Protein: P05067 - metric: balanced_accuracy

Total (train): 859

ROC curve
- AUC_train = 1.000
- AUC_test = 0.970
- Random

Pareto Chart with Highlighted Best Attempt
- Trials
- Pareto Front Trials
- Best Trial (# 1866)

Confusion Matrix

	0	1
0	47	11
1	3	154

Confusion Matrix (nomr.)

	0	1
0	0.81	0.19
1	0.019	0.98

Calibration curve
- train: 0.0058
- test: 0.0497
- perfectly_calibrated

Active/Inactive probability
- Inactive zone
- Active zone
- active (157)
- inactive (58)
- decoy (215)

XGBoost Parameters	Value
alpha	0.263
colsample_bytree	0.9724
gamma	0.403
lambda	4.6353
learning_rate	0.3816
max_depth	8.0
min_child_weight	2.0
n_estimators	802.0
reg_alpha	0.4185
reg_lambda	0.2296
scale_pos_weight	2.6048
subsample	0.8684
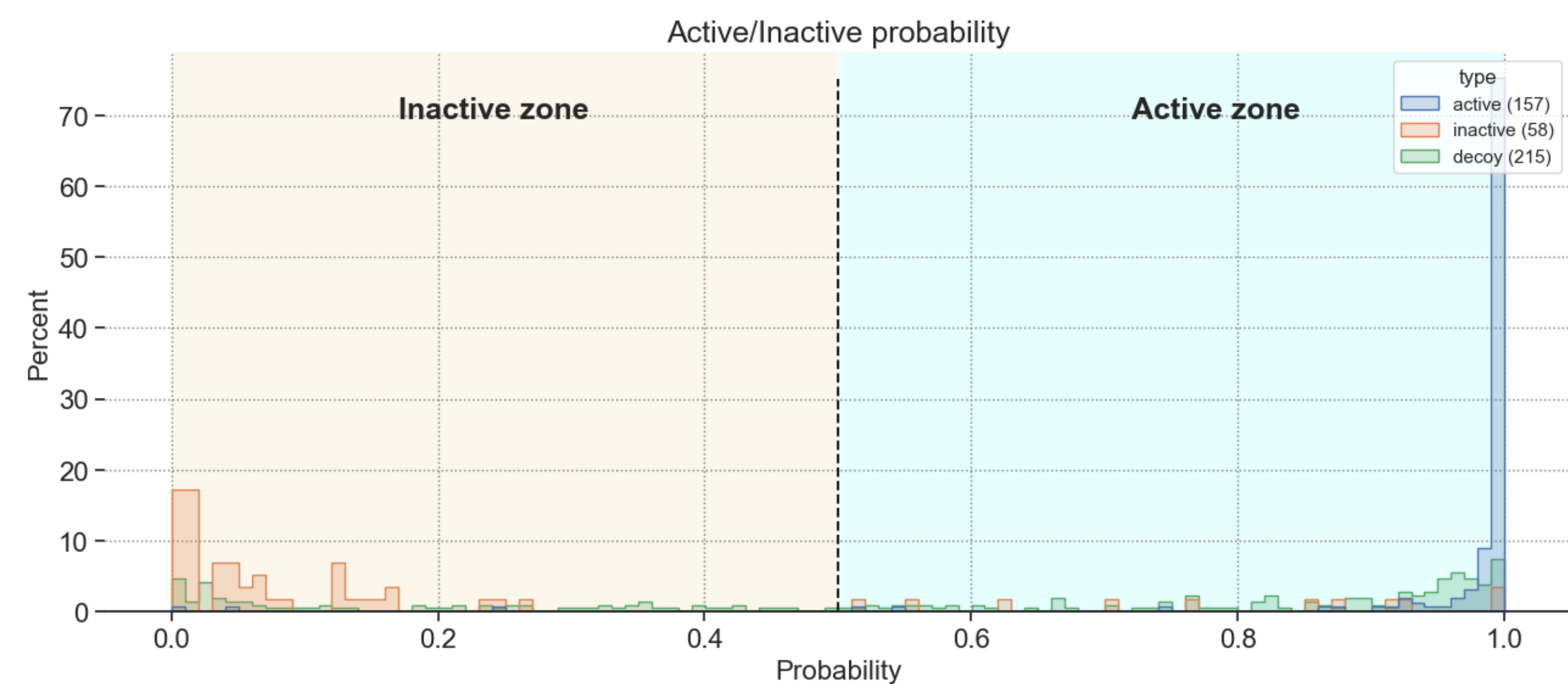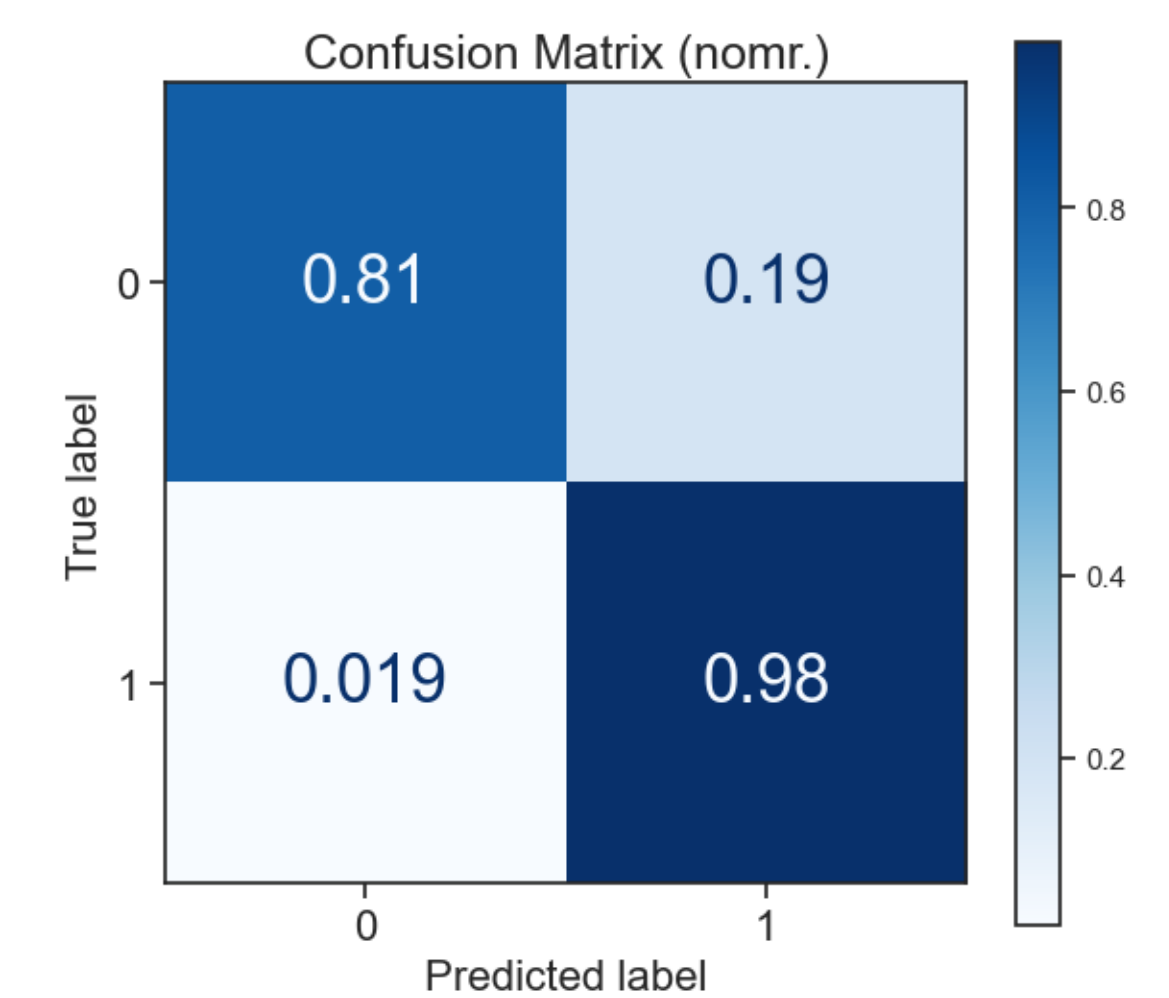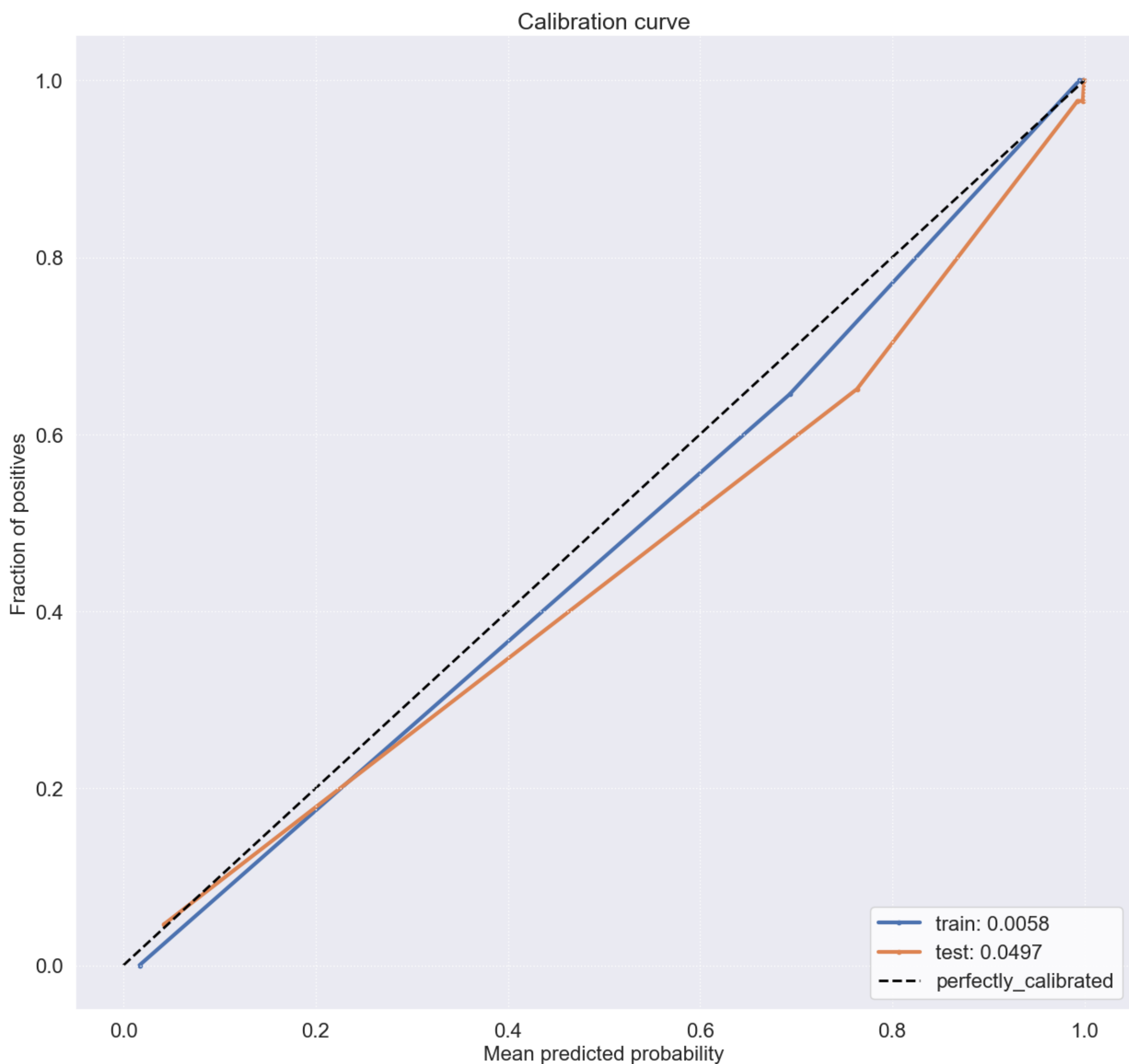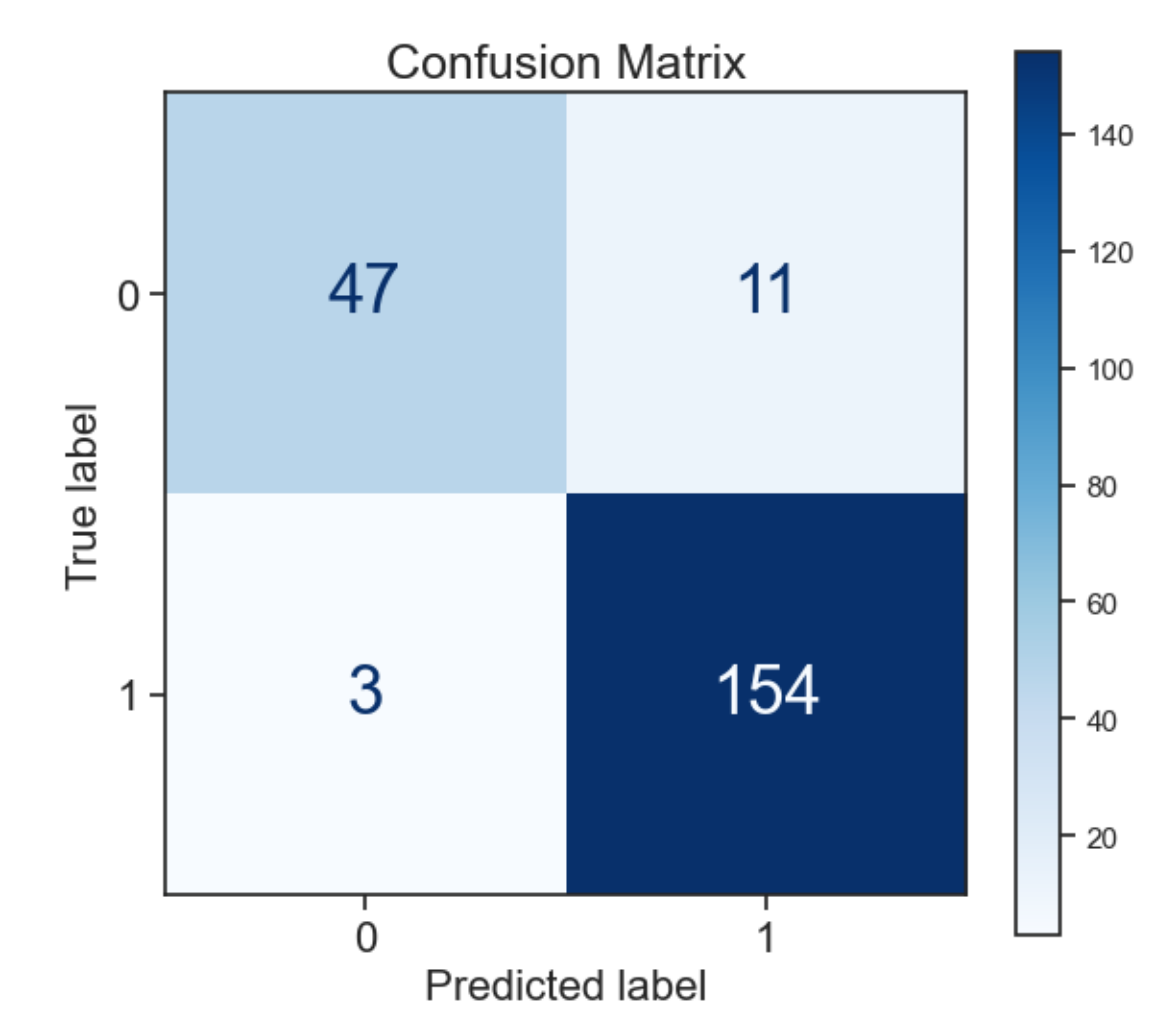
Metric	Train	Test
AUC	1.0	0.9697
accuracy	0.9965	0.9349
balanced accuracy	0.9936	0.8956
recall_(sens)	1.0	0.9809
specificity	0.9871	0.8103
precision	0.9952	0.9333
f1_score	0.9965	0.9333

Protein: P05067 - metric: f1

Total (train): 859

ROC curve

AUC_train = 1.000
AUC_test = 0.969
Random

Pareto Chart with Highlighted Best Attempt

Trials
Pareto Front Trials
Best Trial (# 2665)
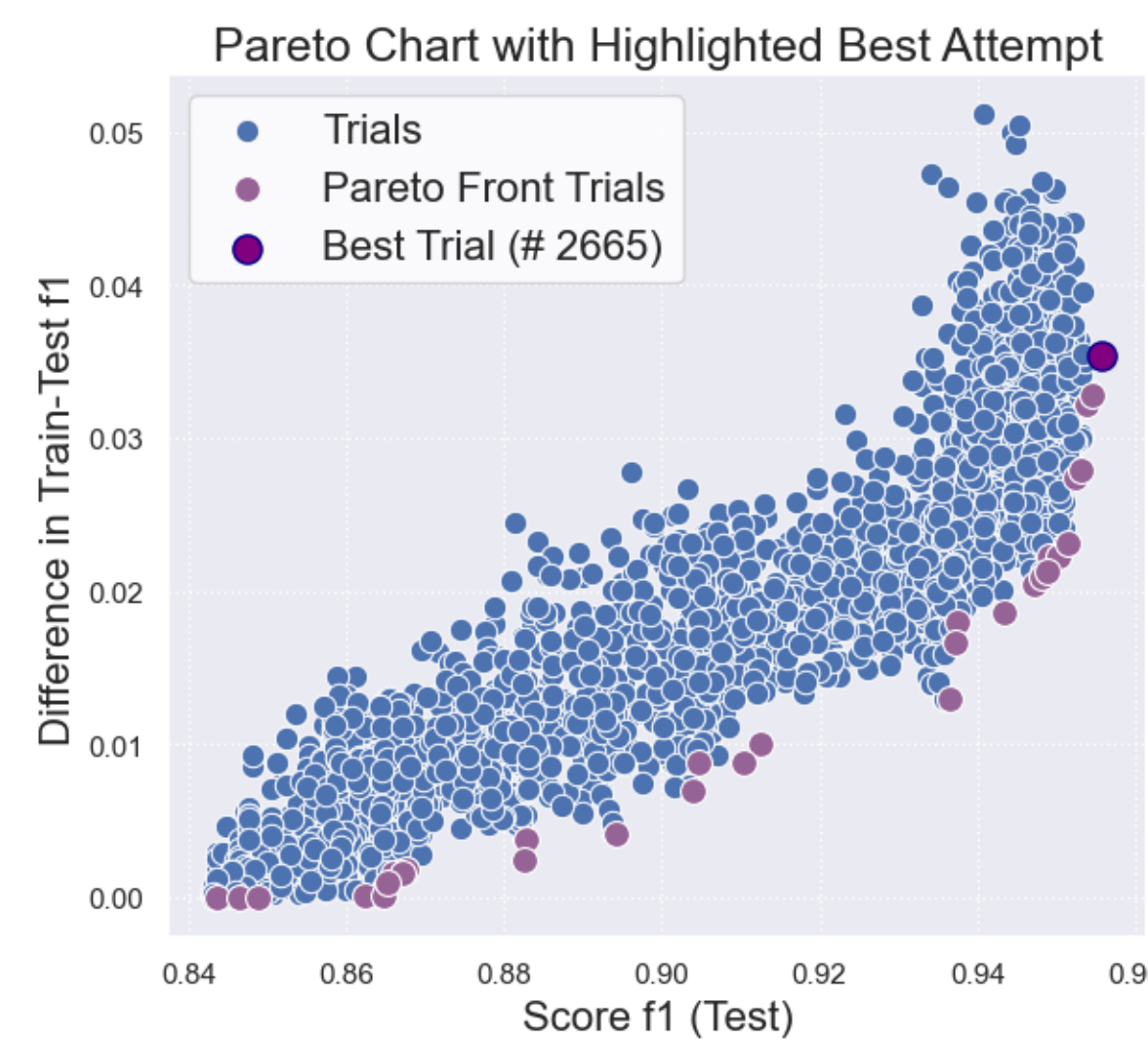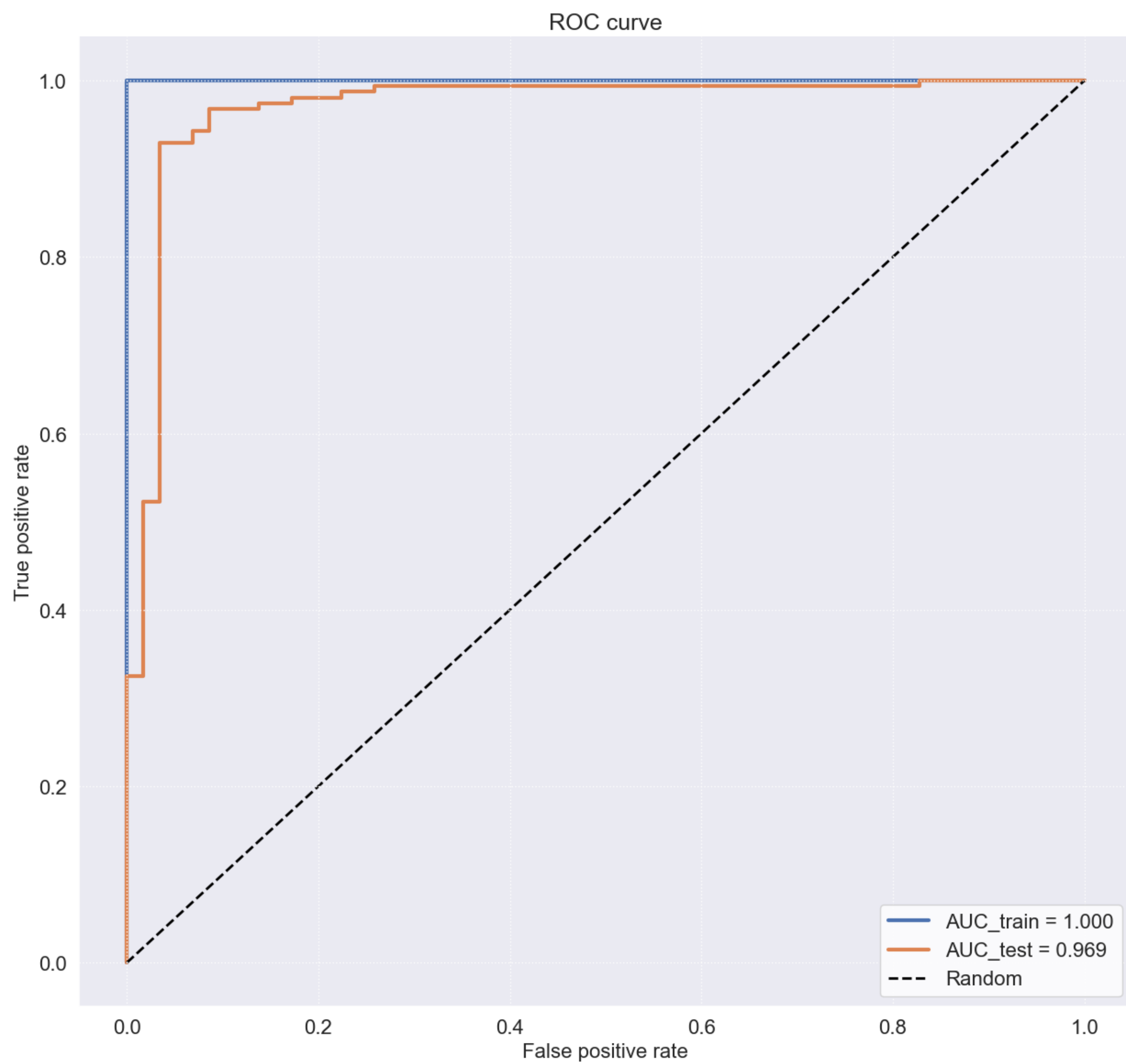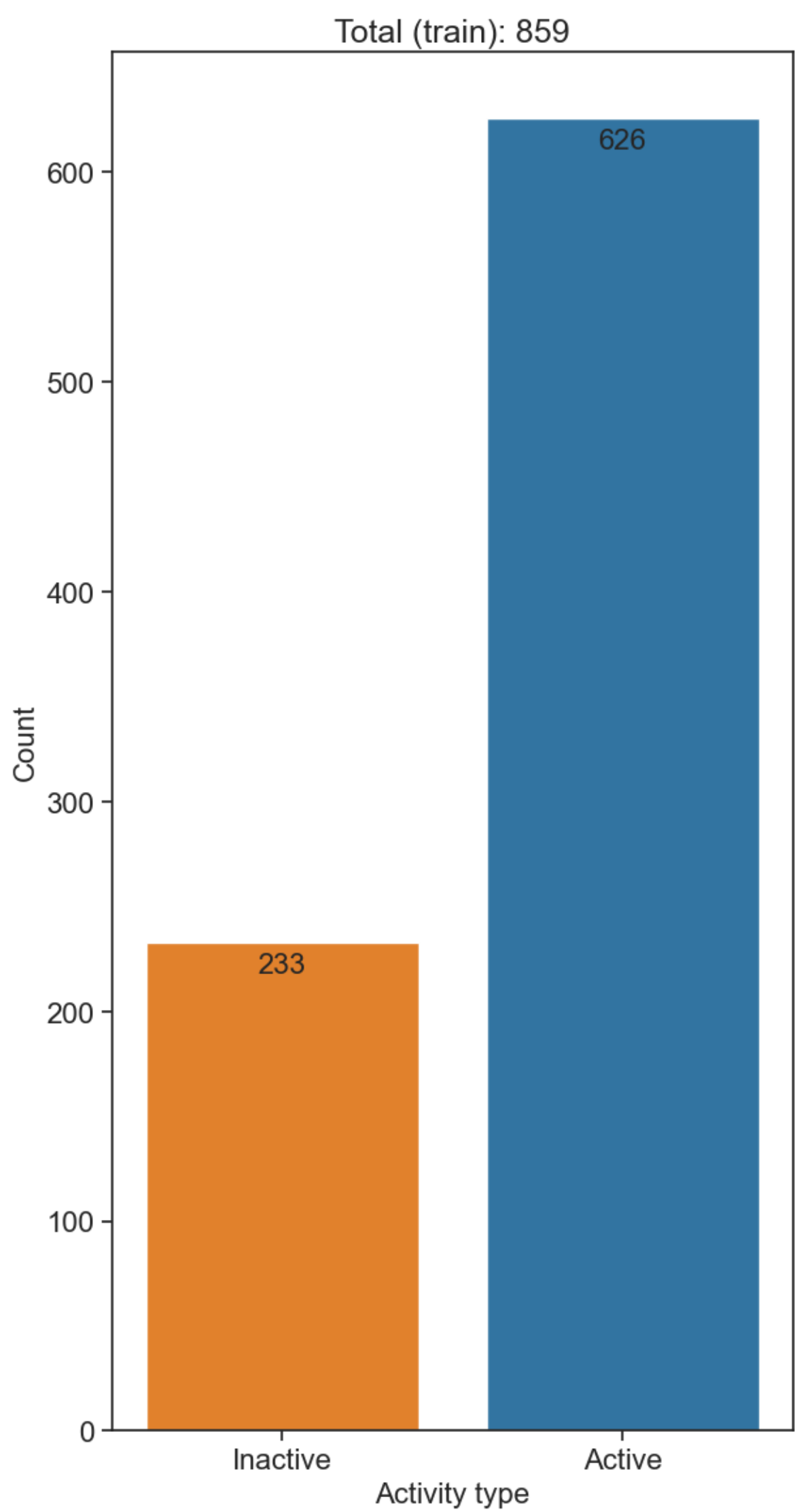
XGBoost Parameters	Value
alpha	2.4641
colsample_bytree	0.938
gamma	0.3353
lambda	3.0646
learning_rate	0.0563
max_depth	8.0
min_child_weight	3.0
n_estimators	973.0
reg_alpha	0.1368
reg_lambda	0.2764
scale_pos_weight	6.636
subsample	0.9612

Confusion Matrix

Calibration curve

train: 0.0102
test: 0.0536
perfectly_calibrated

Confusion Matrix (nomr.)

Active/Inactive probability

Inactive zone
Active zone
type
active (157)
inactive (58)
decoy (215)
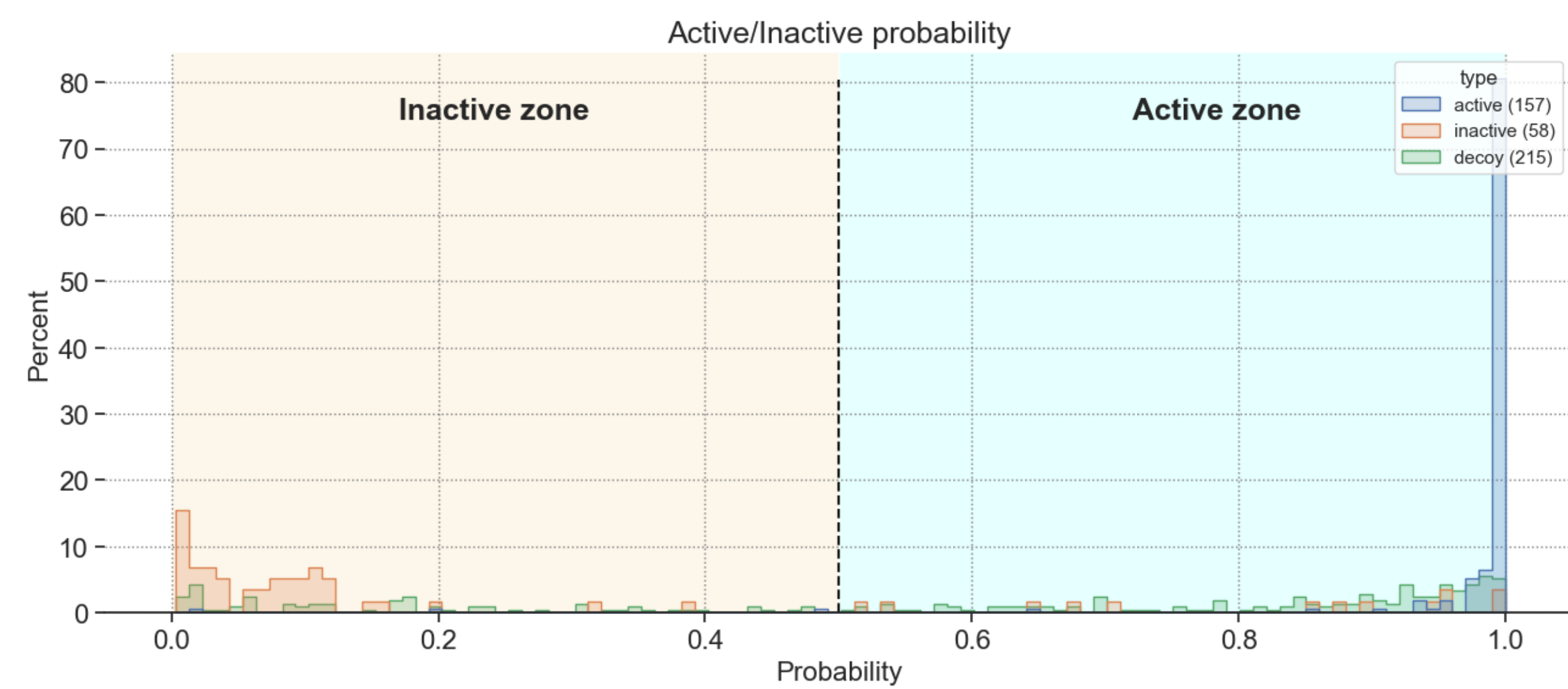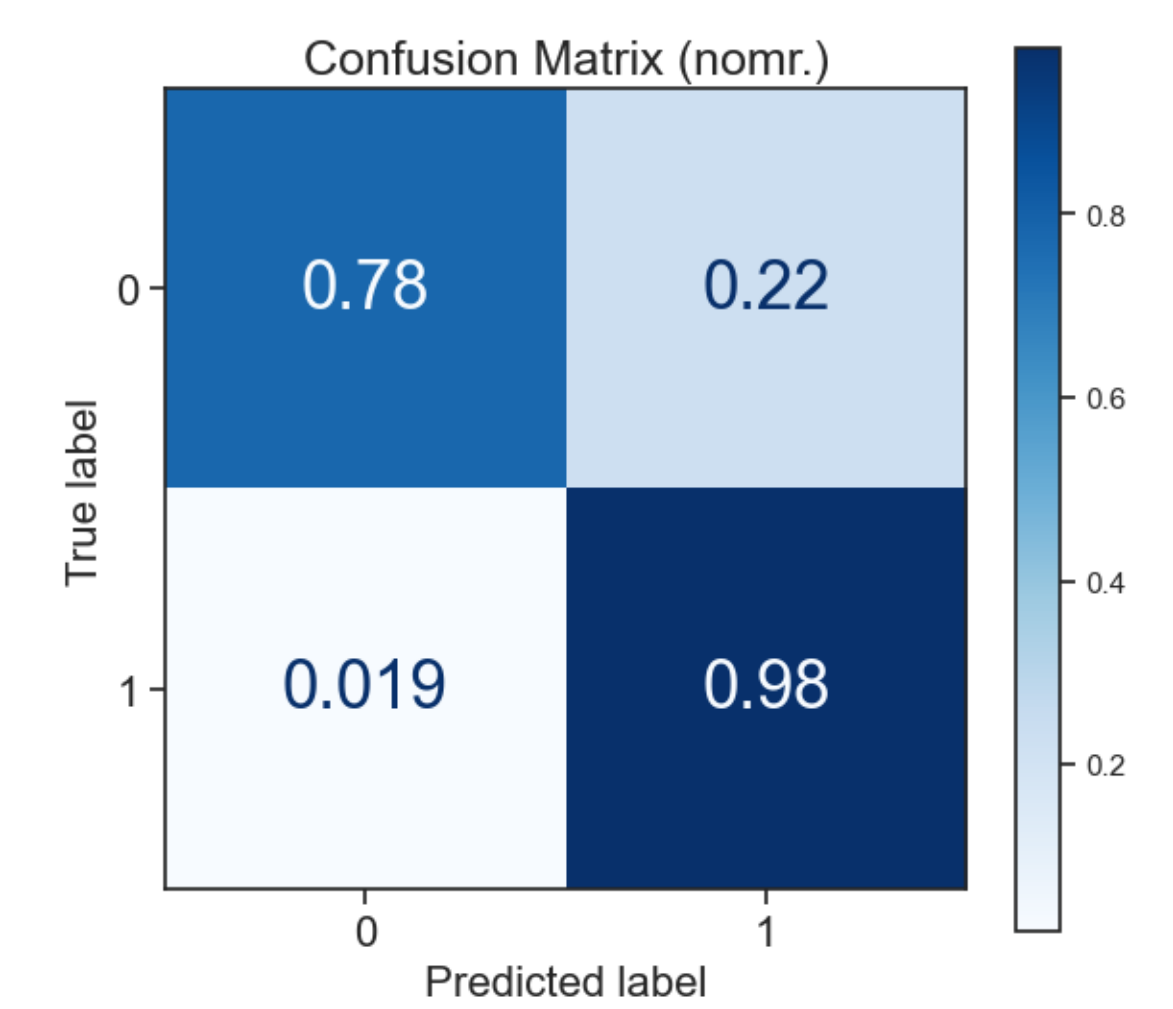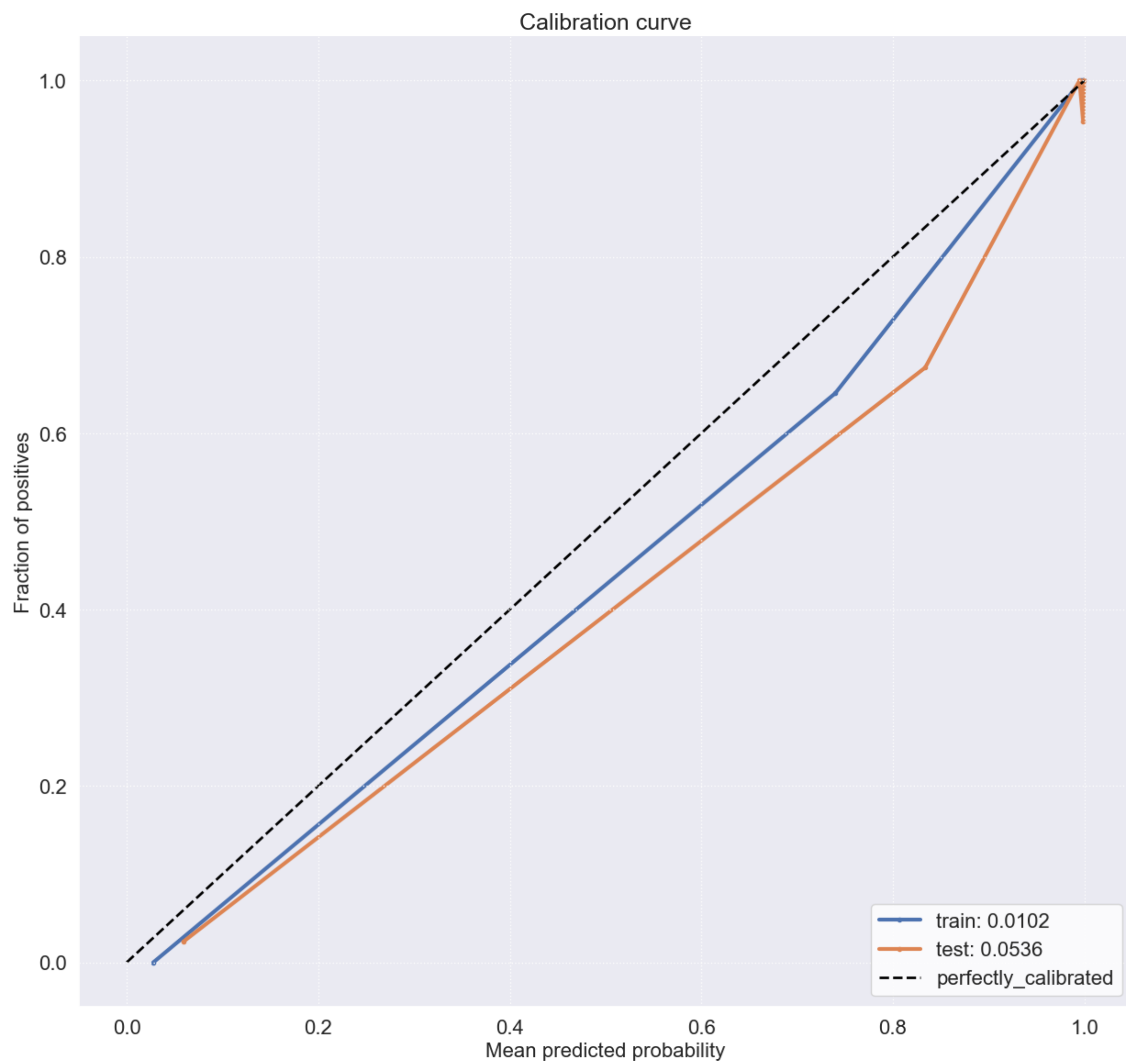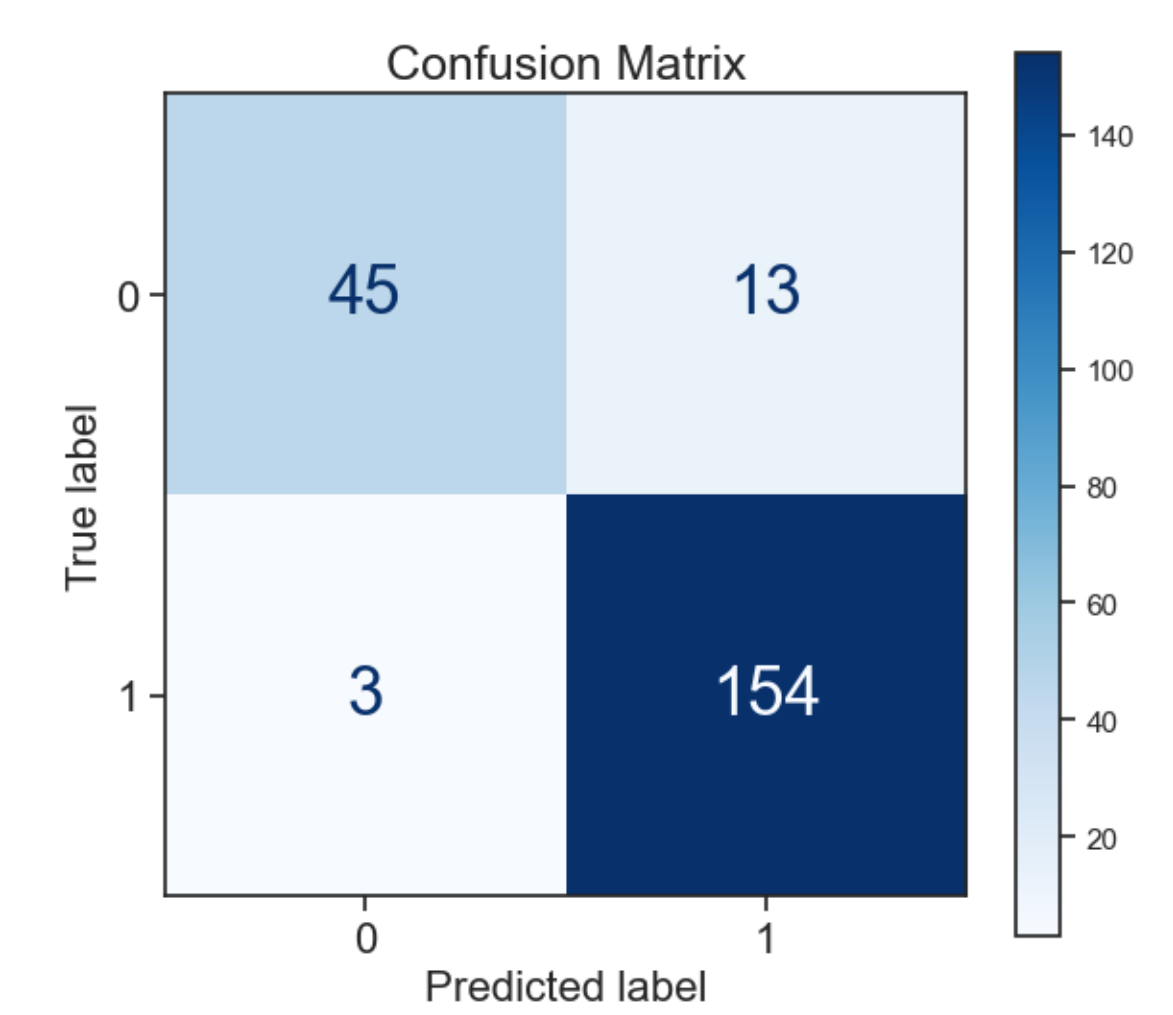
Metric	Train	Test
AUC	1.0	0.9691
accuracy	0.986	0.9256
balanced accuracy	0.9742	0.8784
recall_(sens)	1.0	0.9809
specificity	0.9485	0.7759
precision	0.9812	0.9222
f1_score	0.9859	0.9232

Protein: P05067 - metric: precision

Total (train): 859

ROC curve

	AUC_train = 1.000
	AUC_test = 0.951
	Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 2893)
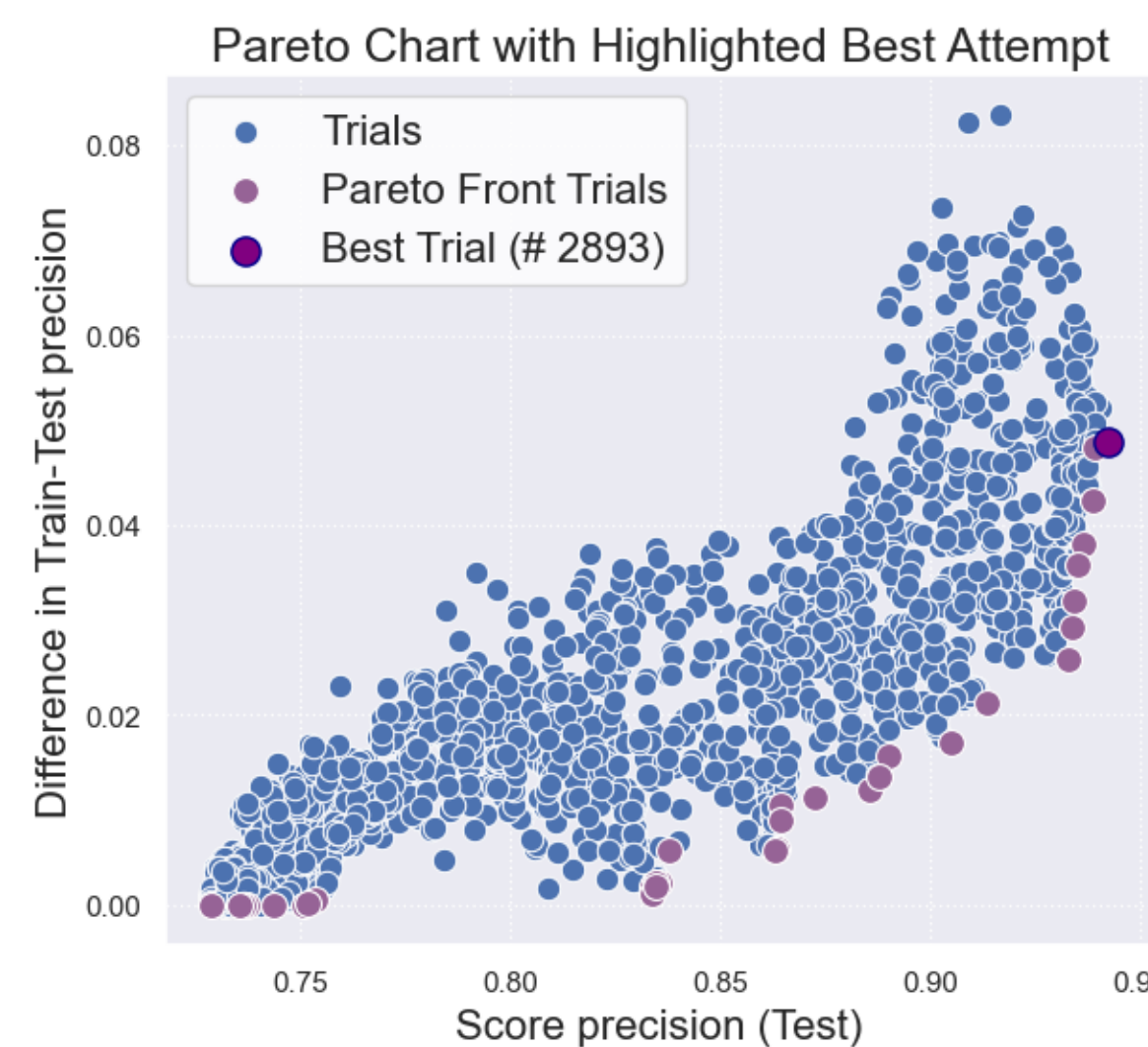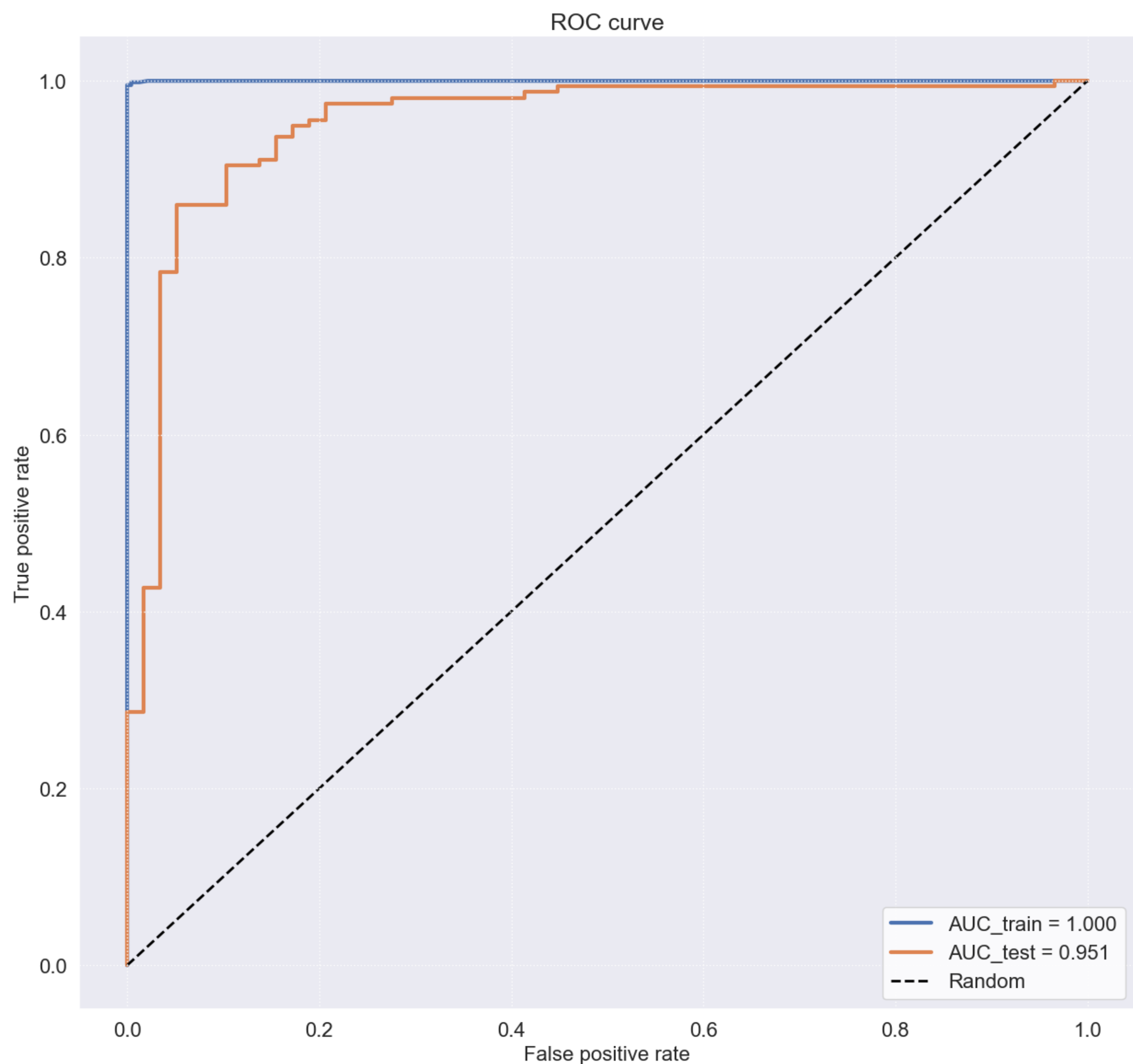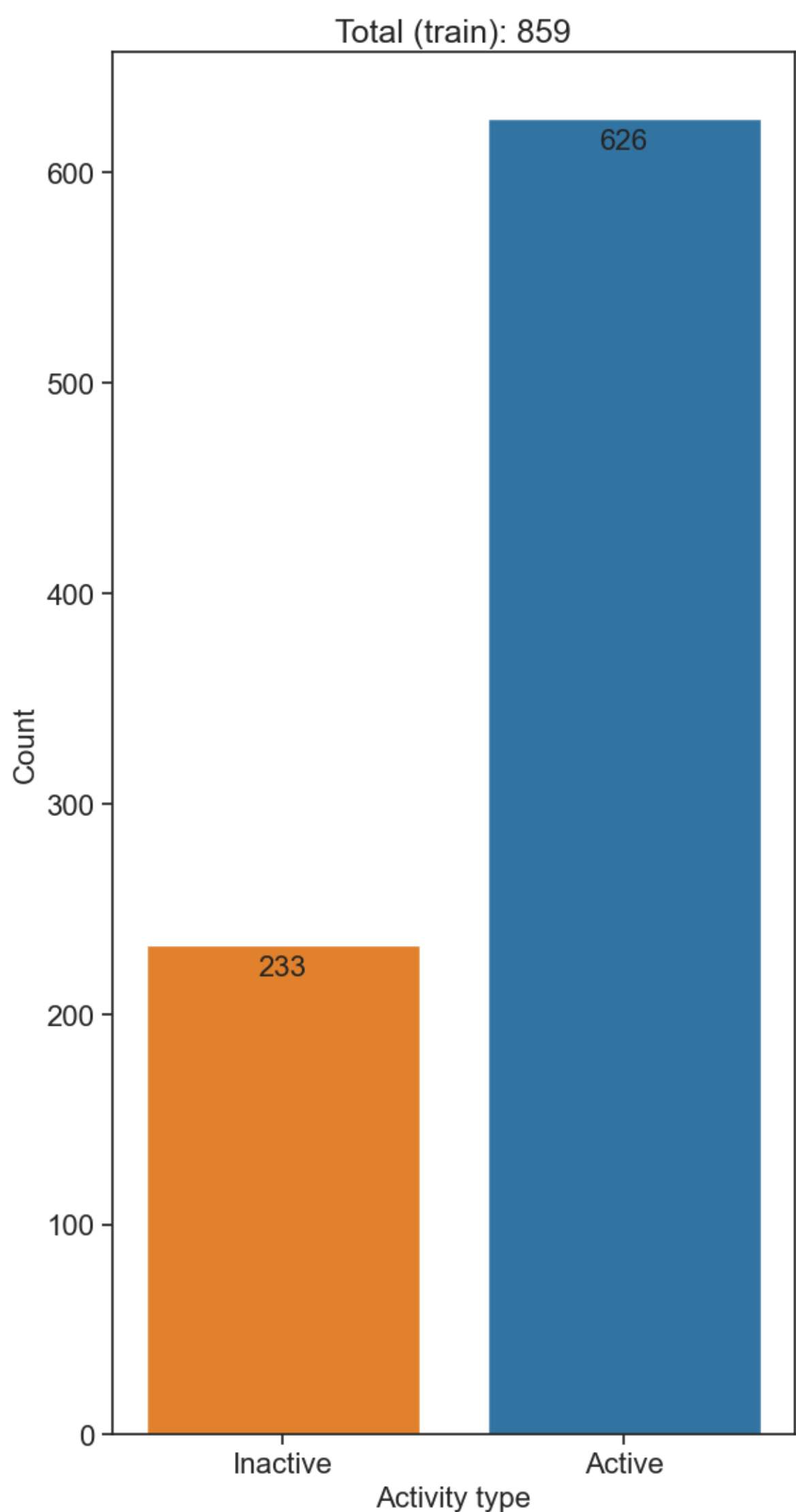
XGBoost Parameters	Value
alpha	4.9234
colsample_bytree	0.8605
gamma	0.5457
lambda	2.9011
learning_rate	0.8991
max_depth	3.0
min_child_weight	2.0
n_estimators	110.0
reg_alpha	0.3692
reg_lambda	0.251
scale_pos_weight	2.6572
subsample	0.7904

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

	train: 0.0072
	test: 0.0695
	perfectly_calibrated

Active/Inactive probability

type
active (157)
inactive (58)
decoy (215)

Inactive zone   Active zone
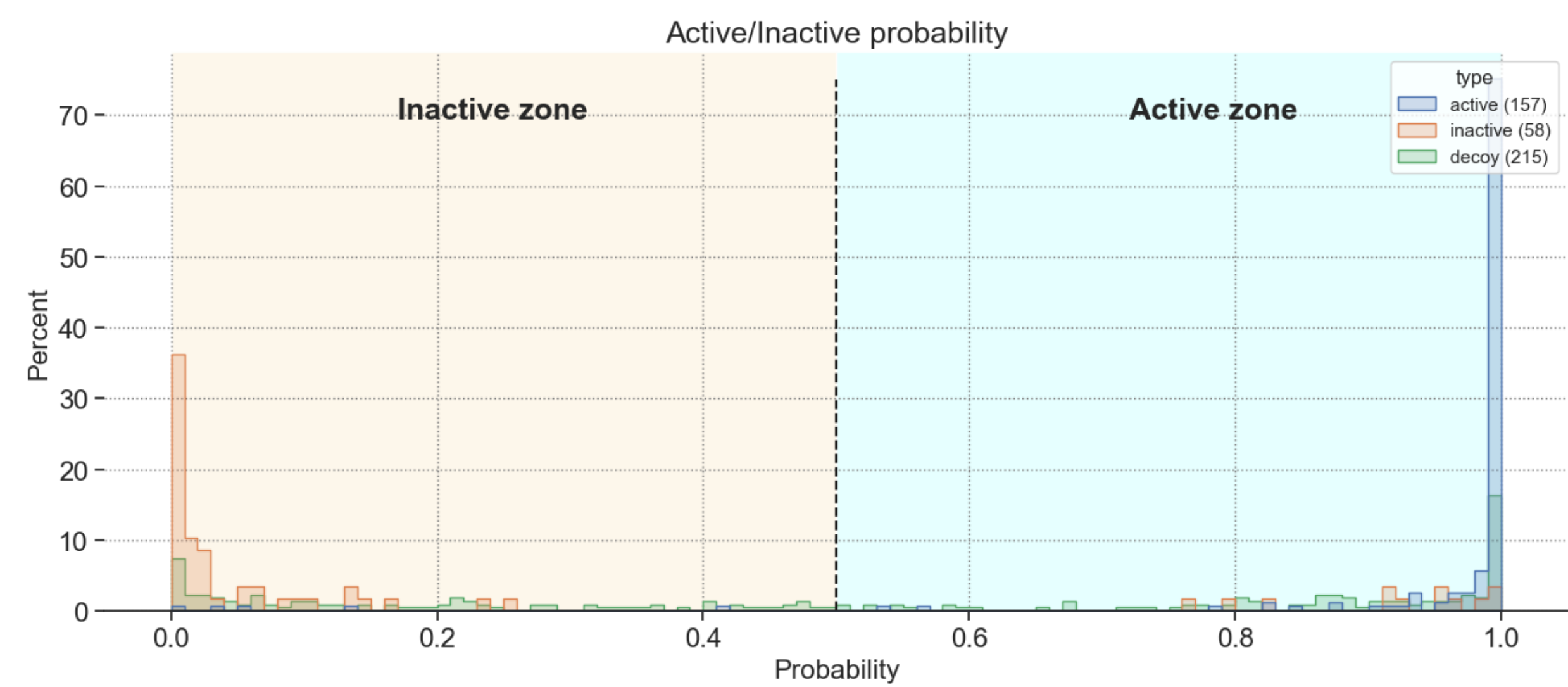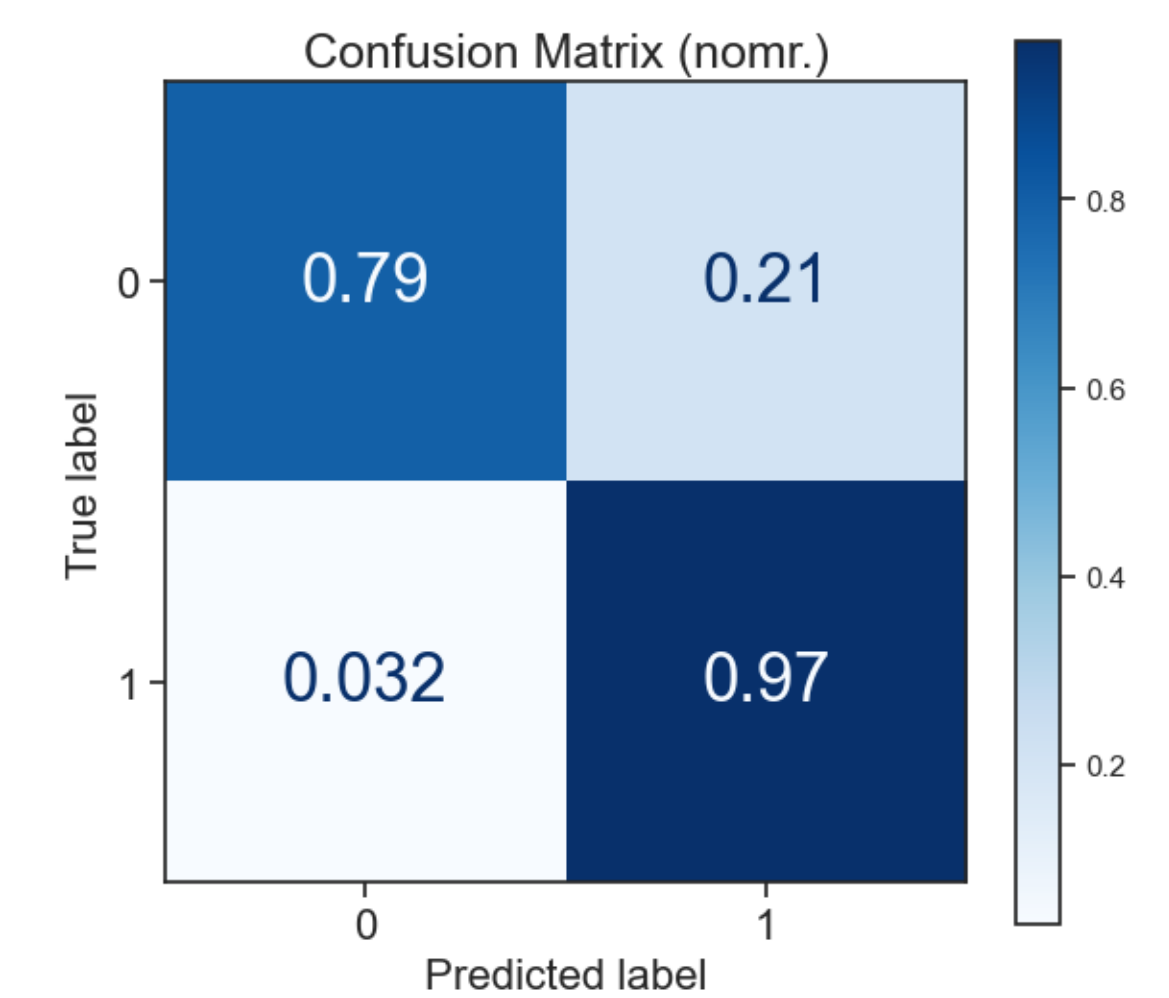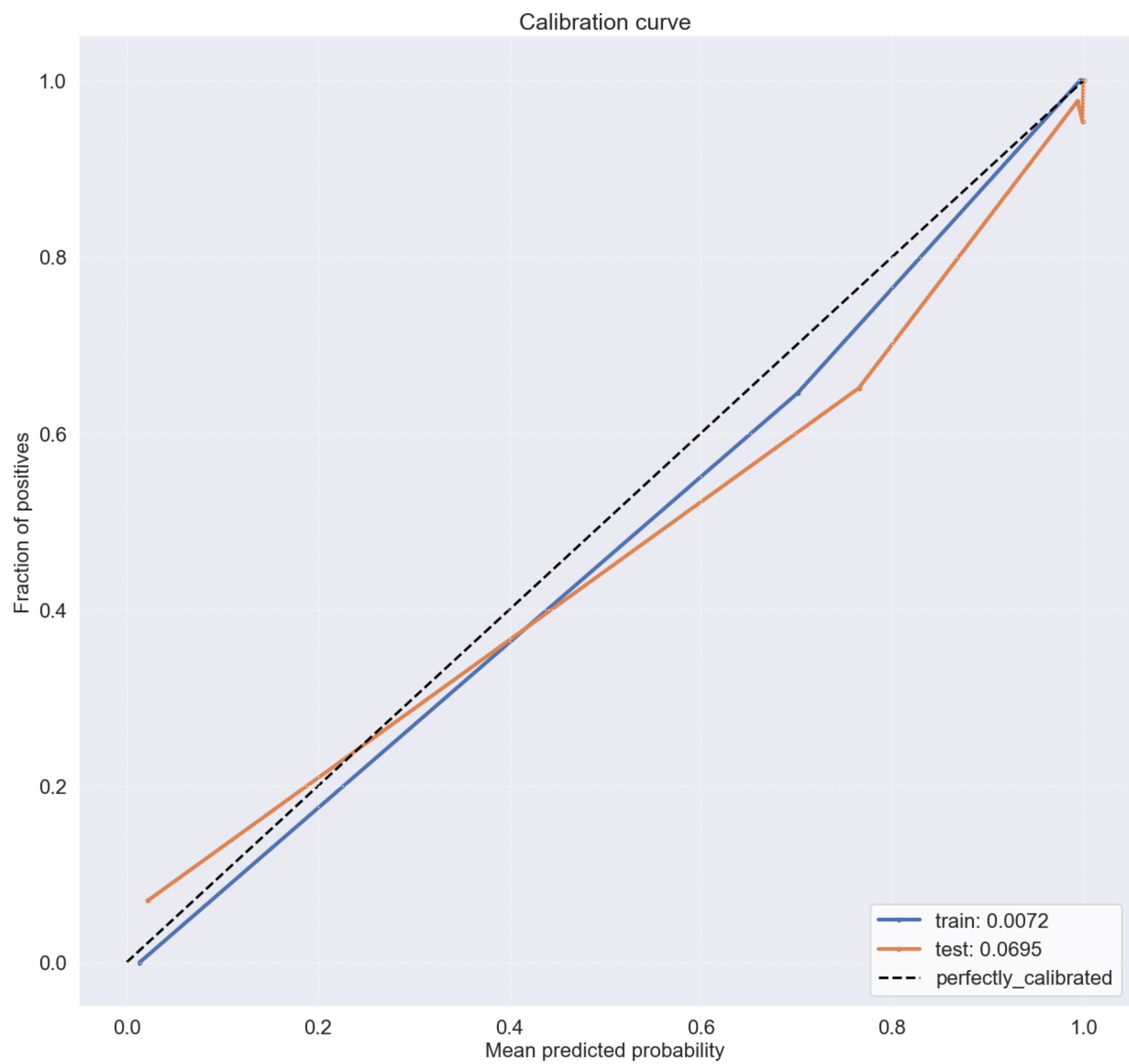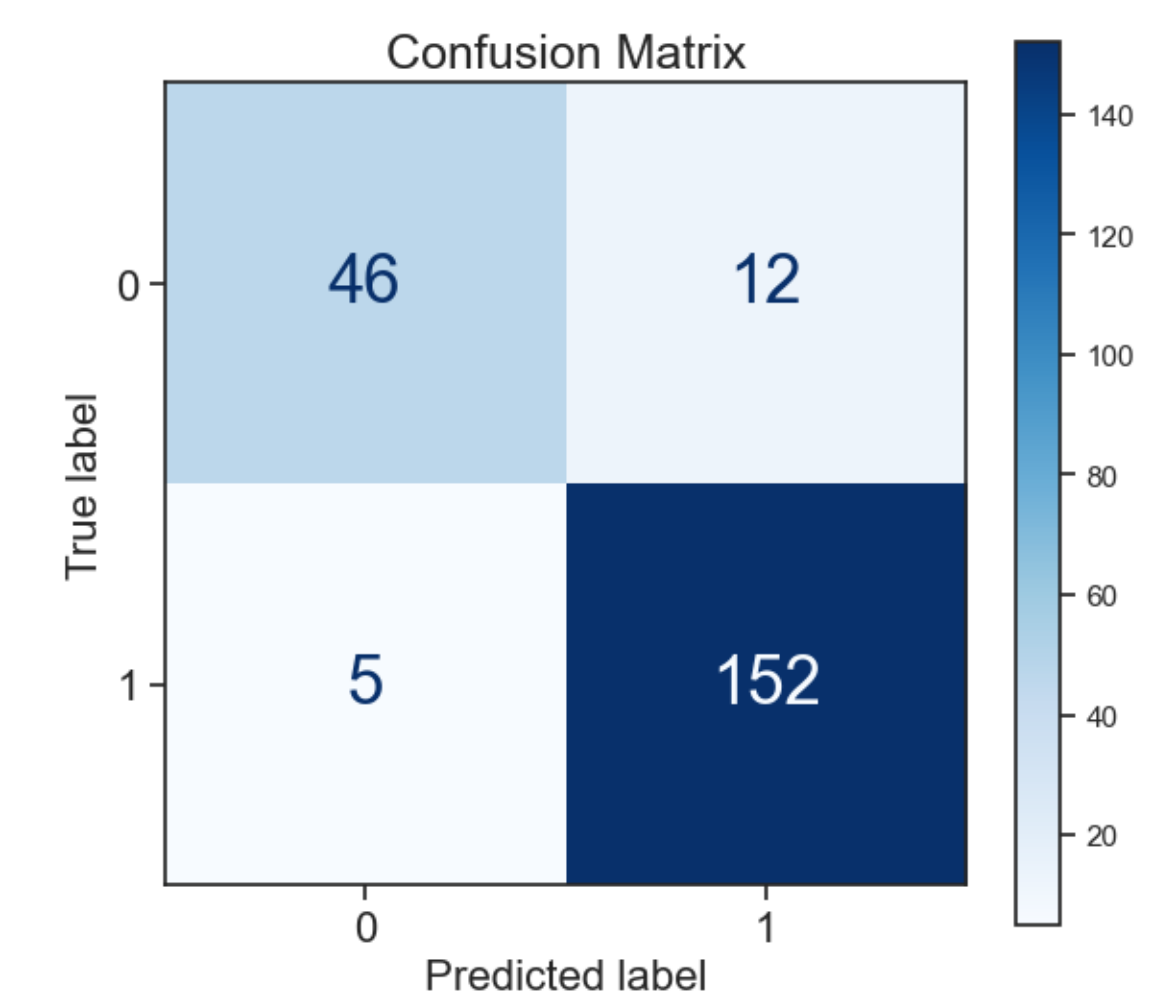
Metric	Train	Test
AUC	1.0	0.9511
accuracy	0.9907	0.9209
balanced accuracy	0.9828	0.8806
recall_(sens)	1.0	0.9682
specificity	0.9657	0.7931
precision	0.9874	0.9268
f1_score	0.9906	0.9193

Protein: P05067 - metric: SoftVotingXGBoostEnsemble
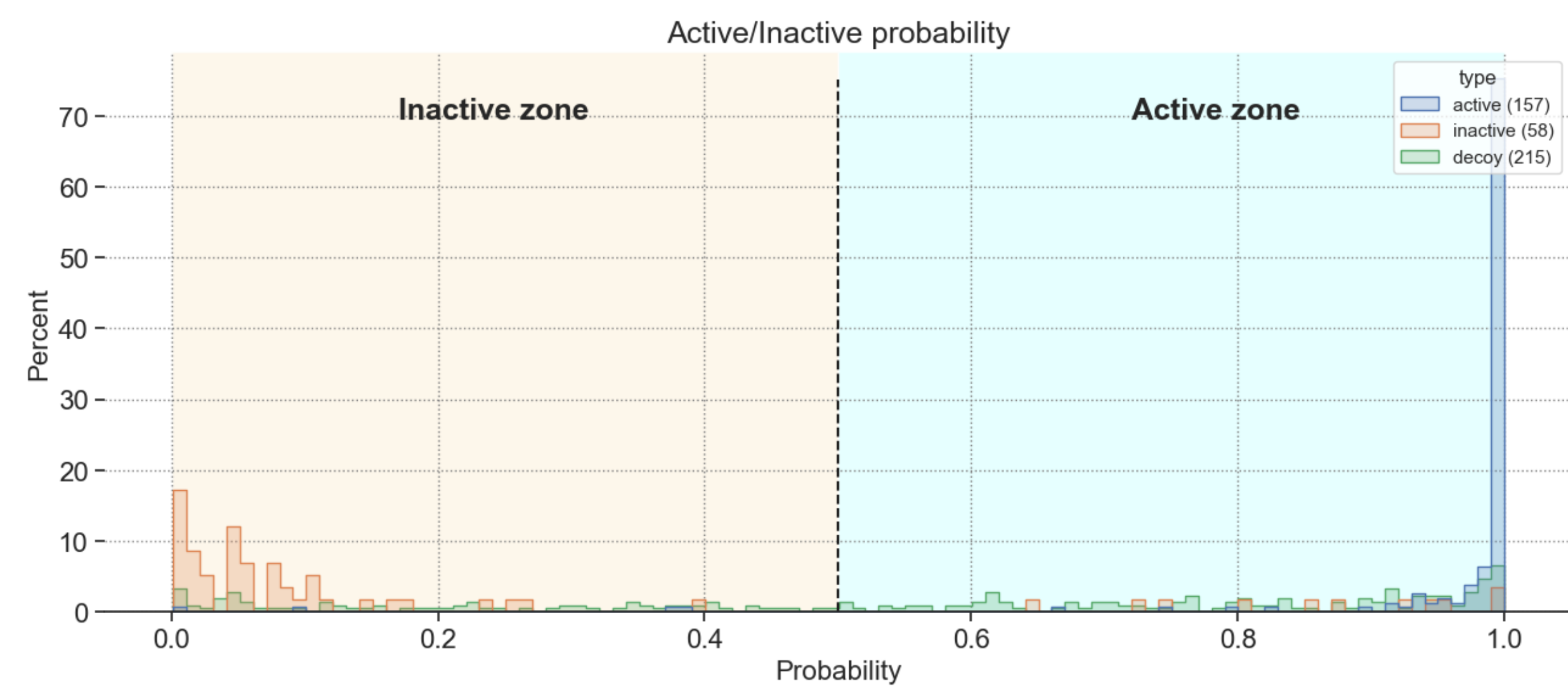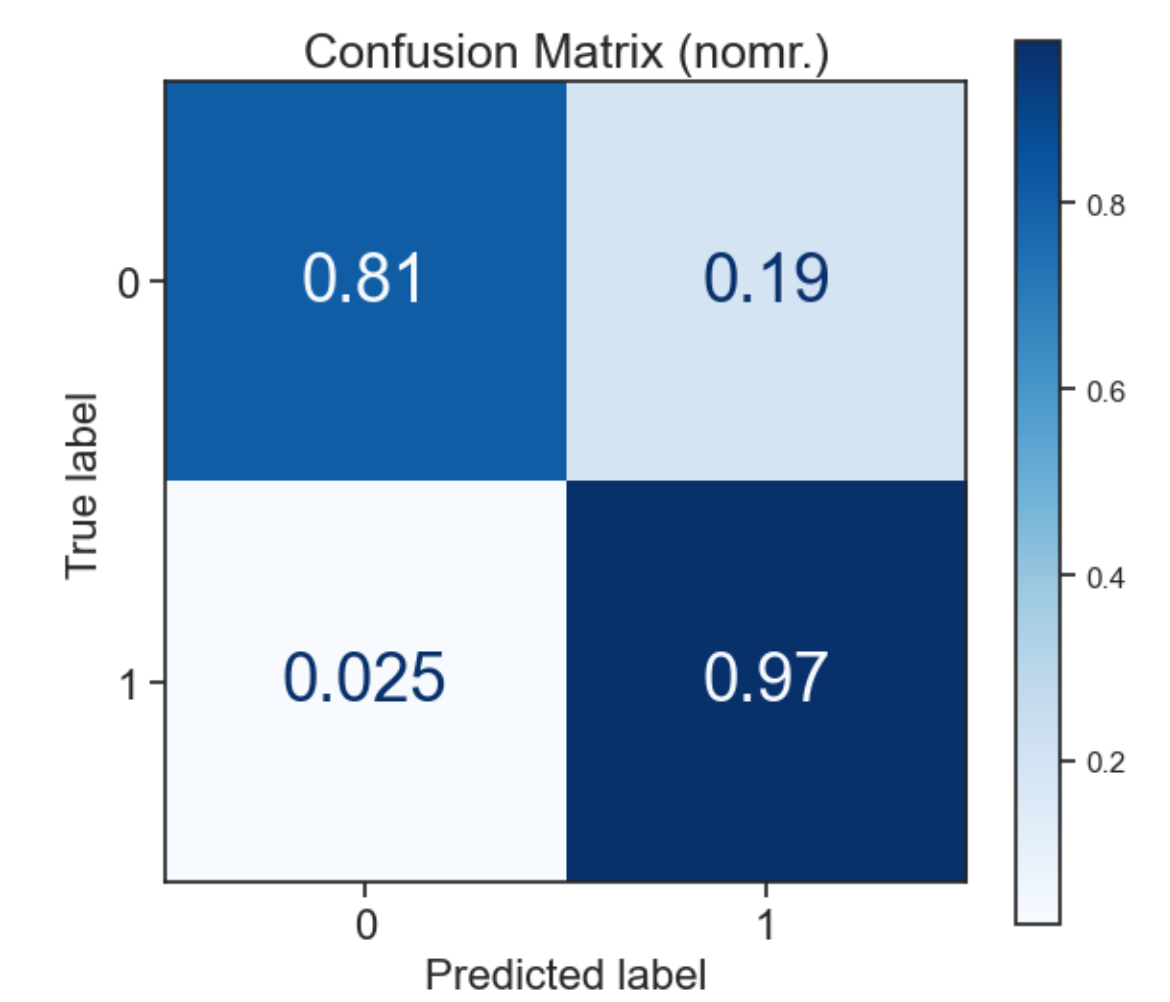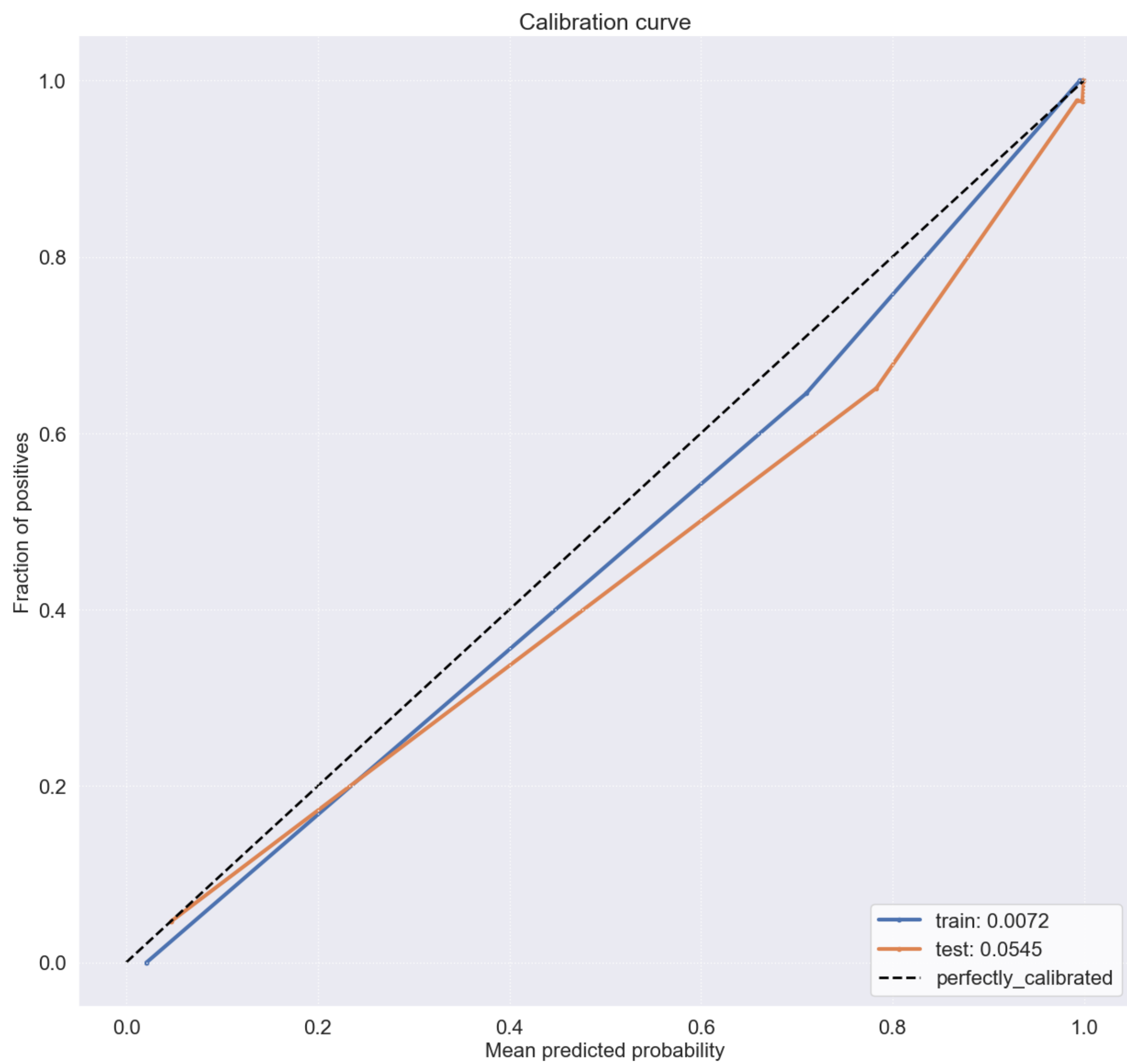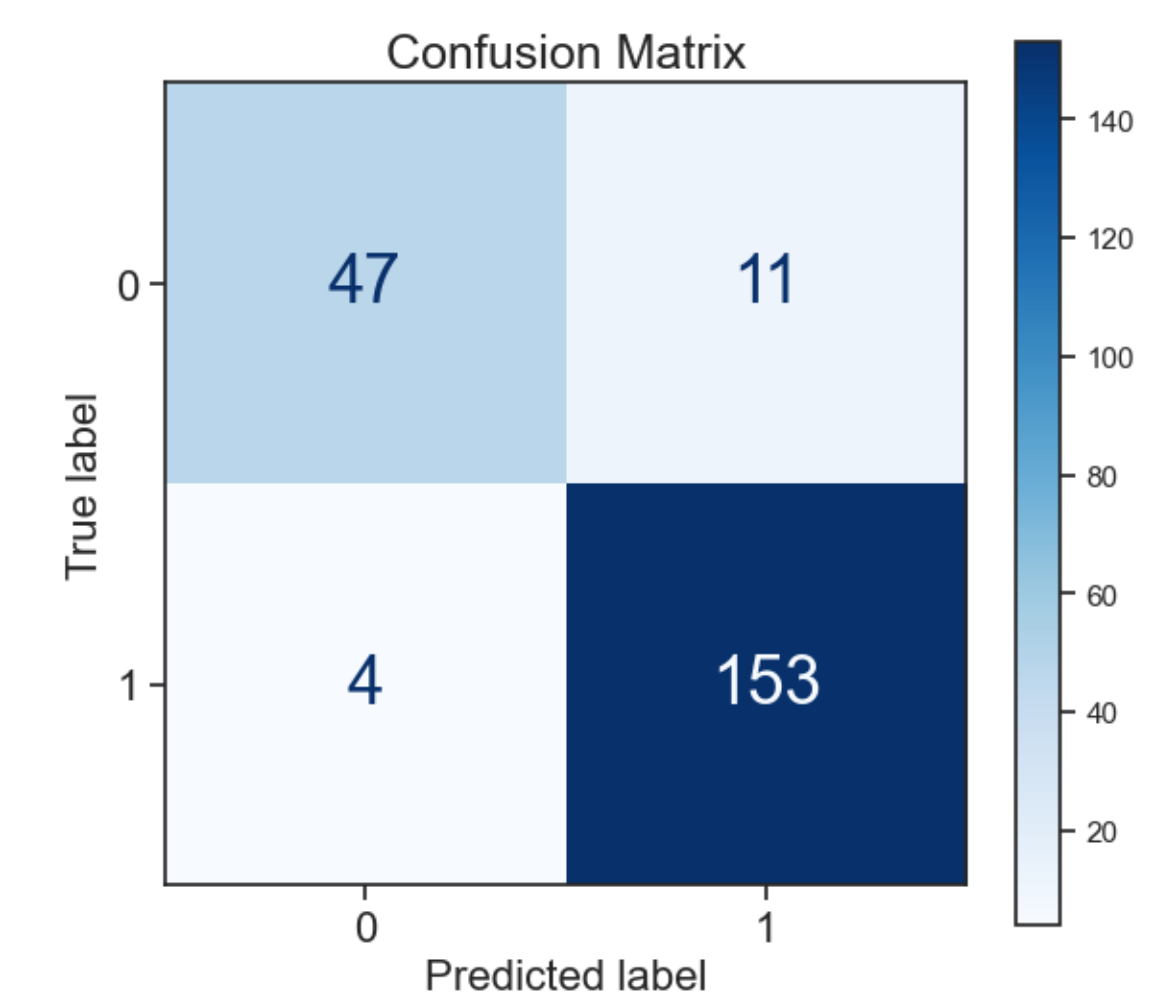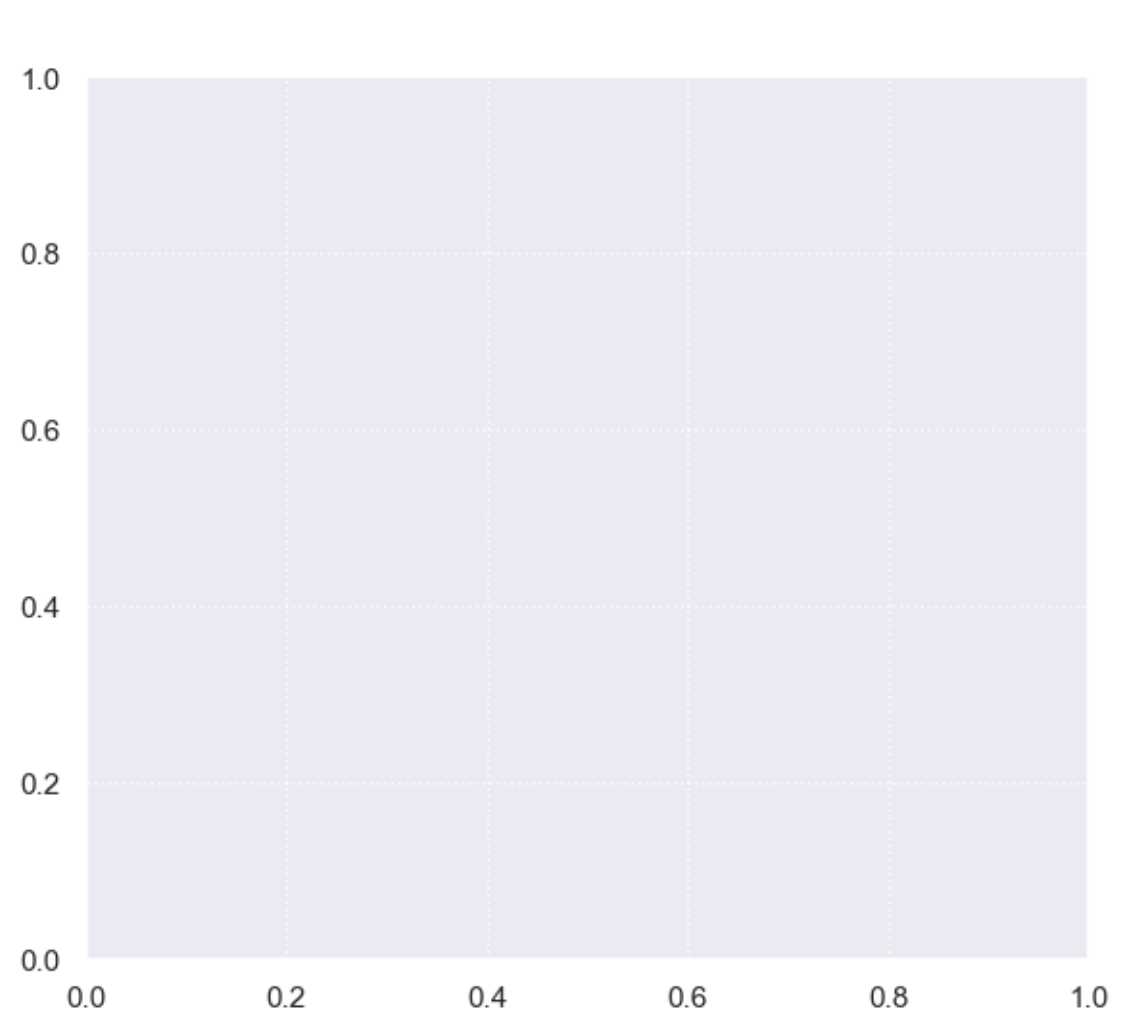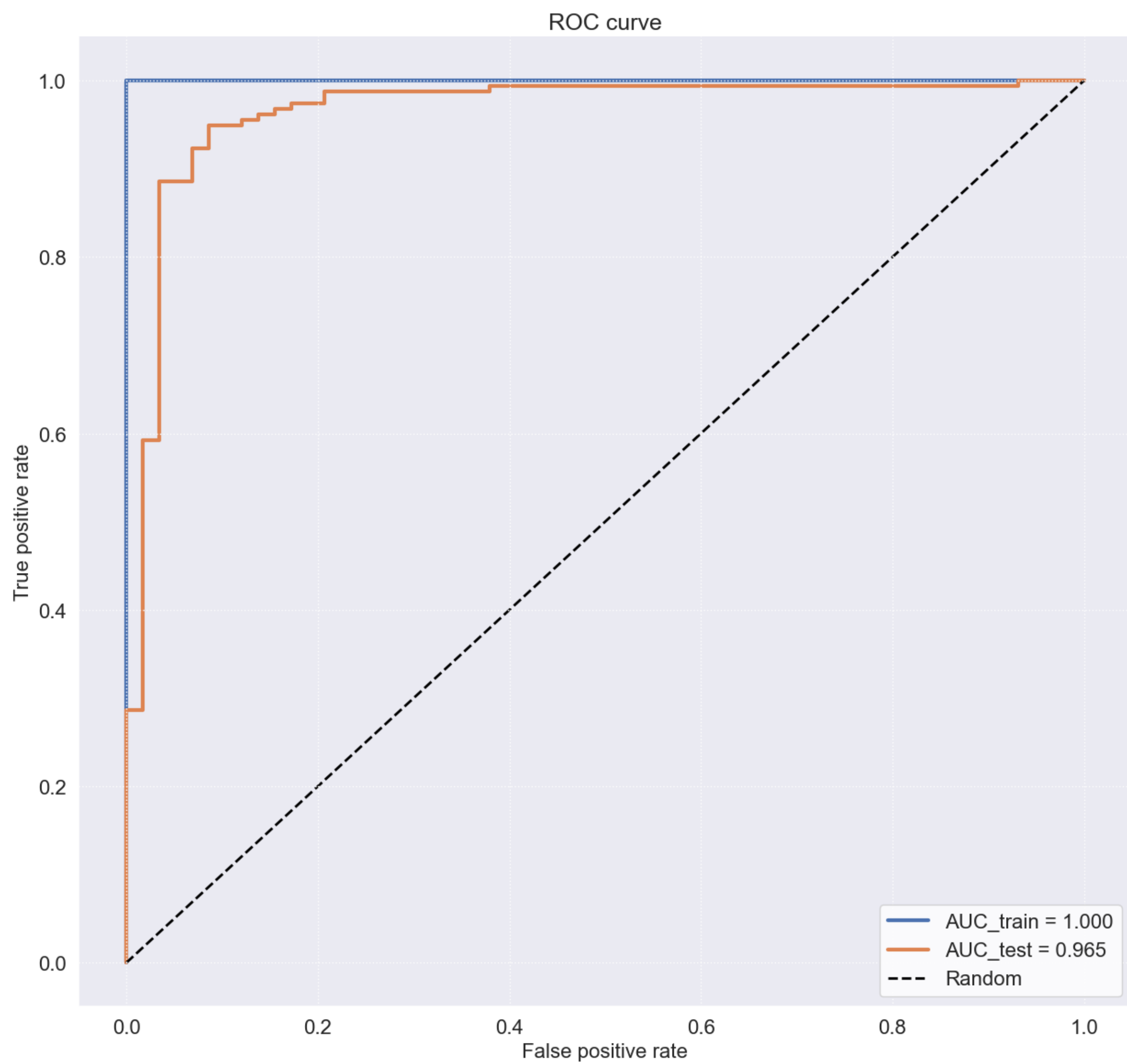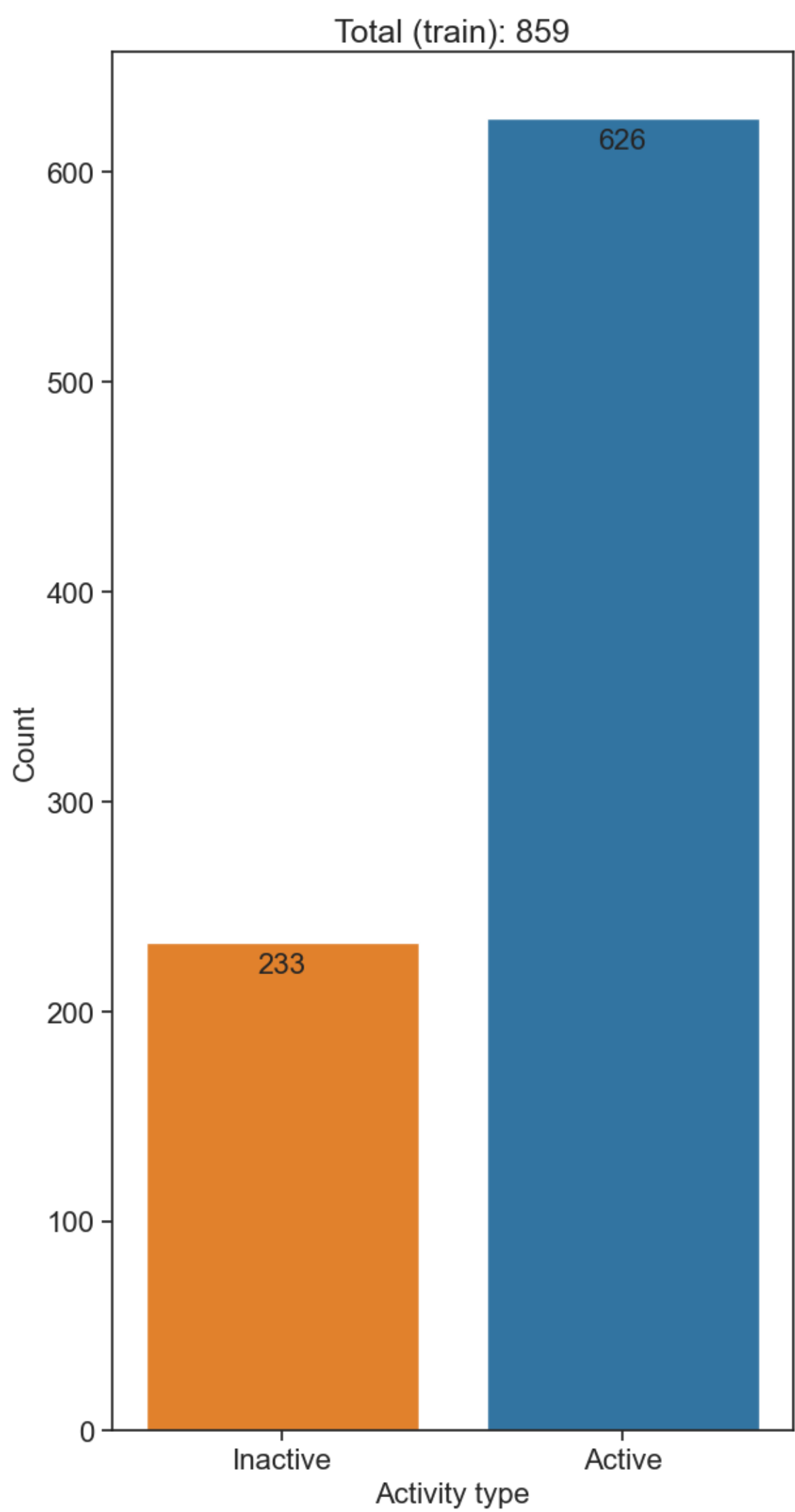
Metric	Train	Test
AUC	1.0	0.9651
accuracy	0.9895	0.9302
balanced accuracy	0.9807	0.8924
recall_(sens)	1.0	0.9745
specificity	0.9614	0.8103
precision	0.9858	0.9329
f1_score	0.9895	0.9288