Protein: P30556 - metric: balanced_accuracy

Total (train): 794

ROC curve

- AUC_train = 0.998
- AUC_test = 0.971
- Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 1301)

Confusion Matrix

	0	1
0	60	9
1	3	127

Confusion Matrix (nomr.)

	0	1
0	0.87	0.13
1	0.023	0.98

Calibration curve

- train: 0.0173
- test: 0.0442
- perfectly_calibrated

Active/Inactive probability

- Inactive zone
- Active zone
- type
  - active (130)
  - inactive (69)
  - decoy (199)

XGBoost Parameters	Value
alpha	1.2623
colsample_bytree	0.5072
gamma	0.6691
lambda	3.6453
learning_rate	0.0415
max_depth	3.0
min_child_weight	2.0
n_estimators	470.0
reg_alpha	0.3109
reg_lambda	0.8643
scale_pos_weight	1.1914
subsample	0.8017
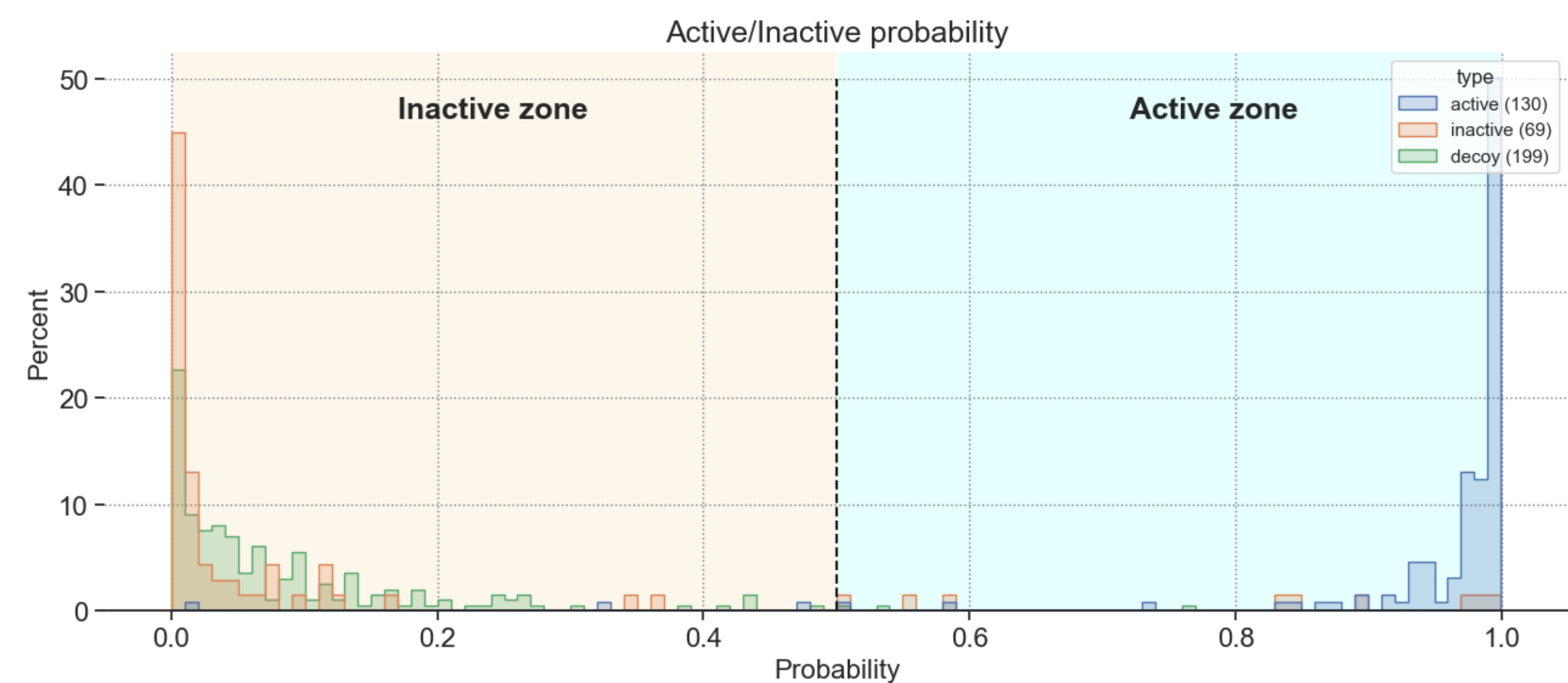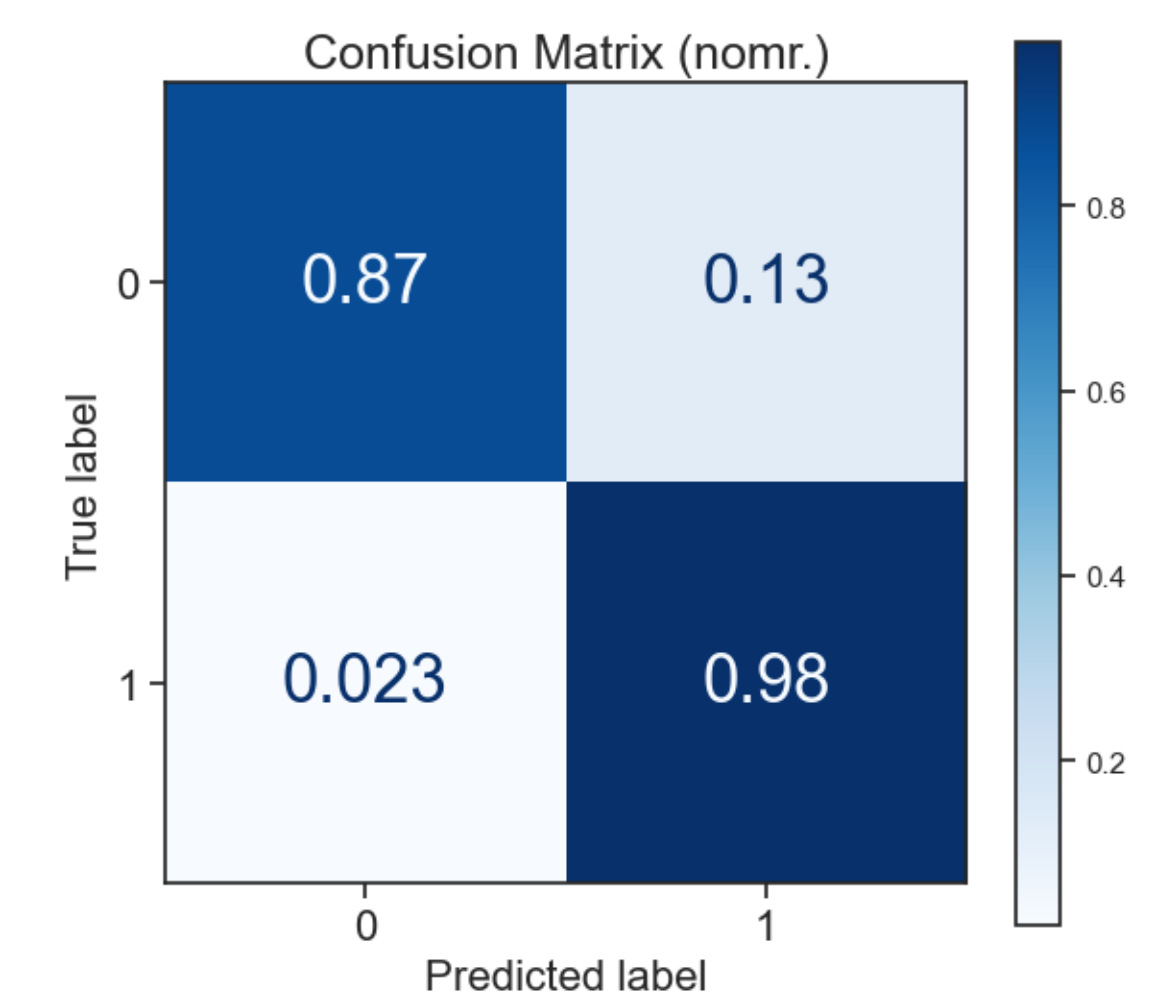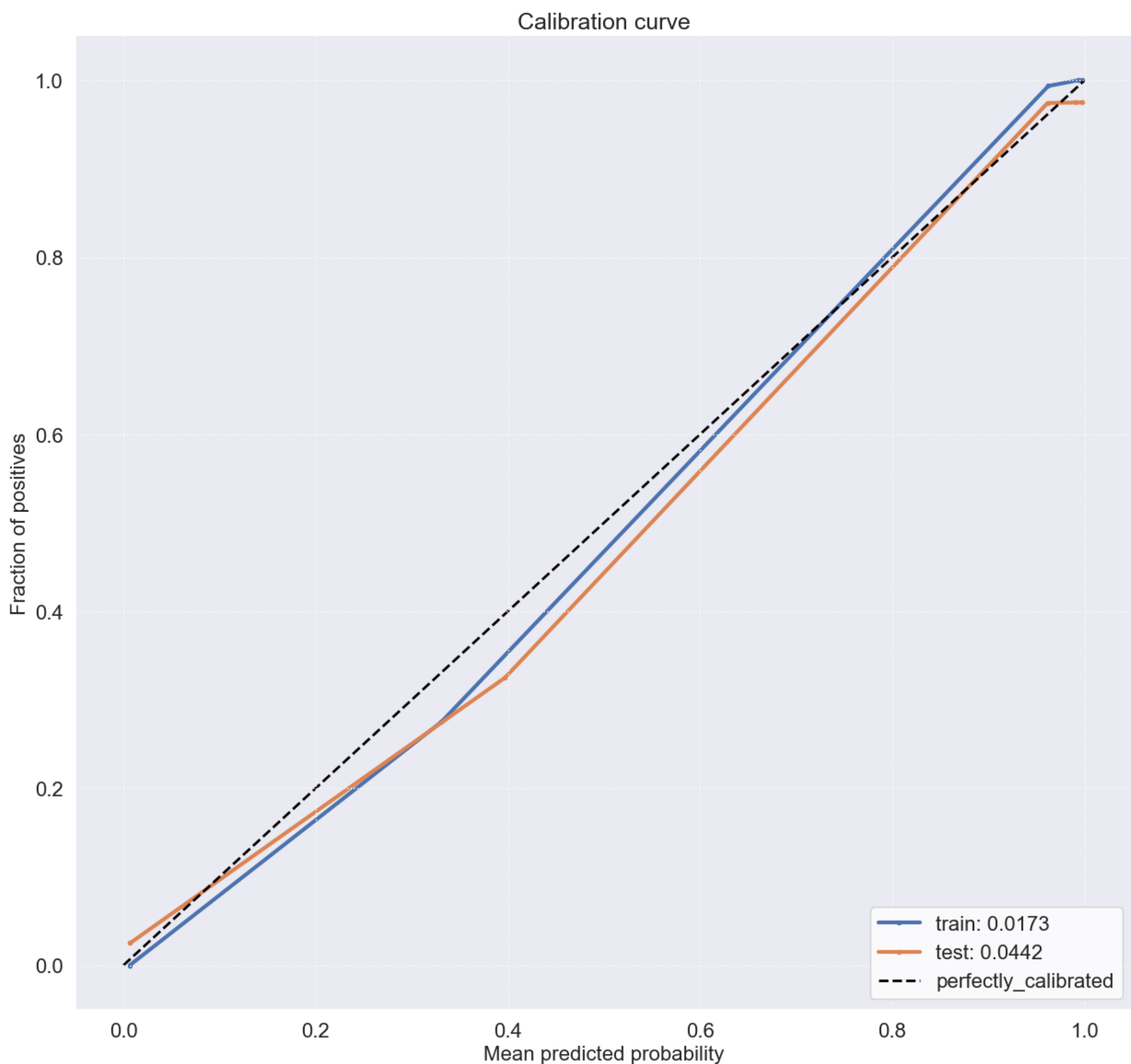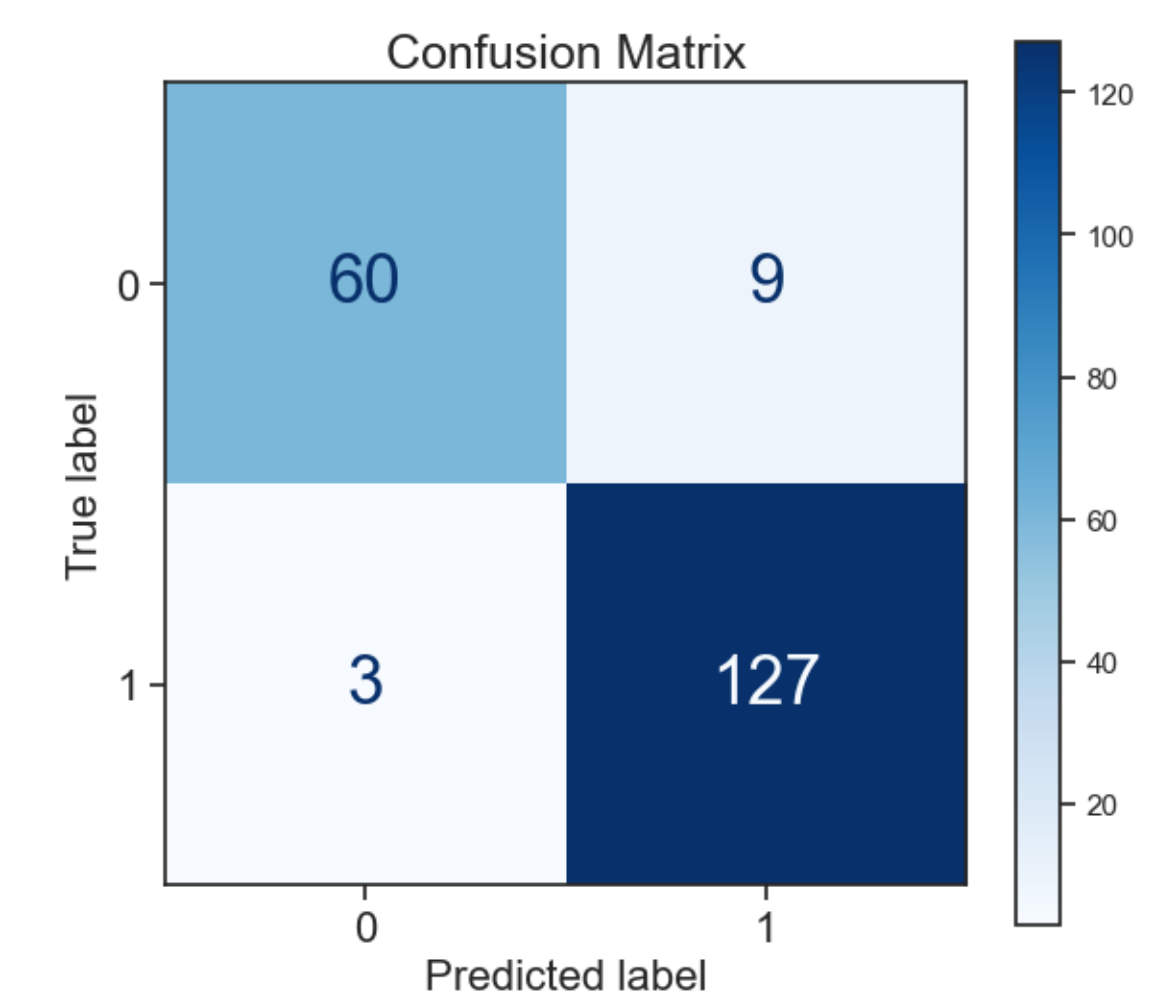
Metric	Train	Test
AUC	0.998	0.9713
accuracy	0.9761	0.9397
balanced accuracy	0.9698	0.9232
recall_(sens)	0.9903	0.9769
specificity	0.9493	0.8696
precision	0.9734	0.9338
f1_score	0.976	0.939

Protein: P30556 - metric: f1

Total (train): 794

ROC curve

AUC_train = 0.999
AUC_test = 0.976
Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 2515)
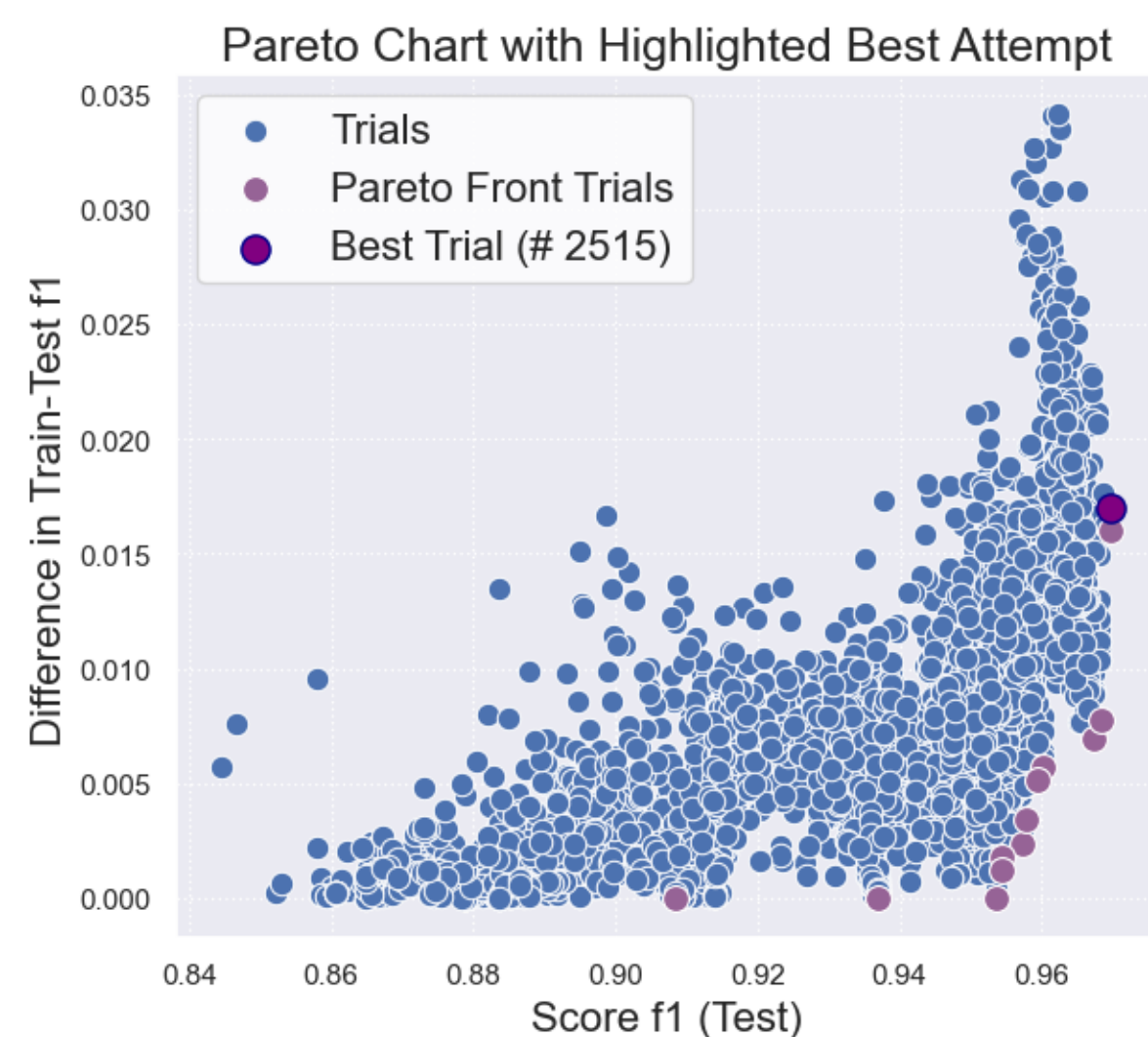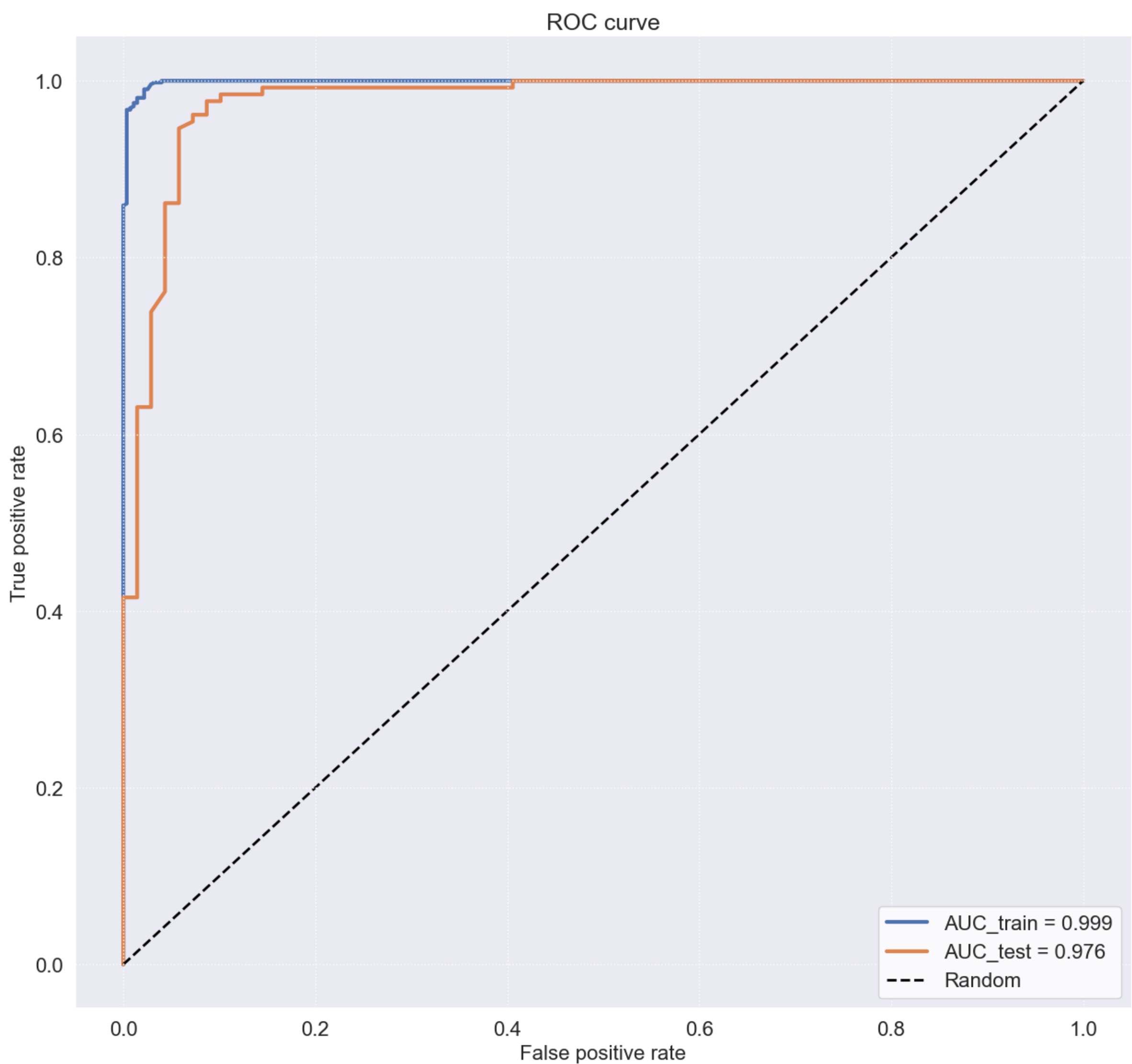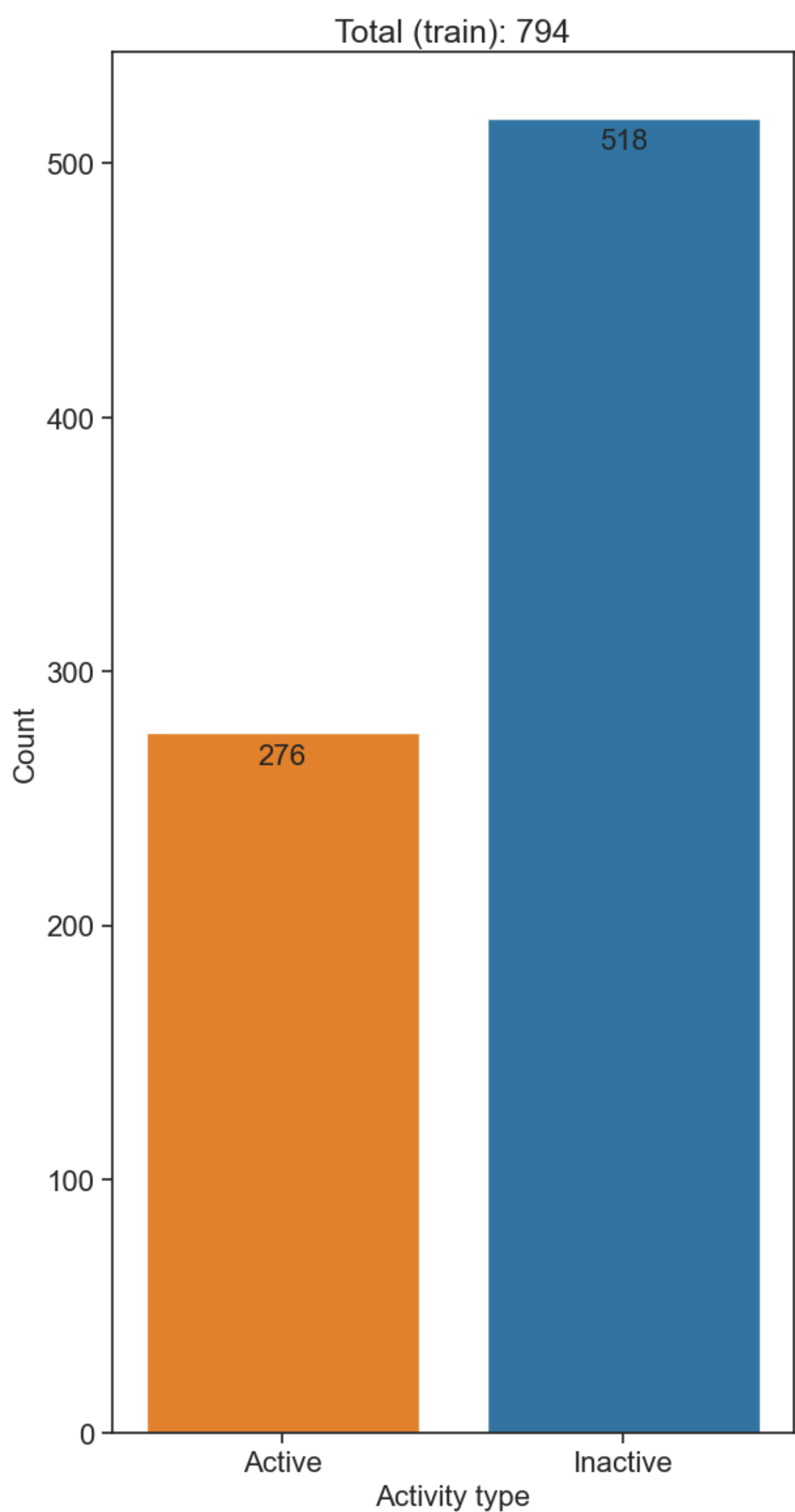
XGBoost Parameters	Value
alpha	0.4348
colsample_bytree	0.8012
gamma	0.0465
lambda	1.2985
learning_rate	0.0322
max_depth	7.0
min_child_weight	3.0
n_estimators	480.0
reg_alpha	0.6542
reg_lambda	0.0976
scale_pos_weight	2.556
subsample	0.8778

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0138
test: 0.0423
perfectly_calibrated

Active/Inactive probability

type
active (130)
inactive (69)
decoy (199)

Inactive zone / Active zone
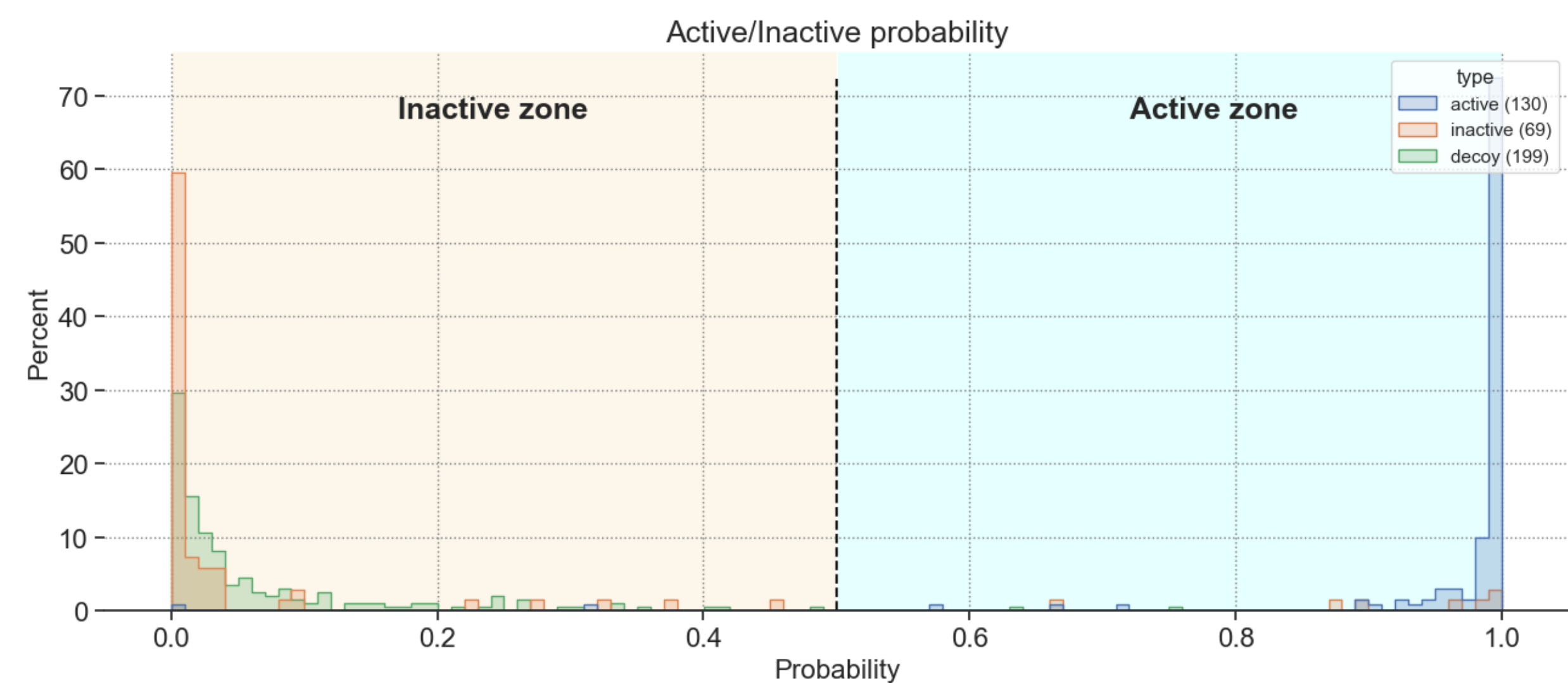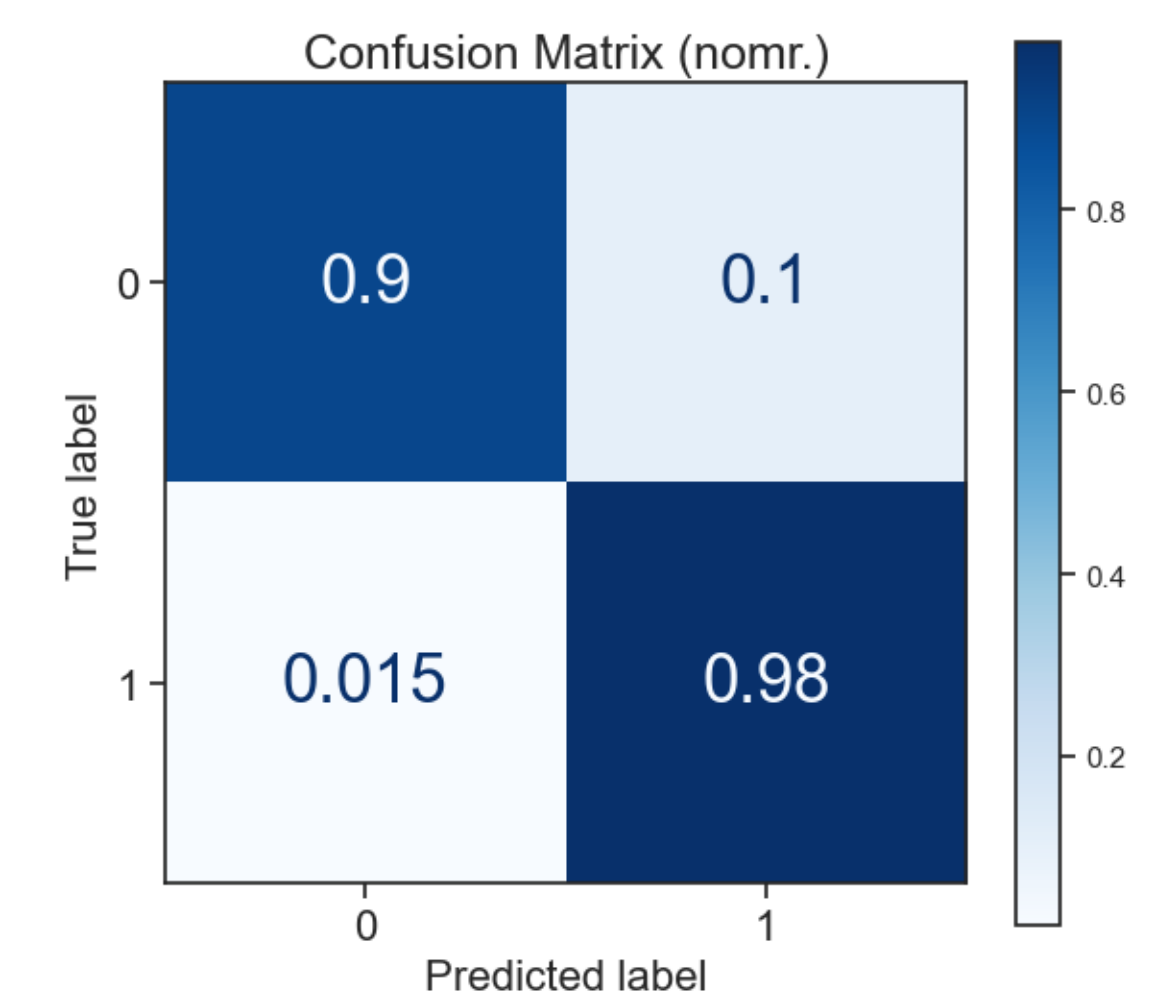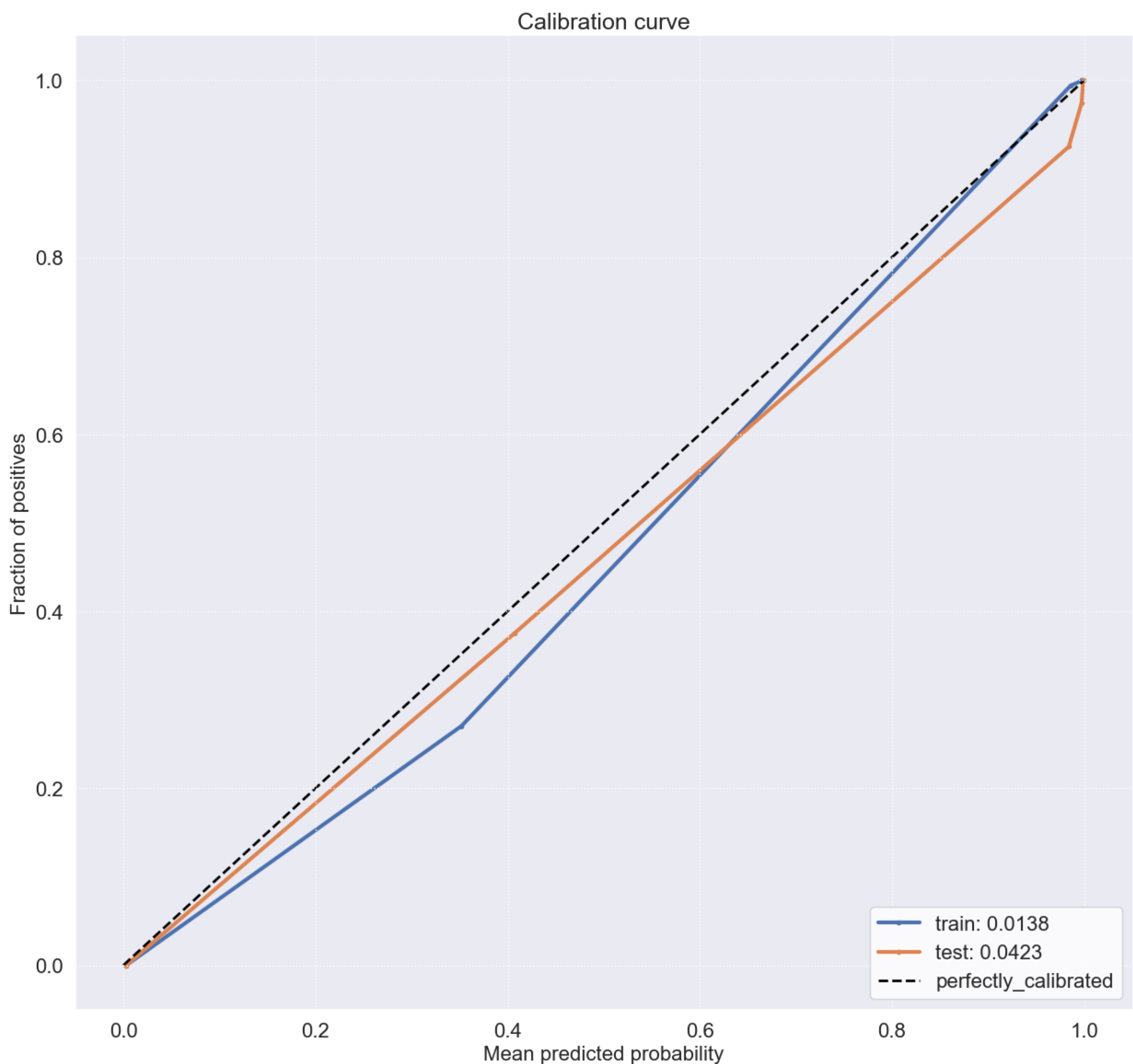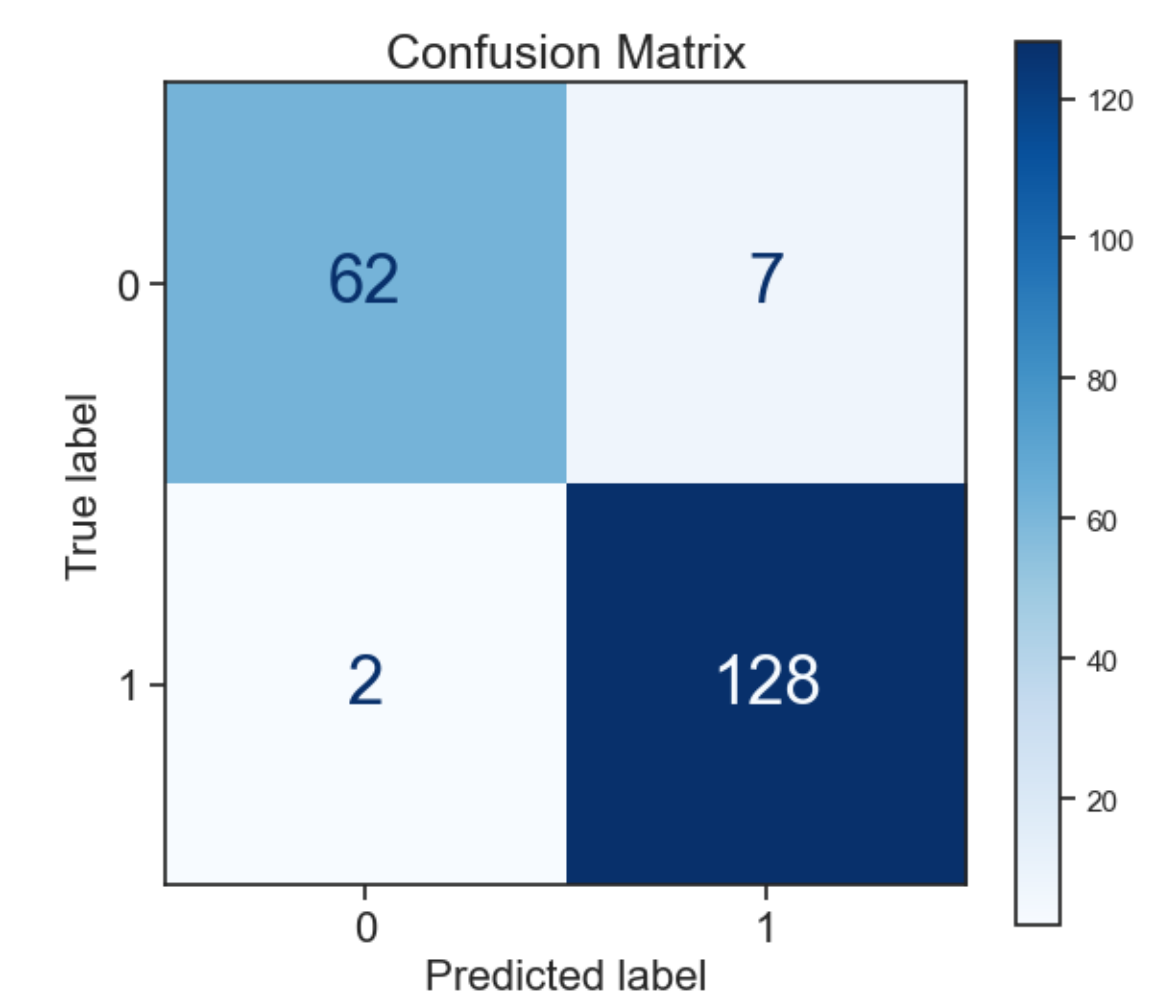
Metric	Train	Test
AUC	0.999	0.9763
accuracy	0.9811	0.9548
balanced accuracy	0.9728	0.9416
recall_(sens)	1.0	0.9846
specificity	0.9457	0.8986
precision	0.9719	0.9481
f1_score	0.981	0.9544

Protein: P30556 - metric: precision

Total (train): 794

ROC curve

- AUC_train = 0.999
- AUC_test = 0.976
- Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 1878)
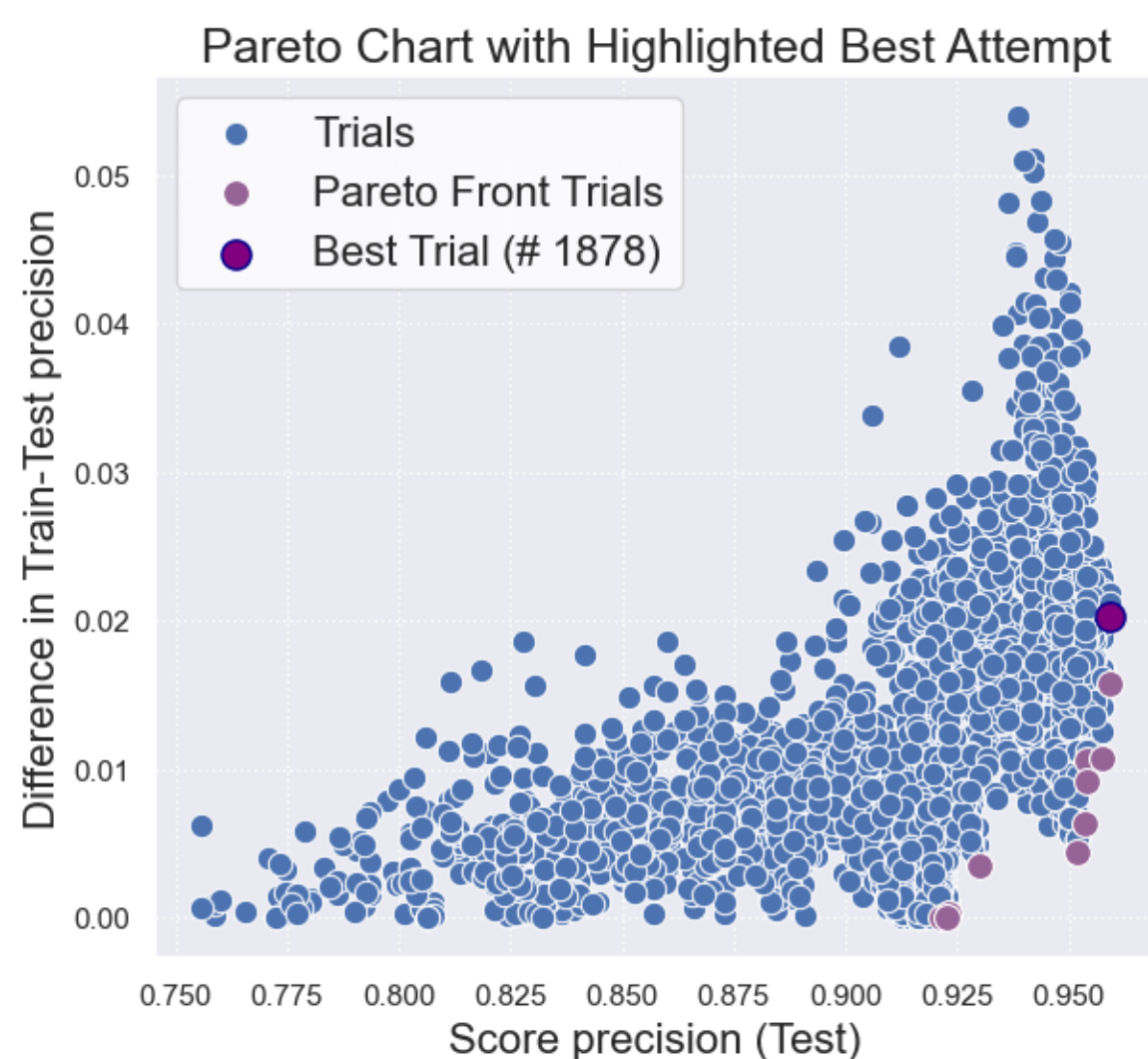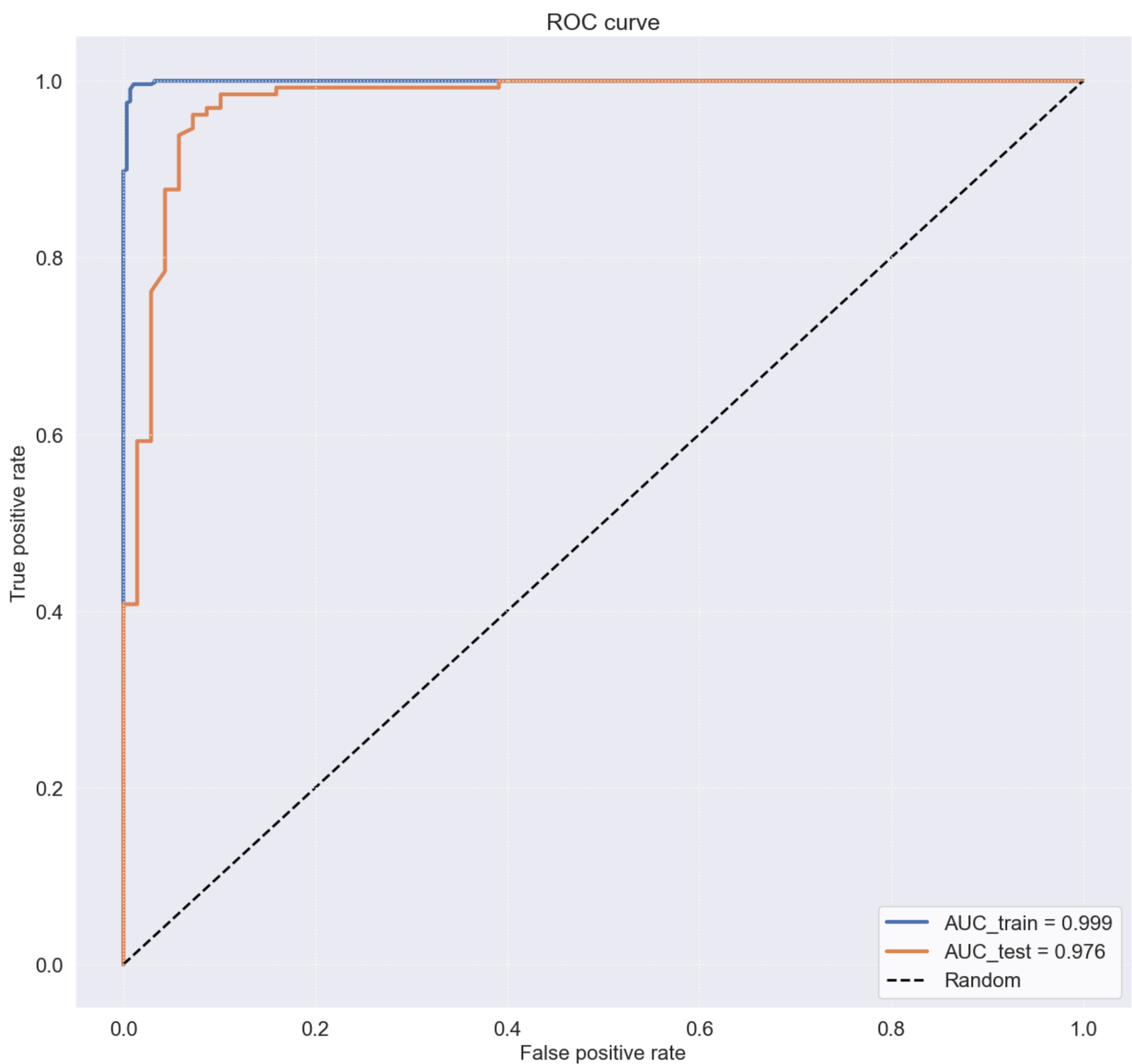
XGBoost Parameters	Value
alpha	2.4309
colsample_bytree	0.5427
gamma	0.1158
lambda	0.3695
learning_rate	0.1247
max_depth	10.0
min_child_weight	3.0
n_estimators	157.0
reg_alpha	0.0054
reg_lambda	0.0535
scale_pos_weight	2.7623
subsample	0.9357

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

- train: 0.0107
- test: 0.043
- perfectly_calibrated

Active/Inactive probability

Inactive zone | Active zone

type
- active (130)
- inactive (69)
- decoy (199)
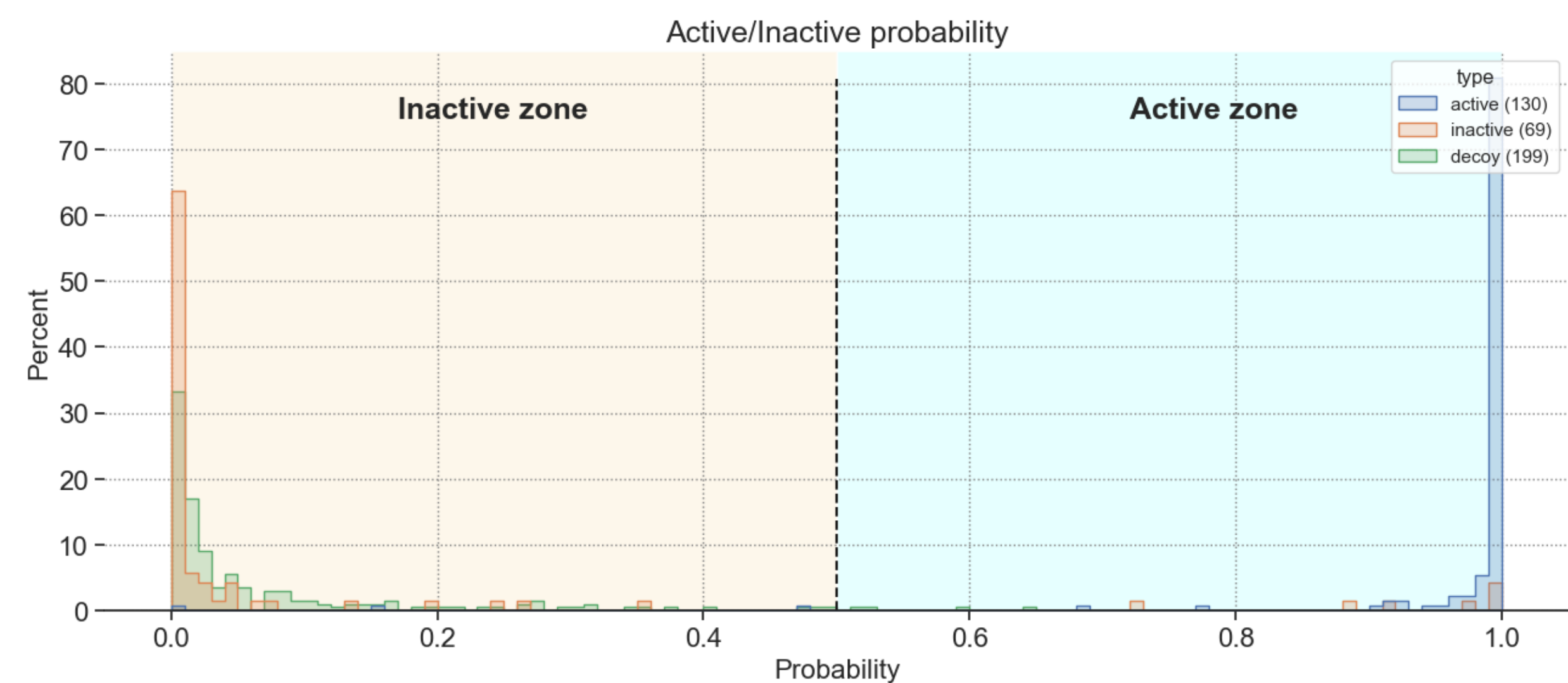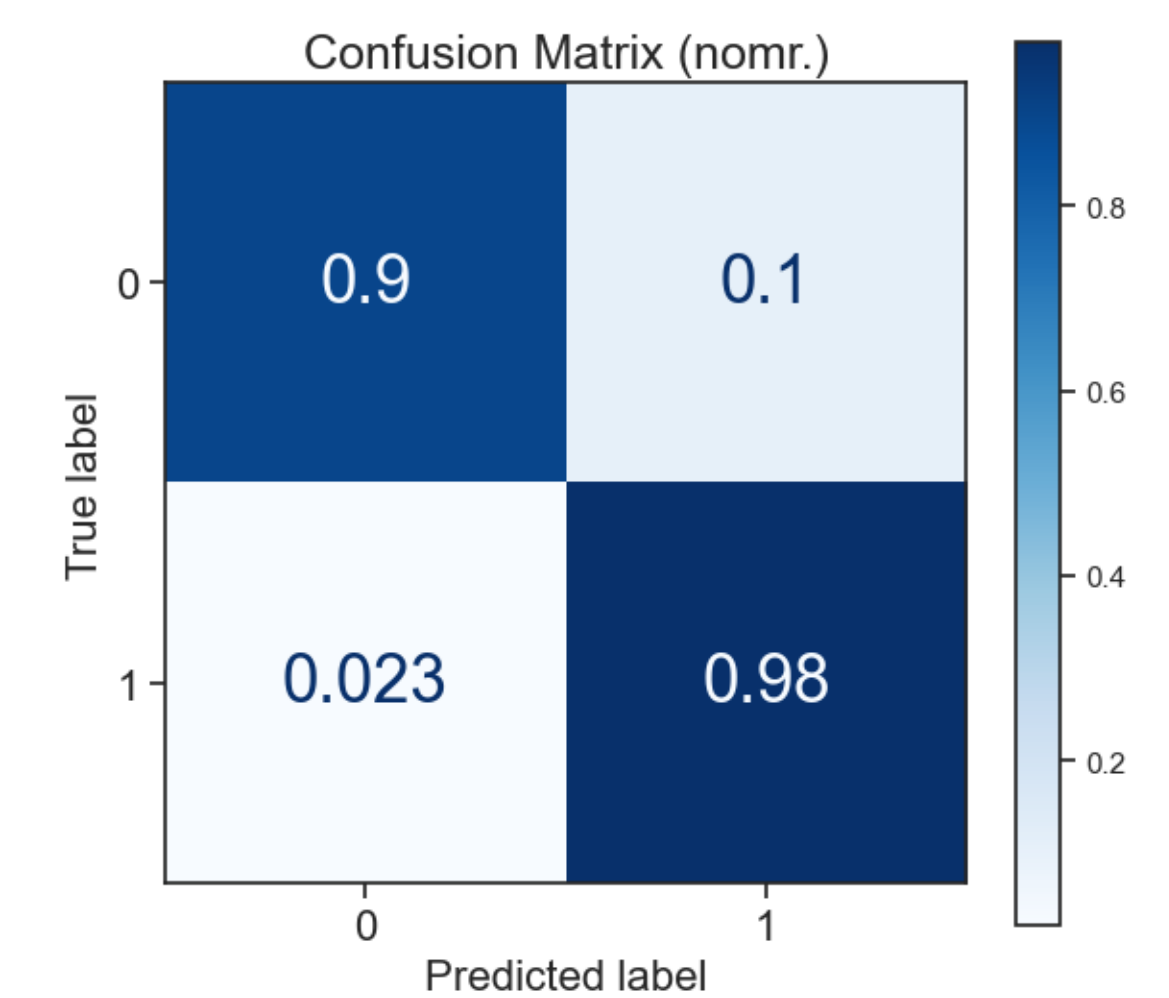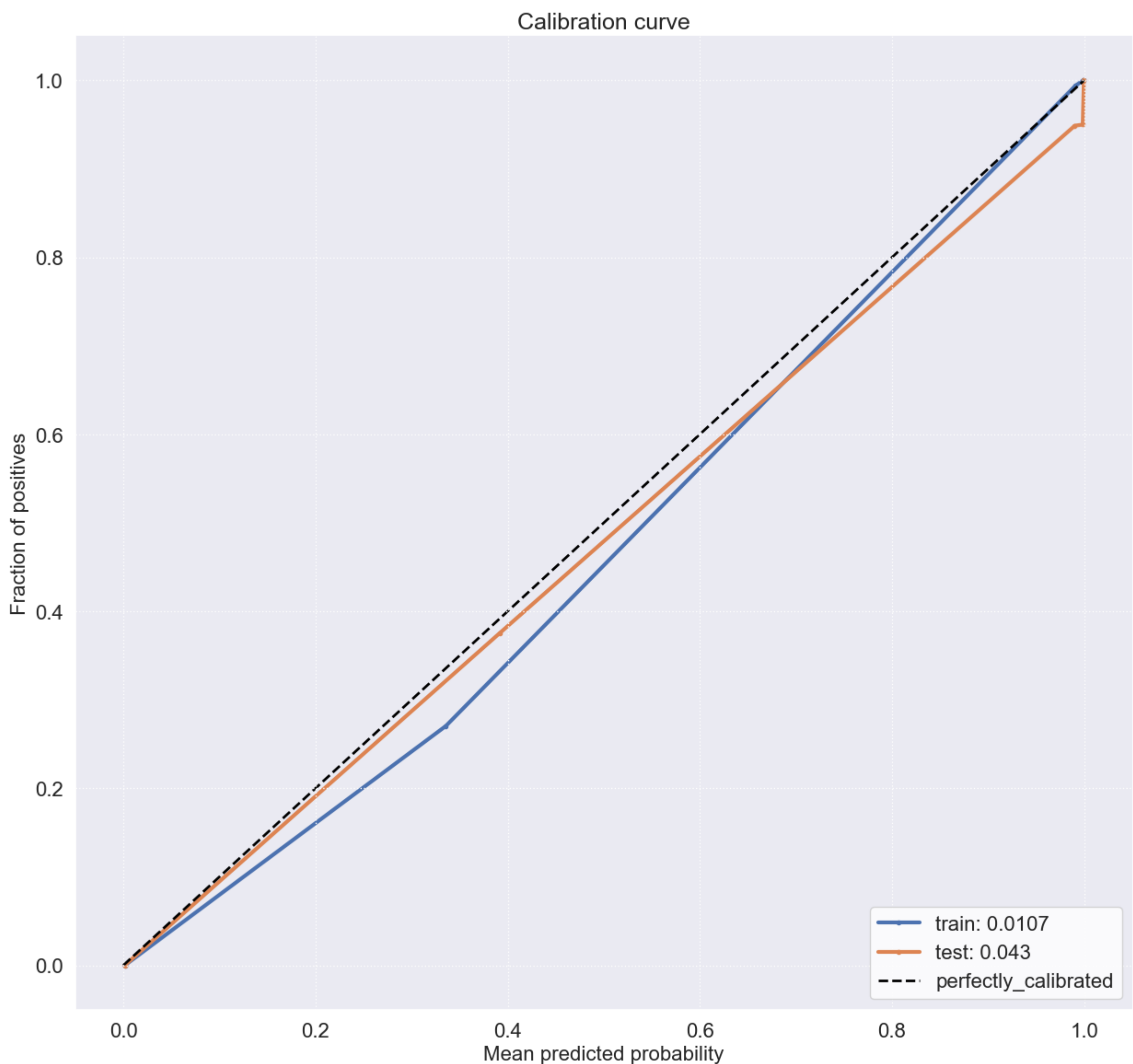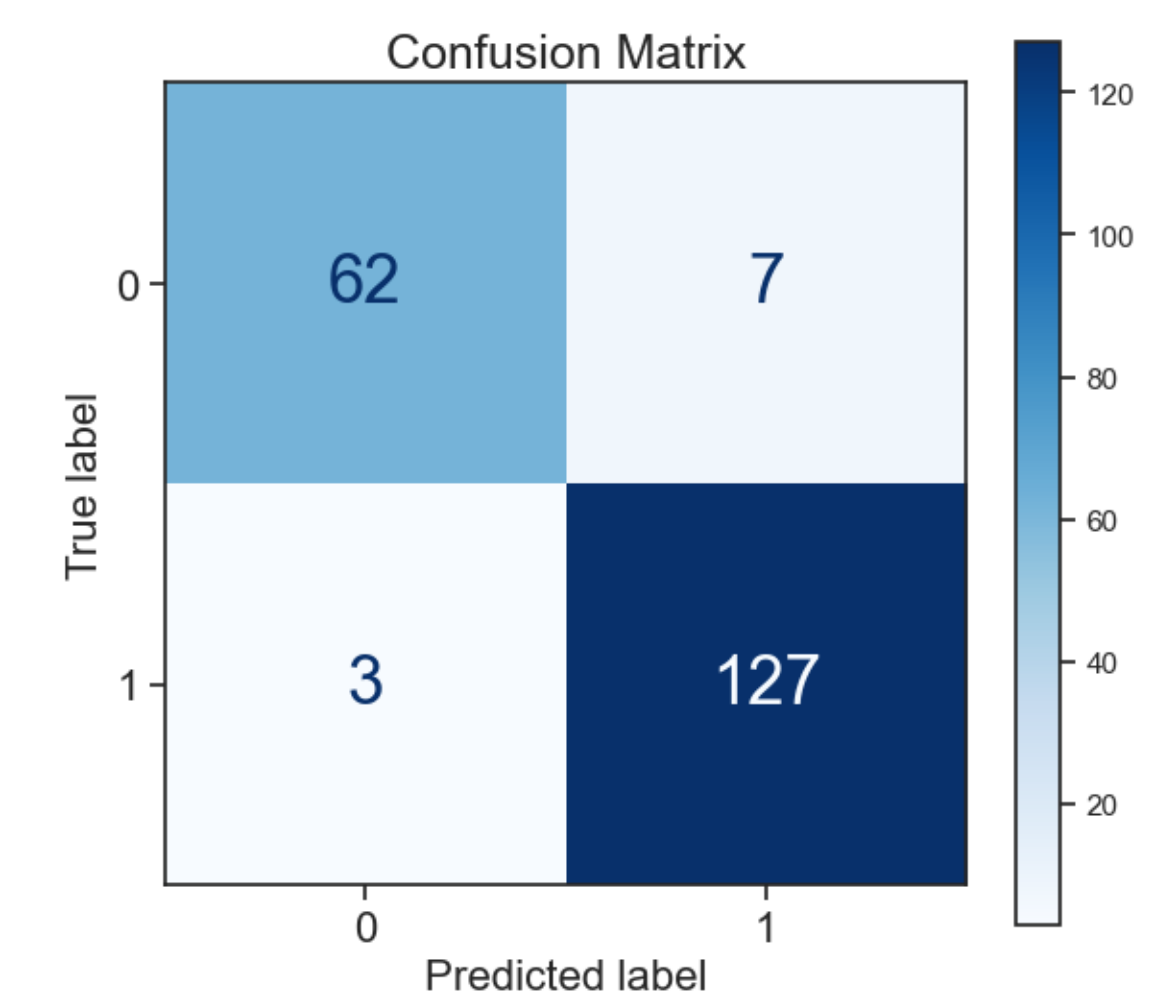
Metric	Train	Test
AUC	0.9994	0.9759
accuracy	0.9849	0.9497
balanced accuracy	0.9783	0.9377
recall_(sens)	1.0	0.9769
specificity	0.9565	0.8986
precision	0.9774	0.9478
f1_score	0.9848	0.9494

Protein: P30556 - metric: SoftVotingXGBoostEnsemble
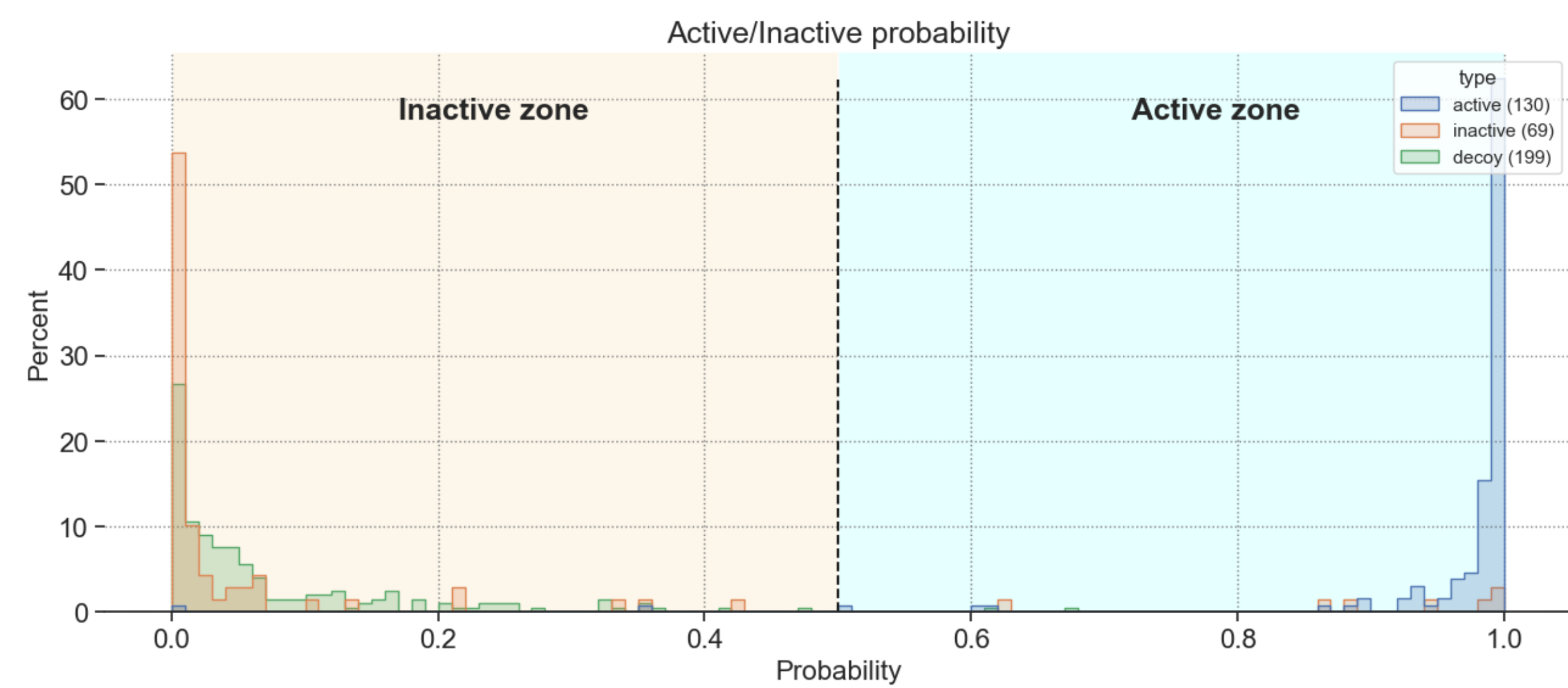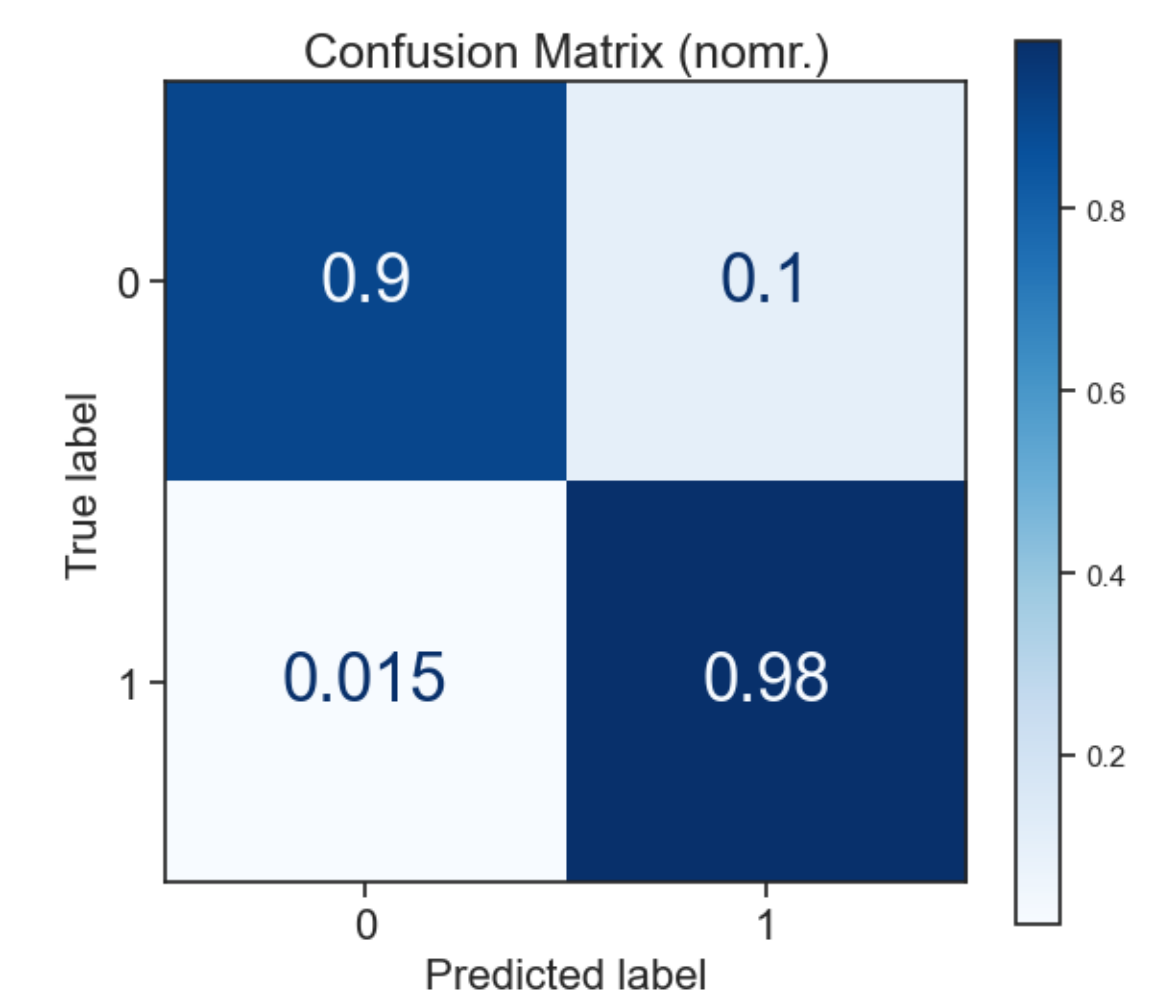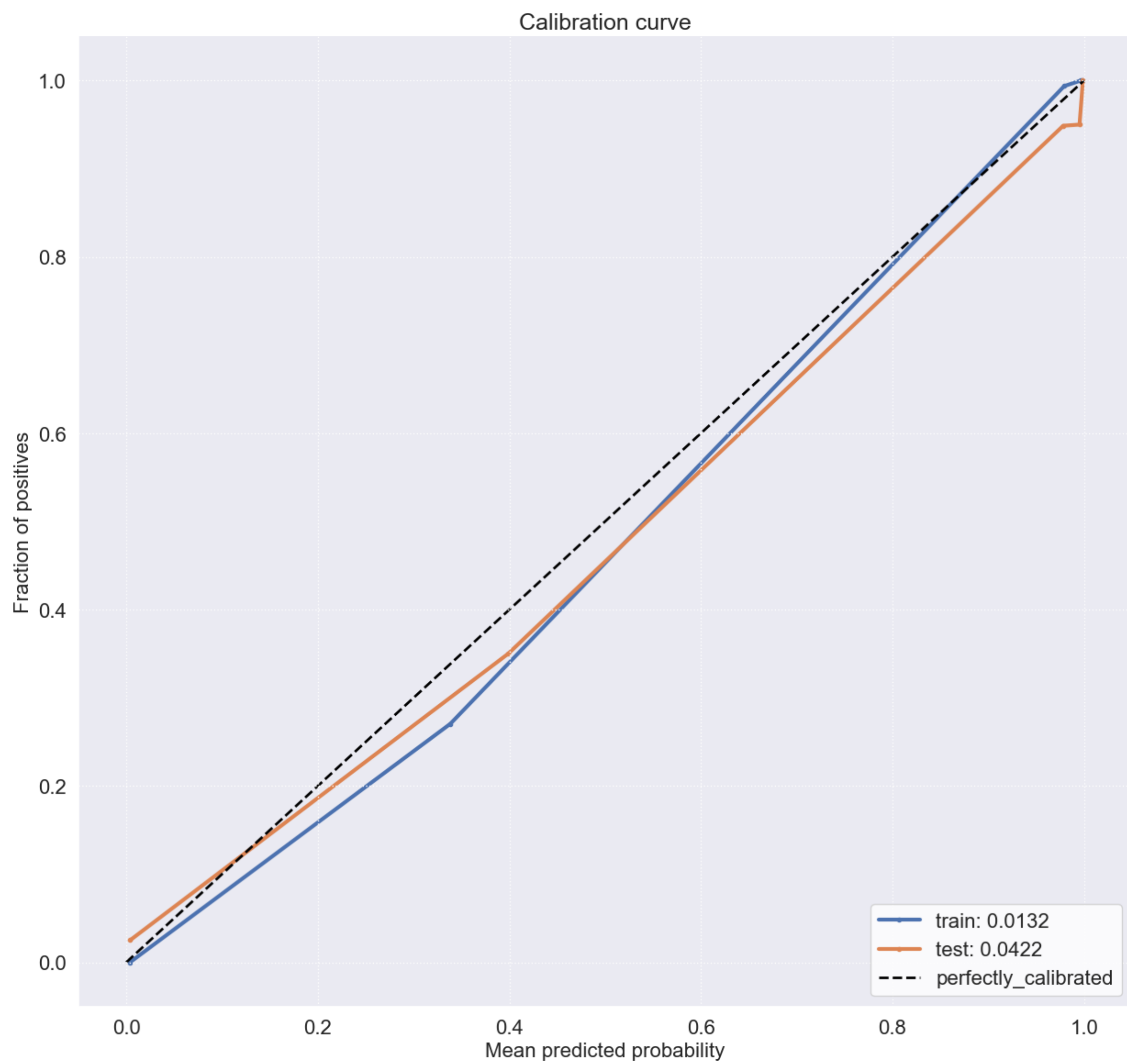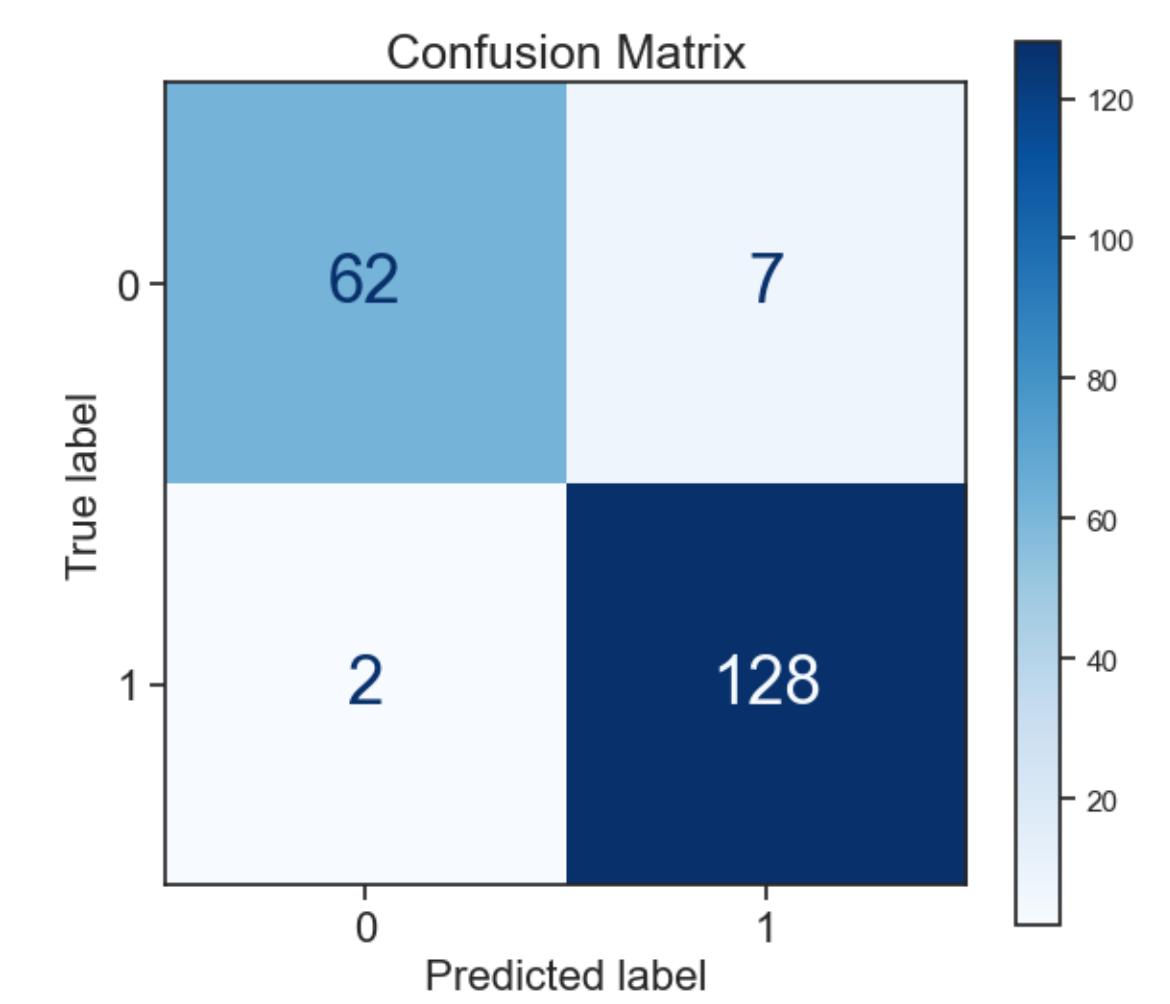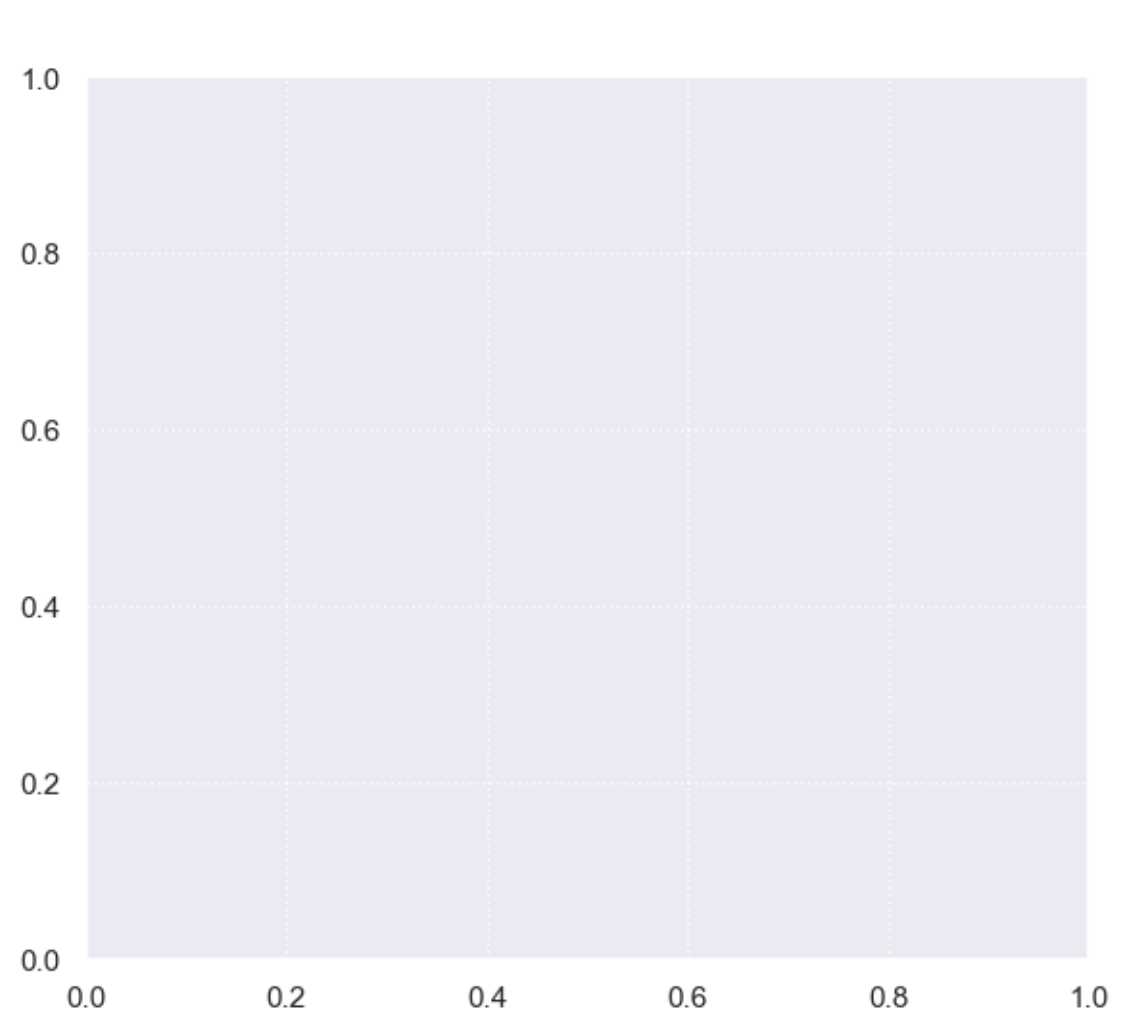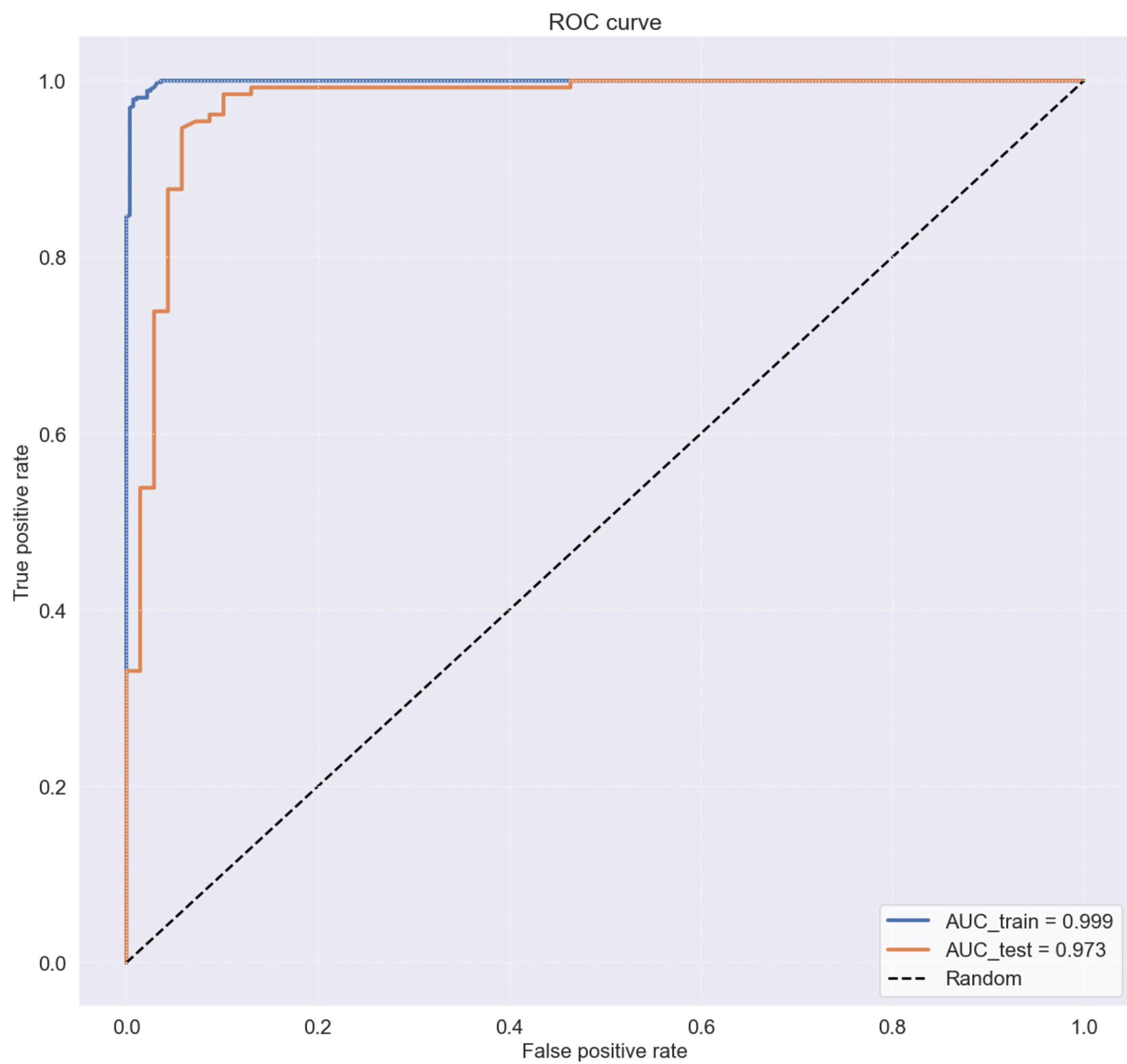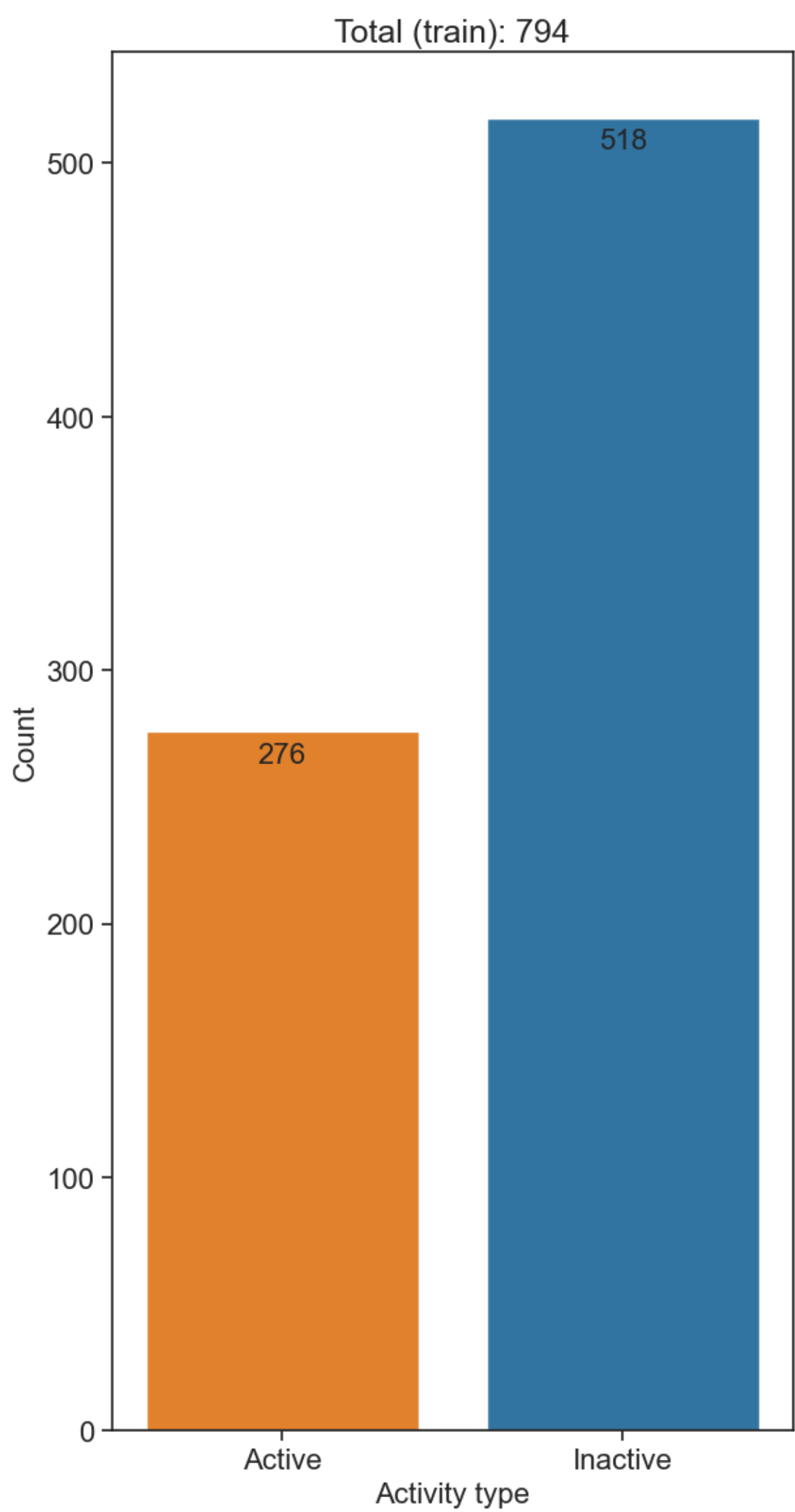
Metric	Train	Test
AUC	0.999	0.9731
accuracy	0.9836	0.9548
balanced accuracy	0.9764	0.9416
recall_(sens)	1.0	0.9846
specificity	0.9529	0.8986
precision	0.9755	0.9481
f1_score	0.9835	0.9544