Protein: P22607 - metric: balanced_accuracy

Total (train): 758

ROC curve
- AUC_train = 1.000
- AUC_test = 0.933
- Random

Pareto Chart with Highlighted Best Attempt
- Trials
- Pareto Front Trials
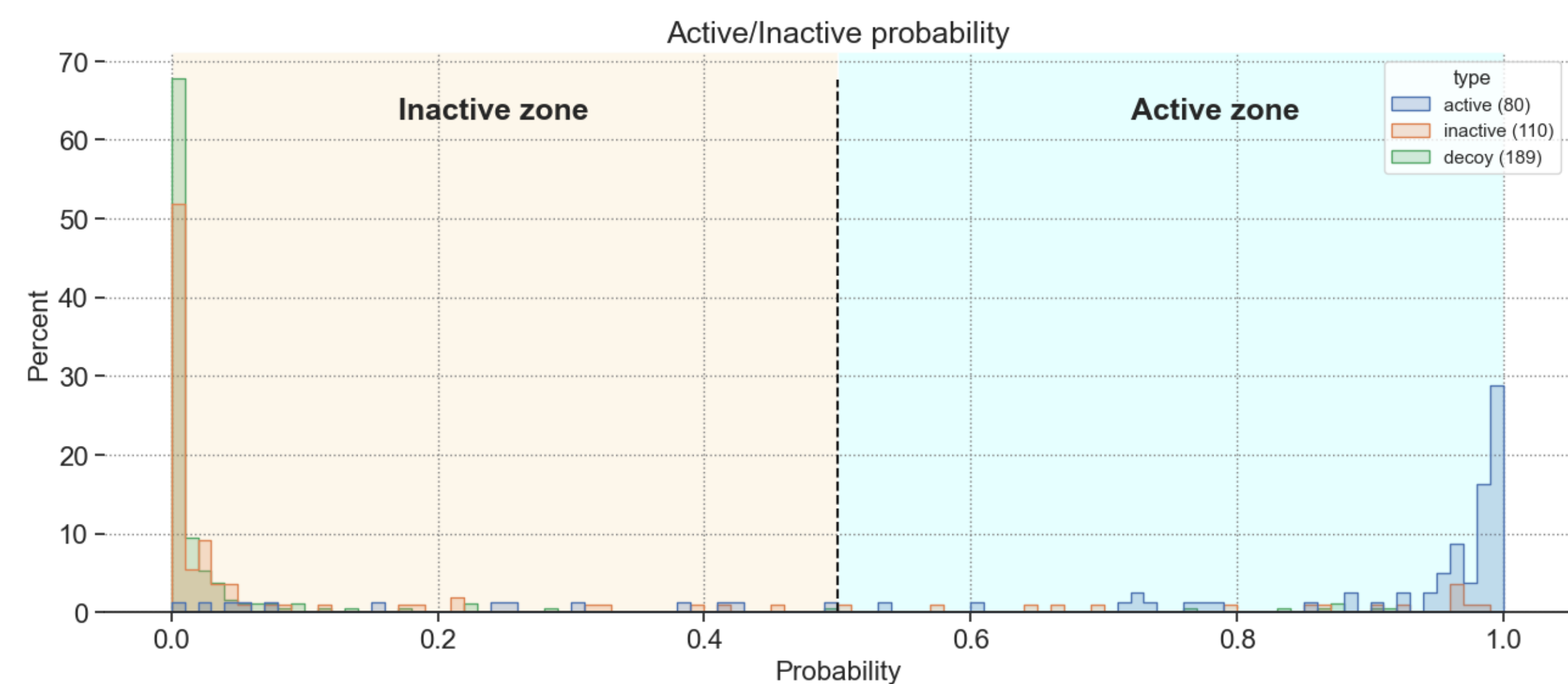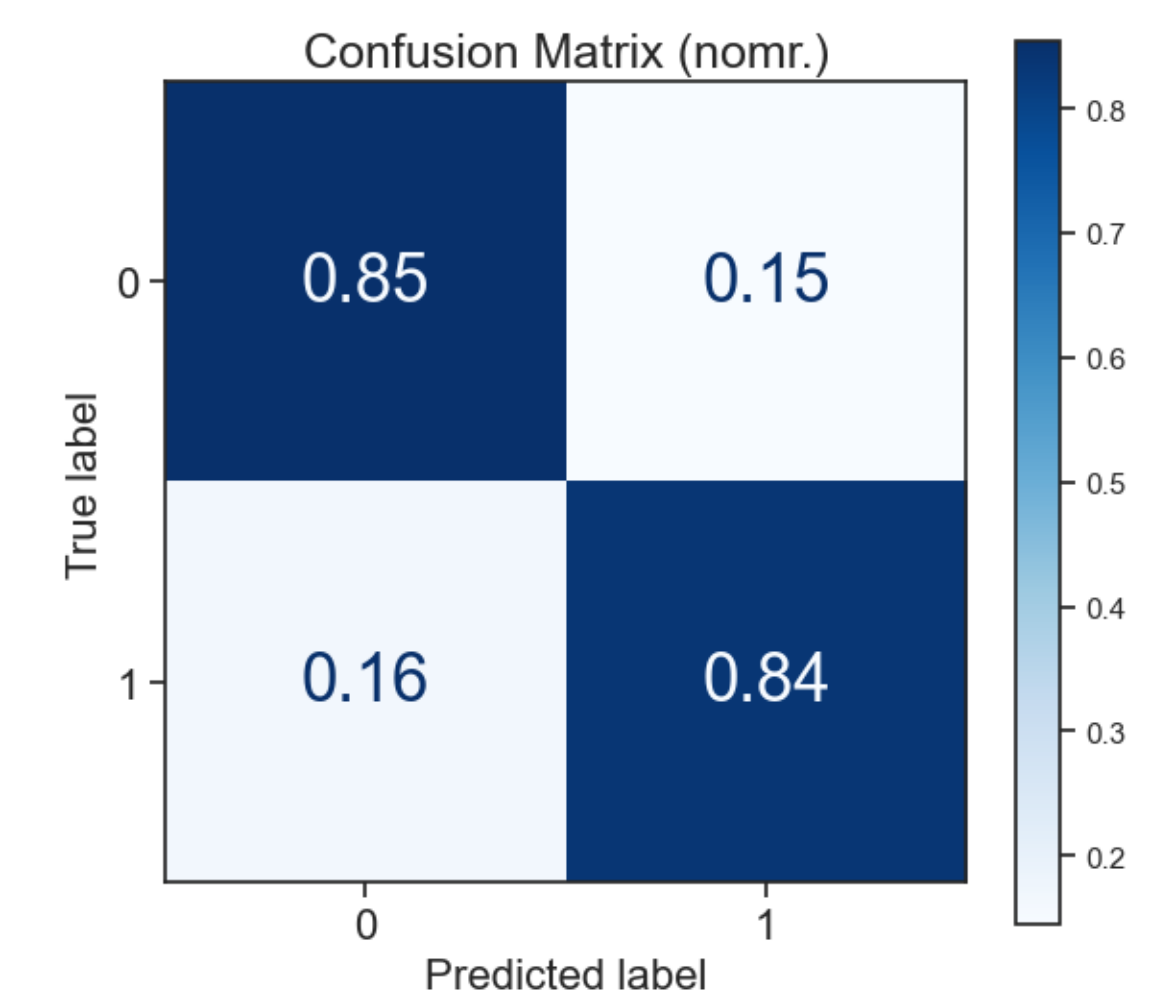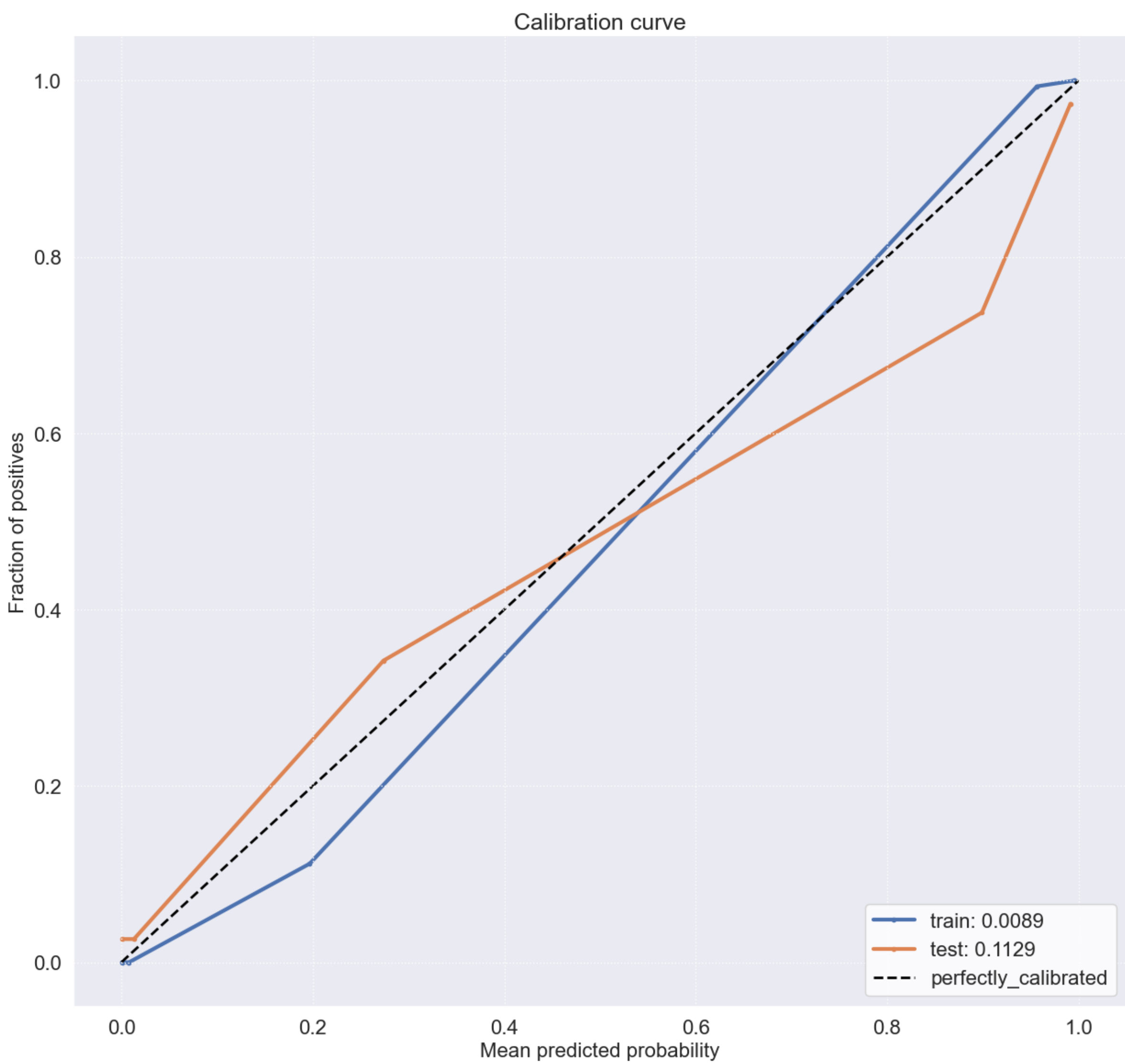- Best Trial (# 726)
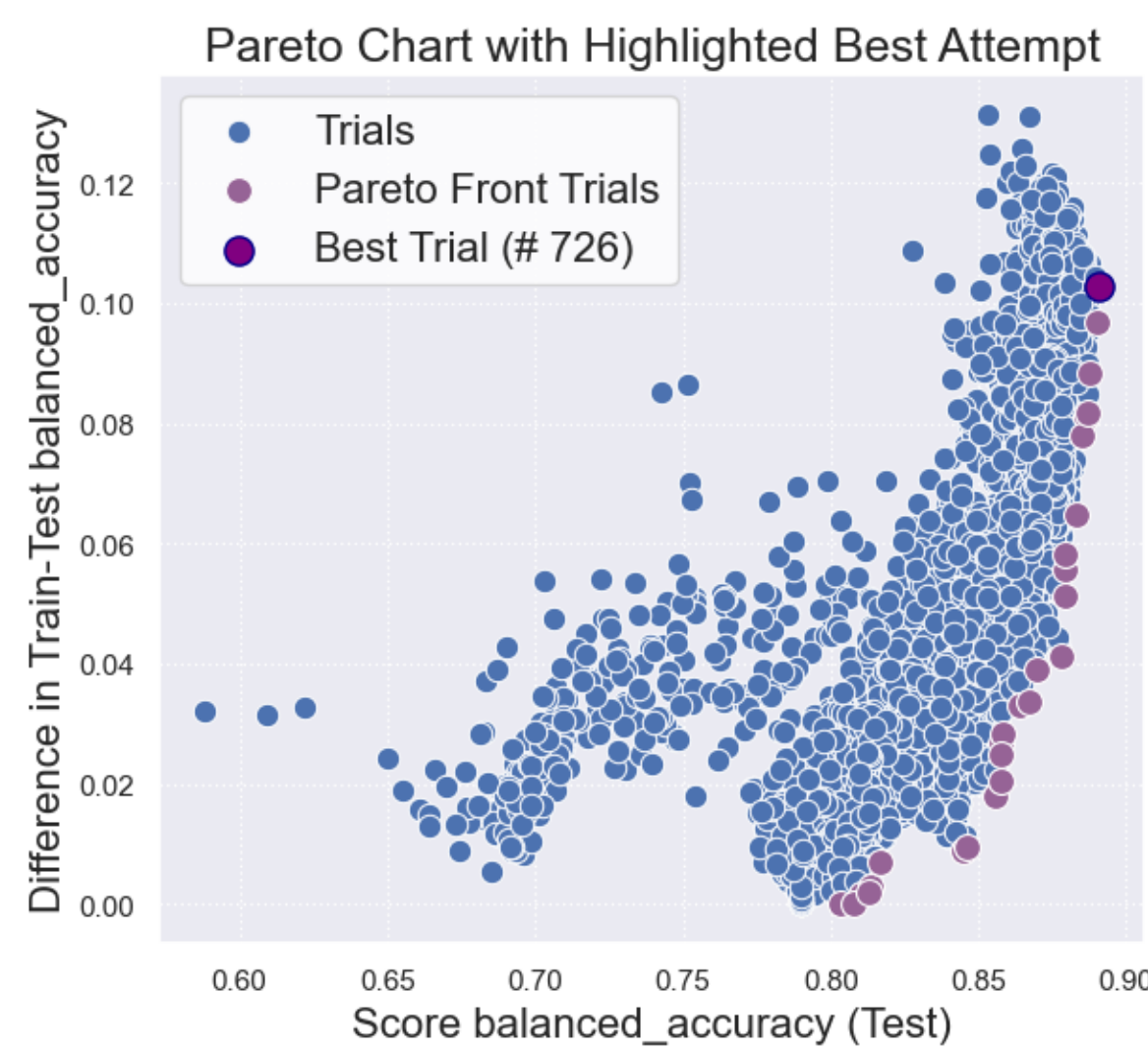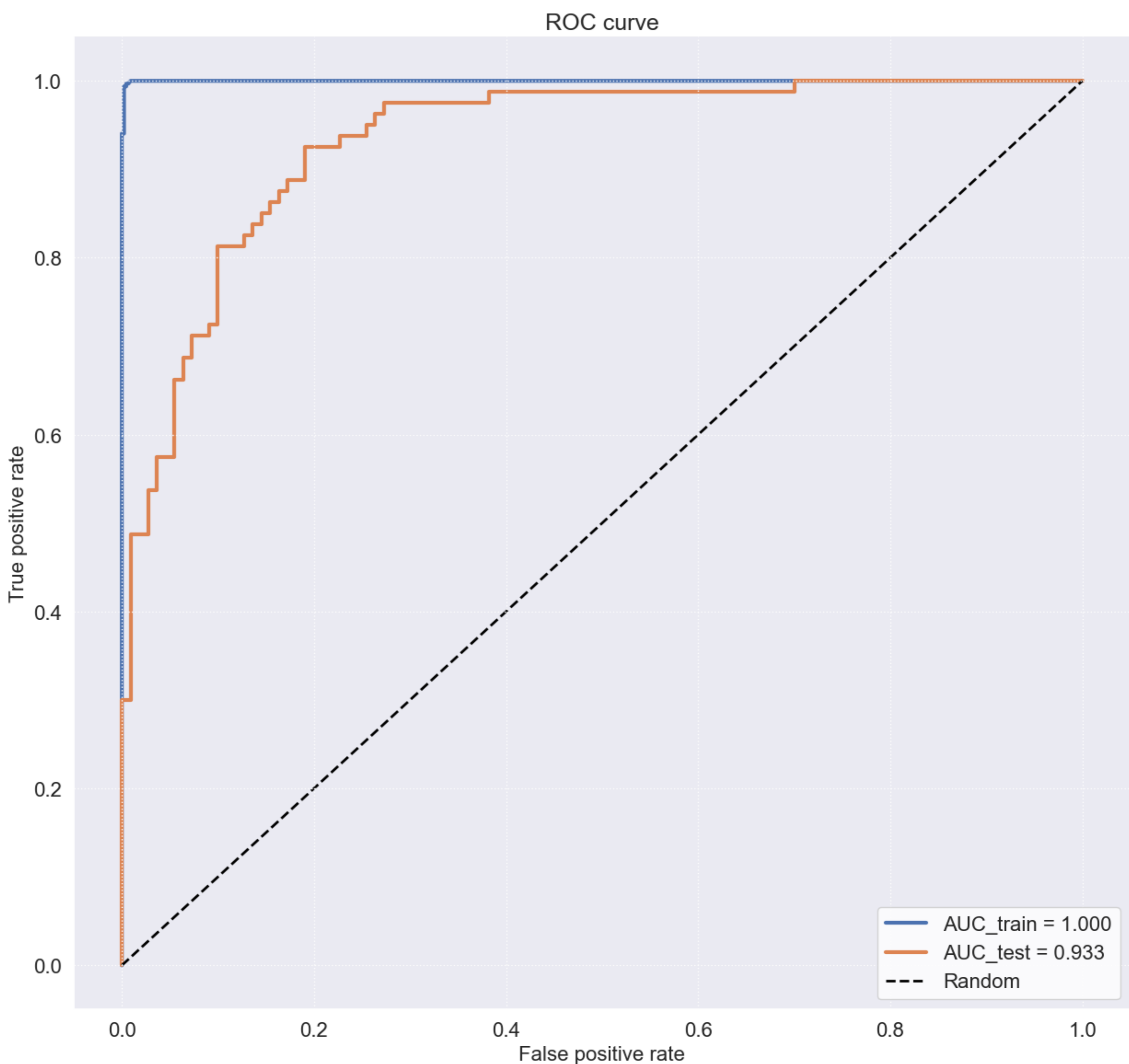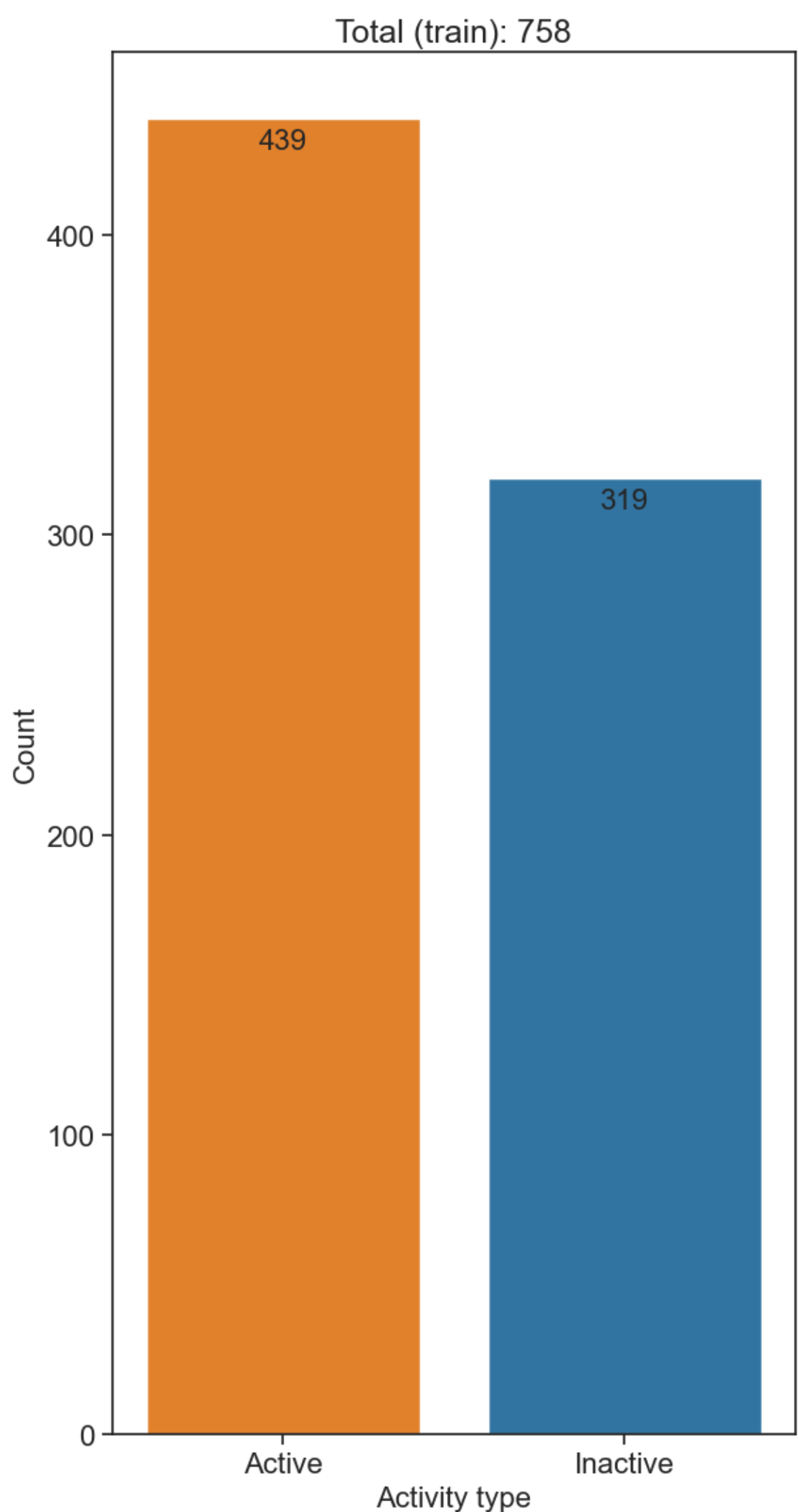
XGBoost Parameters	Value
alpha	1.1616
colsample_bytree	0.7394
gamma	0.1232
lambda	3.7224
learning_rate	0.1222
max_depth	10.0
min_child_weight	2.0
n_estimators	914.0
reg_alpha	0.3072
reg_lambda	0.7652
scale_pos_weight	1.7553
subsample	0.8586

Confusion Matrix

	0	1
0	94	16
1	13	67

Confusion Matrix (nomr.)

	0	1
0	0.85	0.15
1	0.16	0.84

Calibration curve
- train: 0.0089
- test: 0.1129
- perfectly_calibrated

Active/Inactive probability
- Inactive zone
- Active zone
- type
  - active (80)
  - inactive (110)
  - decoy (189)
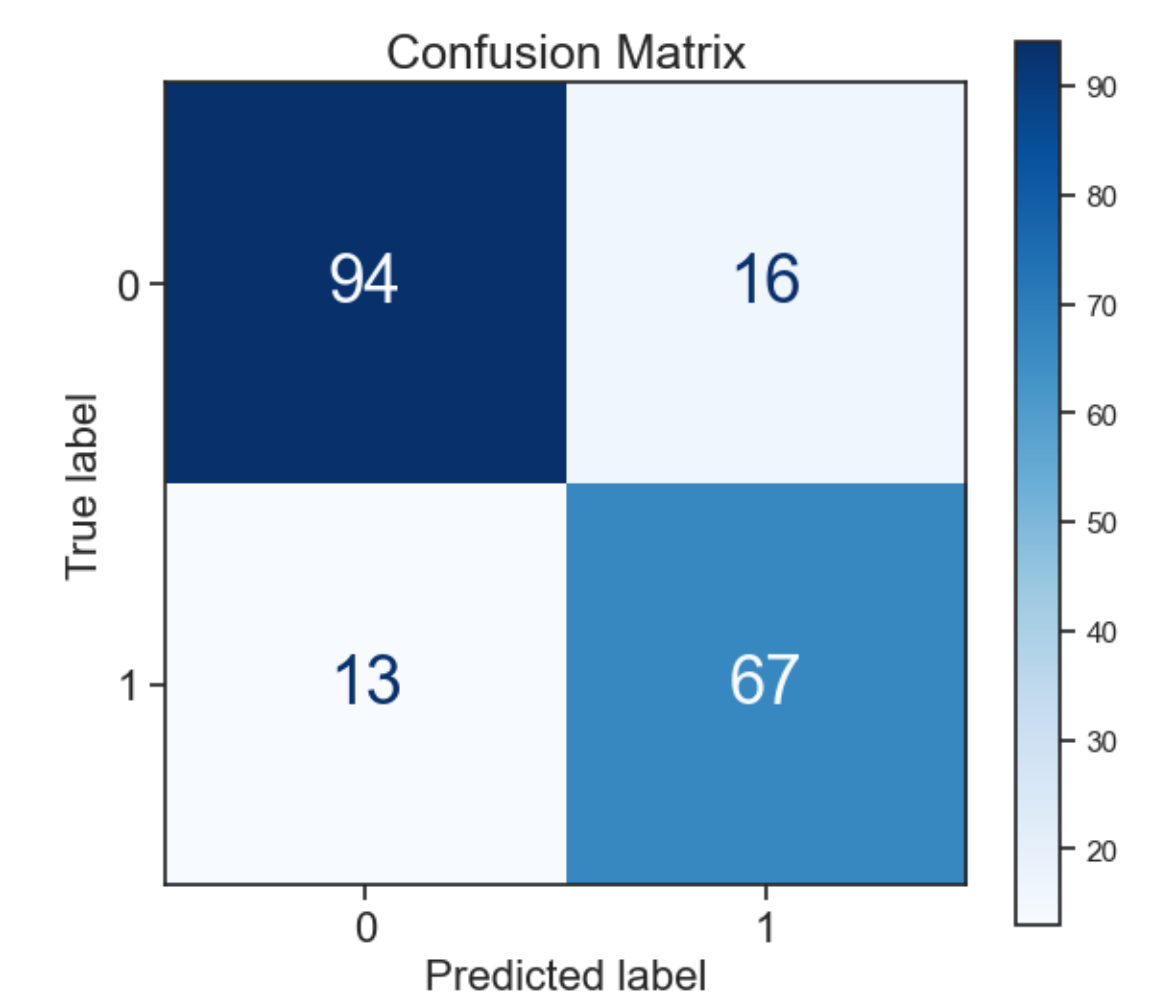
Metric	Train	Test
AUC	0.9998	0.9328
accuracy	0.9921	0.8474
balanced accuracy	0.9932	0.846
recall_(sens)	1.0	0.8375
specificity	0.9863	0.8545
precision	0.9815	0.8072
f1_score	0.9921	0.8477

Protein: P22607 - metric: f1

Total (train): 758

ROC curve

AUC_train = 1.000
AUC_test = 0.920
Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 2074)
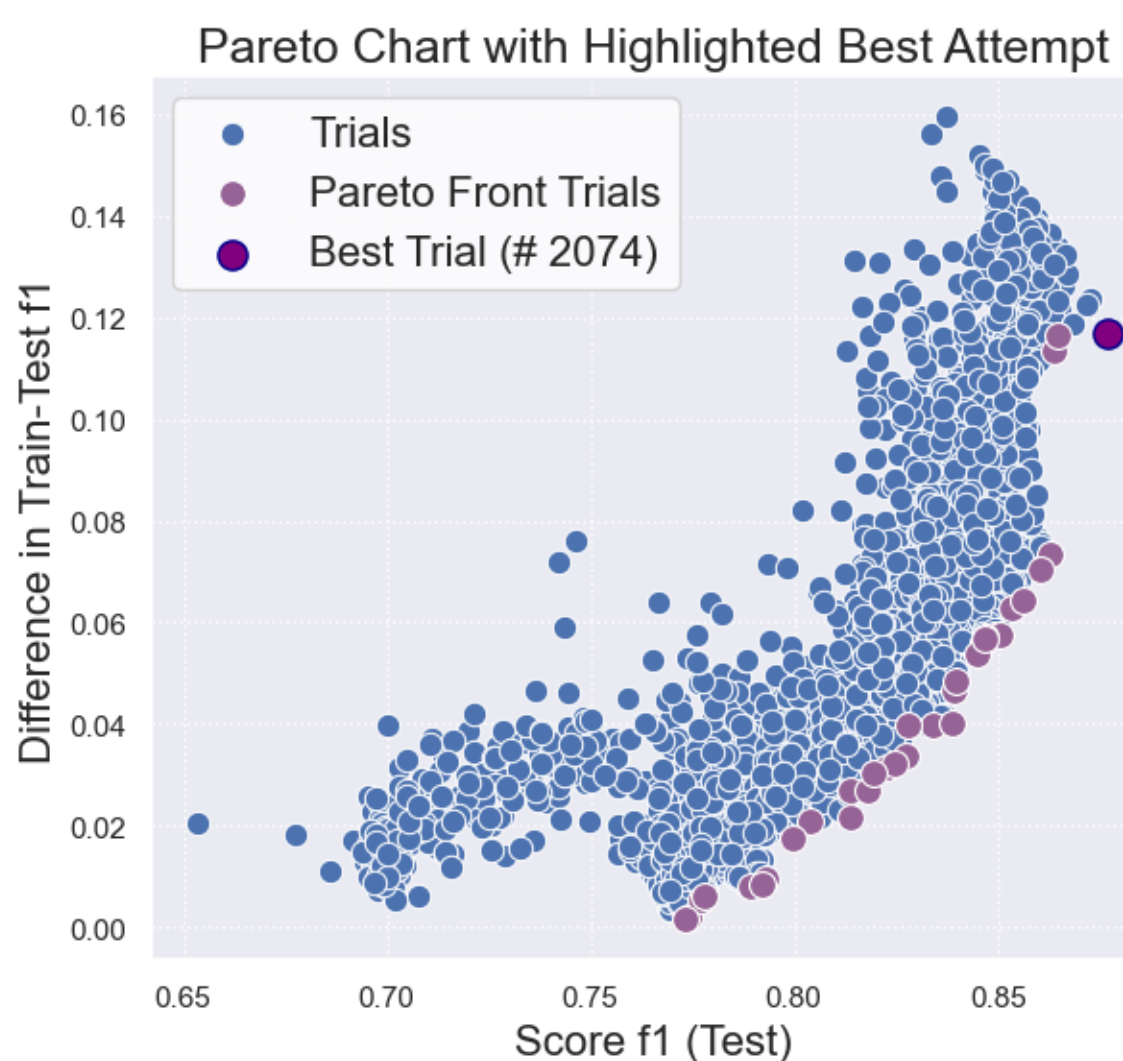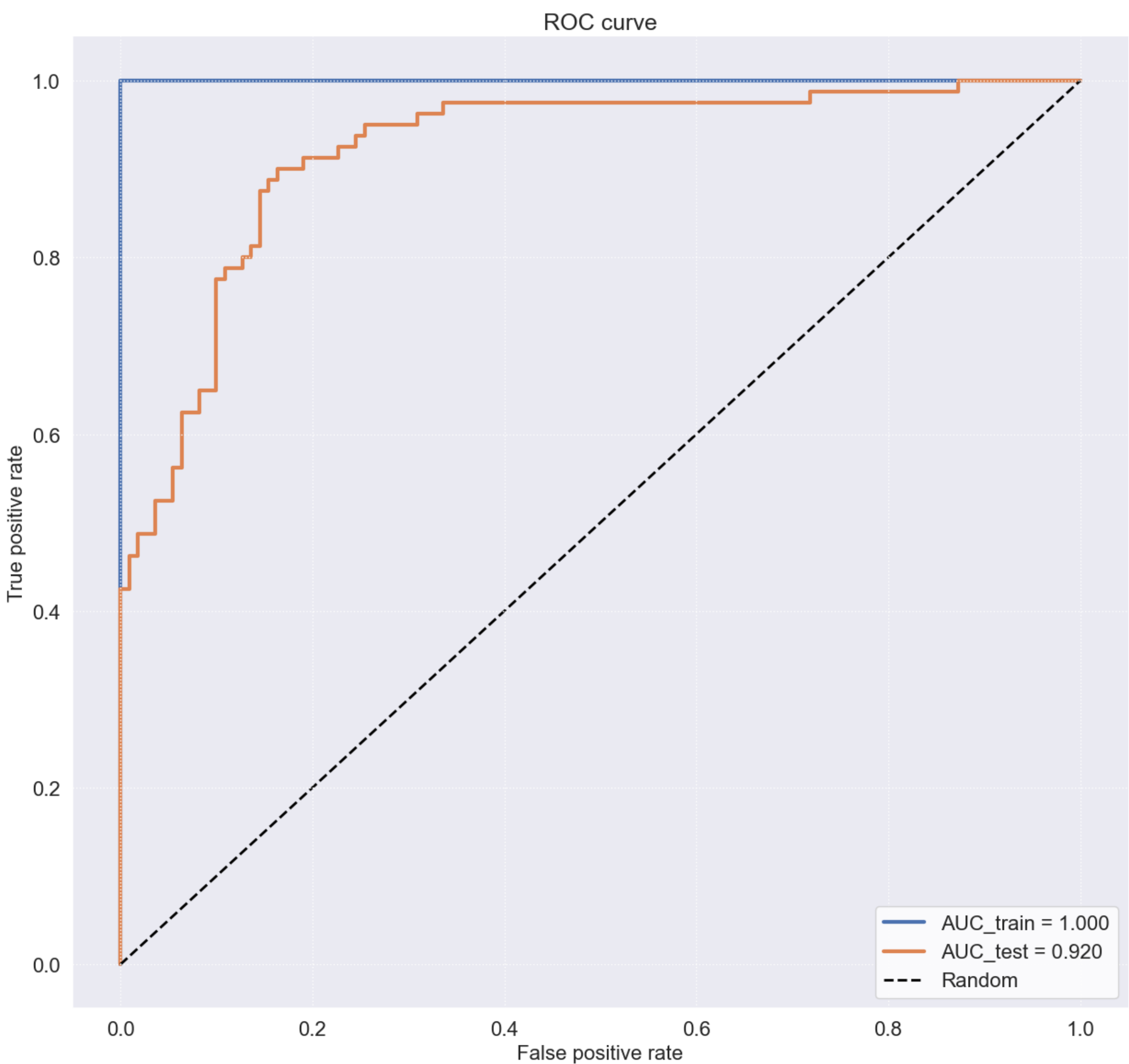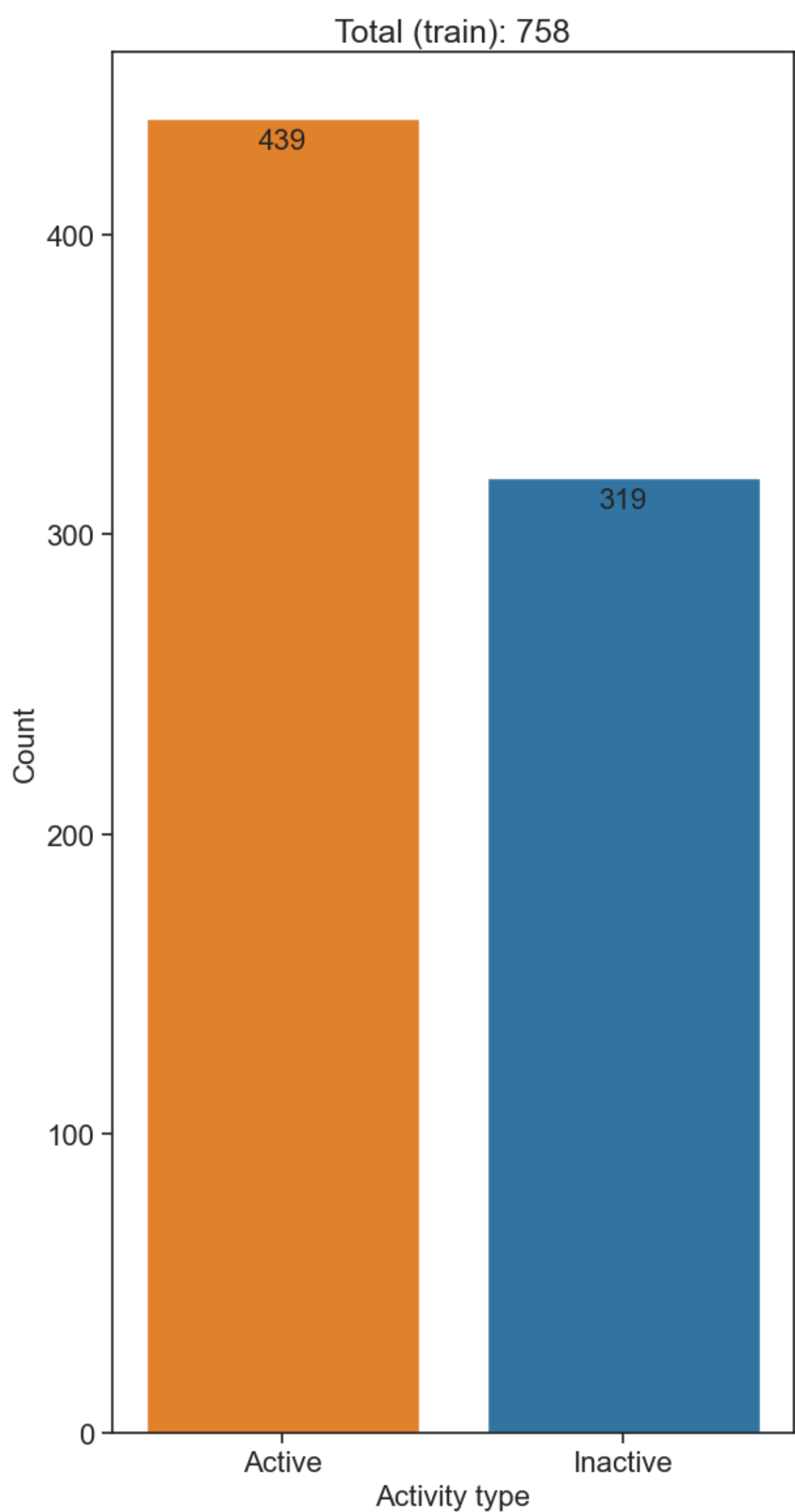
XGBoost Parameters	Value
alpha	2.7046
colsample_bytree	0.7536
gamma	0.4857
lambda	0.6292
learning_rate	0.6855
max_depth	8.0
min_child_weight	1.0
n_estimators	381.0
reg_alpha	0.3886
reg_lambda	0.756
scale_pos_weight	5.1574
subsample	0.7138

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0066
test: 0.1225
perfectly_calibrated

Active/Inactive probability

type
- active (80)
- inactive (110)
- decoy (189)

Inactive zone

Active zone
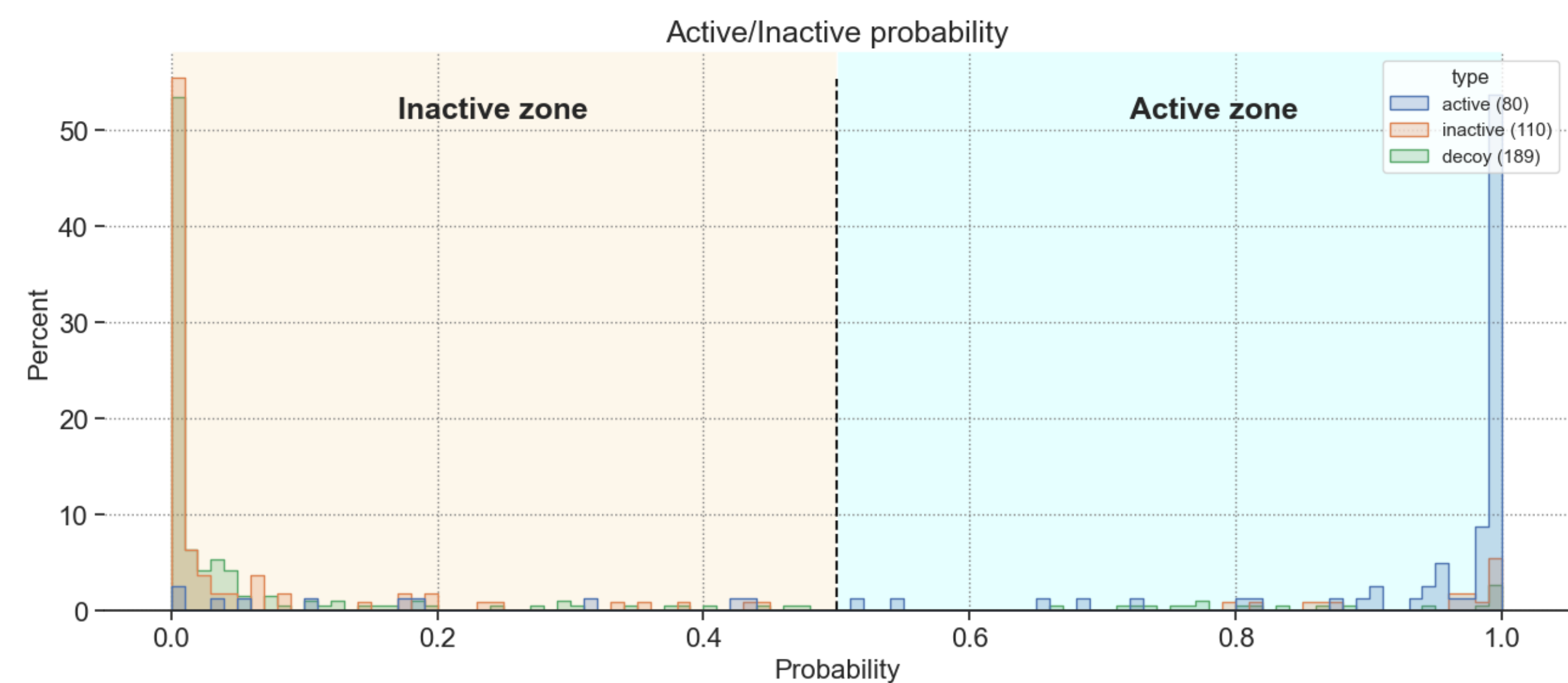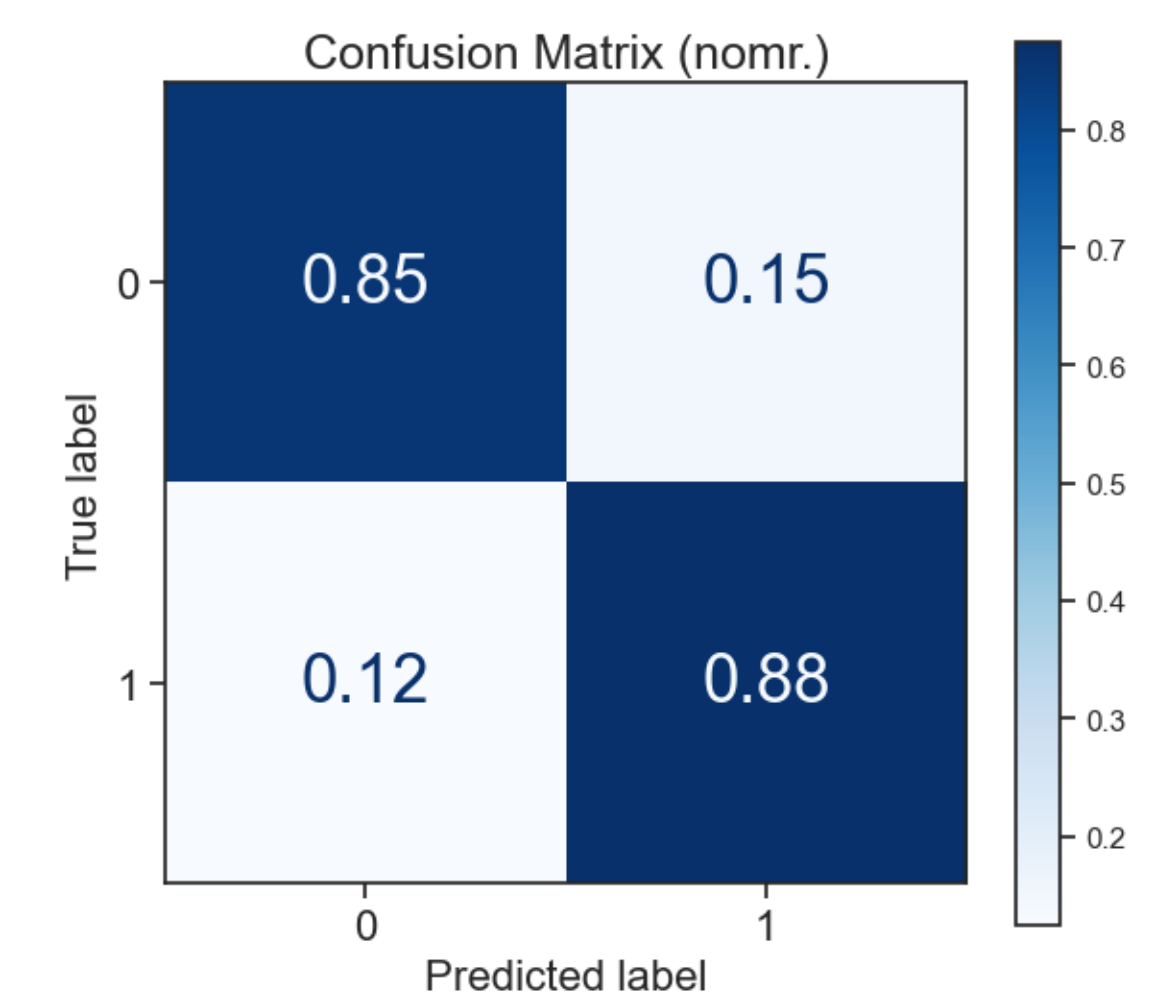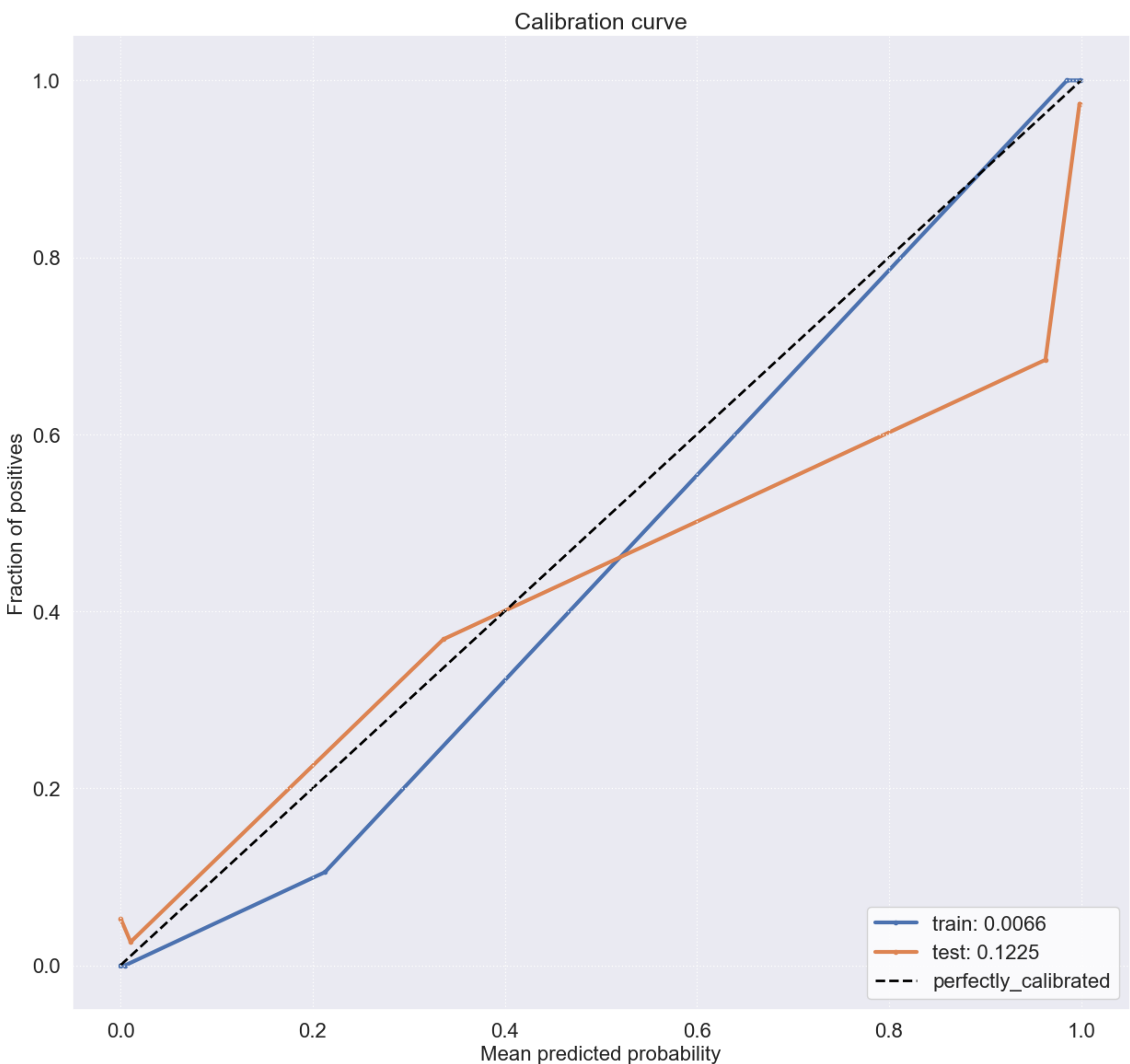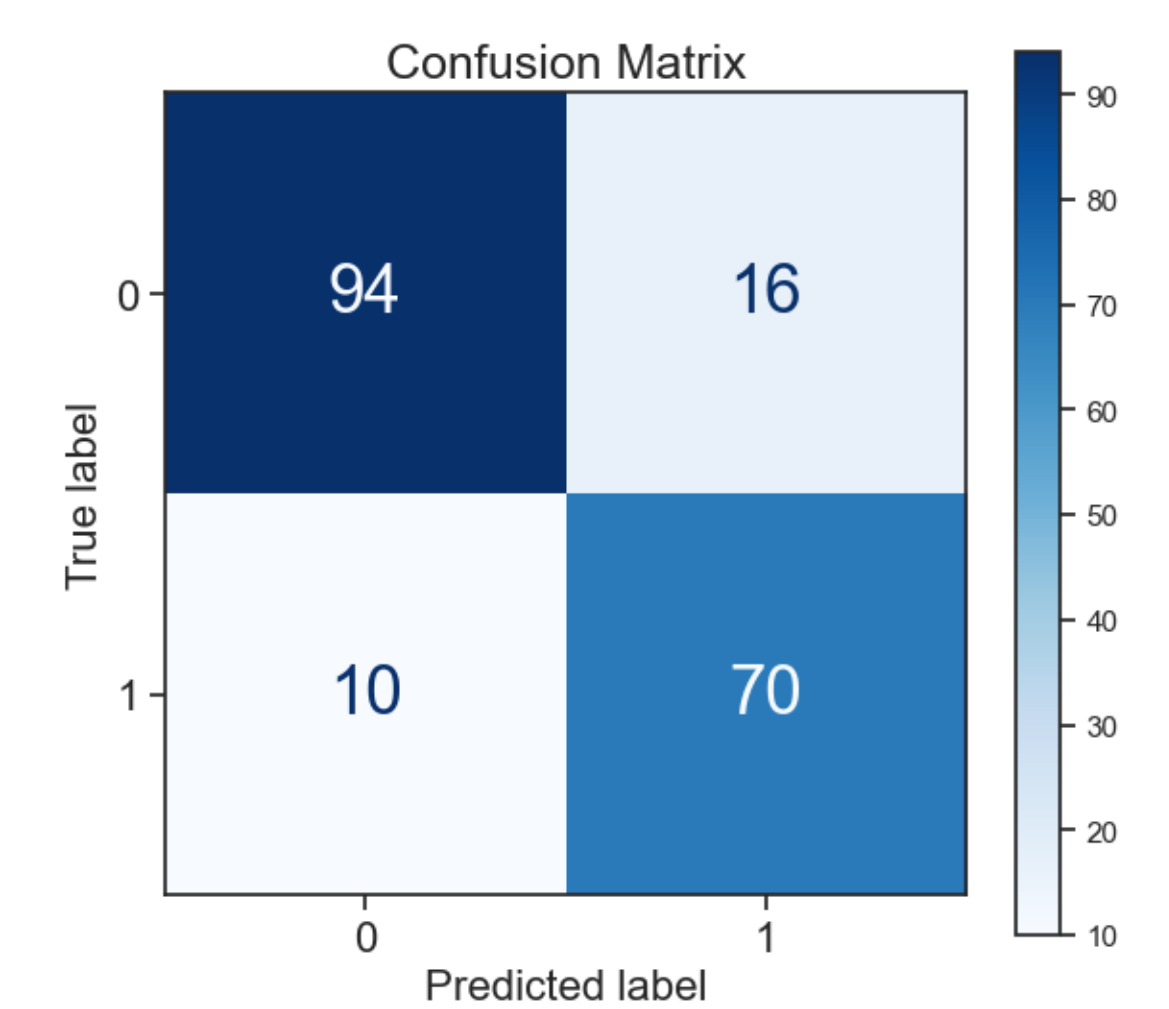
Metric	Train	Test
AUC	1.0	0.9201
accuracy	0.9974	0.8632
balanced accuracy	0.9977	0.8648
recall_(sens)	1.0	0.875
specificity	0.9954	0.8545
precision	0.9938	0.814
f1_score	0.9974	0.8637

Protein: P22607 - metric: precision

Total (train): 758

Activity type: Active 439, Inactive 319

ROC curve

AUC_train = 0.997
AUC_test = 0.912
Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 2701)

Score precision (Test) / Difference in Train-Test precision

Confusion Matrix

	Predicted 0	Predicted 1
True 0	96	14
True 1	14	66

Confusion Matrix (nomr.)

	Predicted 0	Predicted 1
True 0	0.87	0.13
True 1	0.17	0.82

Calibration curve

train: 0.0241
test: 0.1172
perfectly_calibrated

Active/Inactive probability

Inactive zone / Active zone

type:
- active (80)
- inactive (110)
- decoy (189)

XGBoost Parameters	Value
alpha	4.7885
colsample_bytree	0.5162
gamma	0.2684
lambda	4.2322
learning_rate	0.4597
max_depth	3.0
min_child_weight	2.0
n_estimators	937.0
reg_alpha	0.6602
reg_lambda	0.0119
scale_pos_weight	1.017
subsample	0.9137
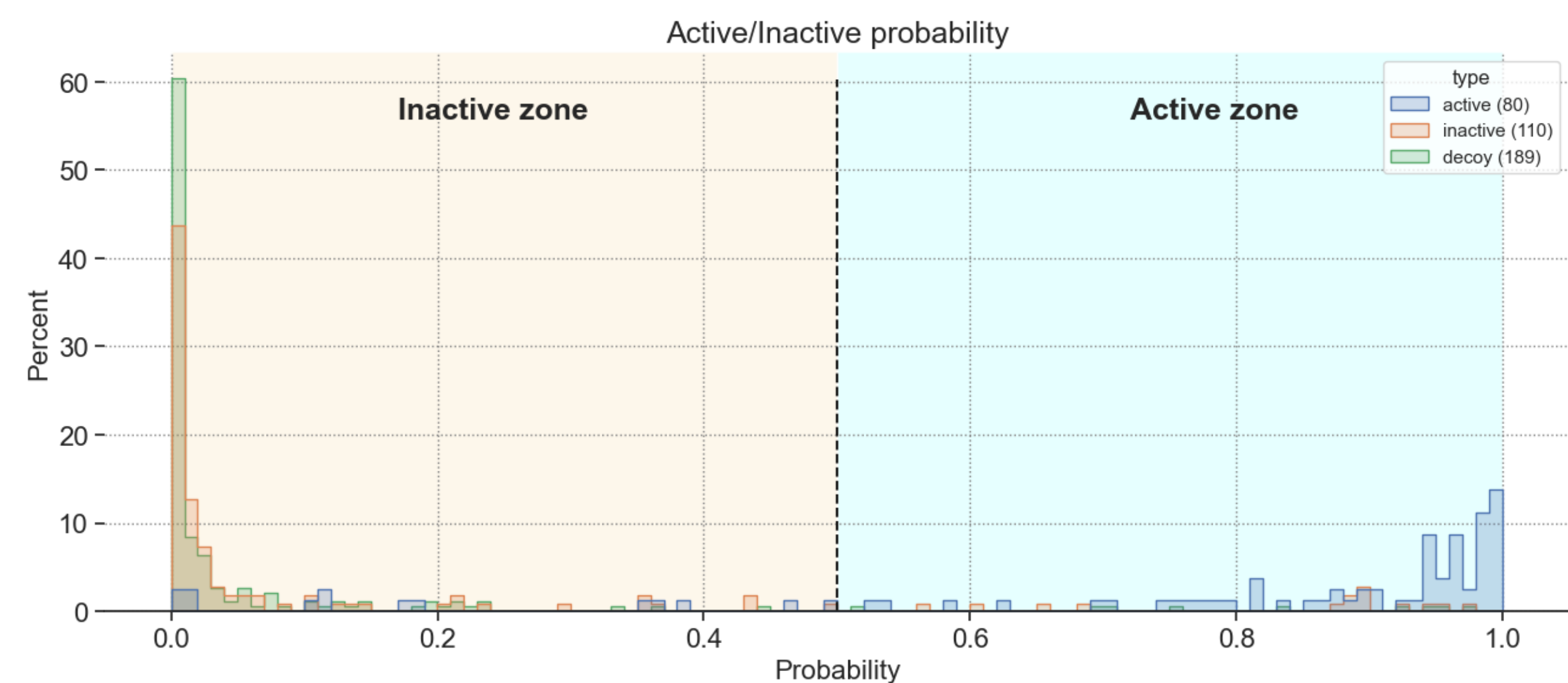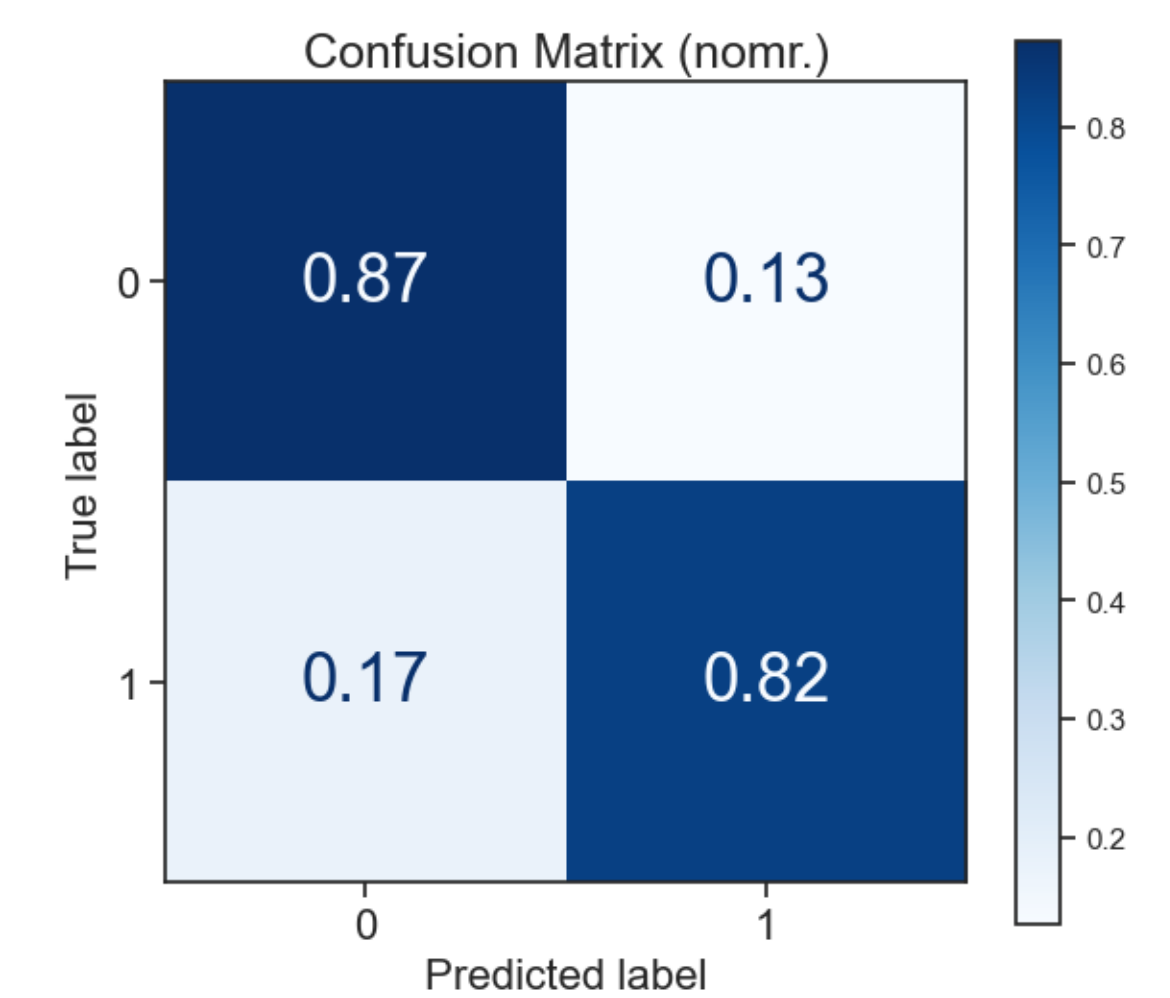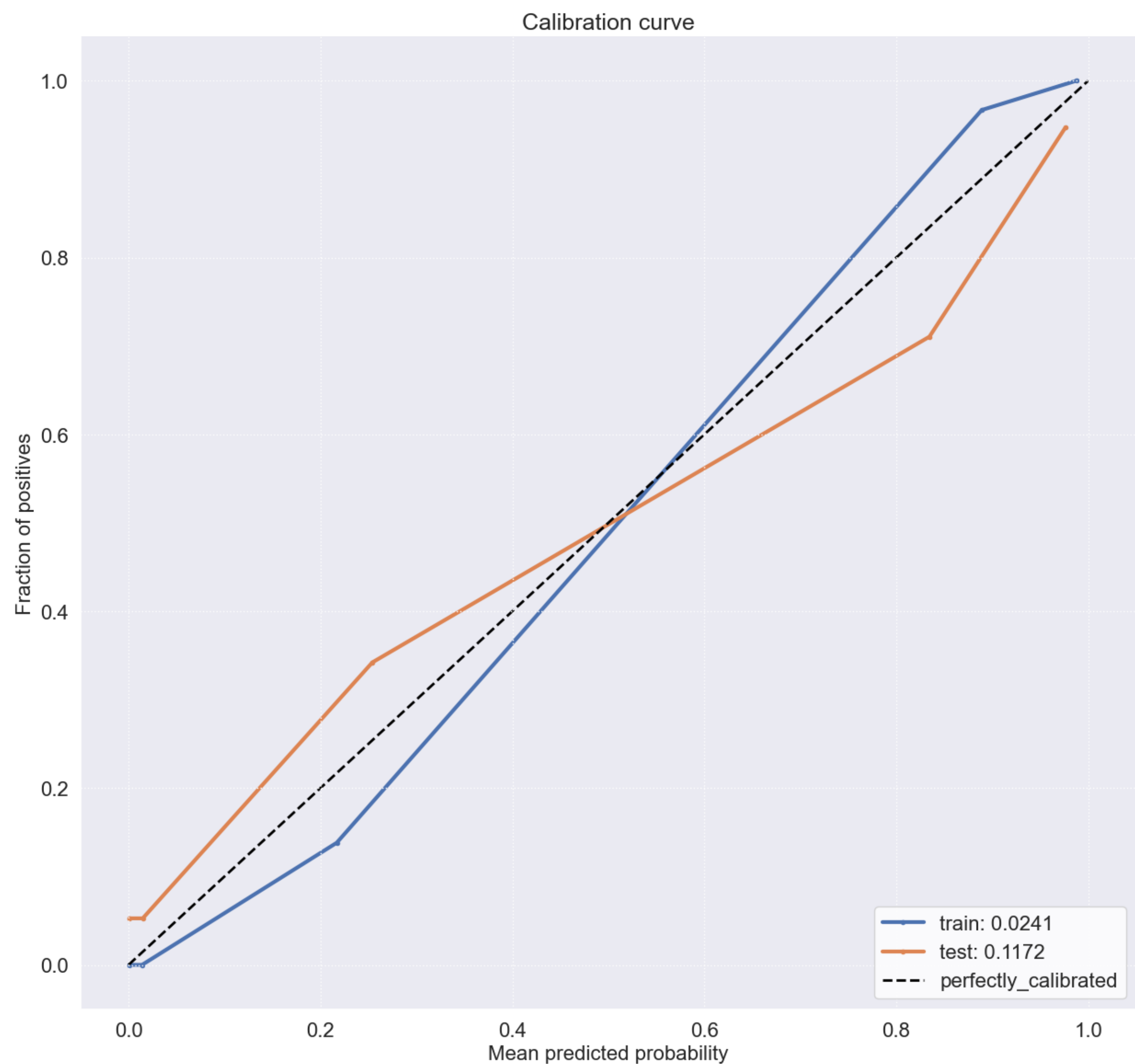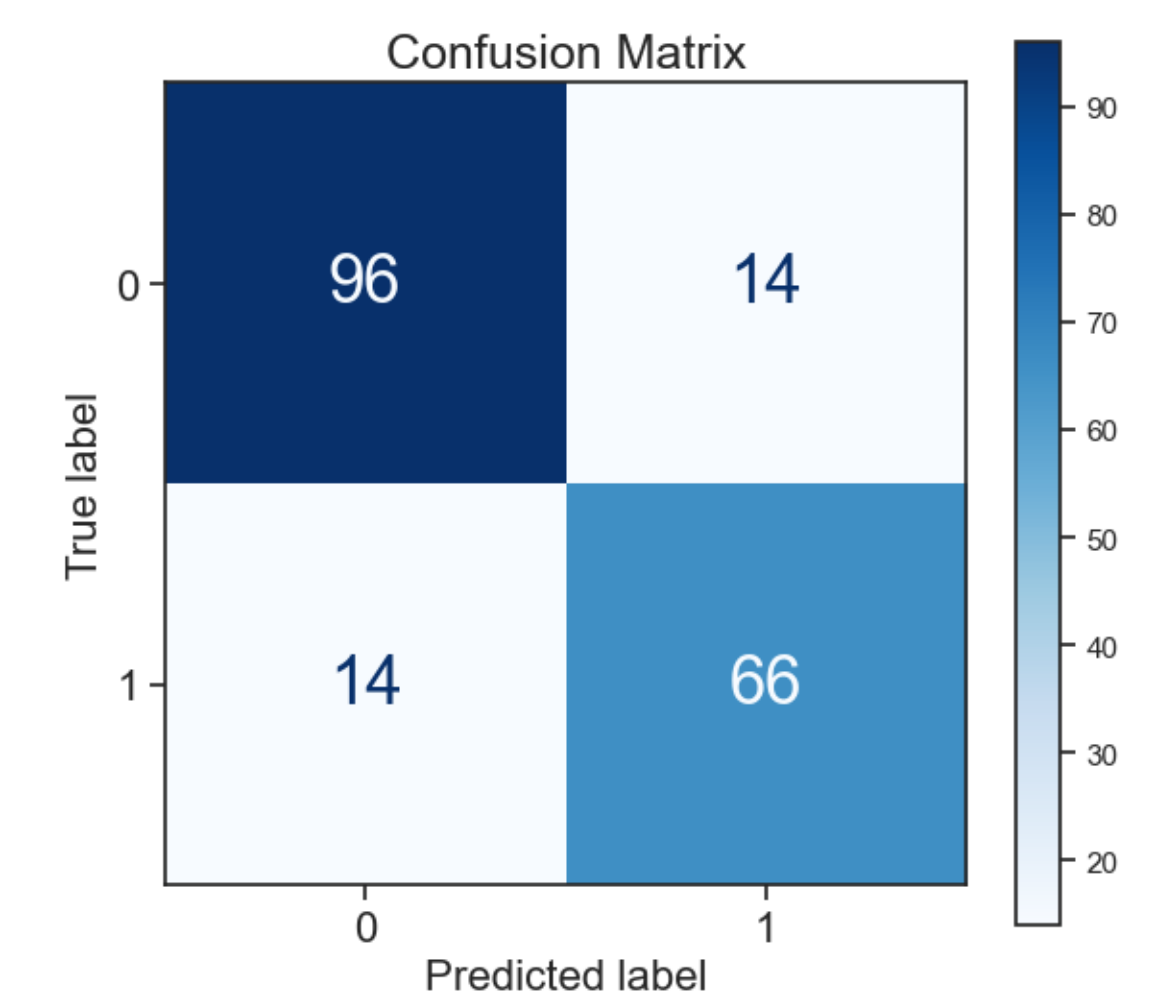
Metric	Train	Test
AUC	0.9968	0.9122
accuracy	0.9749	0.8526
balanced accuracy	0.9741	0.8489
recall_(sens)	0.9687	0.825
specificity	0.9795	0.8727
precision	0.9717	0.825
f1_score	0.9749	0.8526

Protein: P22607 - metric: SoftVotingXGBoostEnsemble
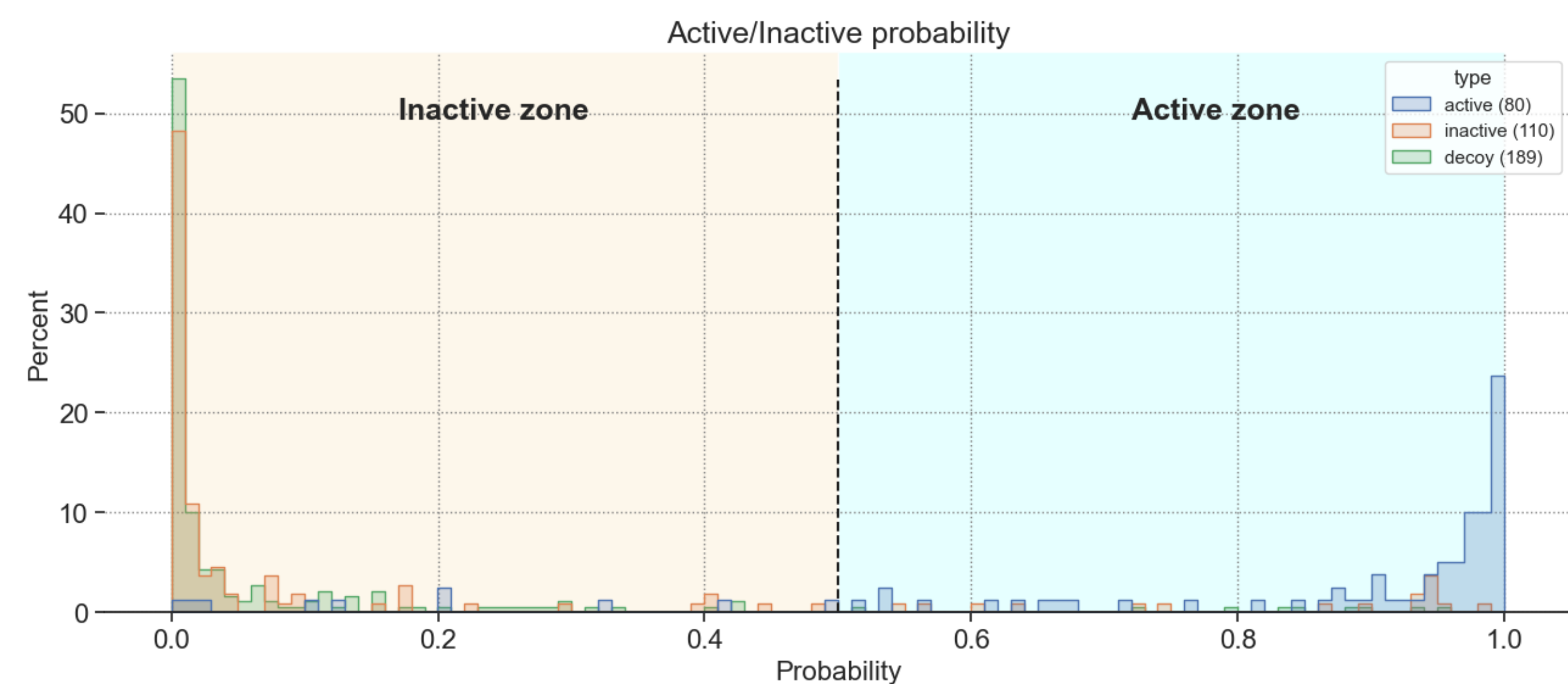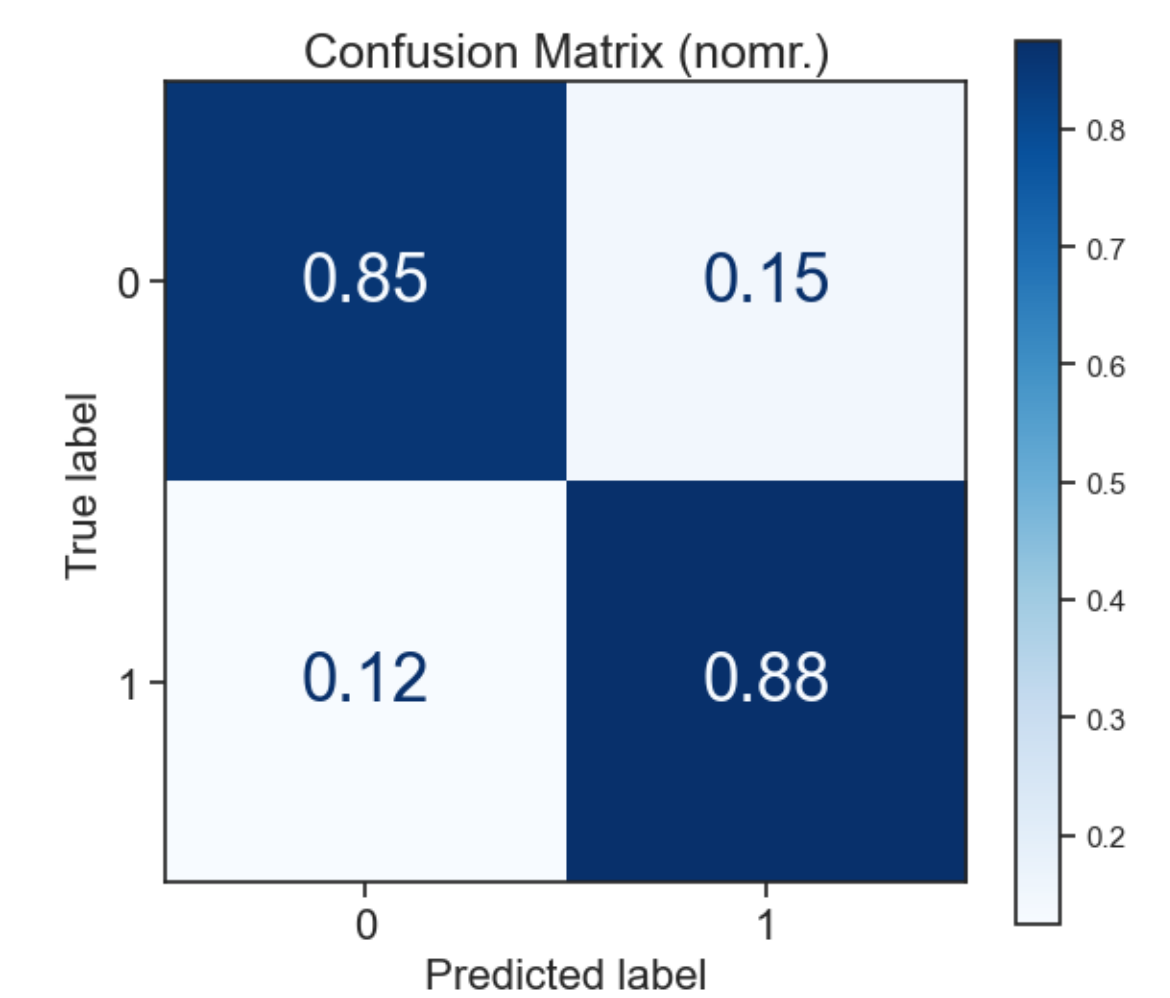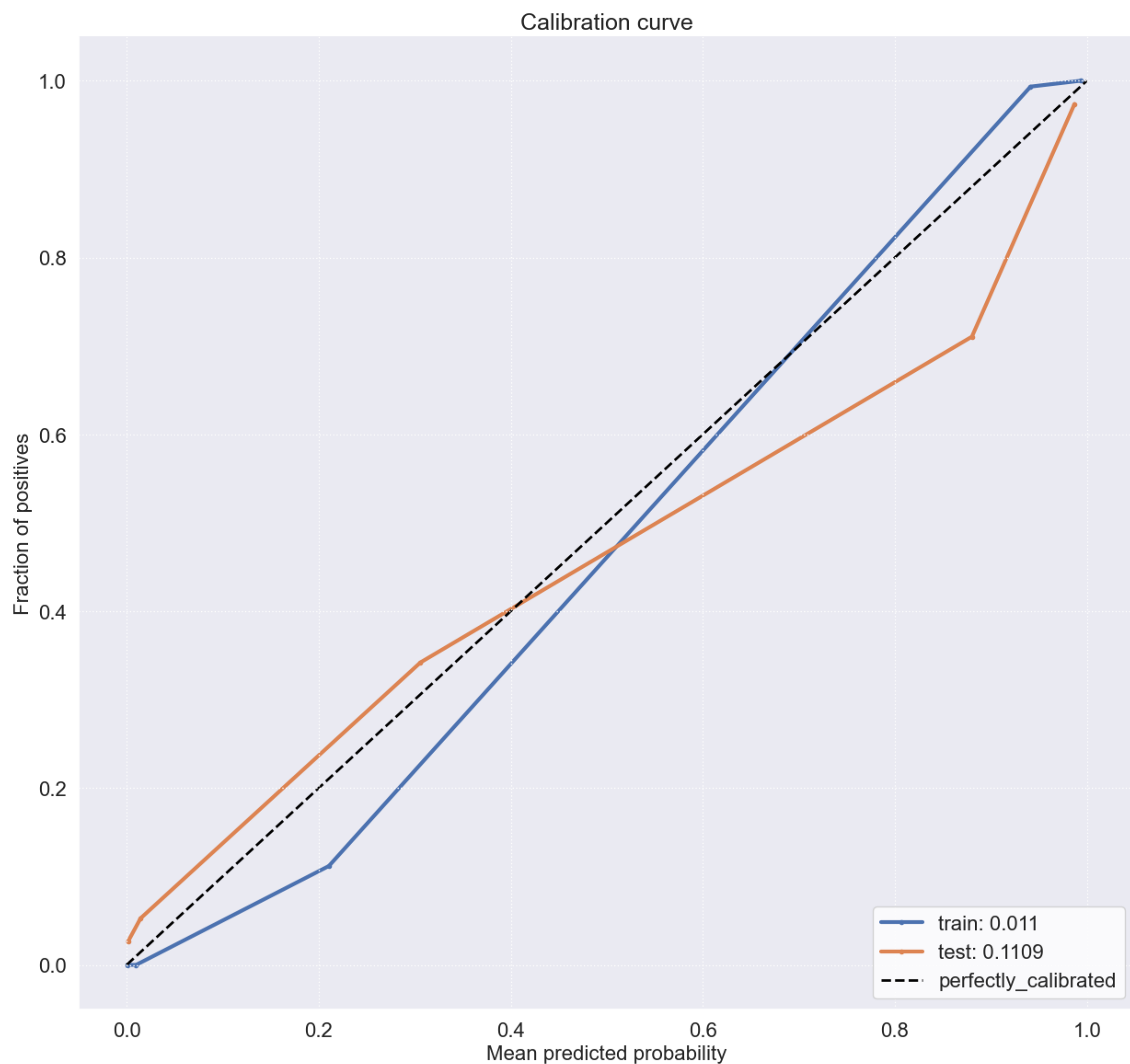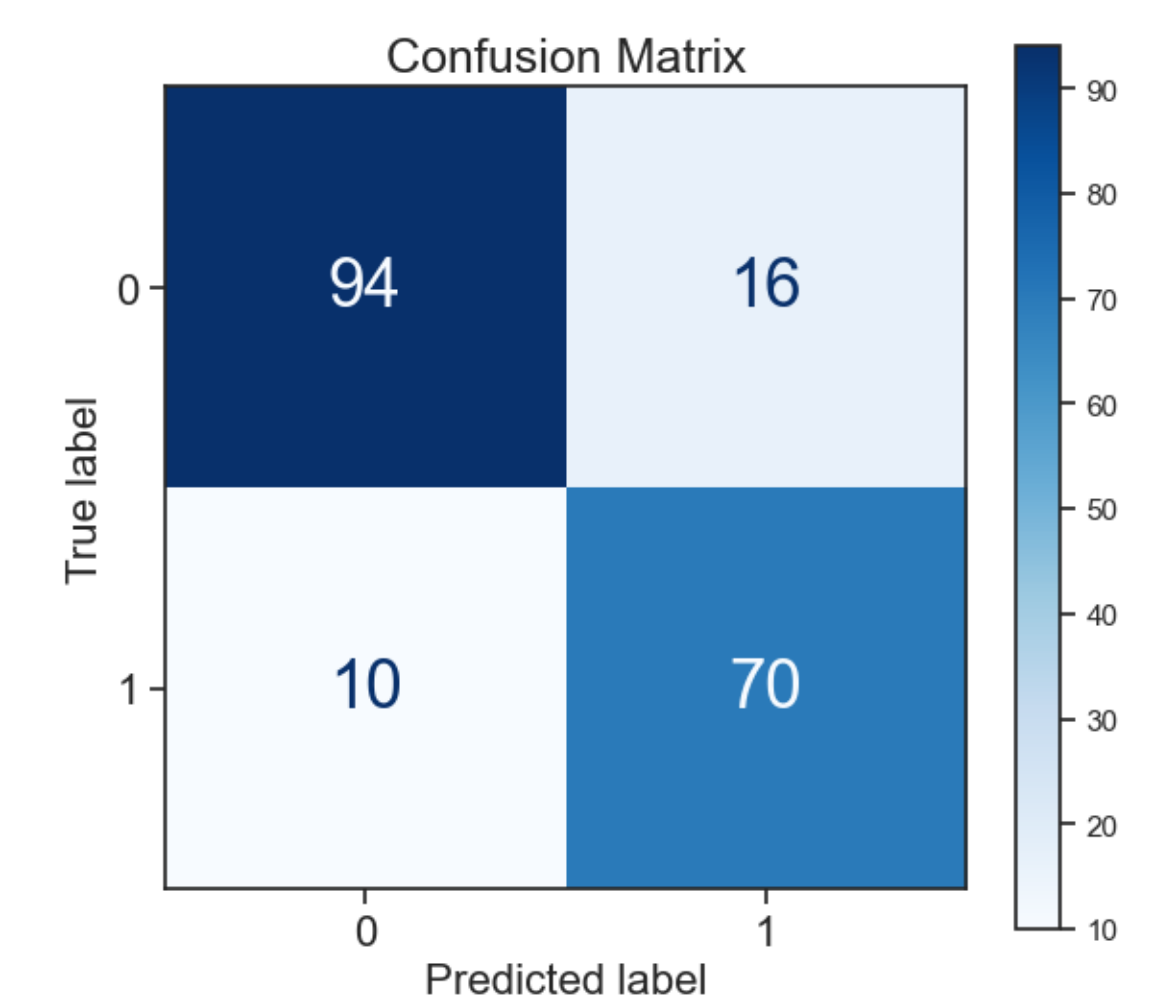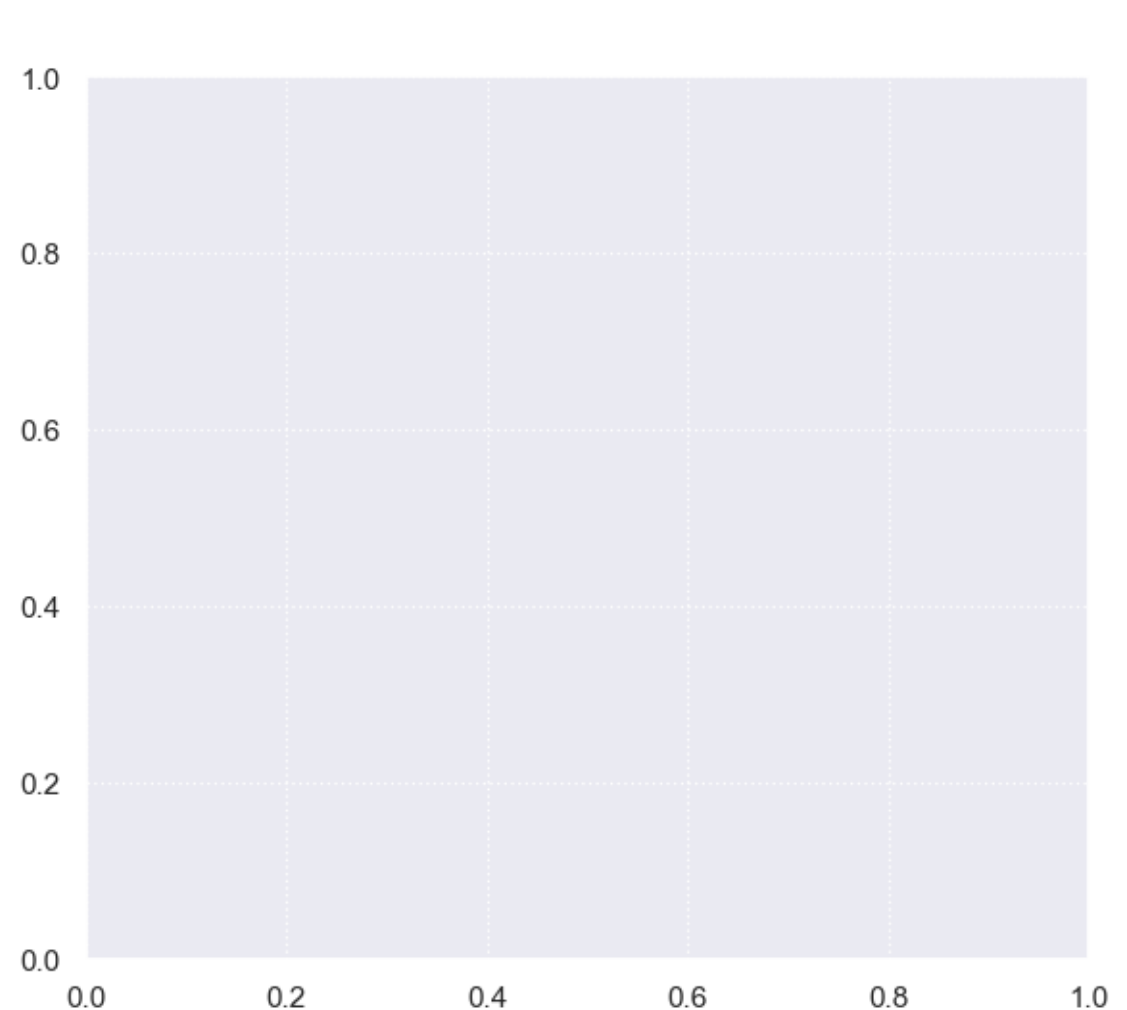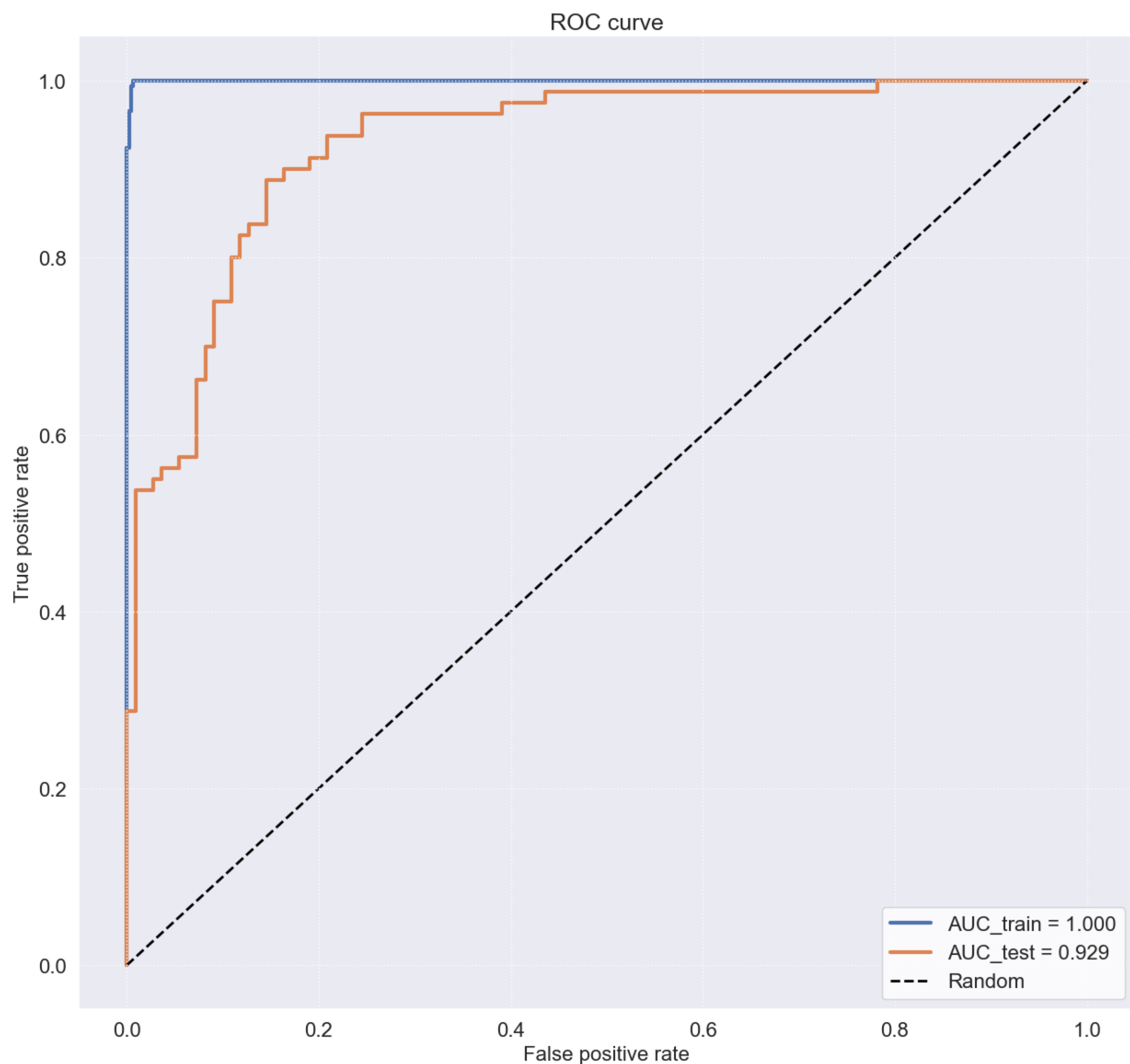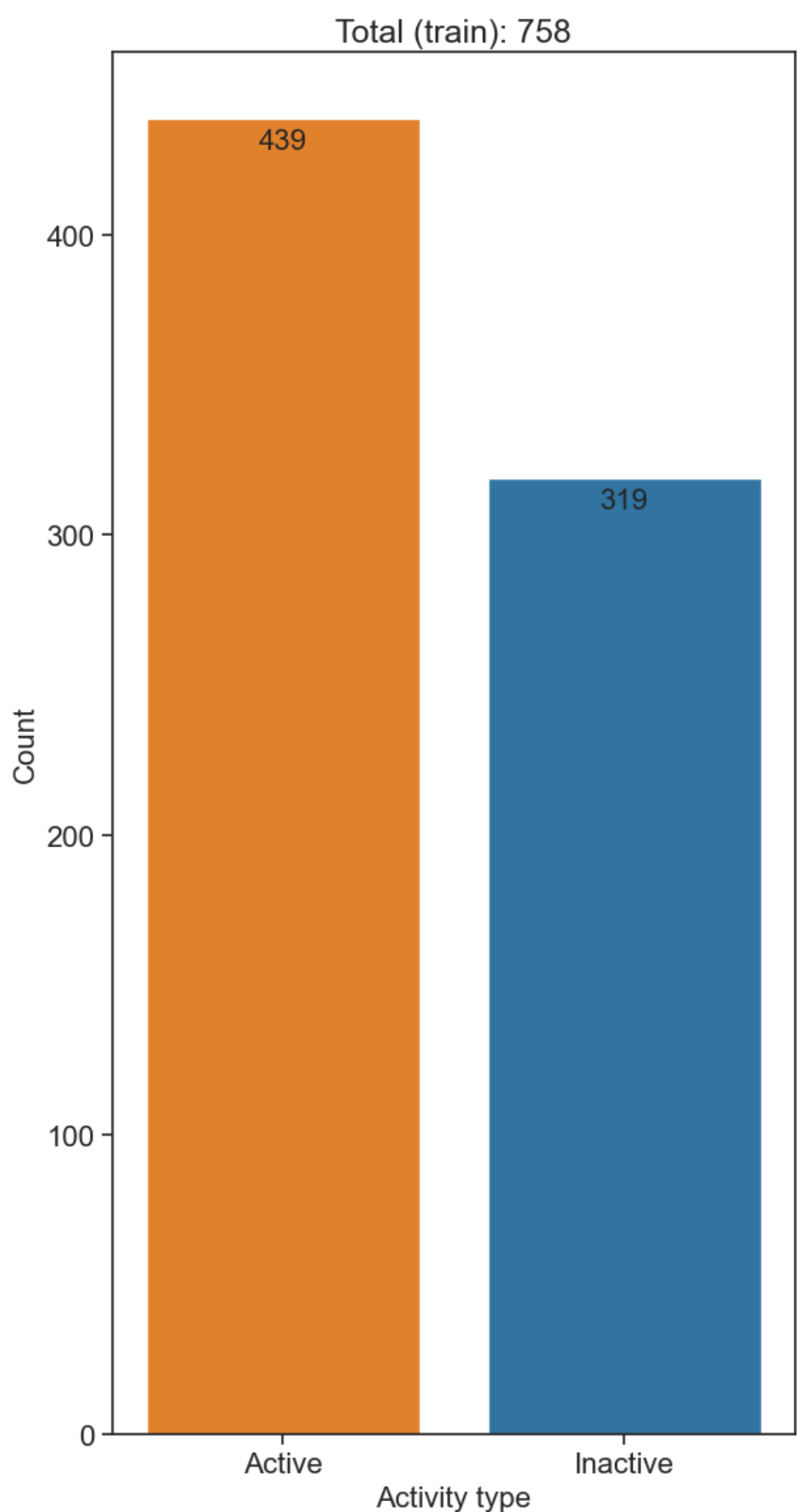
Metric	Train	Test
AUC	0.9997	0.9291
accuracy	0.9947	0.8632
balanced accuracy	0.9954	0.8648
recall_(sens)	1.0	0.875
specificity	0.9909	0.8545
precision	0.9876	0.814
f1_score	0.9947	0.8637