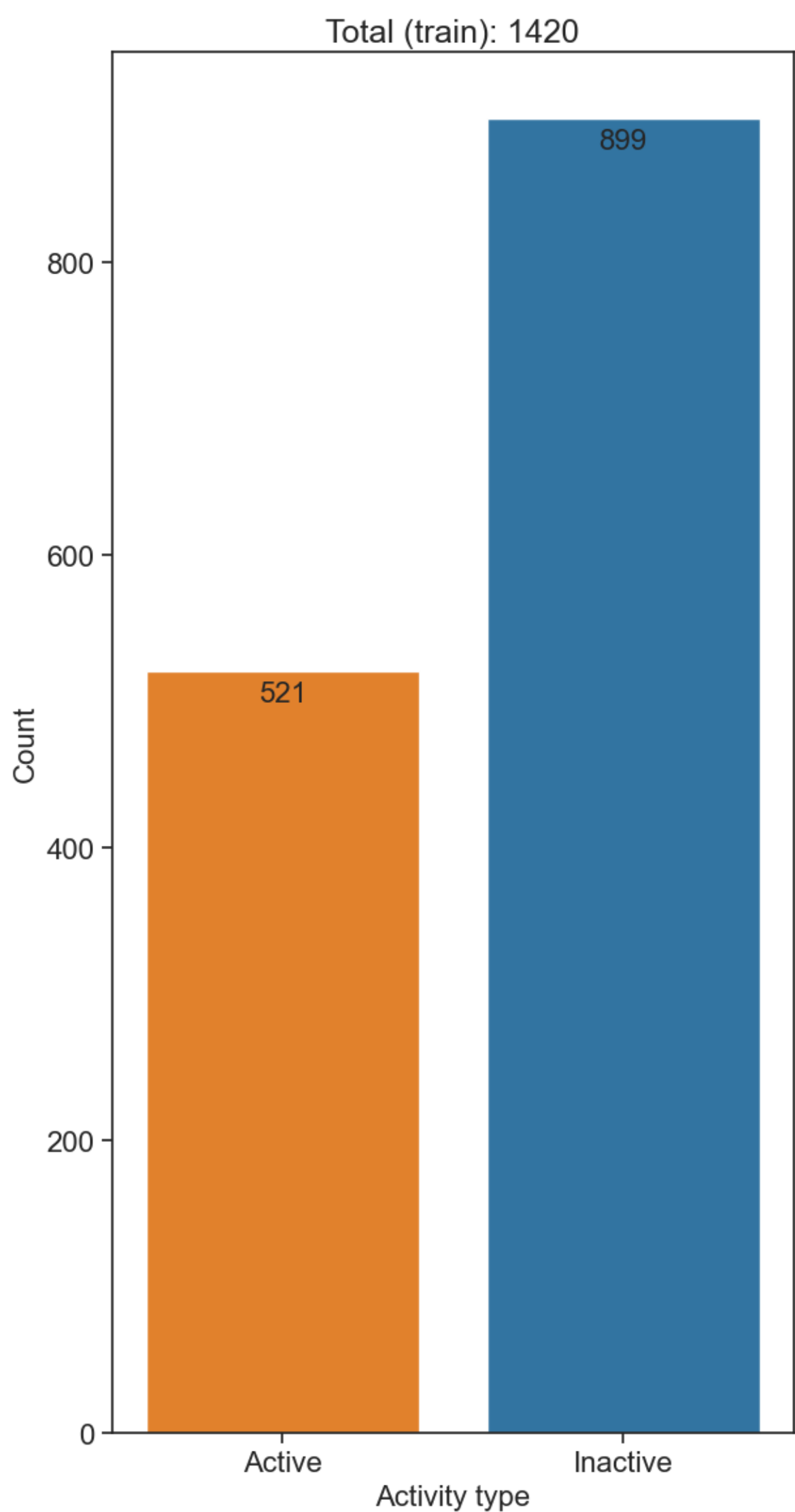
XGBoost Parameters	Value
alpha	0.3034
colsample_bytree	0.8403
gamma	0.2167
lambda	4.7257
learning_rate	0.0472
max_depth	10.0
min_child_weight	1.0
n_estimators	600.0
reg_alpha	0.2927
reg_lambda	0.5981
scale_pos_weight	2.0457
subsample	0.8429

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0094
test: 0.0845
perfectly_calibrated

Active/Inactive probability

type
active (225)
inactive (130)
decoy (354)
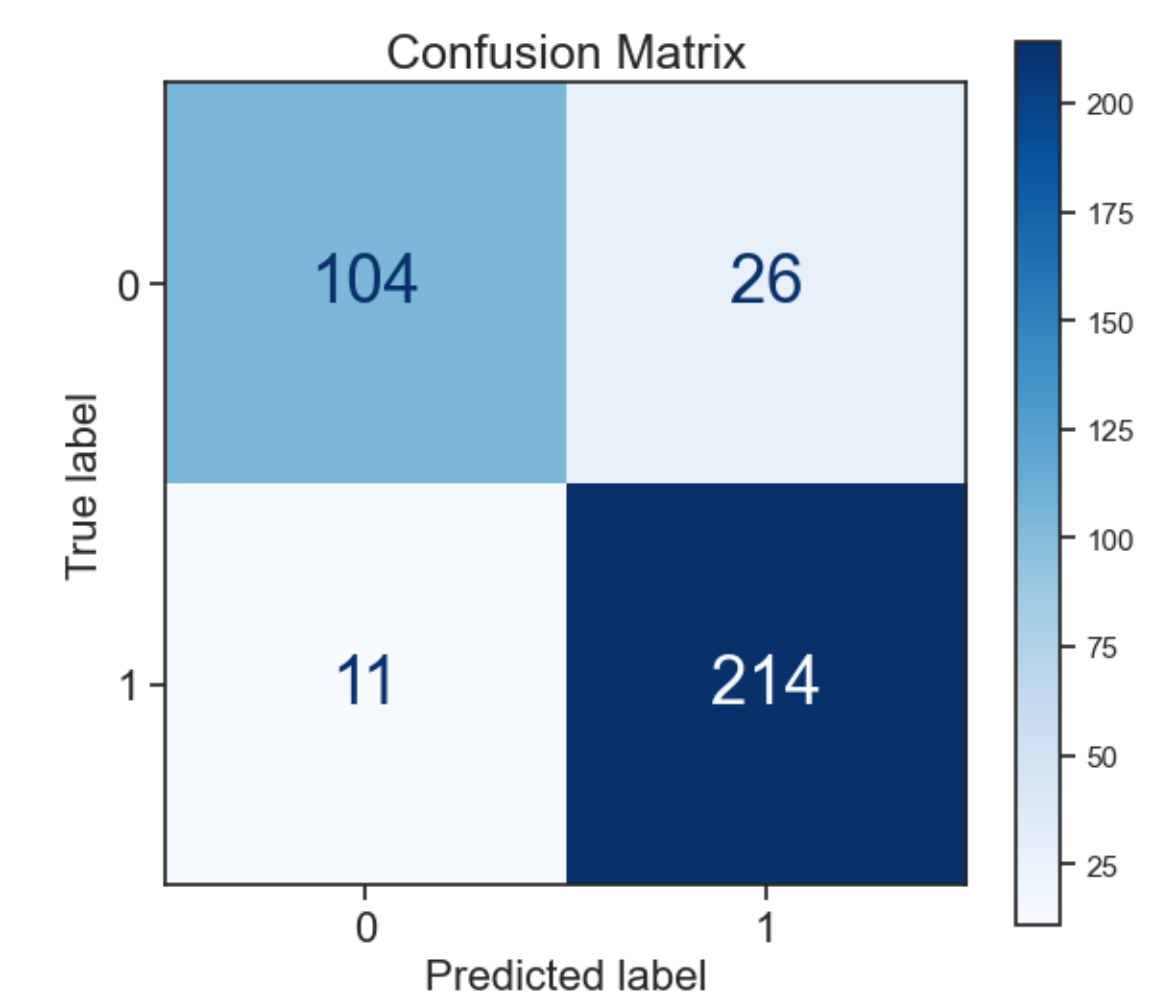
Metric	Train	Test
AUC	0.9999	0.9422
accuracy	0.9908	0.8958
balanced accuracy	0.9875	0.8756
recall_(sens)	1.0	0.9511
specificity	0.975	0.8
precision	0.9857	0.8917
f1_score	0.9908	0.8943

Protein: P04150 - metric: f1

Total (train): 1420

ROC curve

AUC_train = 1.000
AUC_test = 0.941
Random

Pareto Chart with Highlighted Best Attempt

Trials
Pareto Front Trials
Best Trial (# 1682)
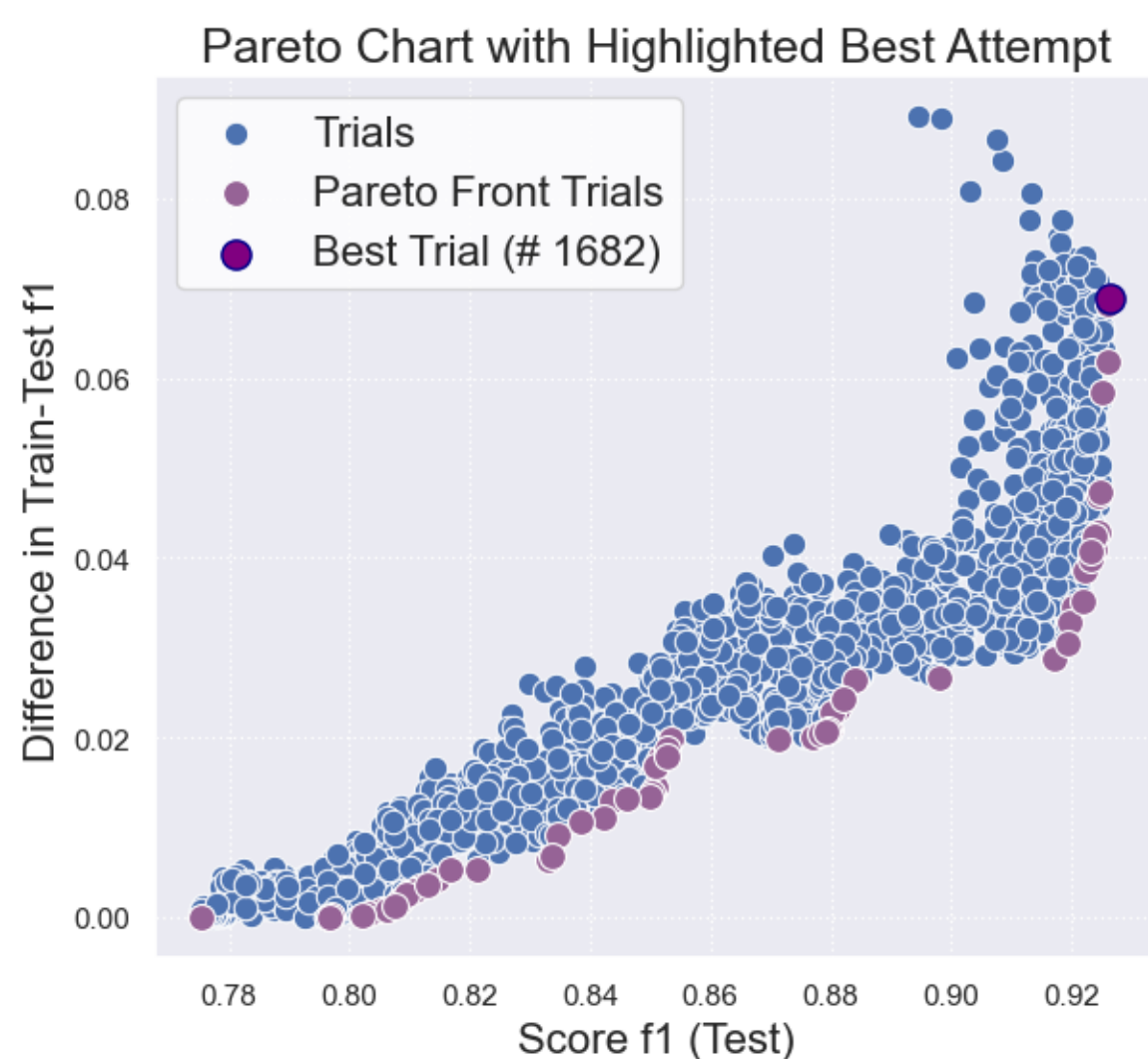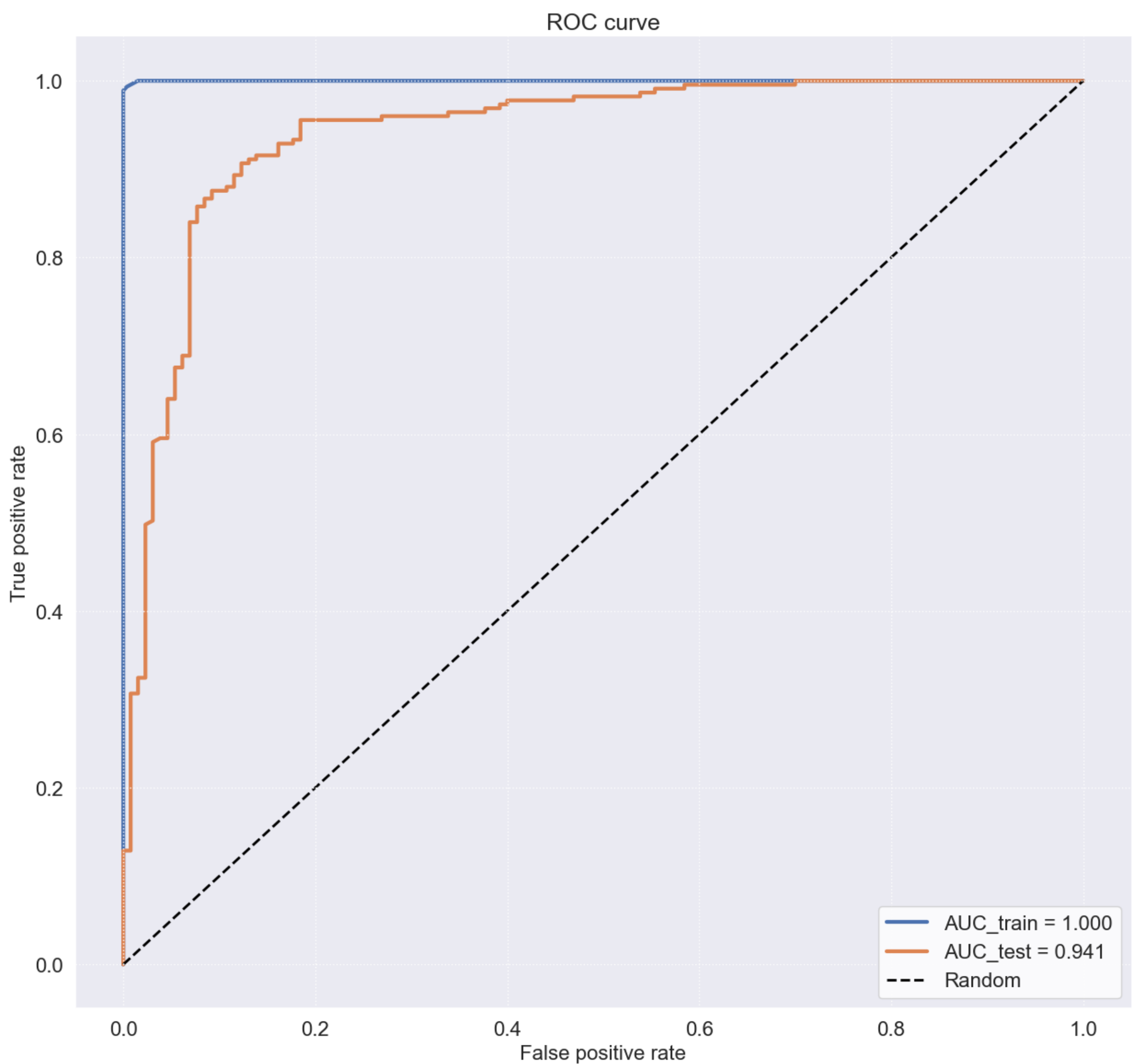
XGBoost Parameters	Value
alpha	3.7009
colsample_bytree	0.6172
gamma	0.0464
lambda	1.5993
learning_rate	0.0563
max_depth	10.0
min_child_weight	1.0
n_estimators	793.0
reg_alpha	0.2877
reg_lambda	0.1742
scale_pos_weight	3.9924
subsample	0.6152

Confusion Matrix

Calibration curve

train: 0.0065
test: 0.0917
perfectly_calibrated

Confusion Matrix (nomr.)

Active/Inactive probability

Inactive zone
Active zone

type
active (225)
inactive (130)
decoy (354)
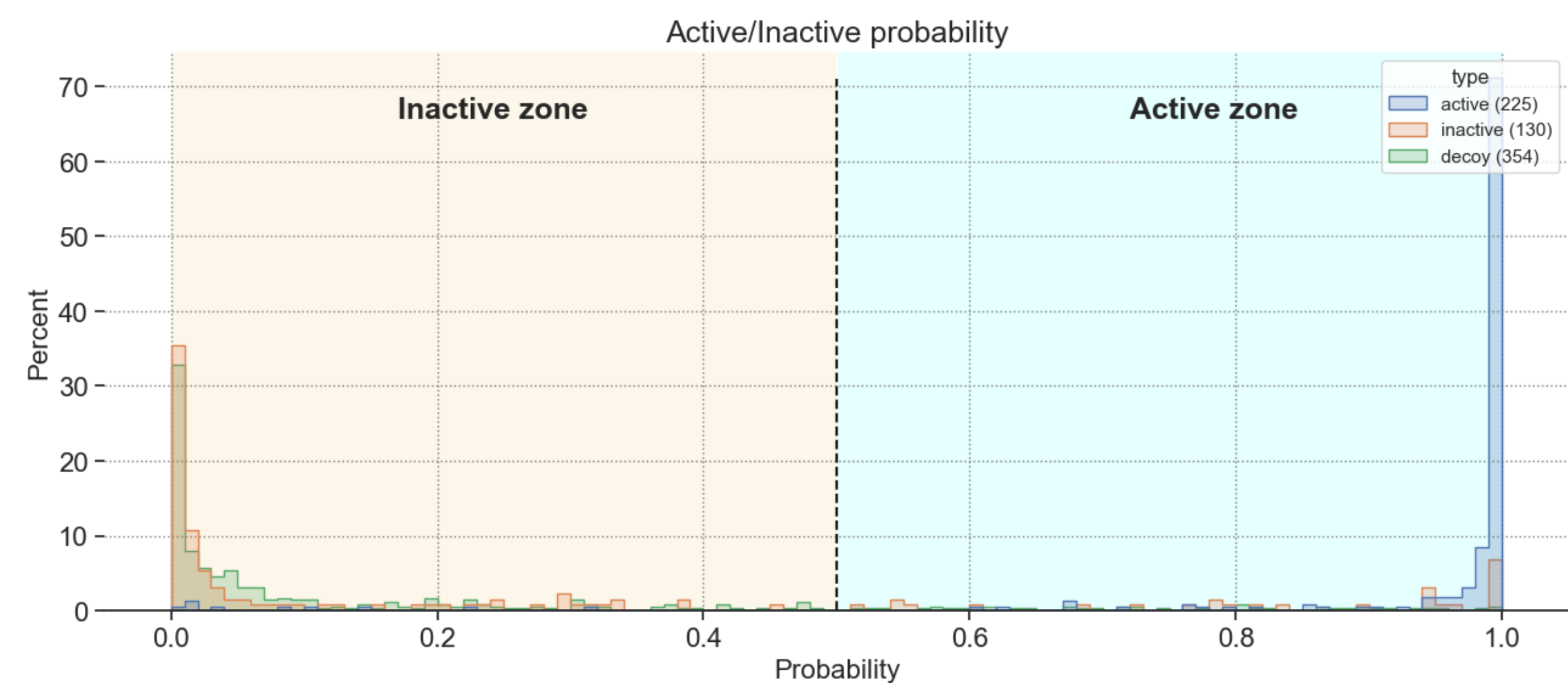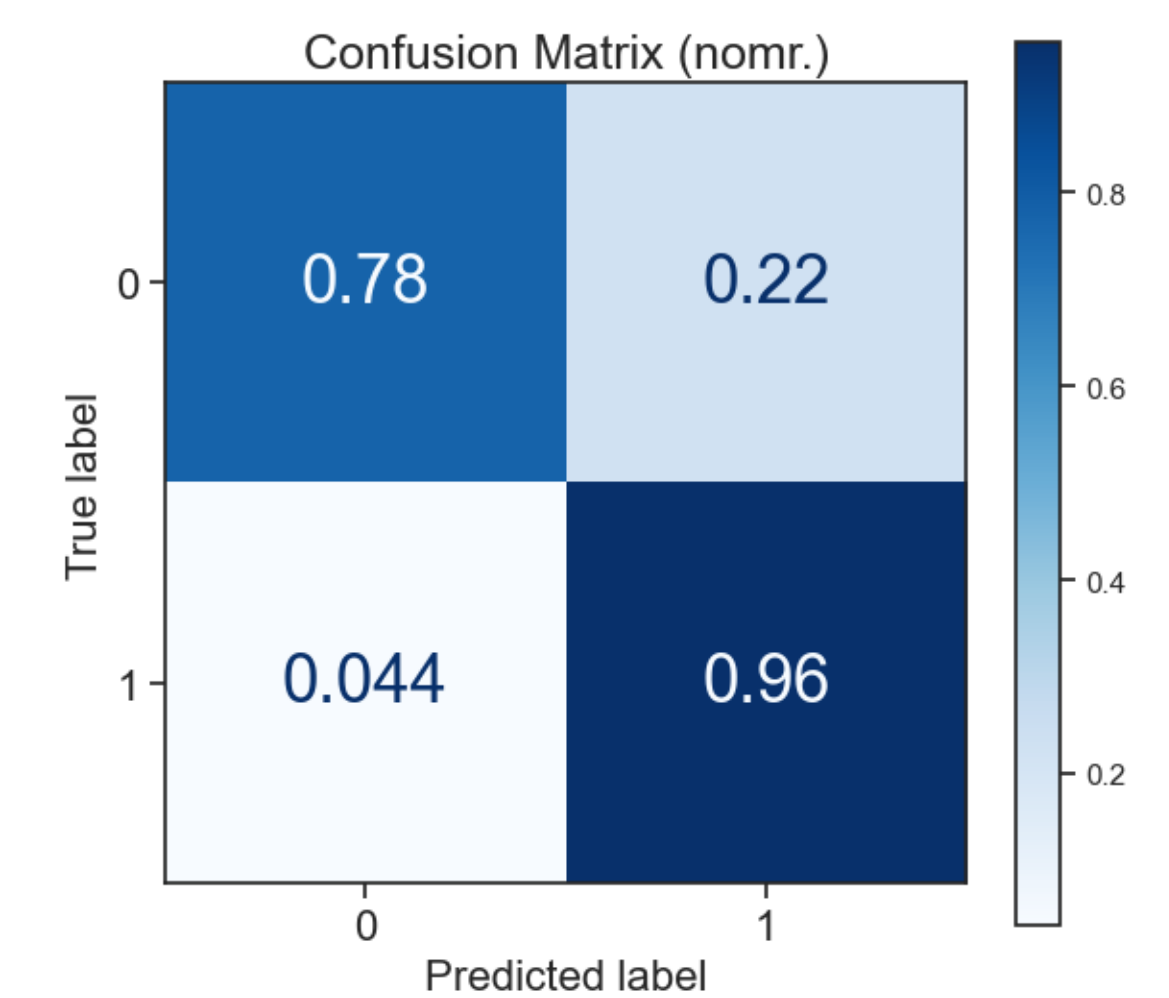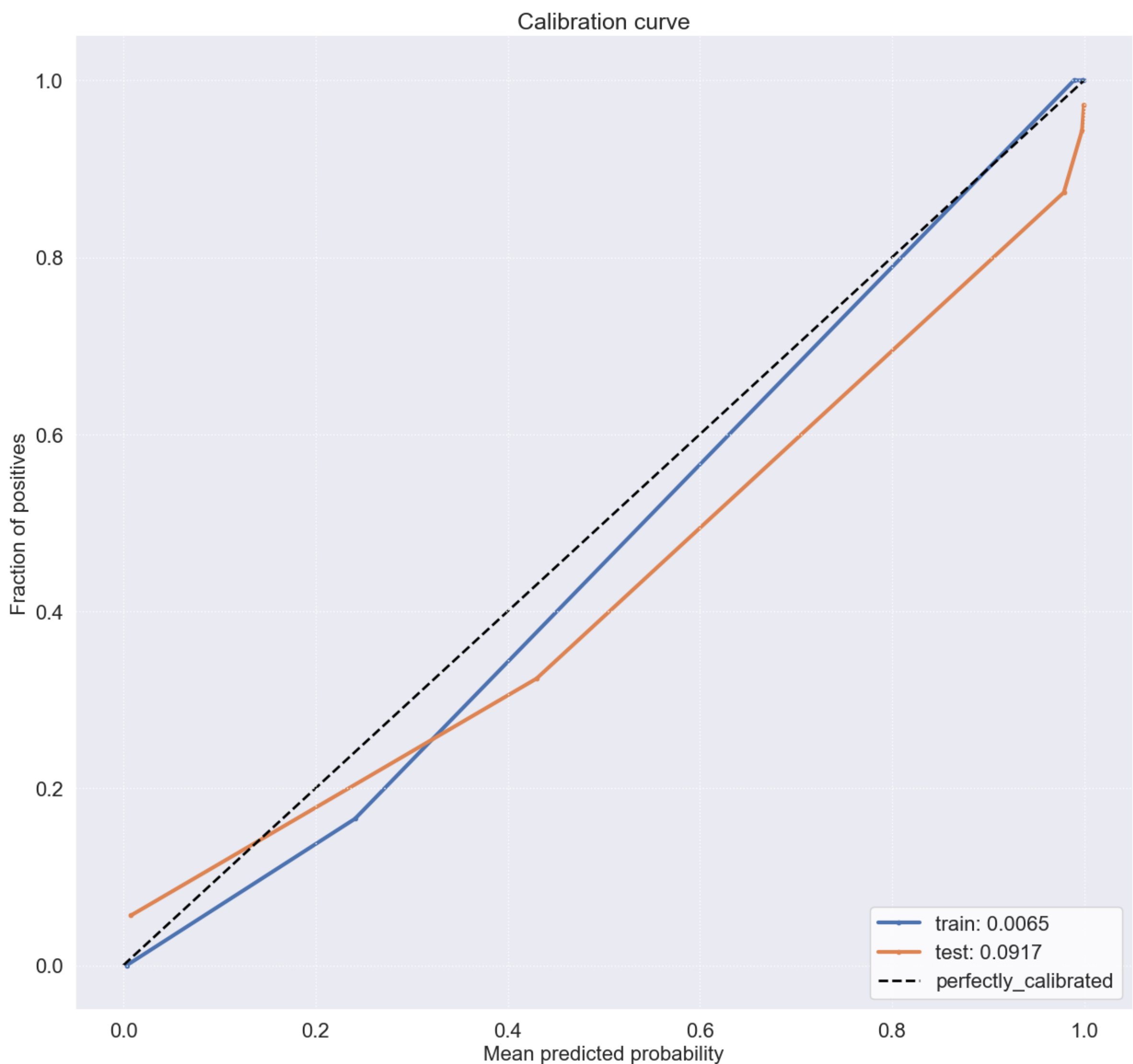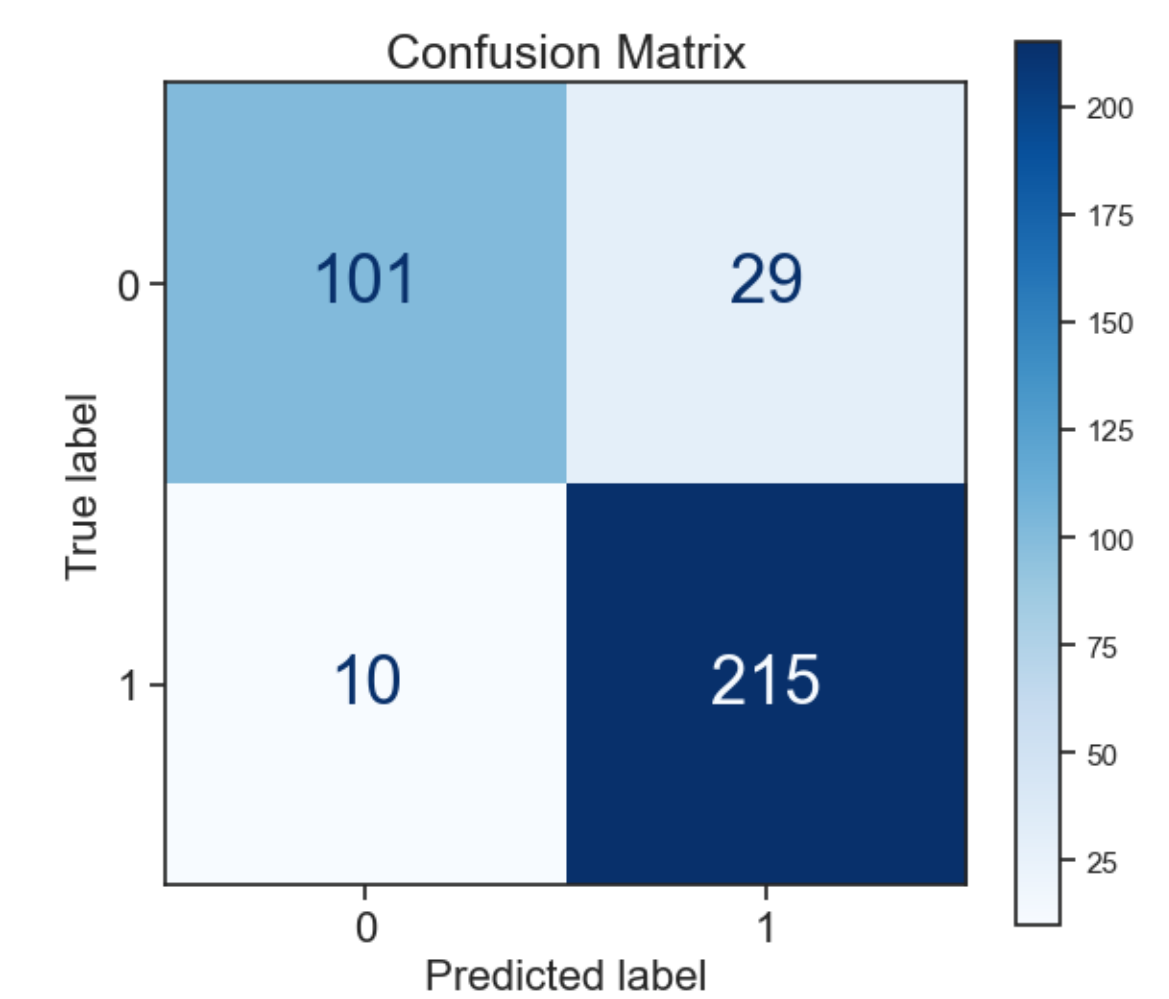
Metric	Train	Test
AUC	0.9999	0.9407
accuracy	0.993	0.8901
balanced accuracy	0.9904	0.8662
recall_(sens)	1.0	0.9556
specificity	0.9808	0.7769
precision	0.989	0.8811
f1_score	0.9929	0.888

Protein: P04150 - metric: precision

Total (train): 1420

ROC curve

AUC_train = 0.996
AUC_test = 0.936
Random

Pareto Chart with Highlighted Best Attempt

Trials
Pareto Front Trials
Best Trial (# 1080)
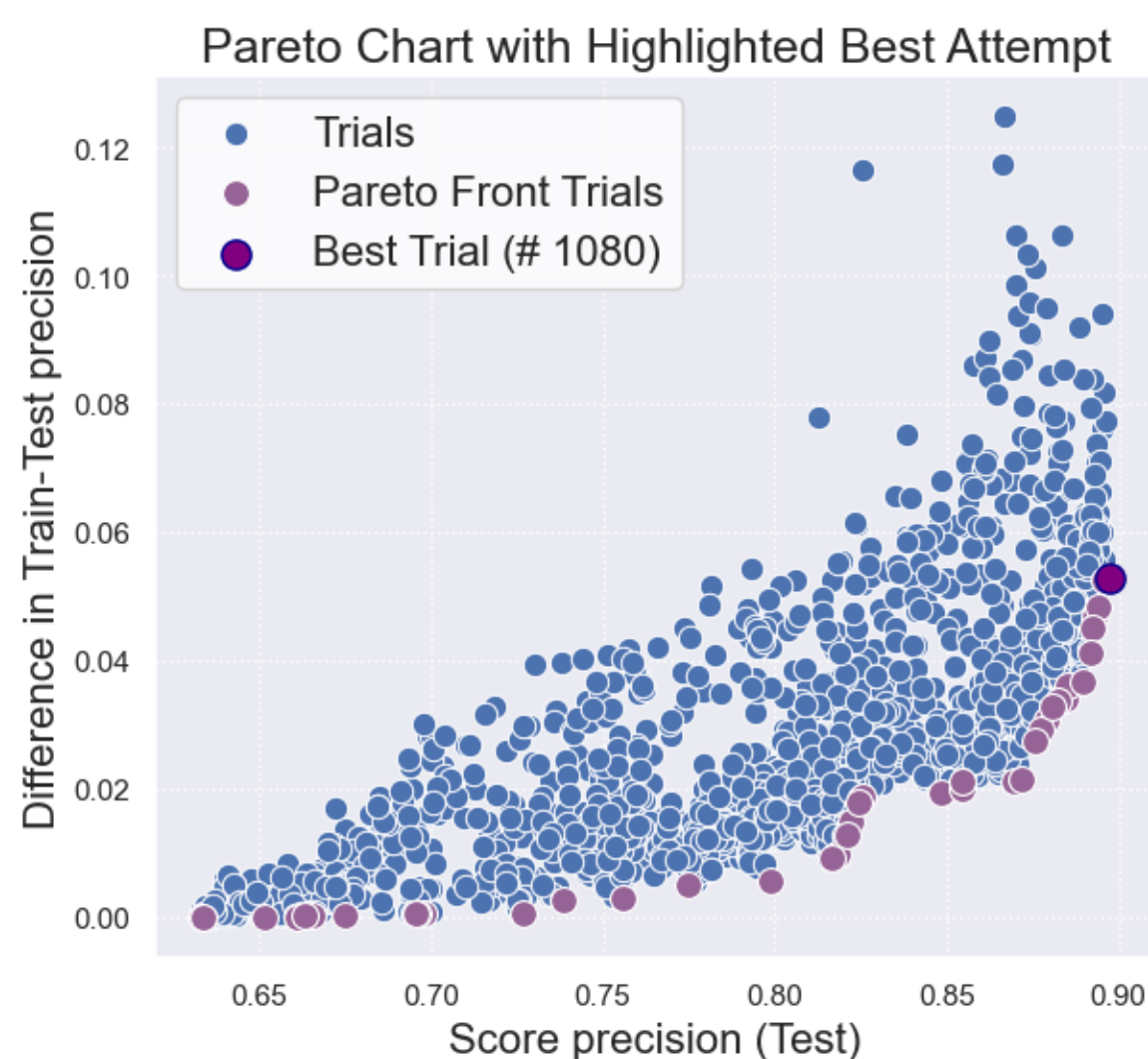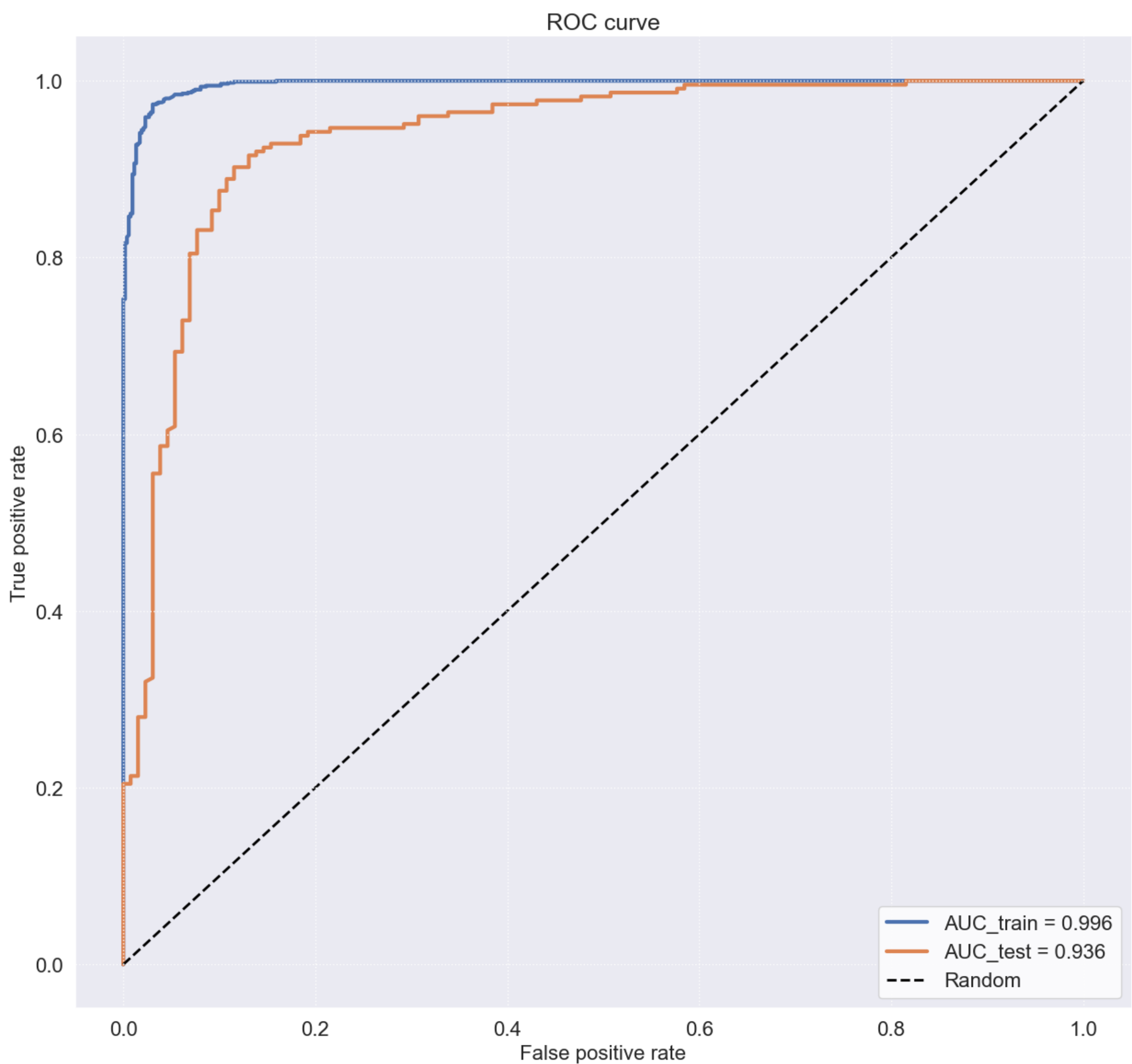
XGBoost Parameters	Value
alpha	0.9455
colsample_bytree	0.6514
gamma	0.886
lambda	2.6787
learning_rate	0.0346
max_depth	8.0
min_child_weight	2.0
n_estimators	862.0
reg_alpha	0.2388
reg_lambda	0.942
scale_pos_weight	1.3901
subsample	0.761

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0293
test: 0.0893
perfectly_calibrated

Active/Inactive probability

Inactive zone
Active zone

type
active (225)
inactive (130)
decoy (354)
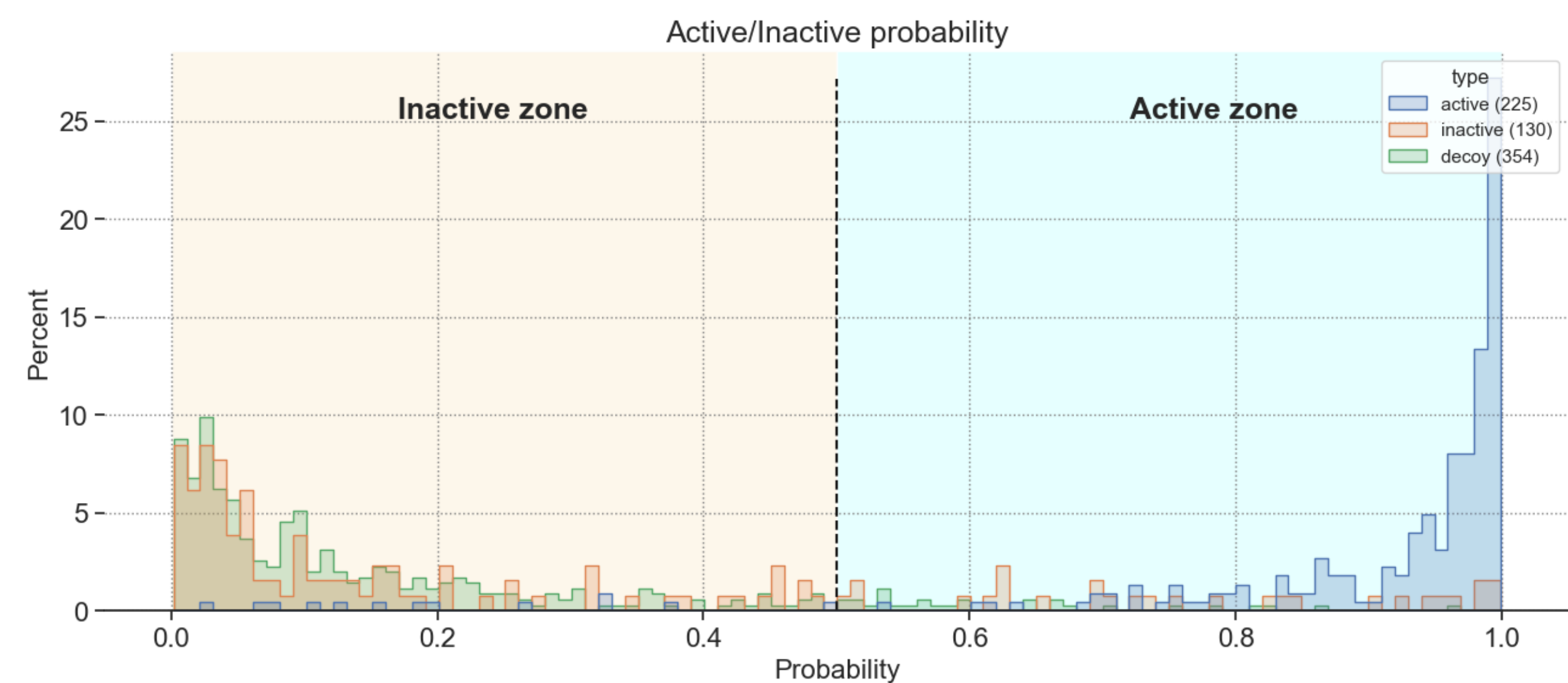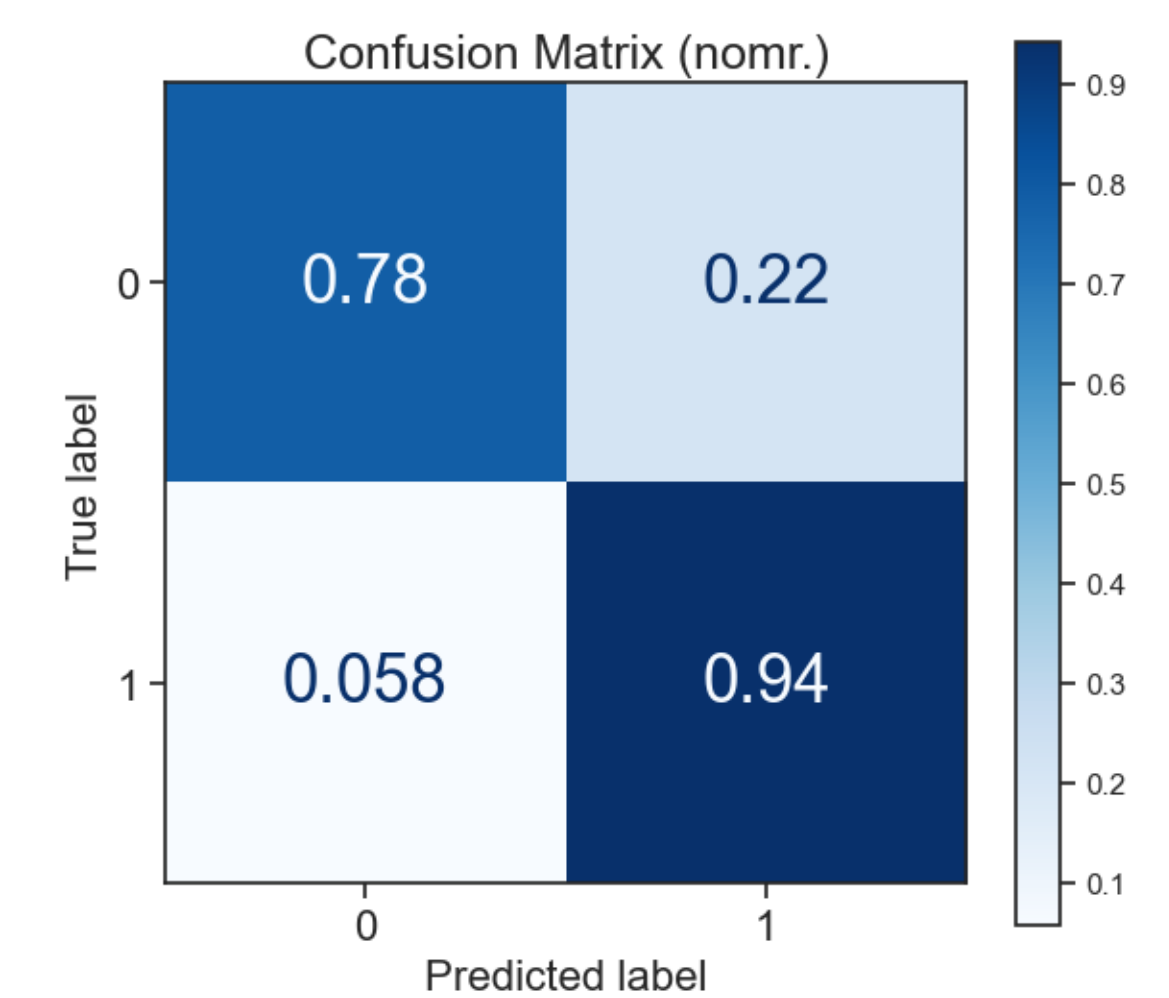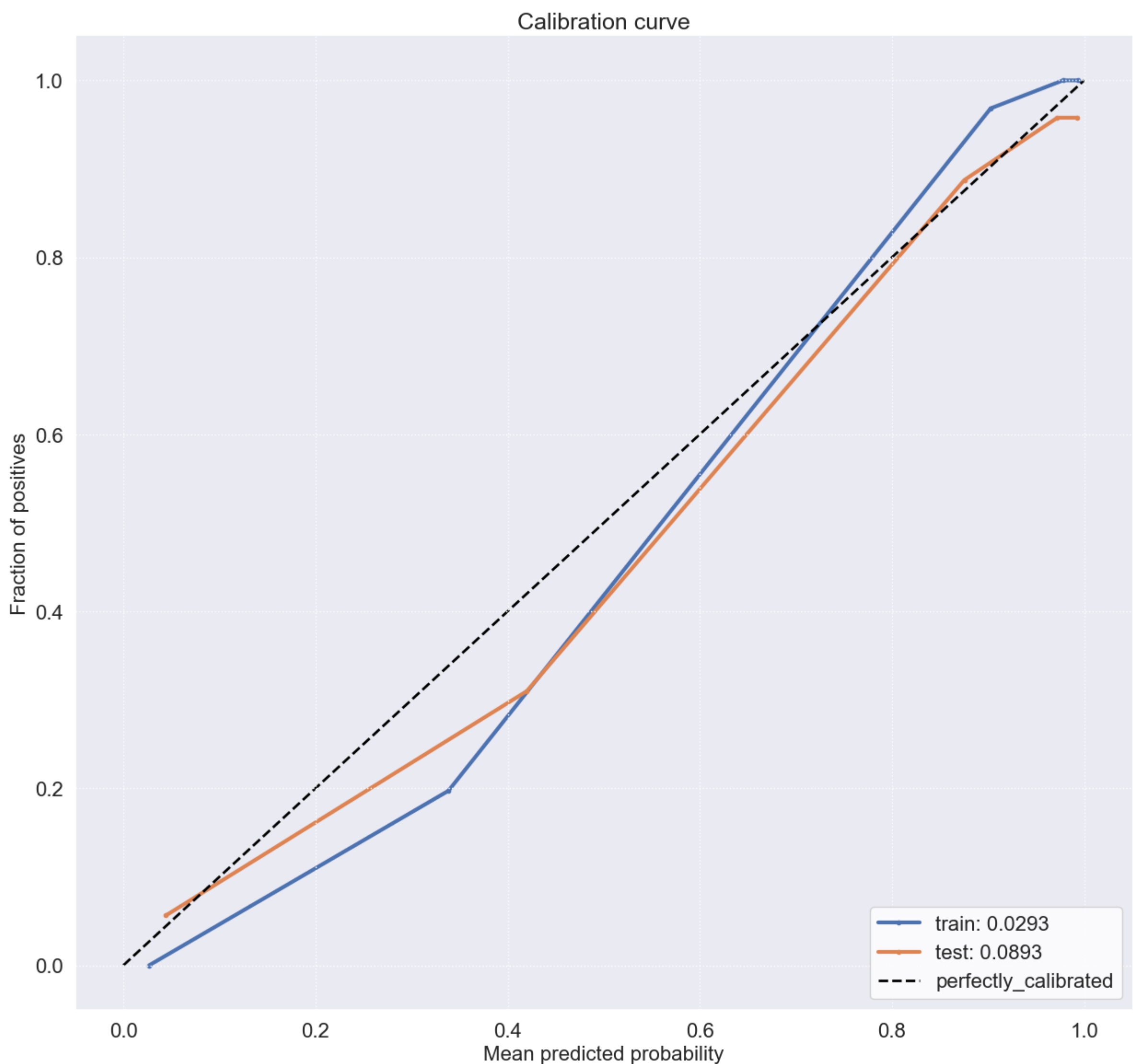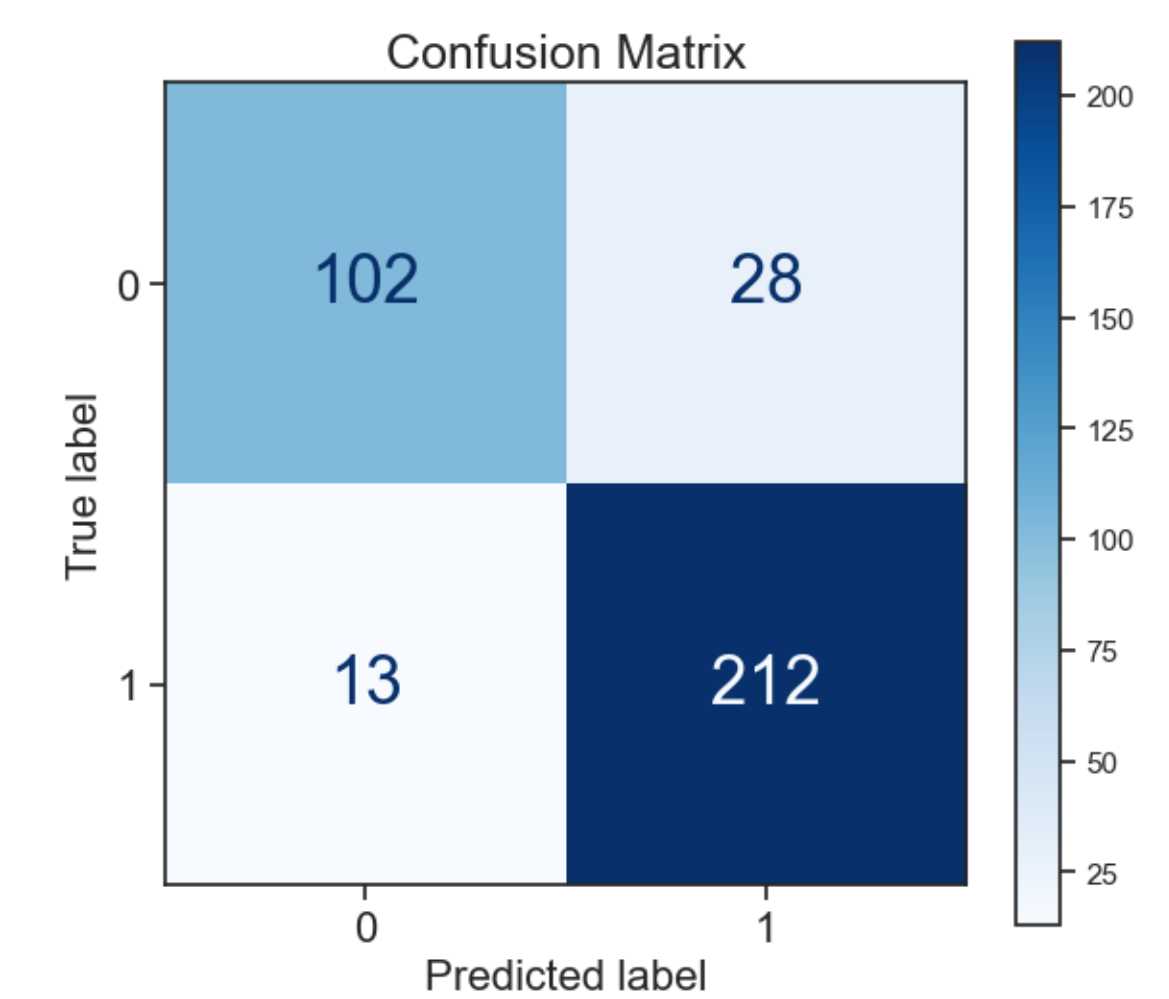
Metric	Train	Test
AUC	0.9959	0.9361
accuracy	0.9599	0.8845
balanced accuracy	0.9473	0.8634
recall_(sens)	0.9944	0.9422
specificity	0.9002	0.7846
precision	0.945	0.8833
f1_score	0.9594	0.8828

Protein: P04150 - metric: SoftVotingXGBoostEnsemble
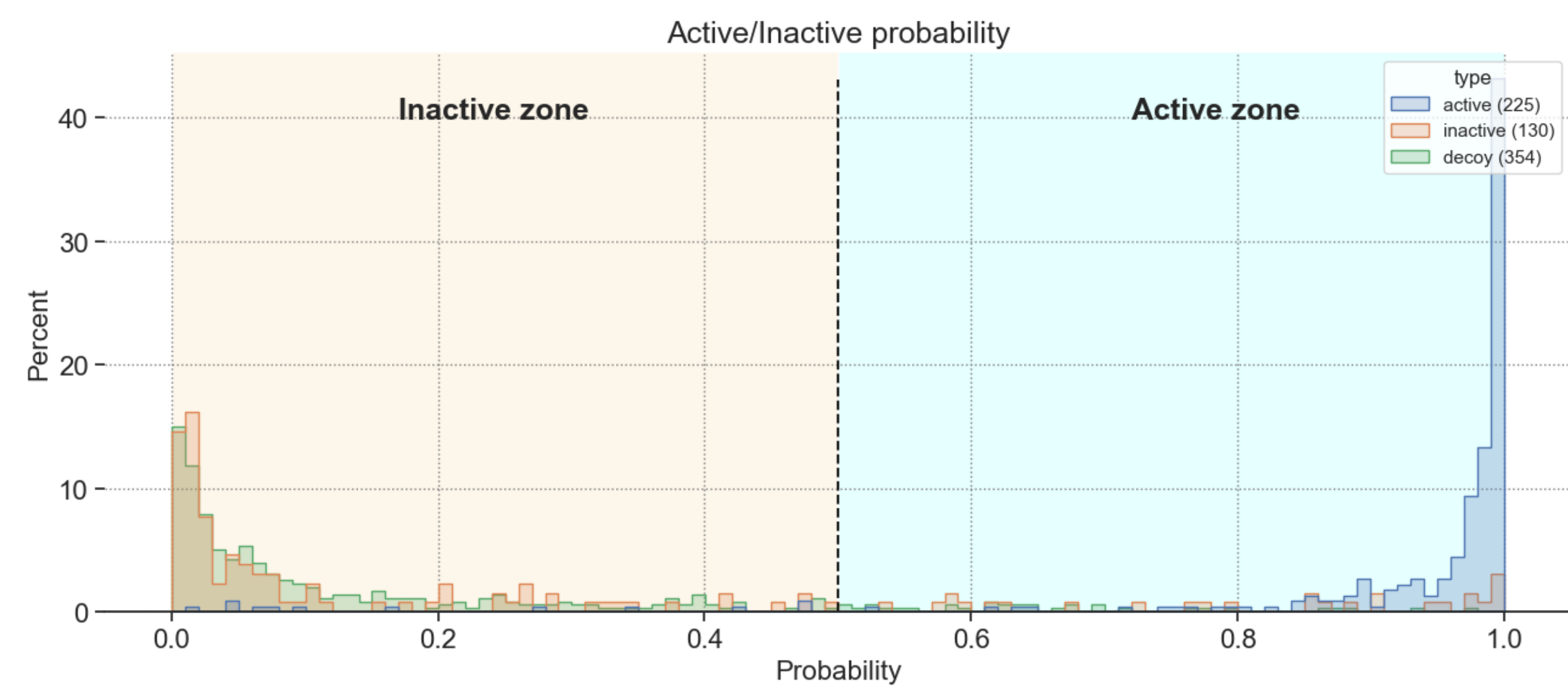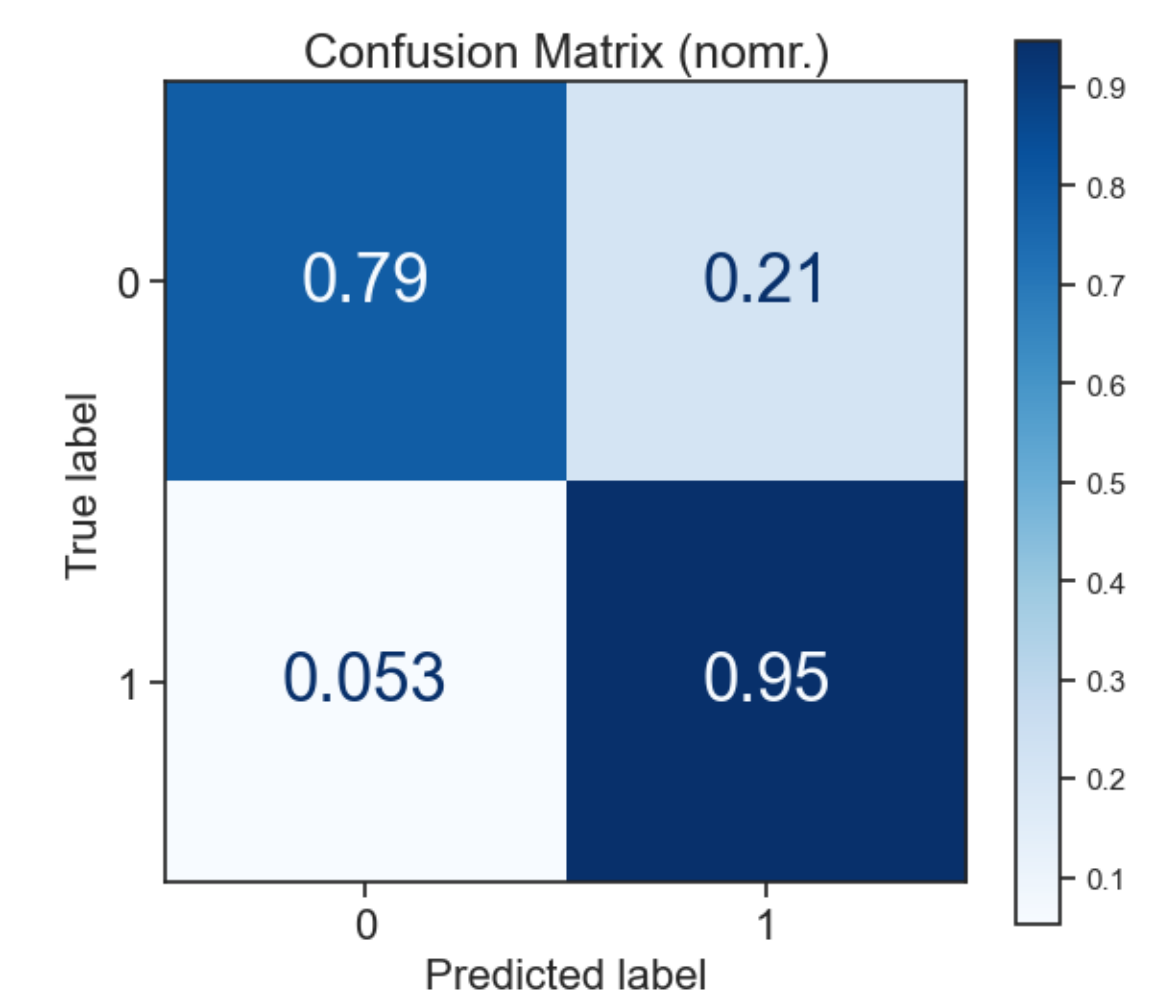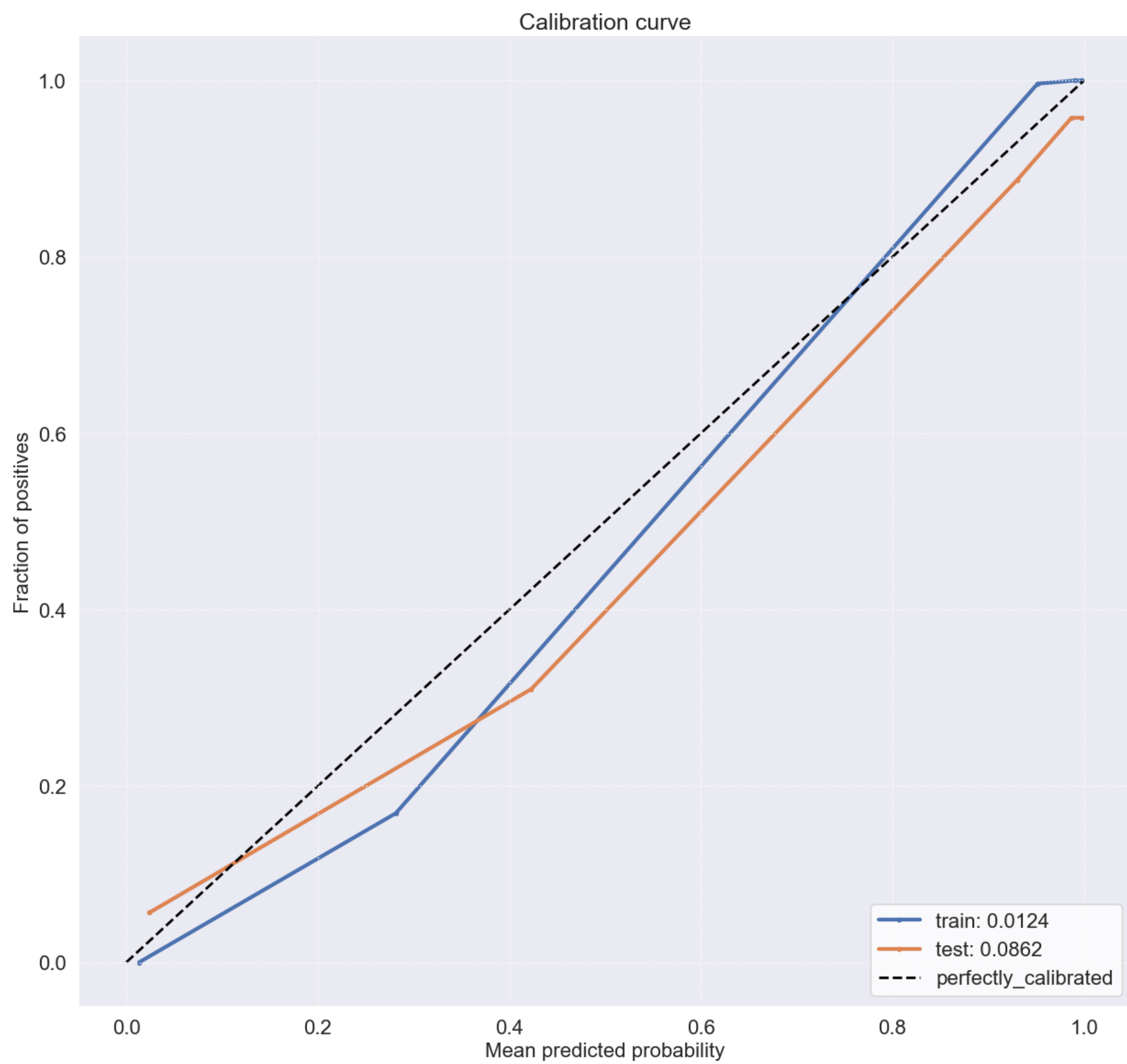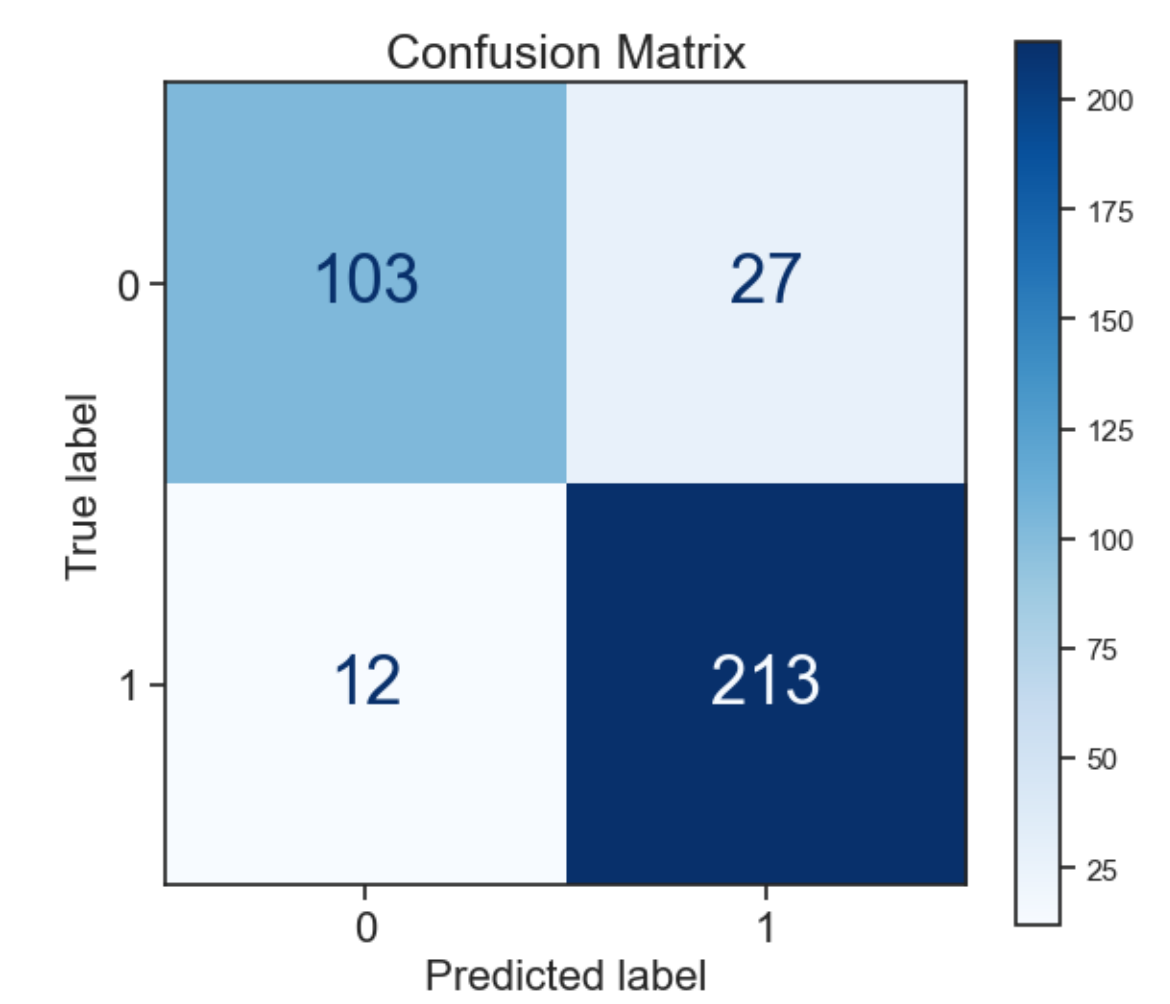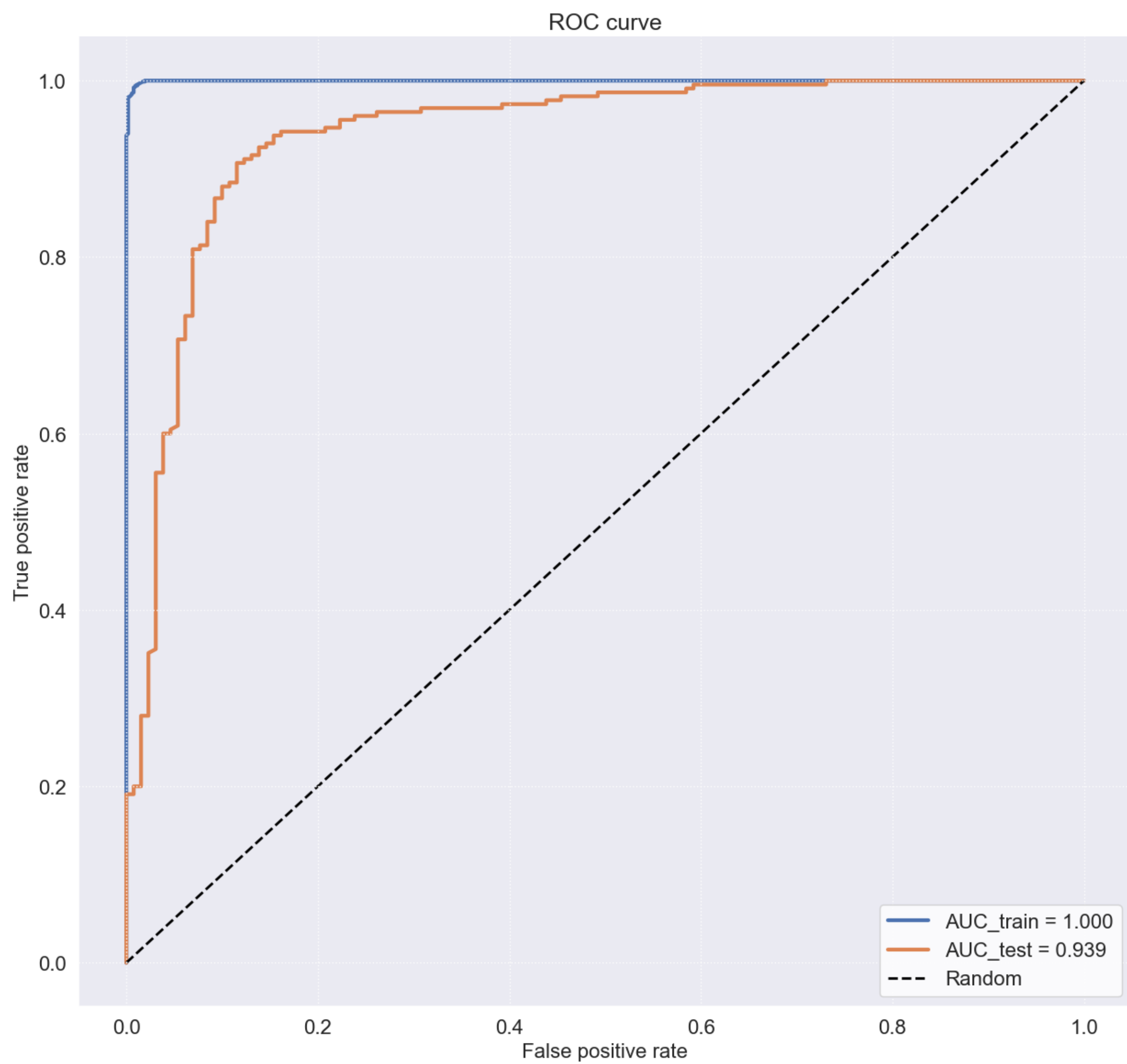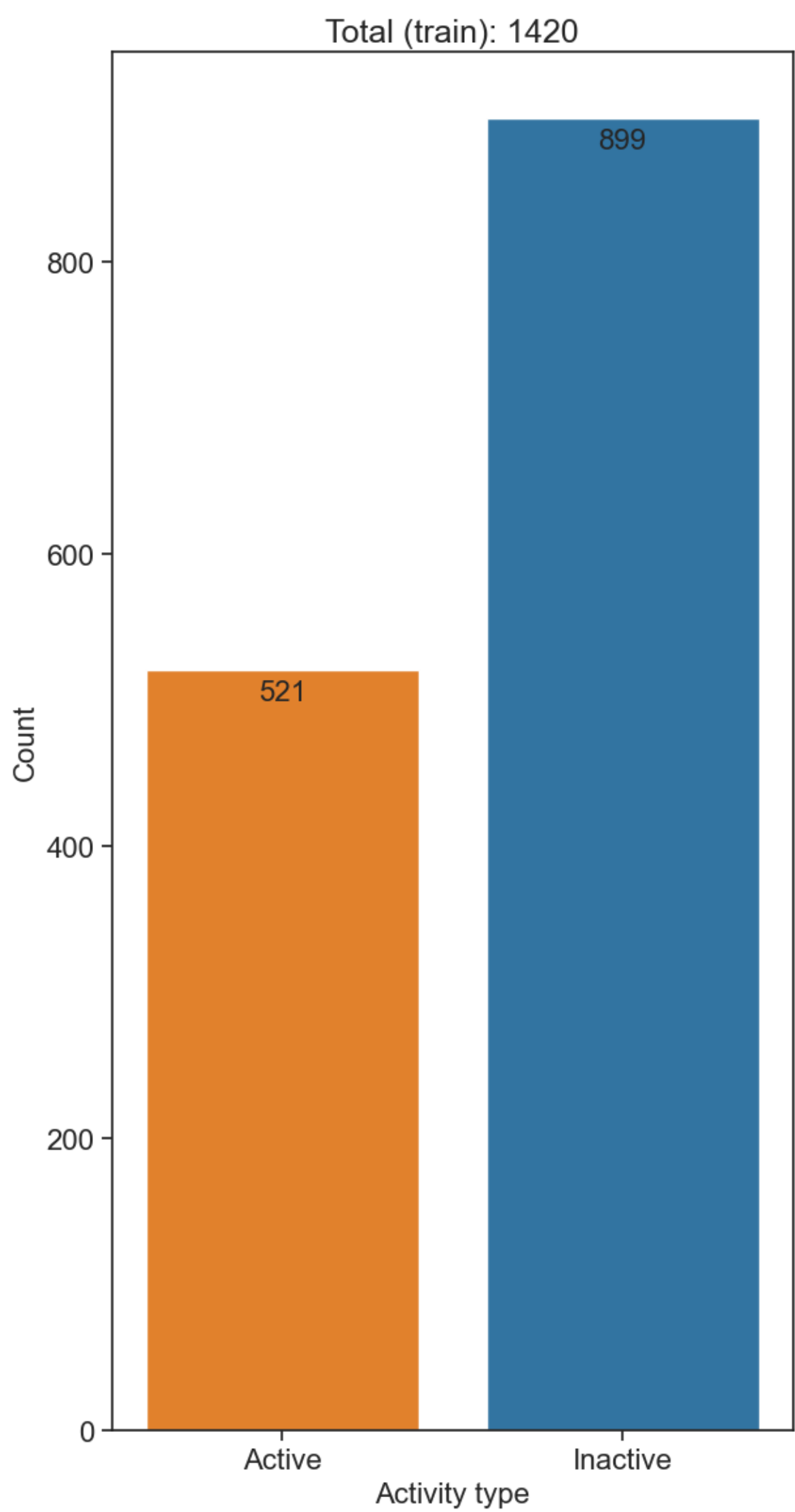
Metric	Train	Test
AUC	0.9997	0.9391
accuracy	0.9887	0.8901
balanced accuracy	0.9846	0.8695
recall_(sens)	1.0	0.9467
specificity	0.9693	0.7923
precision	0.9825	0.8875
f1_score	0.9887	0.8885