Protein: P21917 - metric: balanced_accuracy

Total (train): 1423

ROC curve
- AUC_train = 0.996
- AUC_test = 0.929
- Random

Pareto Chart with Highlighted Best Attempt
- Trials
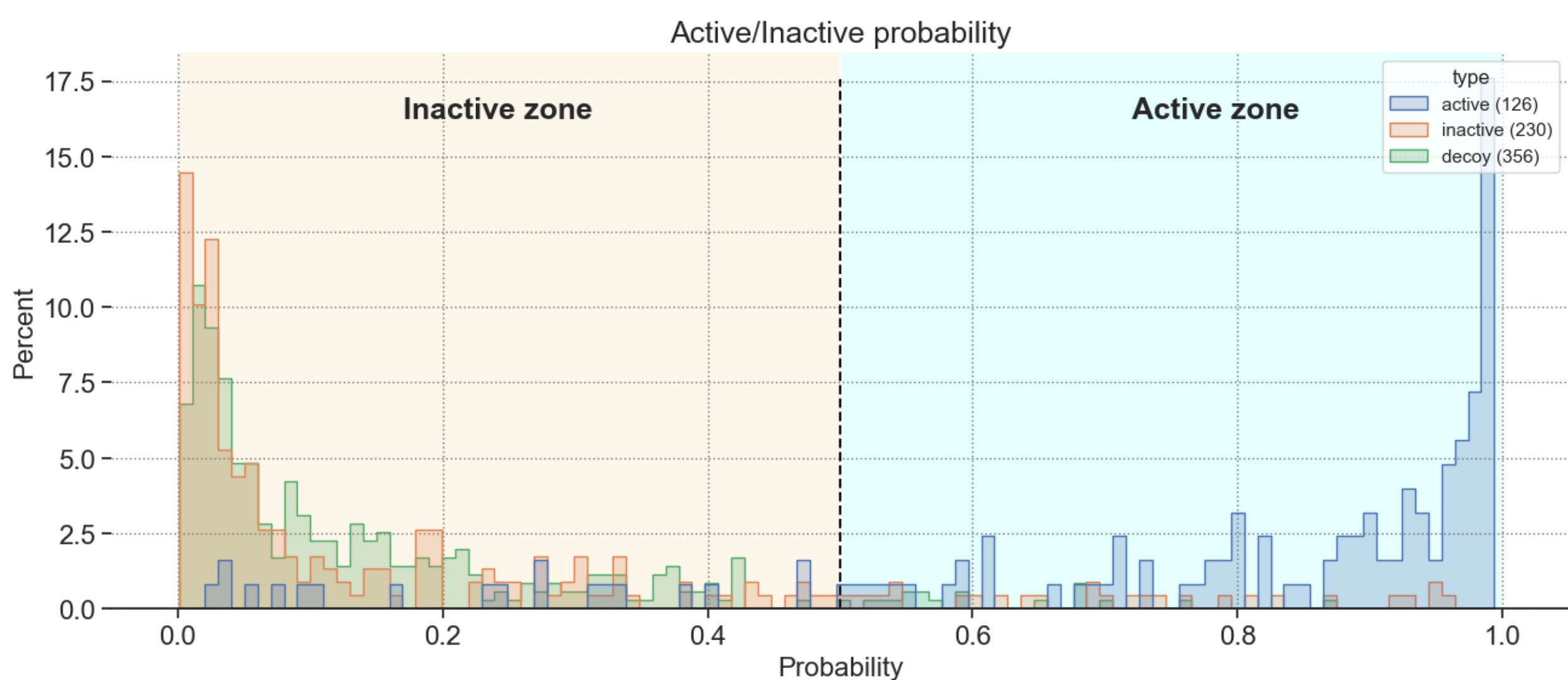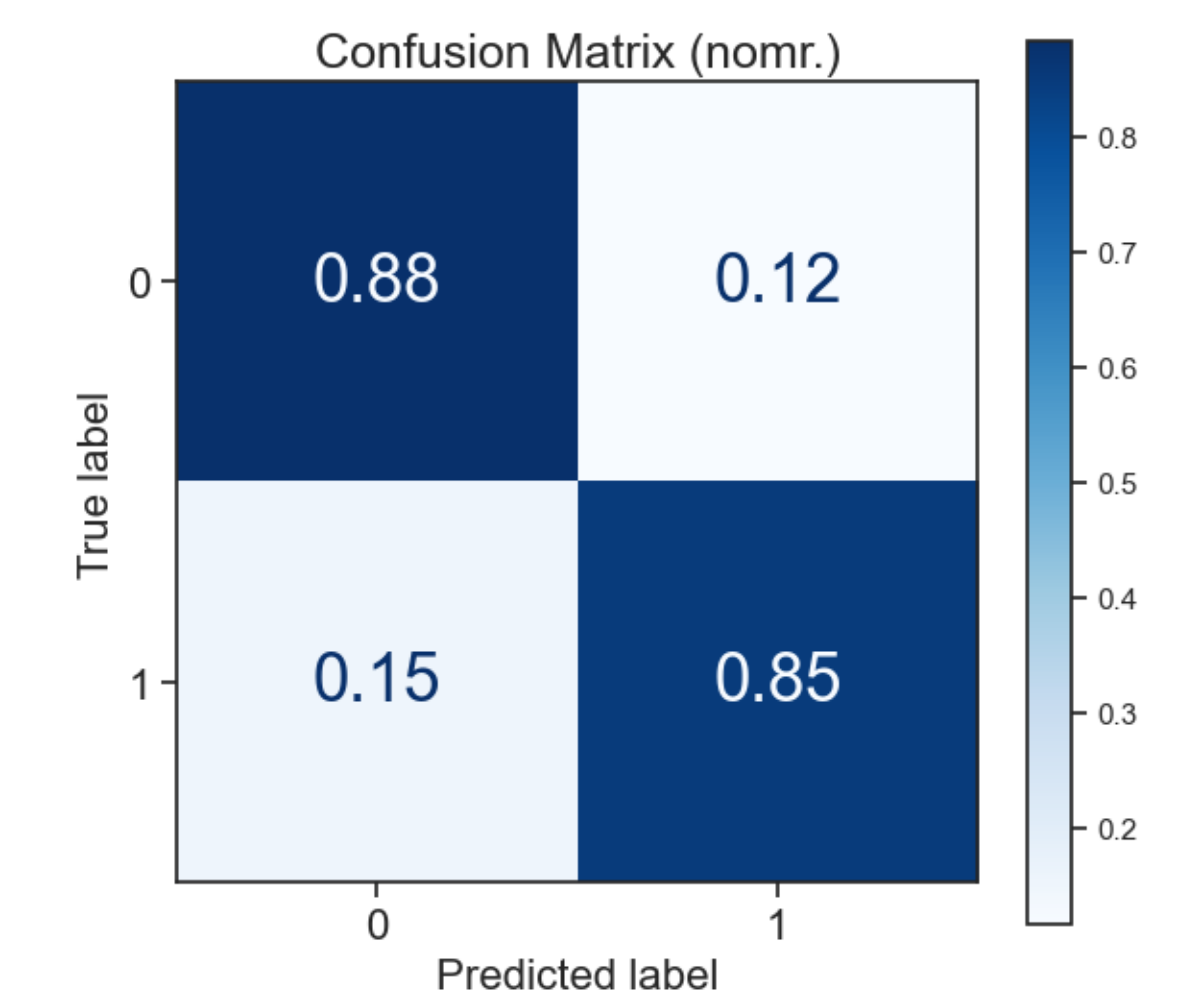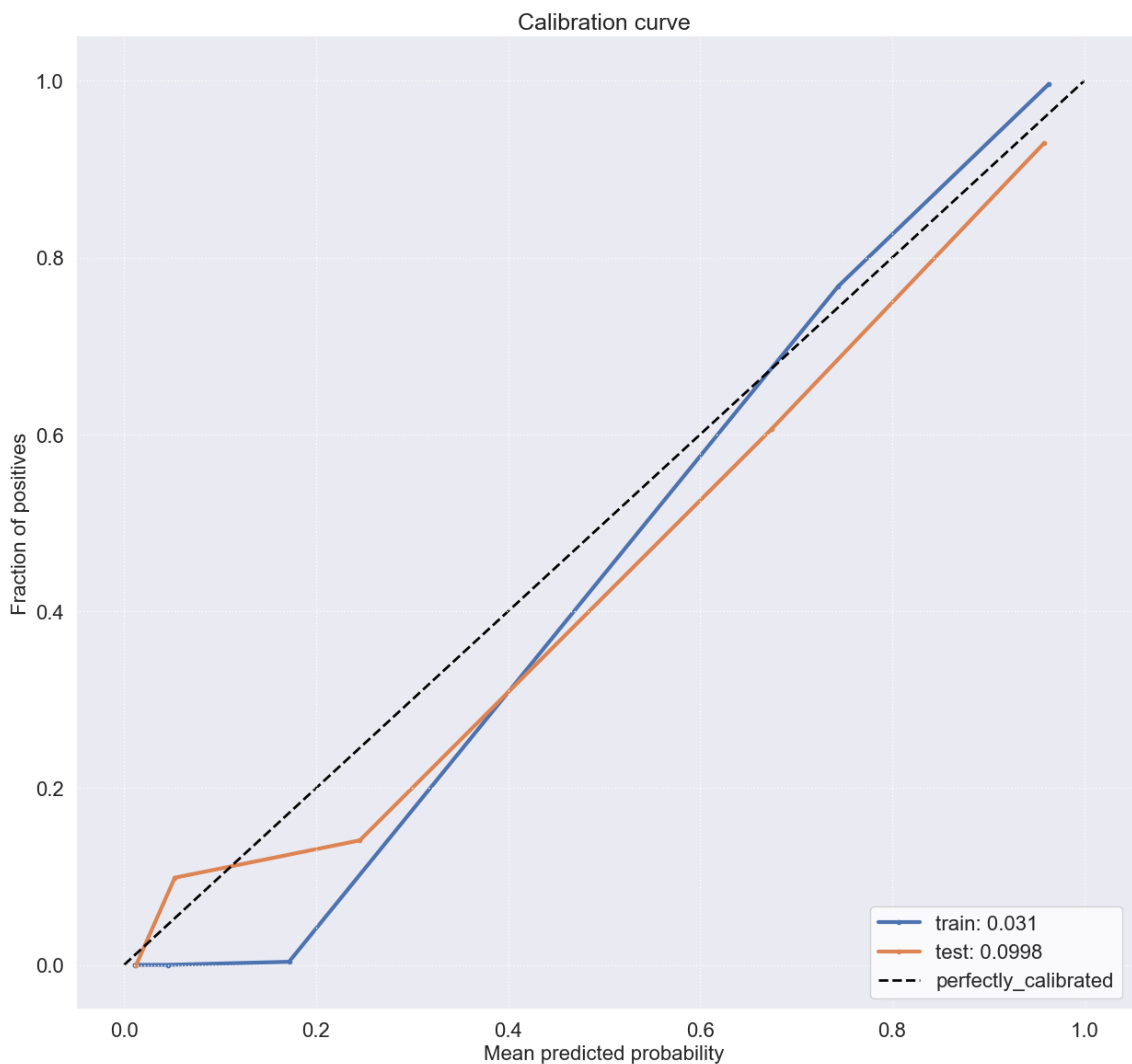- Pareto Front Trials
- Best Trial (# 921)
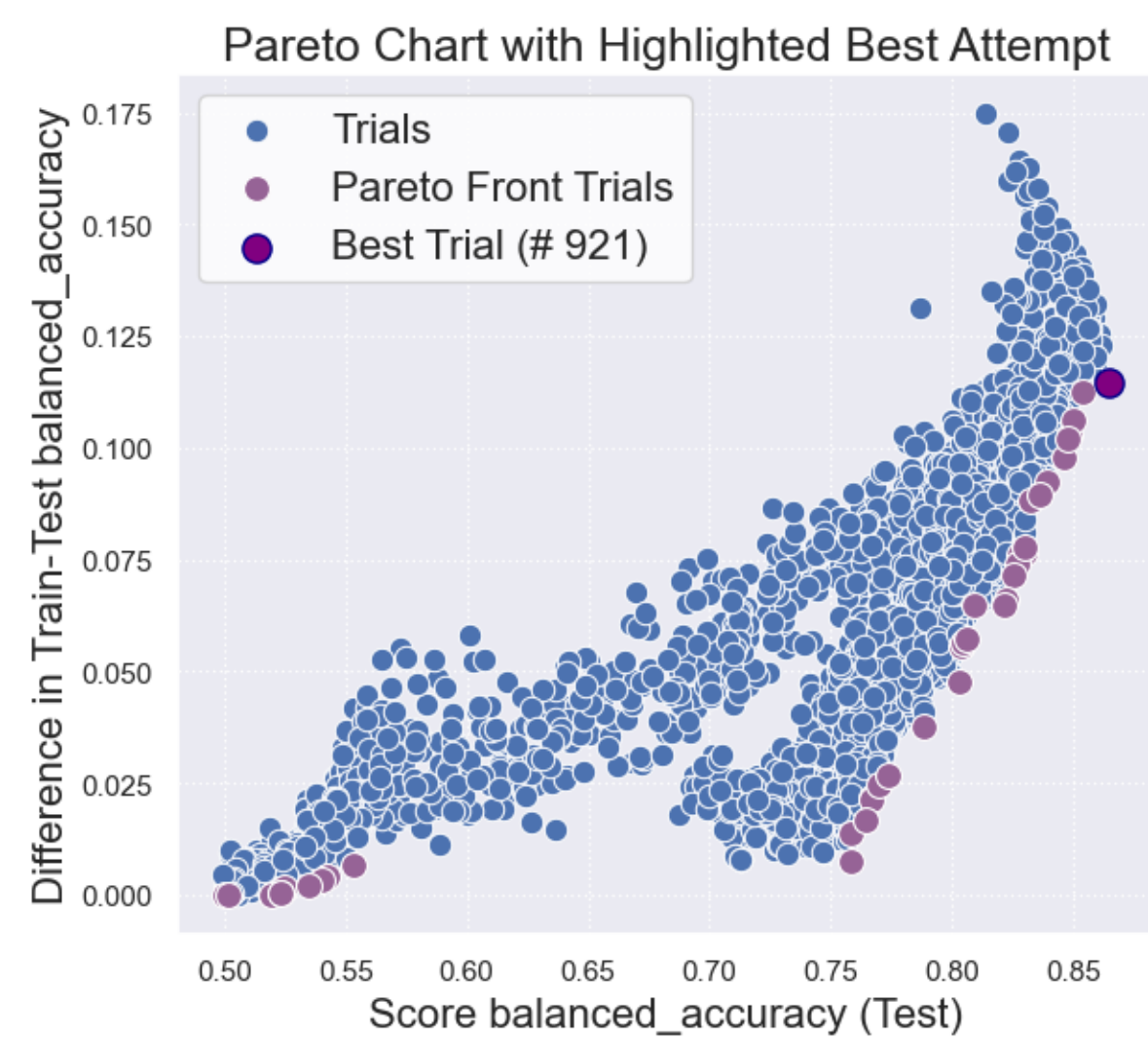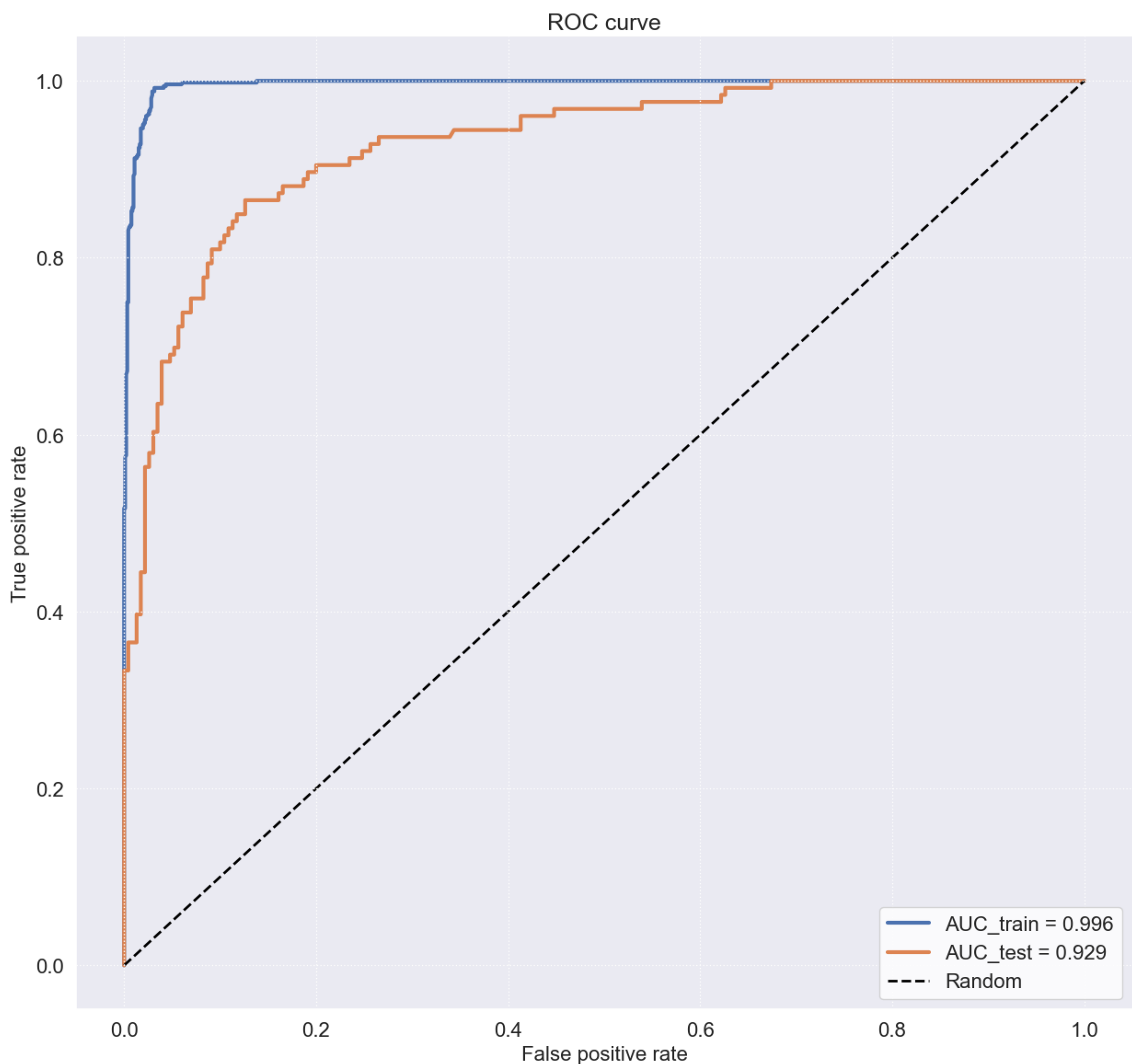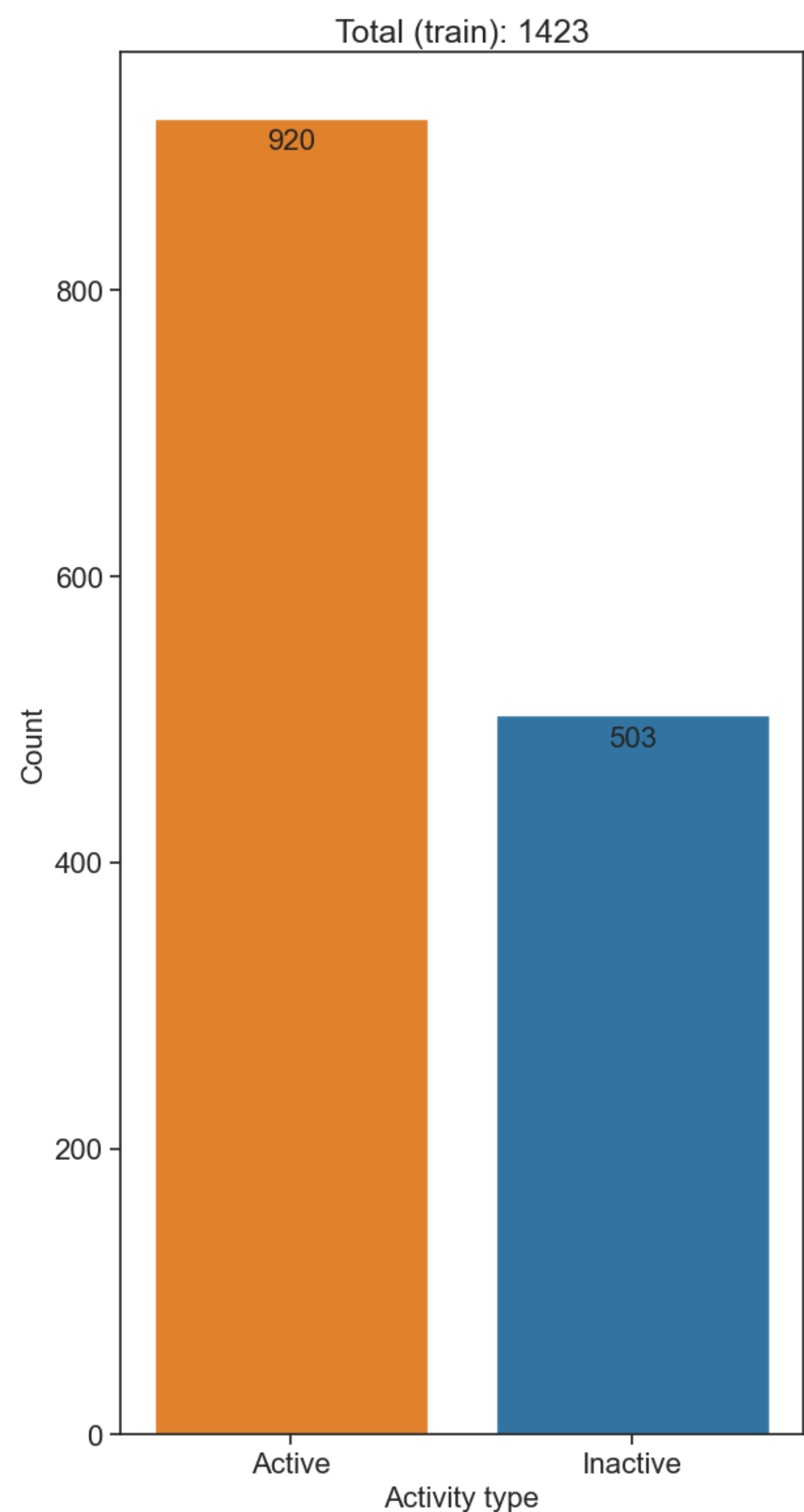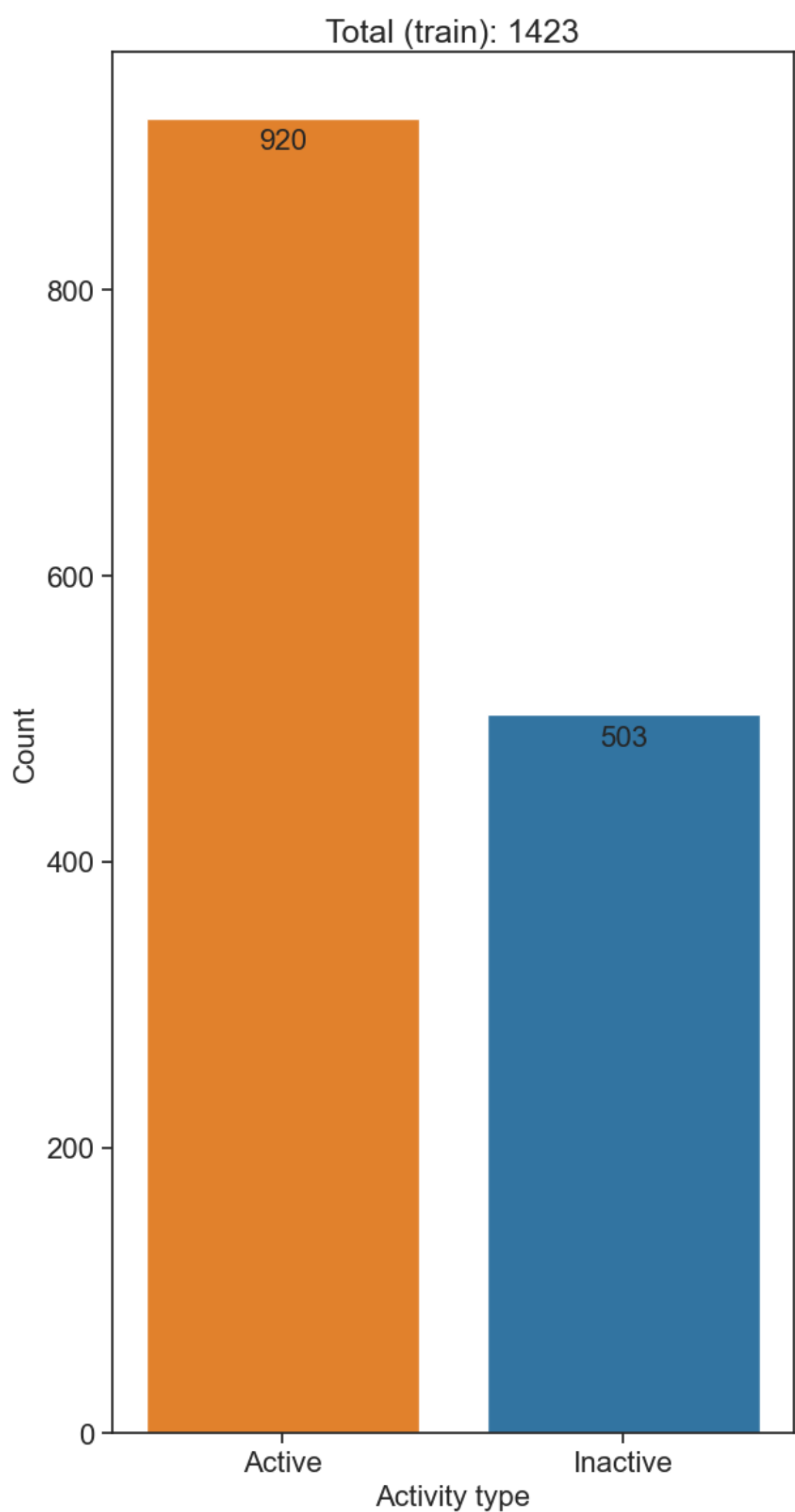
XGBoost Parameters	Value
alpha	0.5253
colsample_bytree	0.7453
gamma	0.2129
lambda	1.1119
learning_rate	0.0687
max_depth	4.0
min_child_weight	1.0
n_estimators	673.0
reg_alpha	0.8551
reg_lambda	0.7983
scale_pos_weight	1.9249
subsample	0.6125

Confusion Matrix

	0	1
0	203	27
1	19	107

Confusion Matrix (nomr.)

	0	1
0	0.88	0.12
1	0.15	0.85

Calibration curve
- train: 0.031
- test: 0.0998
- perfectly_calibrated

Active/Inactive probability
- active (126)
- inactive (230)
- decoy (356)

Inactive zone | Active zone
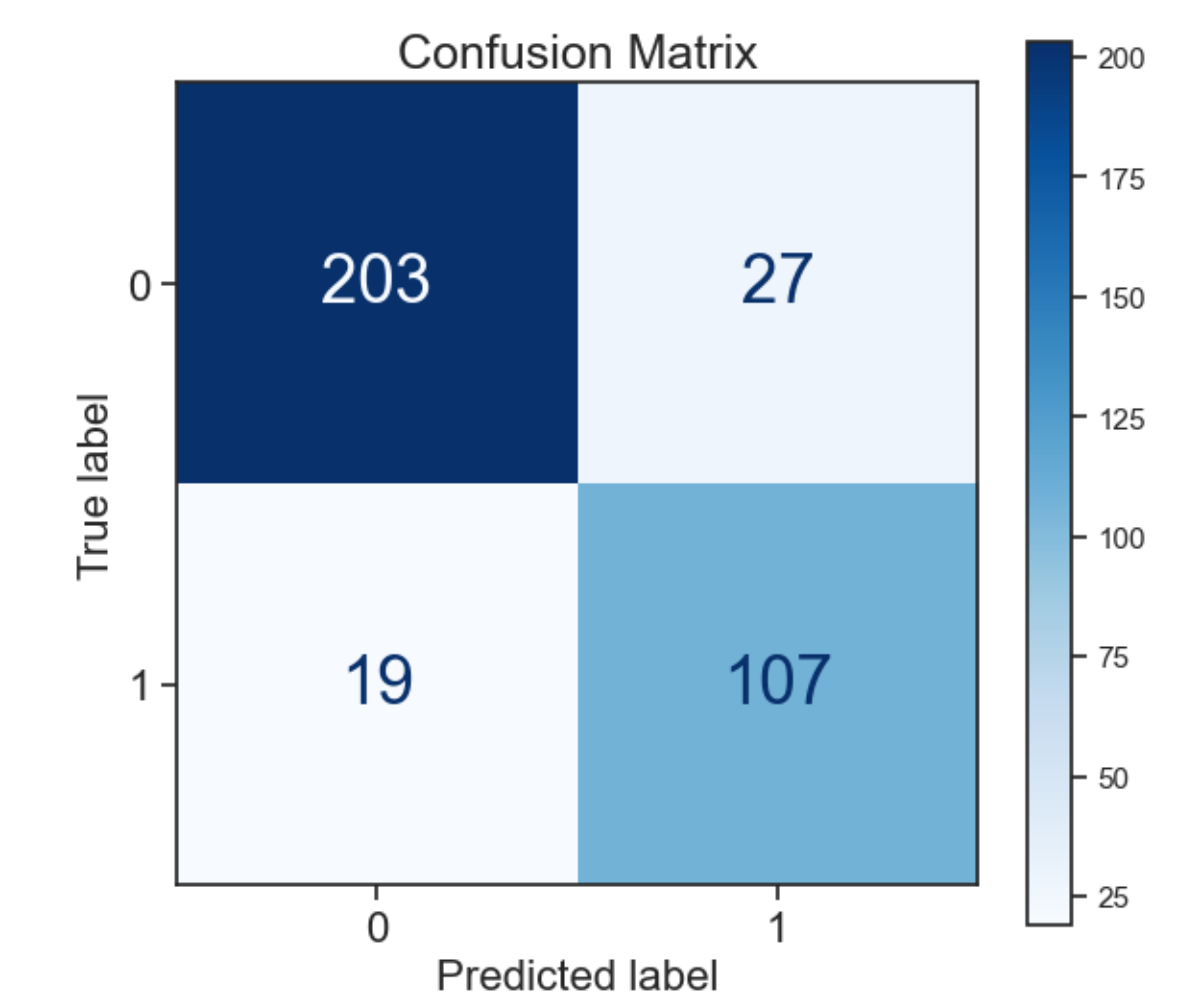
Metric	Train	Test
AUC	0.9961	0.9291
accuracy	0.9761	0.8708
balanced accuracy	0.9797	0.8659
recall_(sens)	0.992	0.8492
specificity	0.9674	0.8826
precision	0.9433	0.7985
f1_score	0.9762	0.8716

Protein: P21917 - metric: f1

Total (train): 1423

ROC curve

AUC_train = 0.999
AUC_test = 0.933
Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 899)
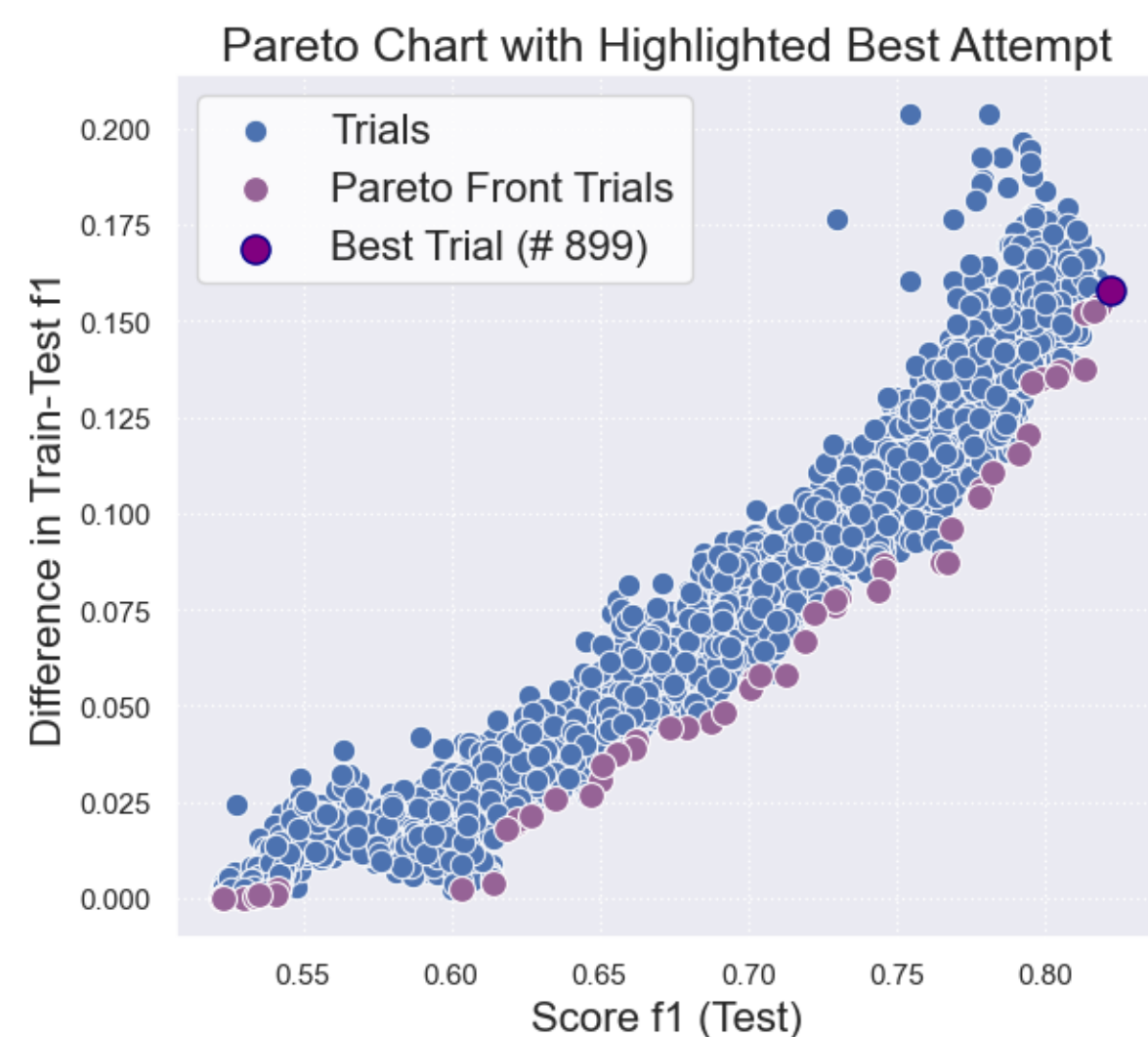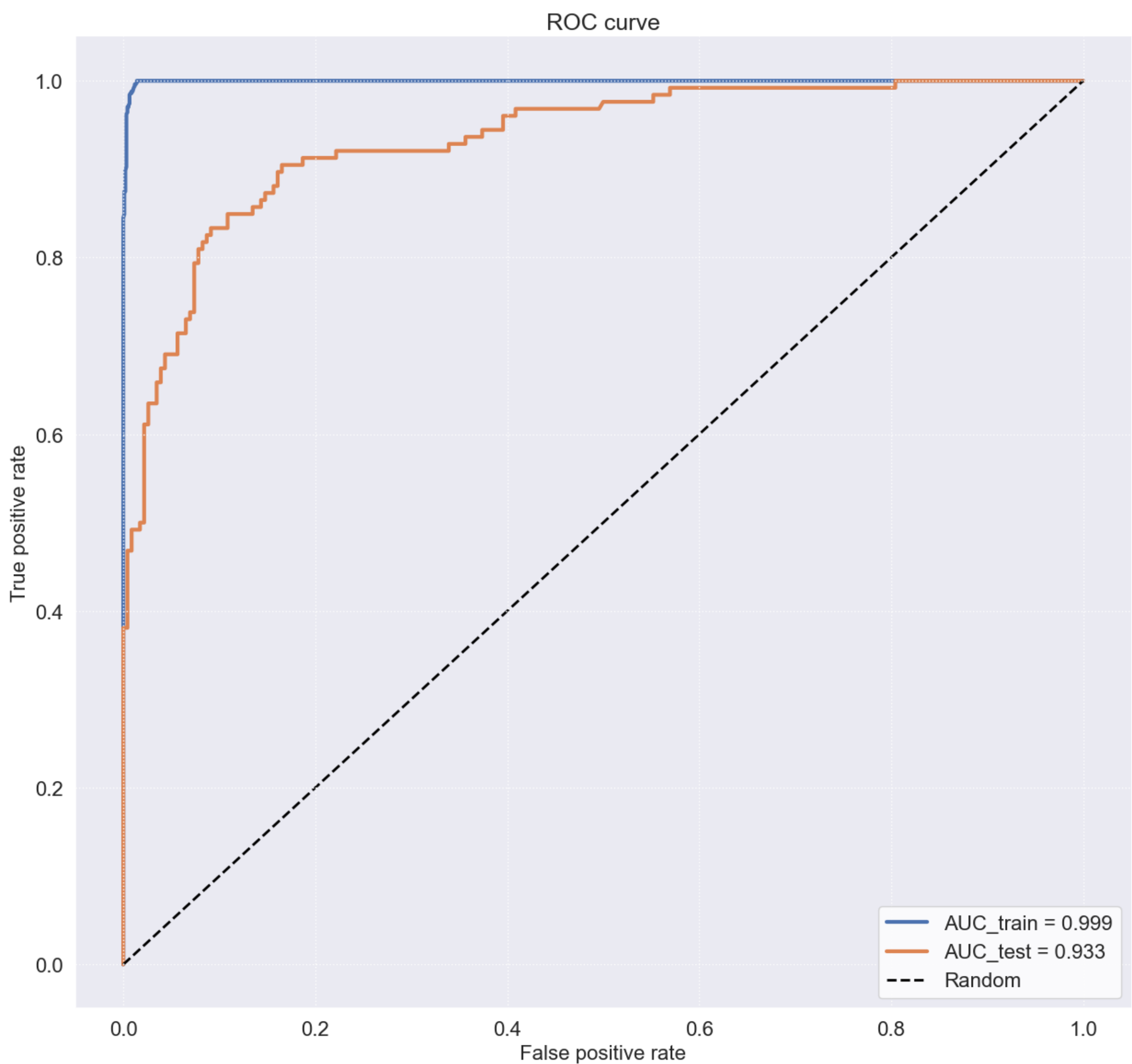
XGBoost Parameters	Value
alpha	2.3709
colsample_bytree	0.8334
gamma	0.1988
lambda	1.0098
learning_rate	0.1147
max_depth	9.0
min_child_weight	2.0
n_estimators	245.0
reg_alpha	0.2669
reg_lambda	0.1769
scale_pos_weight	3.2783
subsample	0.7374

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0155
test: 0.0982
perfectly_calibrated

Active/Inactive probability

type
- active (126)
- inactive (230)
- decoy (356)

Inactive zone   Active zone
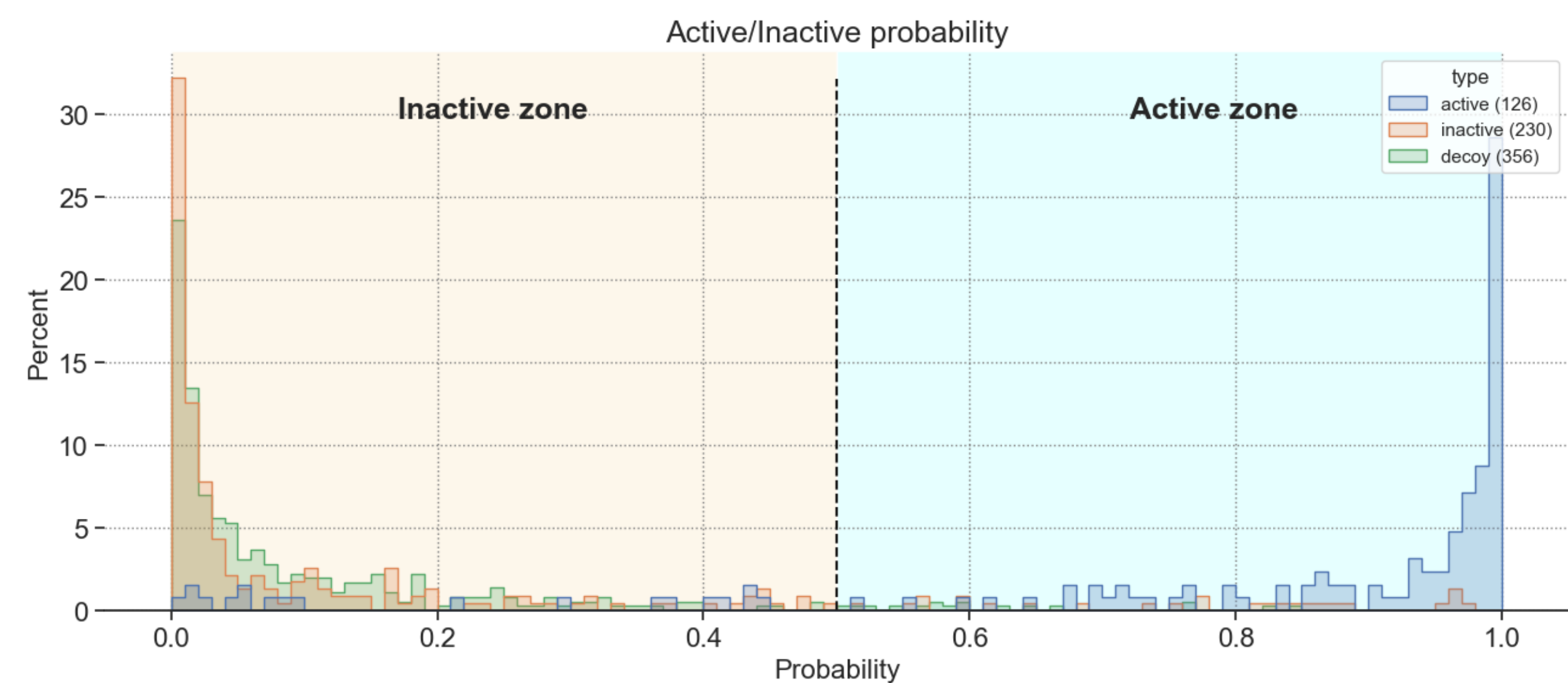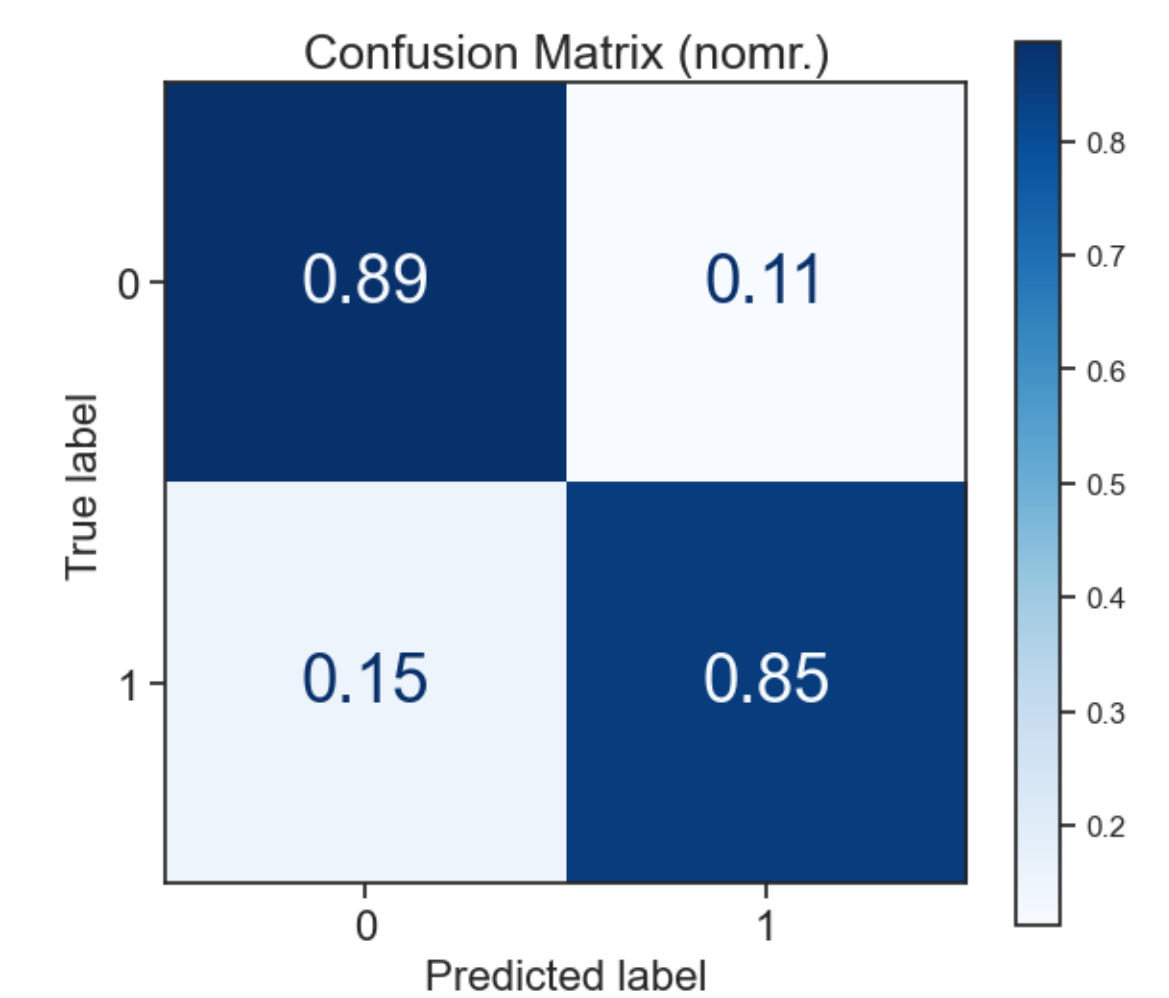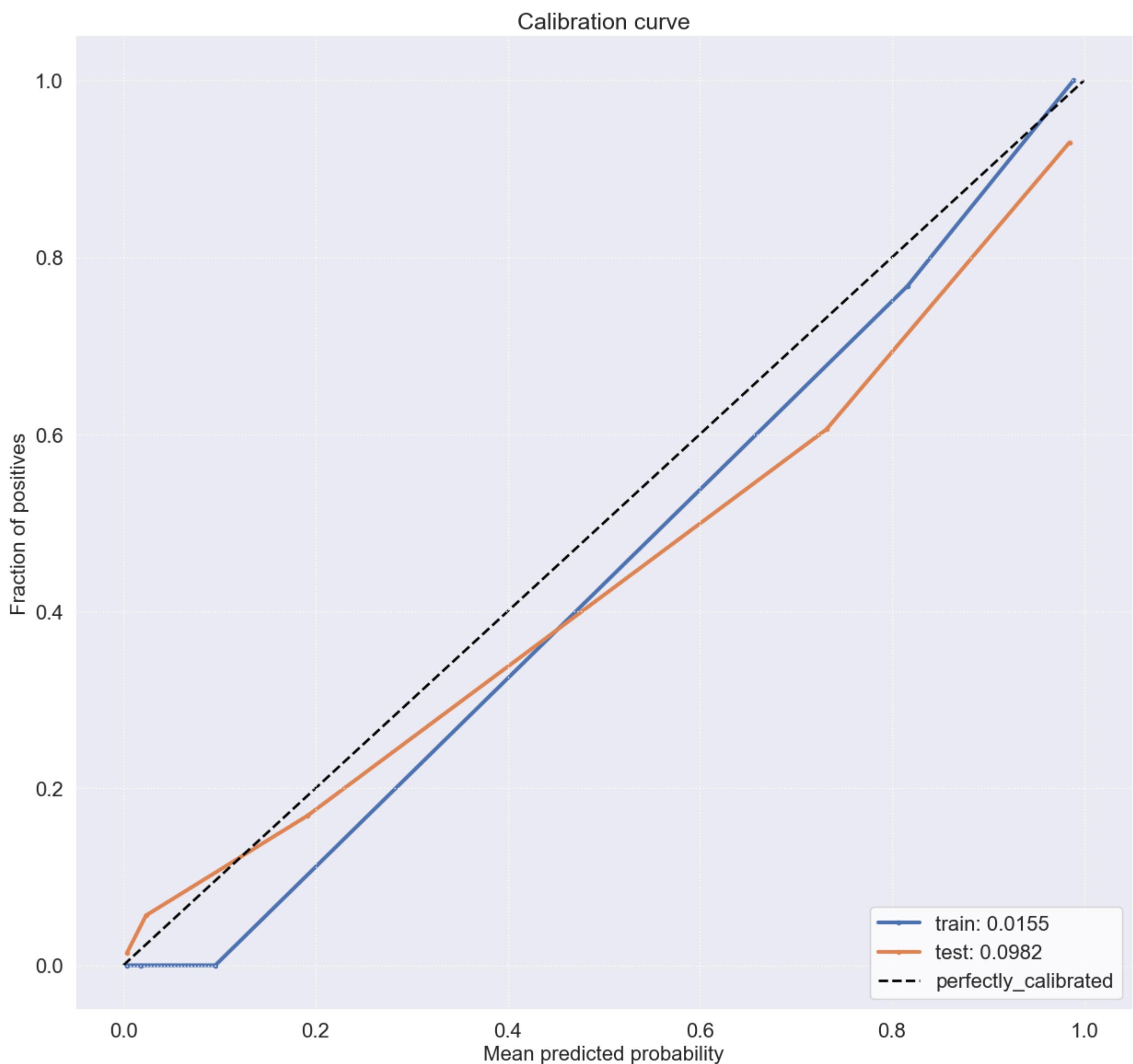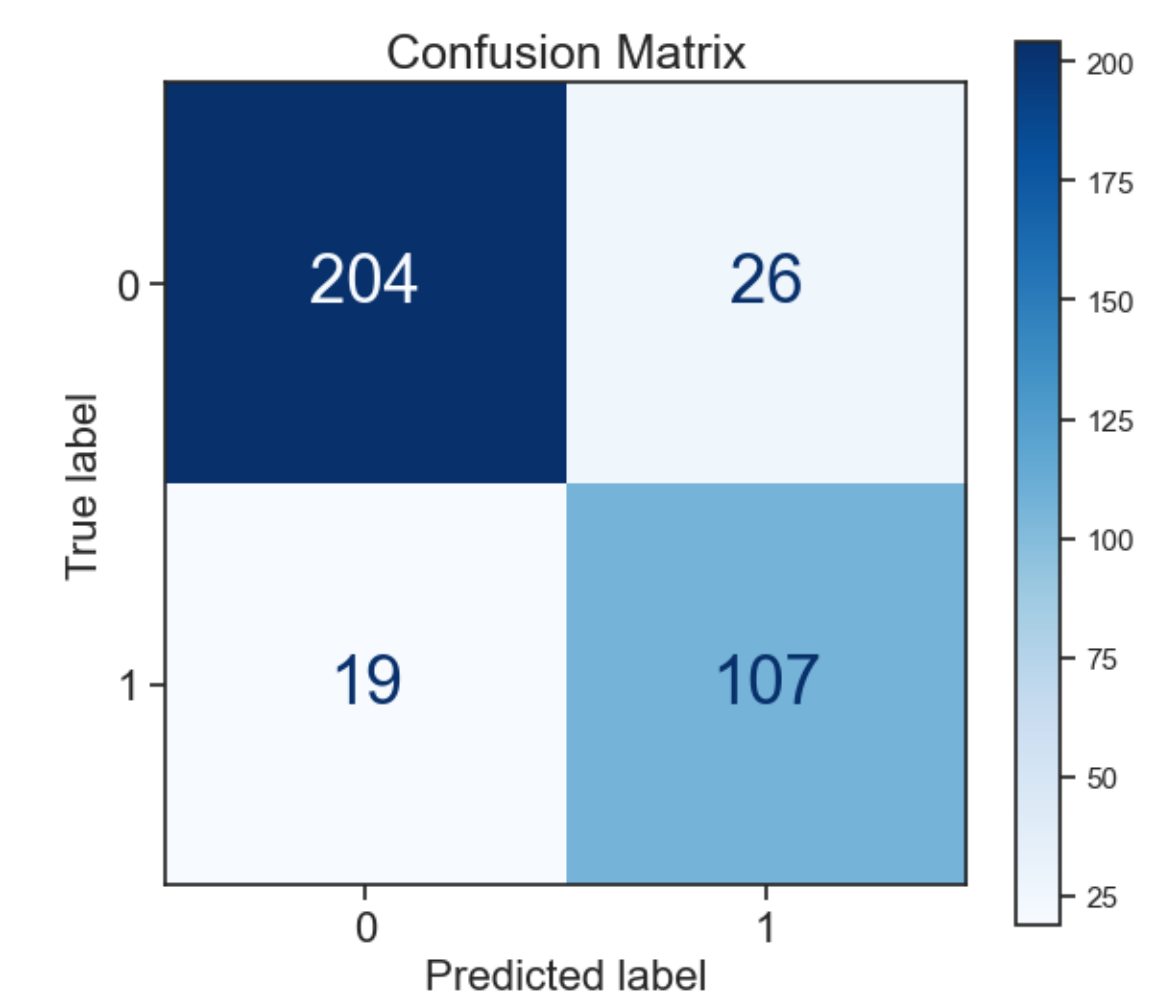
Metric	Train	Test
AUC	0.9994	0.933
accuracy	0.9838	0.8736
balanced accuracy	0.9875	0.8681
recall_(sens)	1.0	0.8492
specificity	0.975	0.887
precision	0.9563	0.8045
f1_score	0.9839	0.8743

Protein: P21917 - metric: precision

Total (train): 1423

ROC curve

- AUC_train = 1.000
- AUC_test = 0.927
- Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 2742)
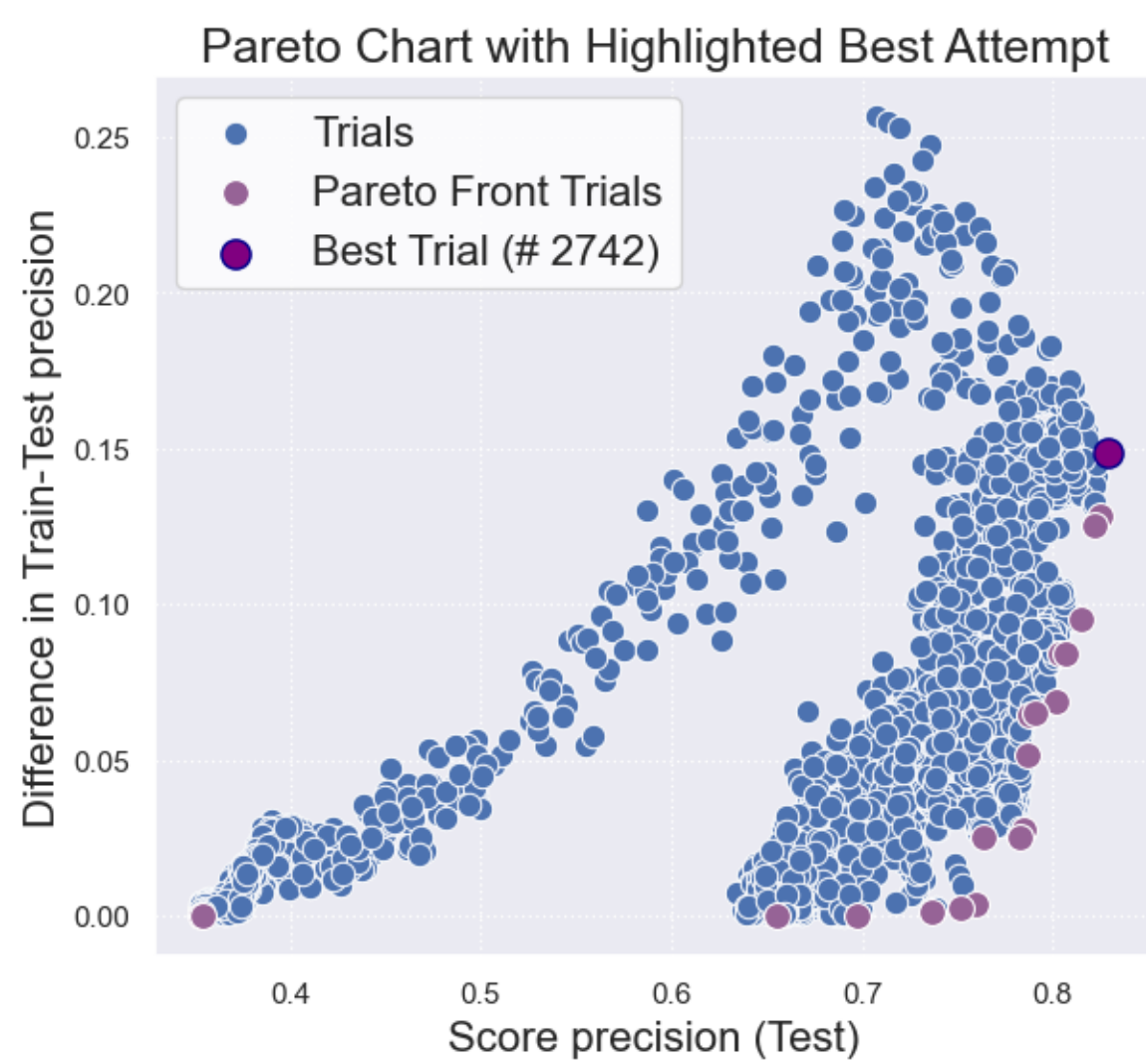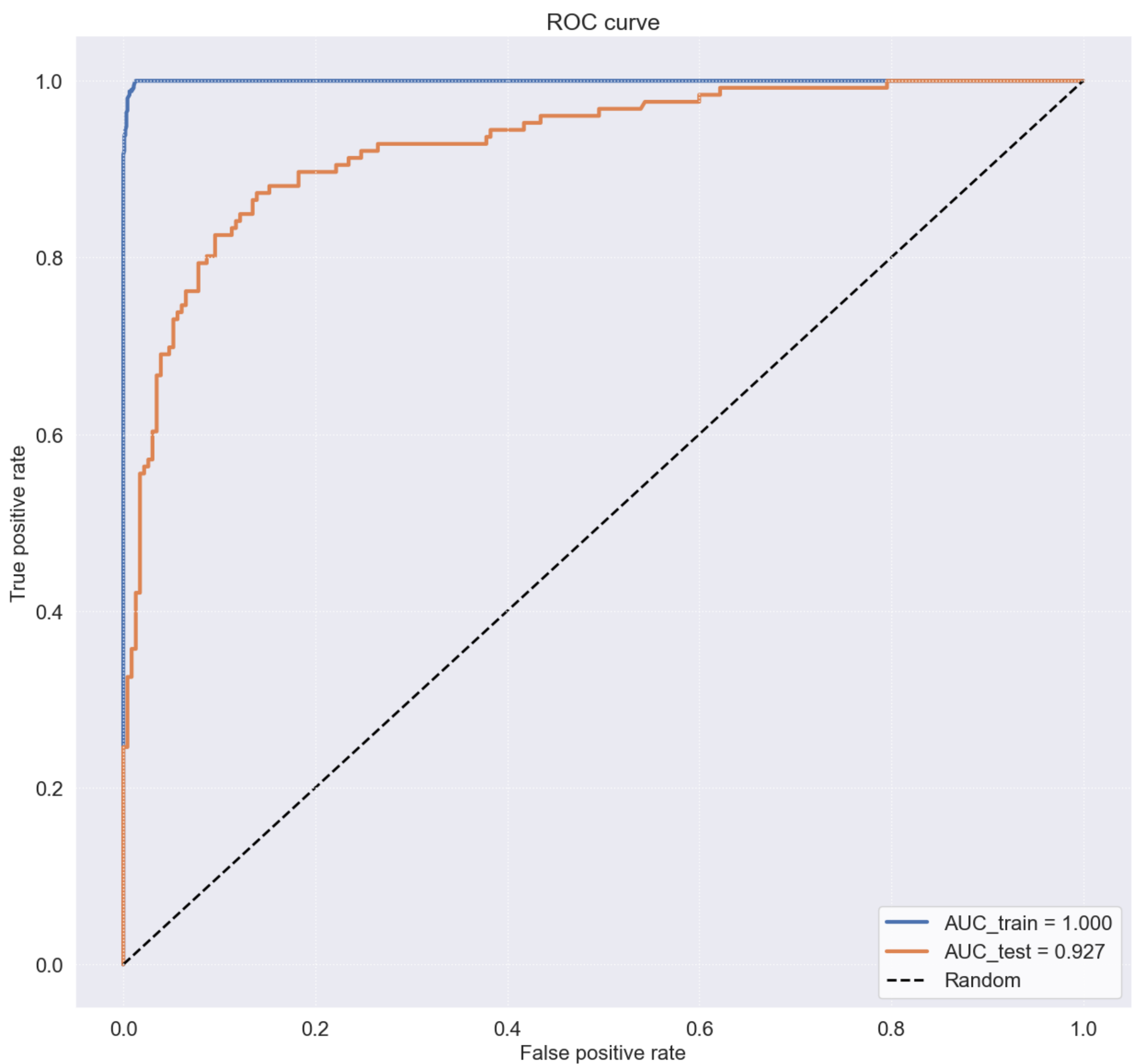
XGBoost Parameters	Value
alpha	0.7002
colsample_bytree	0.565
gamma	0.1708
lambda	4.2069
learning_rate	0.0885
max_depth	10.0
min_child_weight	2.0
n_estimators	753.0
reg_alpha	0.2822
reg_lambda	0.0791
scale_pos_weight	1.3728
subsample	0.82

Confusion Matrix

	0	1
0	208	22
1	23	103

Confusion Matrix (nomr.)

	0	1
0	0.9	0.096
1	0.18	0.82

Calibration curve

- train: 0.0104
- test: 0.098
- perfectly_calibrated

Active/Inactive probability

Inactive zone | Active zone

type
- active (126)
- inactive (230)
- decoy (356)
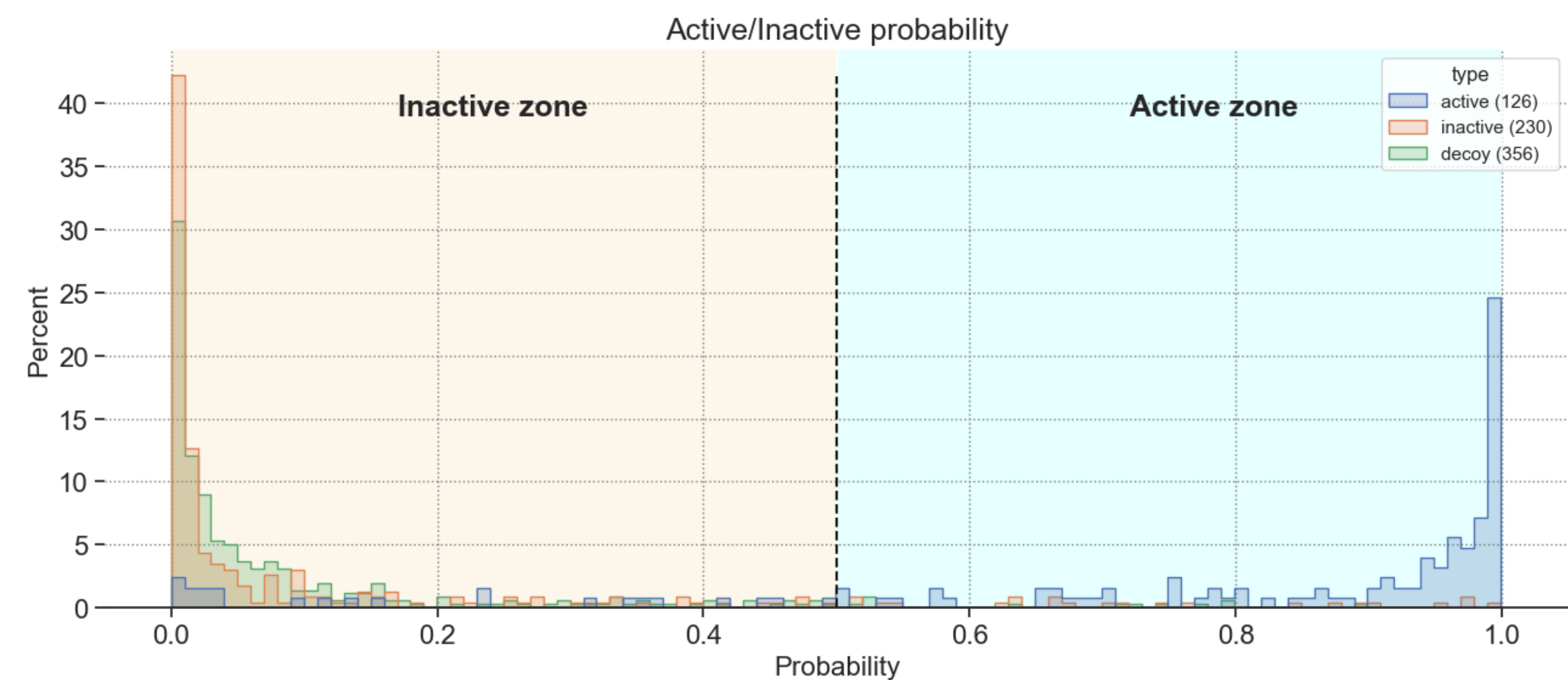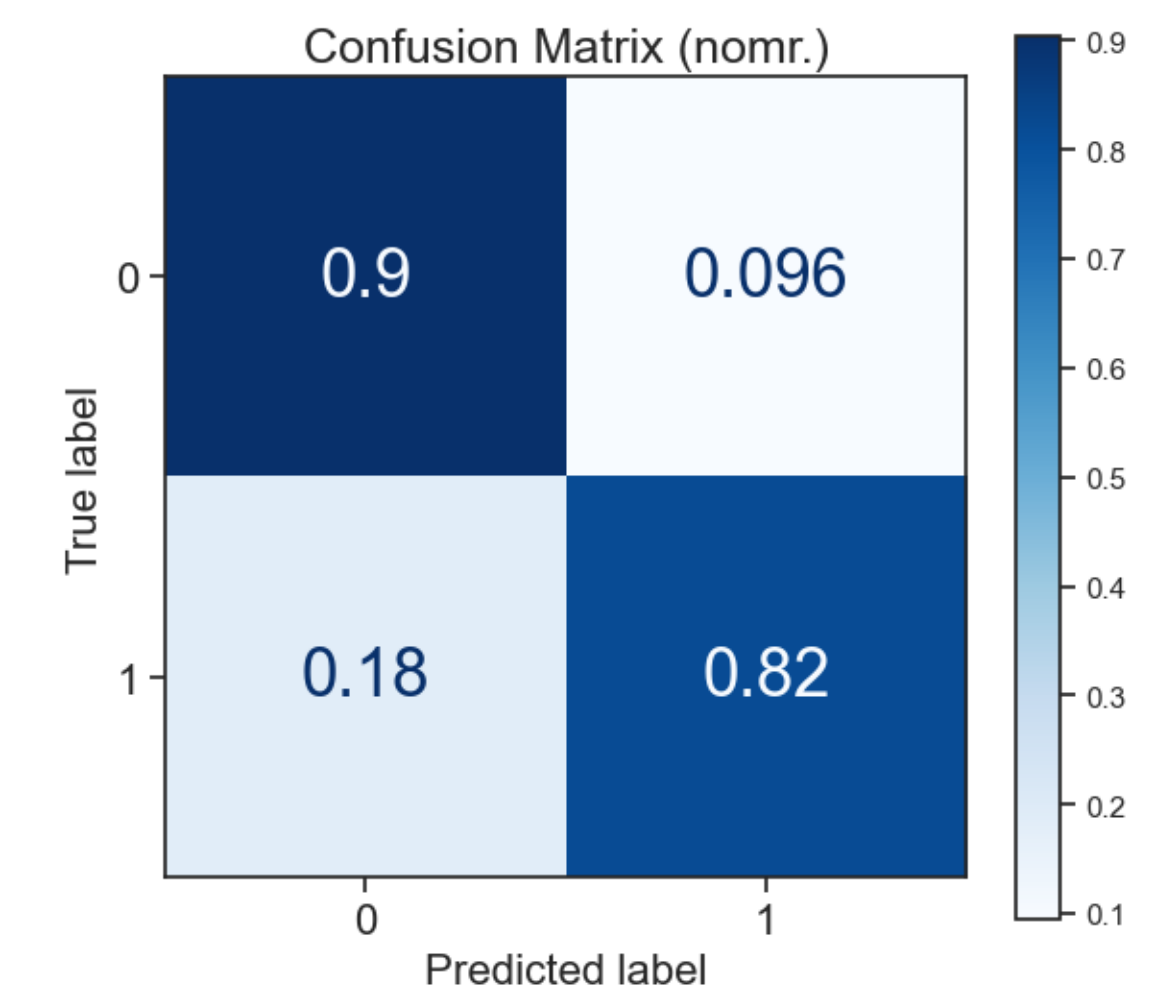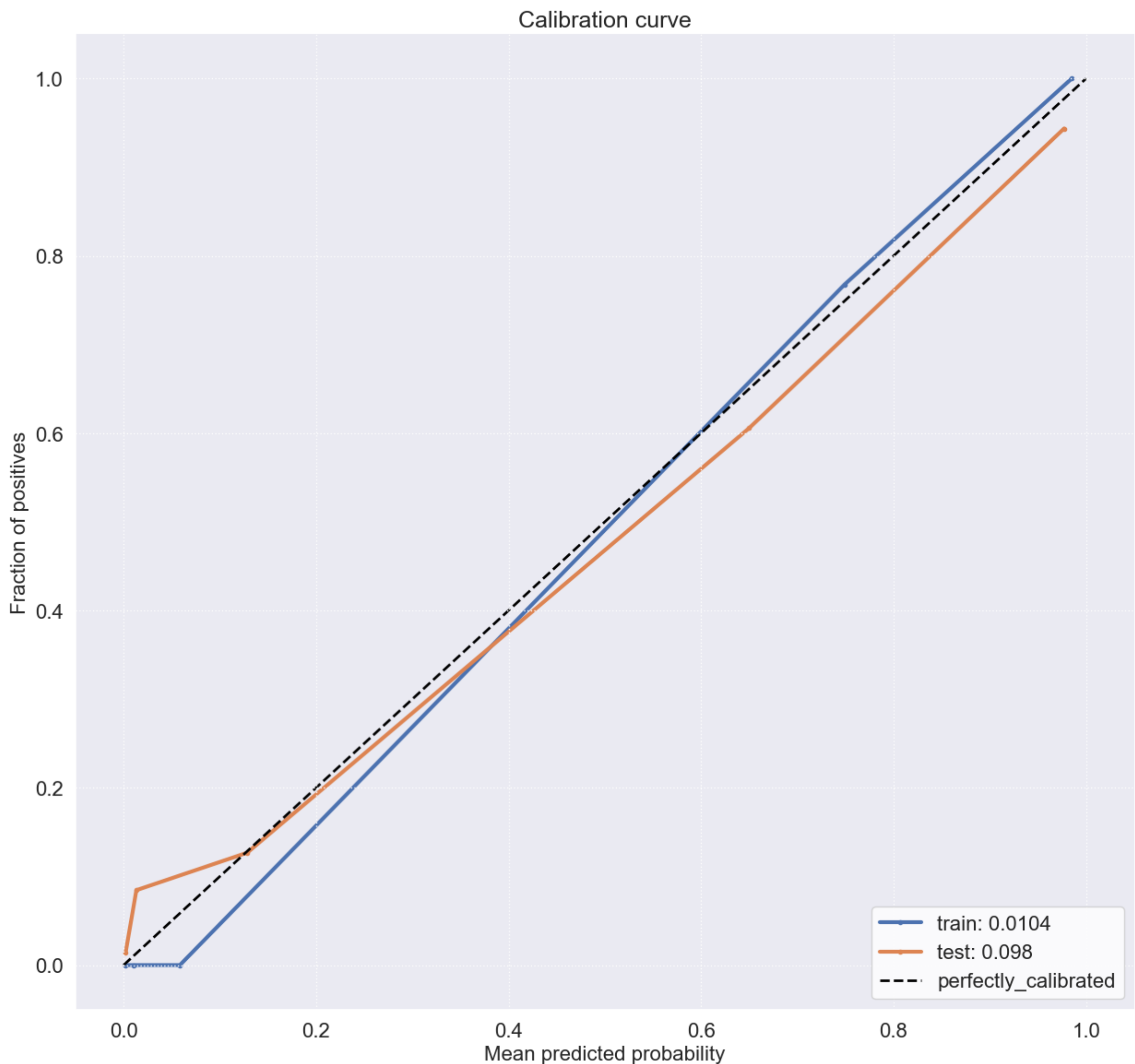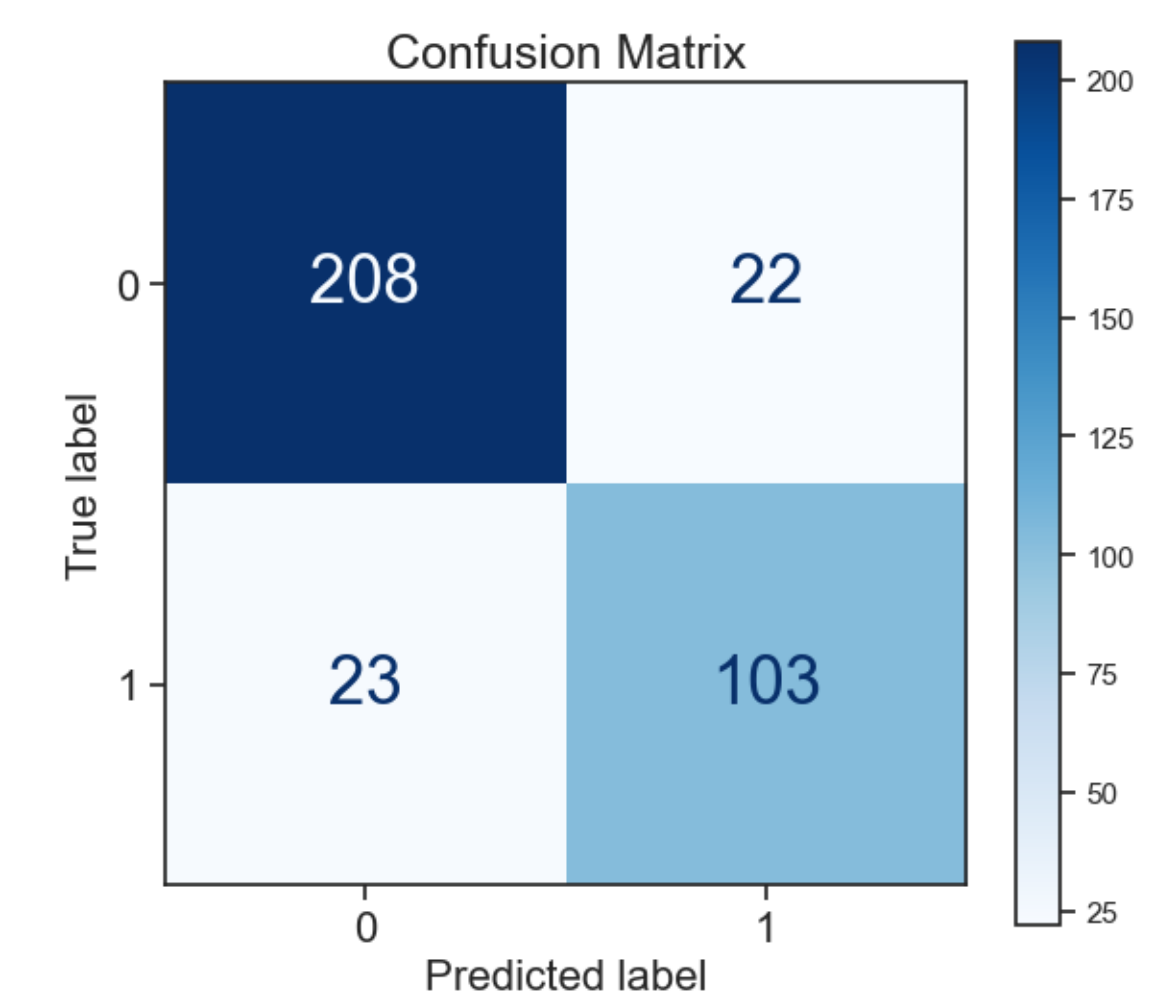
Metric	Train	Test
AUC	0.9997	0.9274
accuracy	0.9909	0.8736
balanced accuracy	0.9929	0.8609
recall_(sens)	1.0	0.8175
specificity	0.9859	0.9043
precision	0.9748	0.824
f1_score	0.9909	0.8735

Protein: P21917 - metric: SoftVotingXGBoostEnsemble
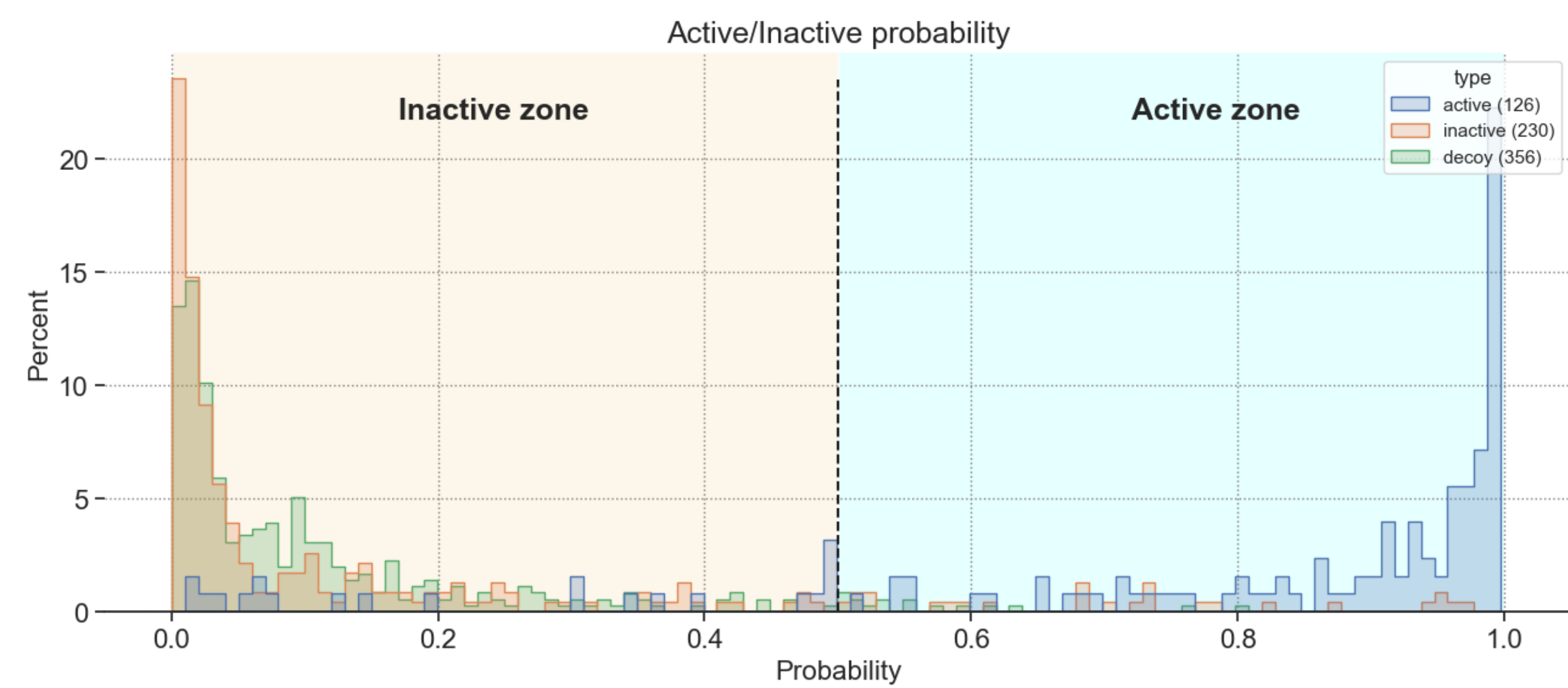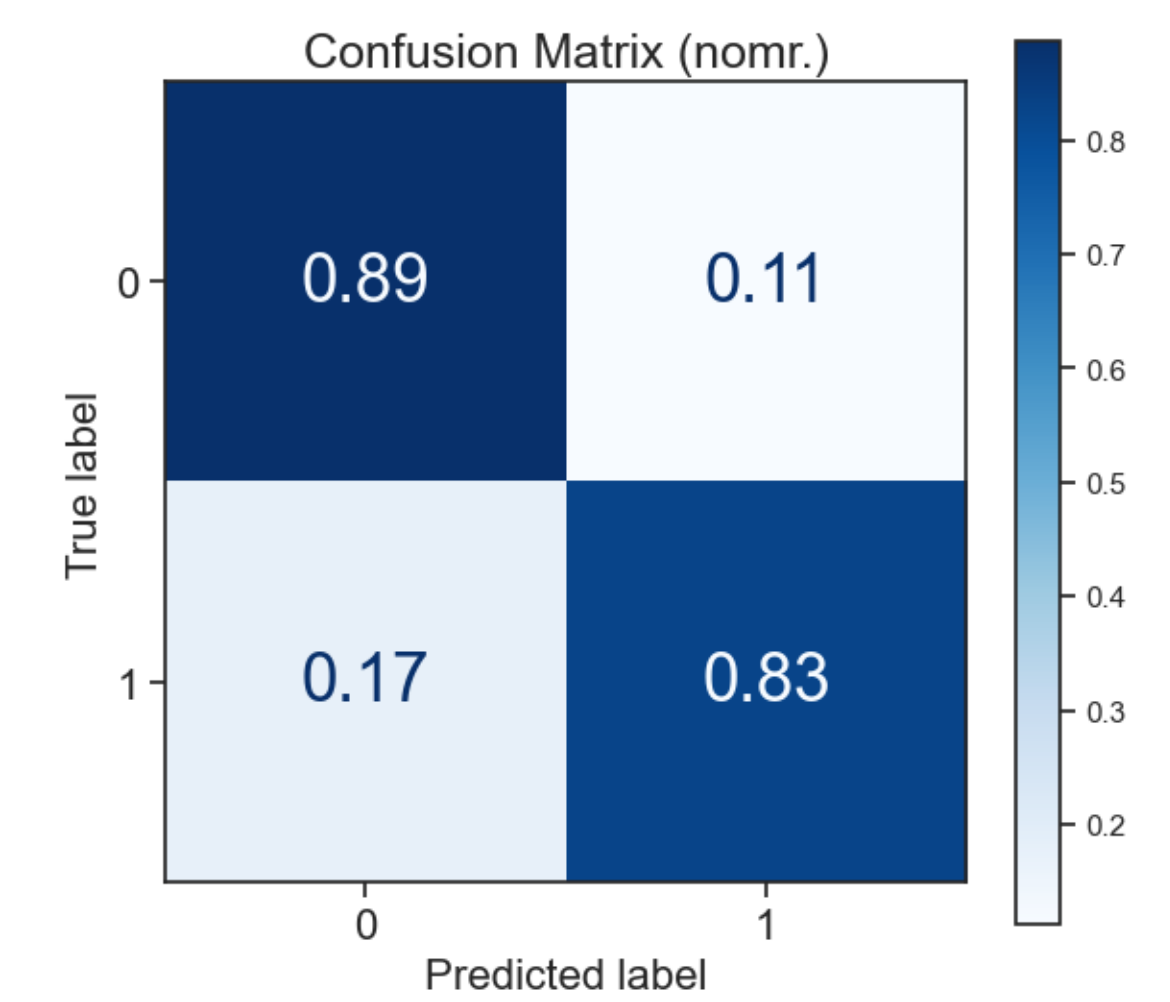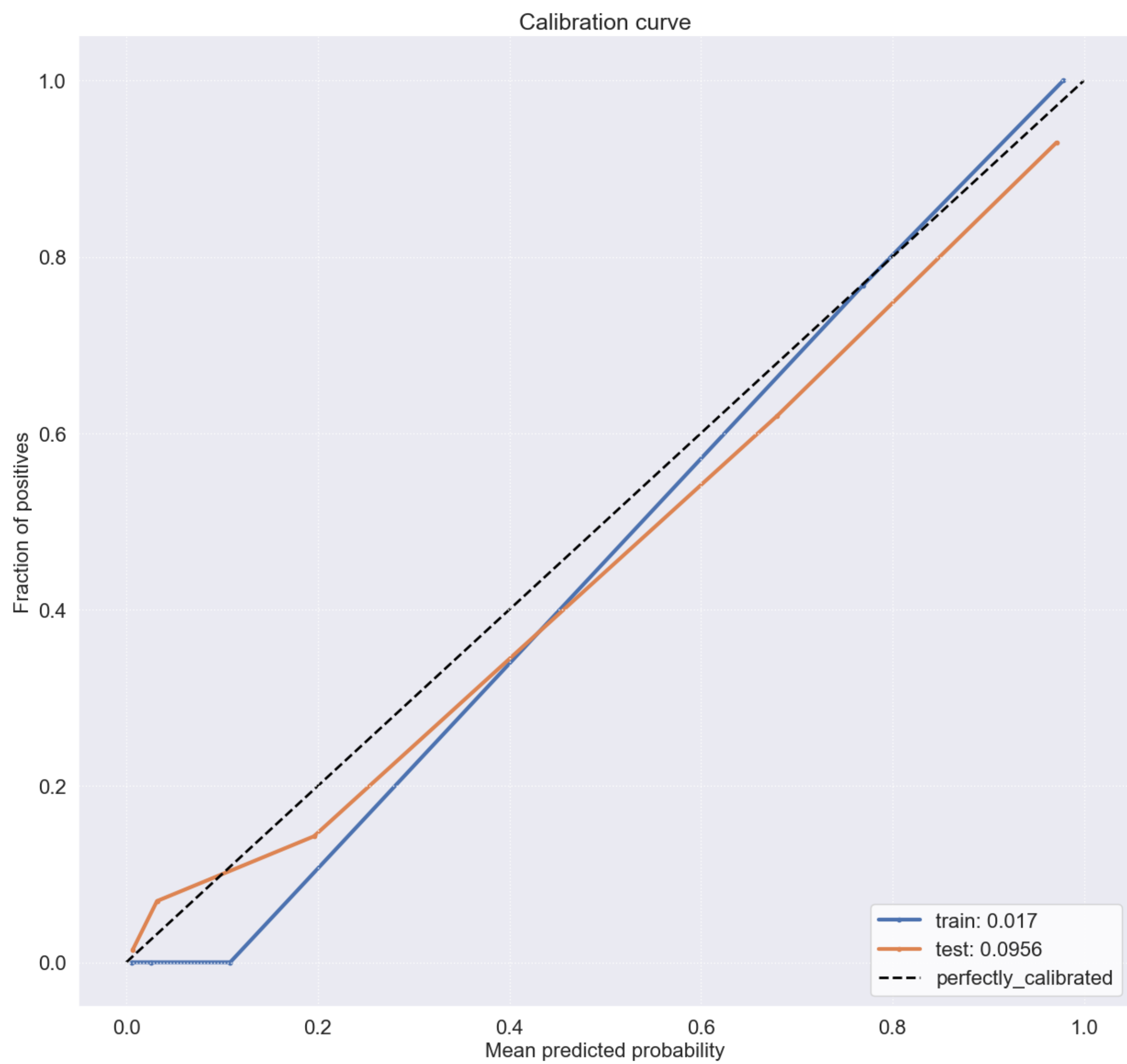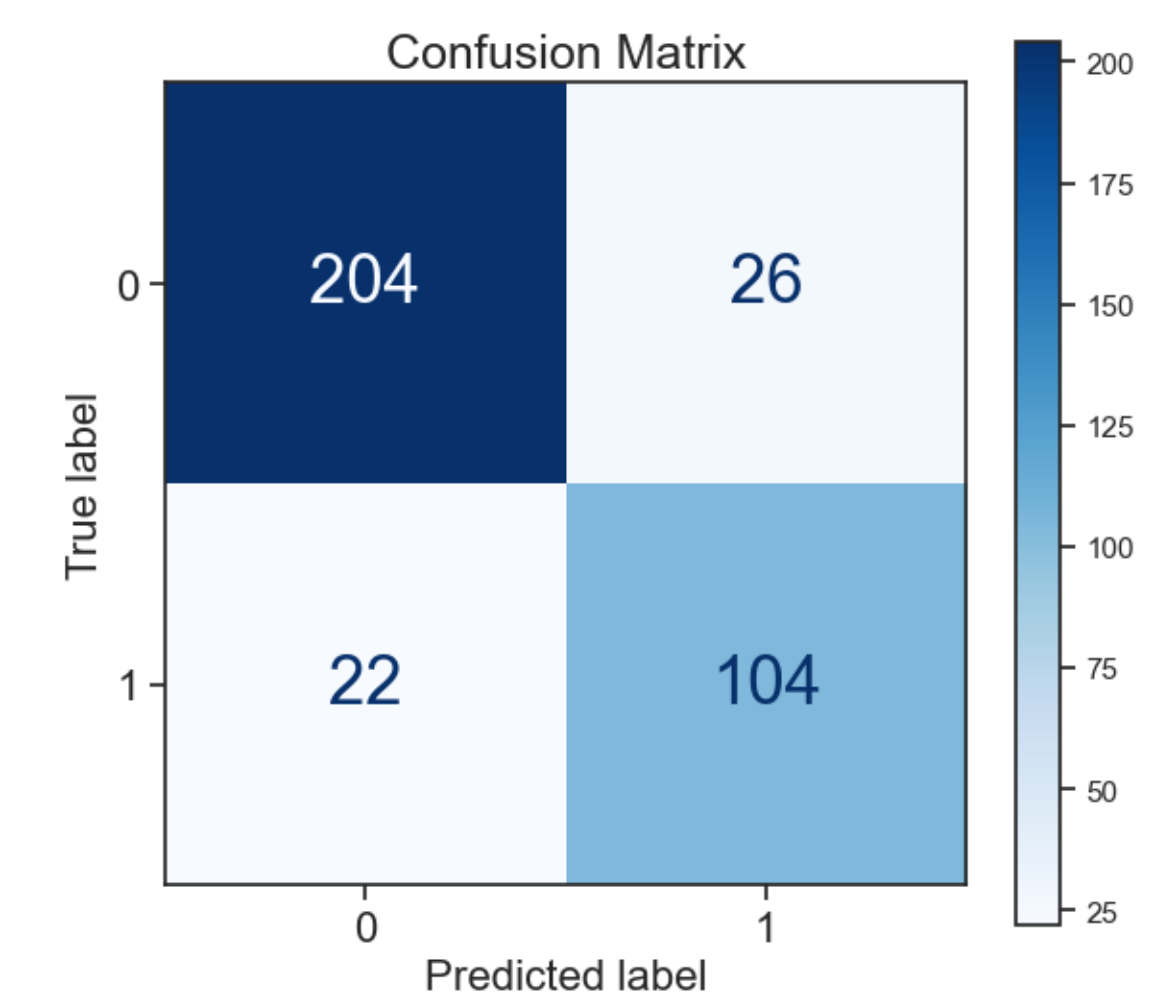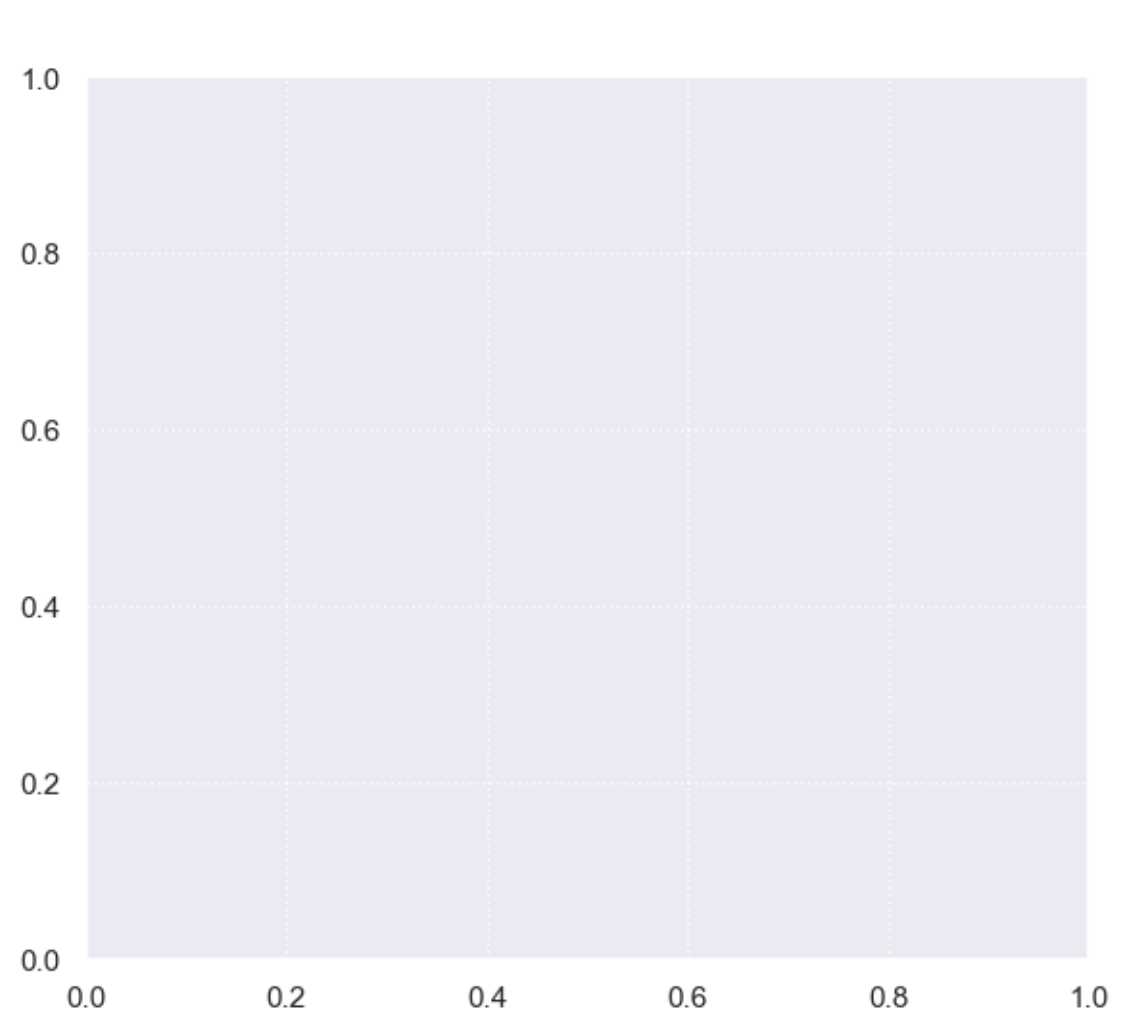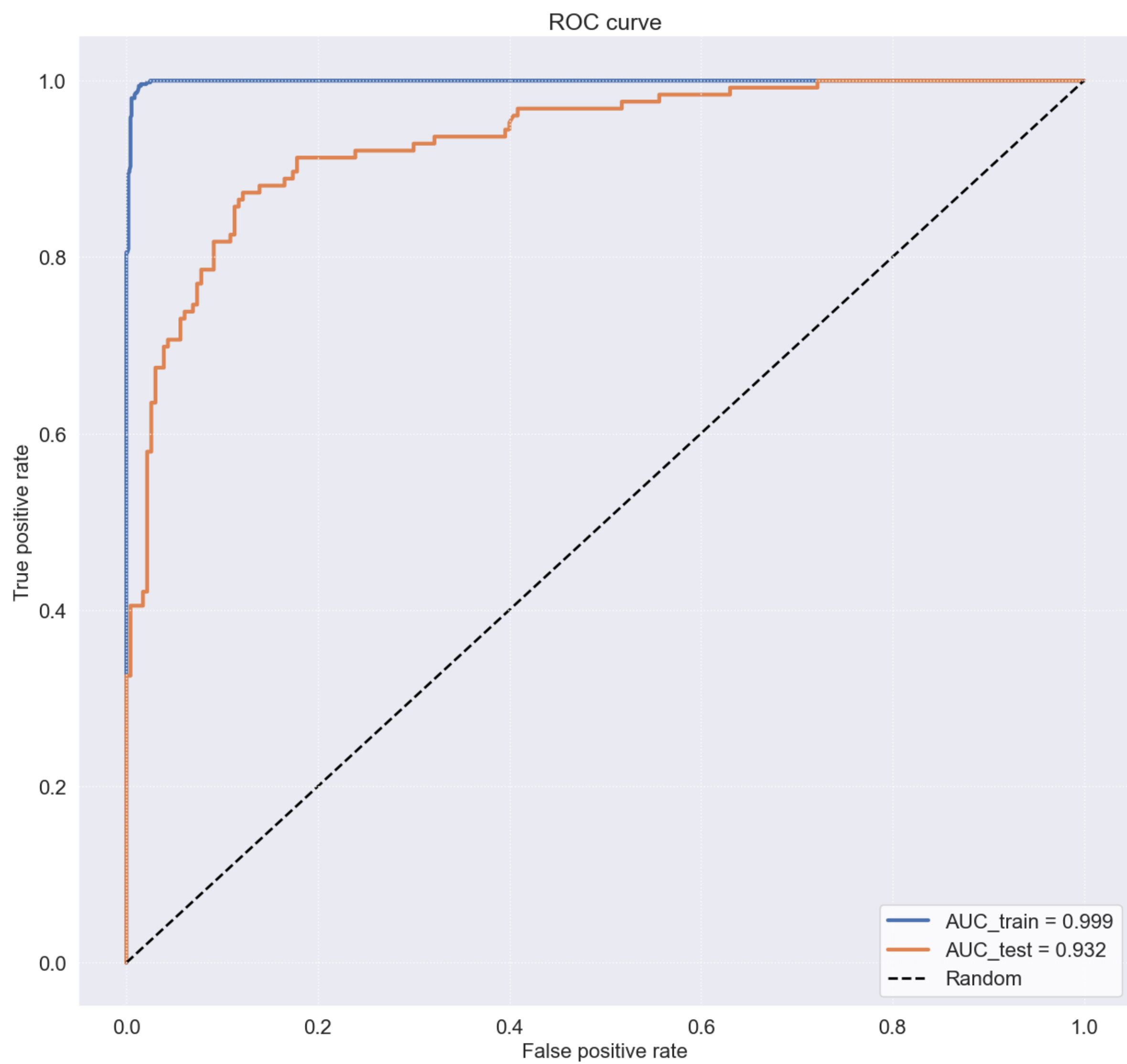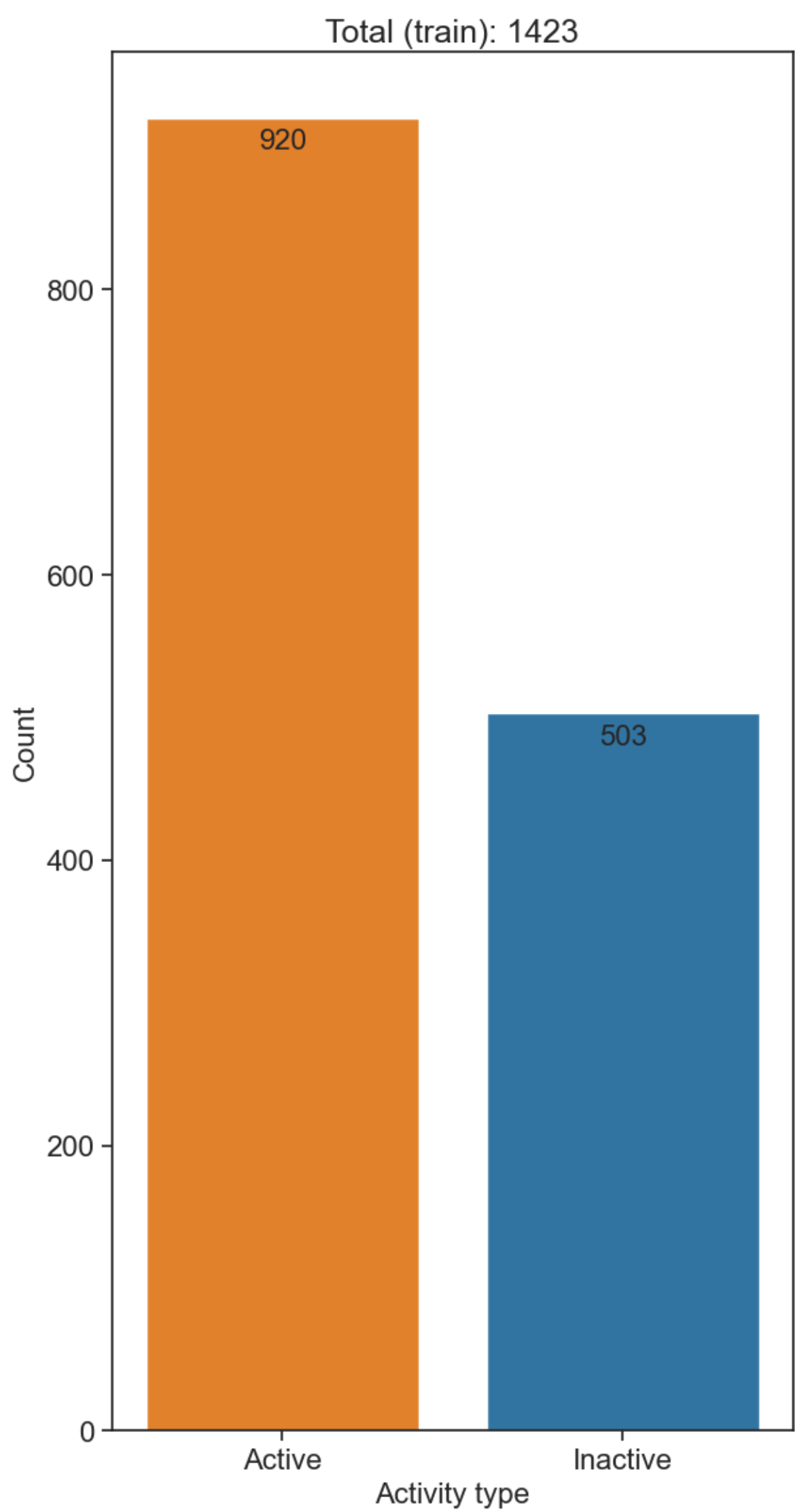
Metric	Train	Test
AUC	0.9991	0.9319
accuracy	0.9831	0.8652
balanced accuracy	0.987	0.8562
recall_(sens)	1.0	0.8254
specificity	0.9739	0.887
precision	0.9545	0.8
f1_score	0.9832	0.8656