Protein: P35368 - metric: balanced_accuracy

Total (train): 664 — bar chart: Inactive 166, Active 498

ROC curve — AUC_train = 0.999, AUC_test = 0.875, Random

Confusion Matrix

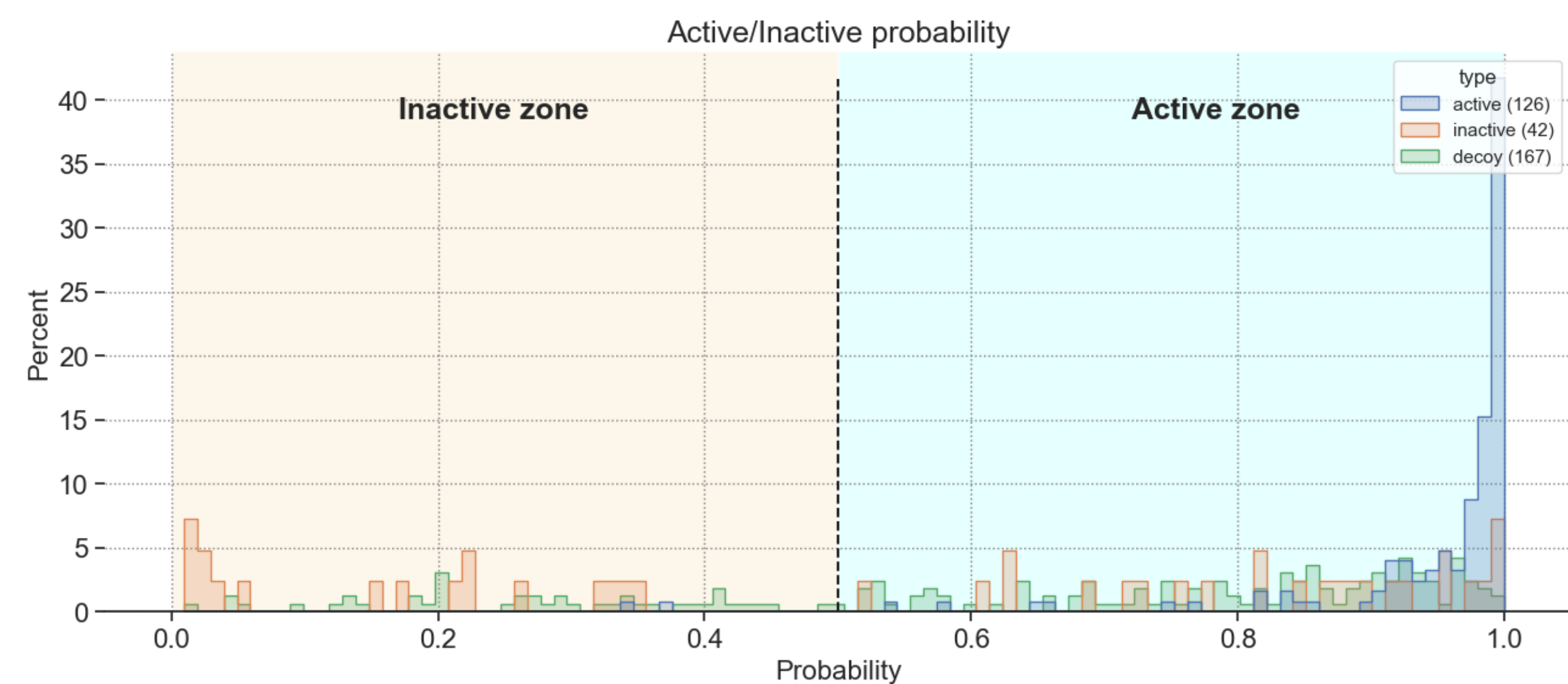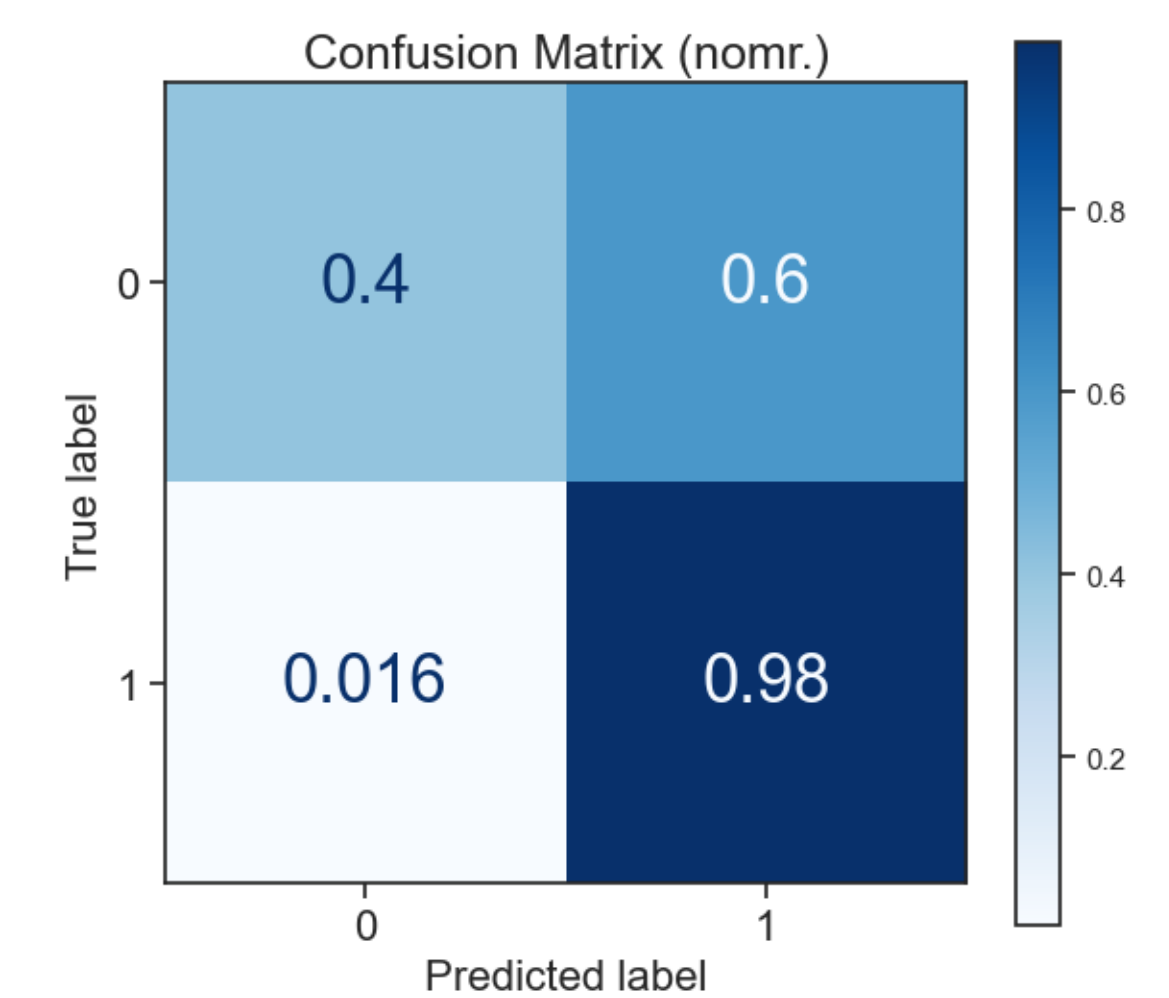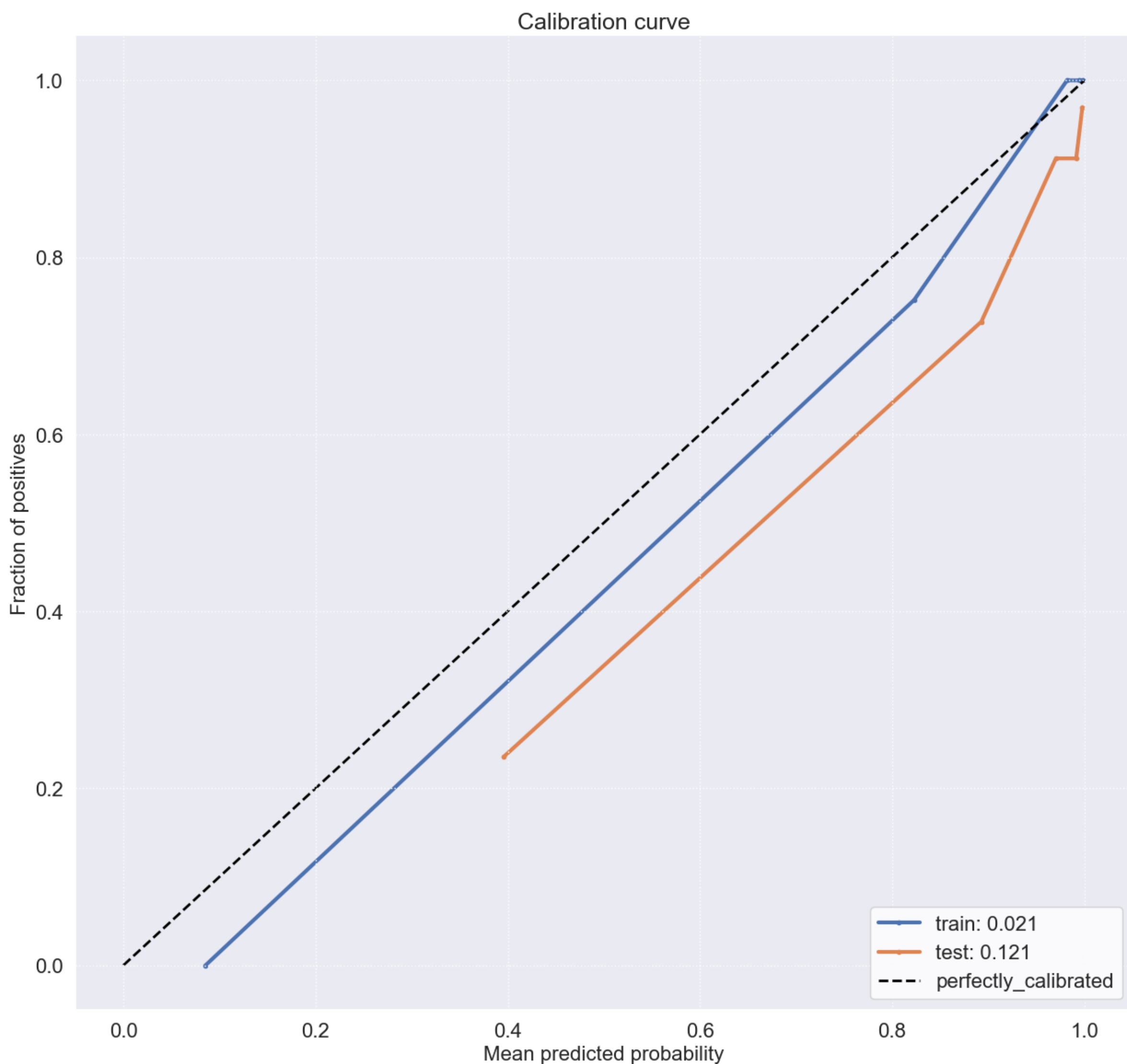|	Predicted 0	Predicted 1
True 0	17	25
True 1	2	124

Confusion Matrix (nomr.)

	Predicted 0	Predicted 1
True 0	0.4	0.6
True 1	0.016	0.98

Calibration curve — train: 0.021, test: 0.121, perfectly_calibrated

Pareto Chart with Highlighted Best Attempt — Trials, Pareto Front Trials, Best Trial (# 1169)

XGBoost Parameters	Value
alpha	4.993
colsample_bytree	0.7539
gamma	0.7071
lambda	2.064
learning_rate	0.0521
max_depth	8.0
min_child_weight	2.0
n_estimators	846.0
reg_alpha	0.0242
reg_lambda	0.1078
scale_pos_weight	2.4873
subsample	0.8912

Active/Inactive probability — Inactive zone, Active zone; type: active (126), inactive (42), decoy (167)
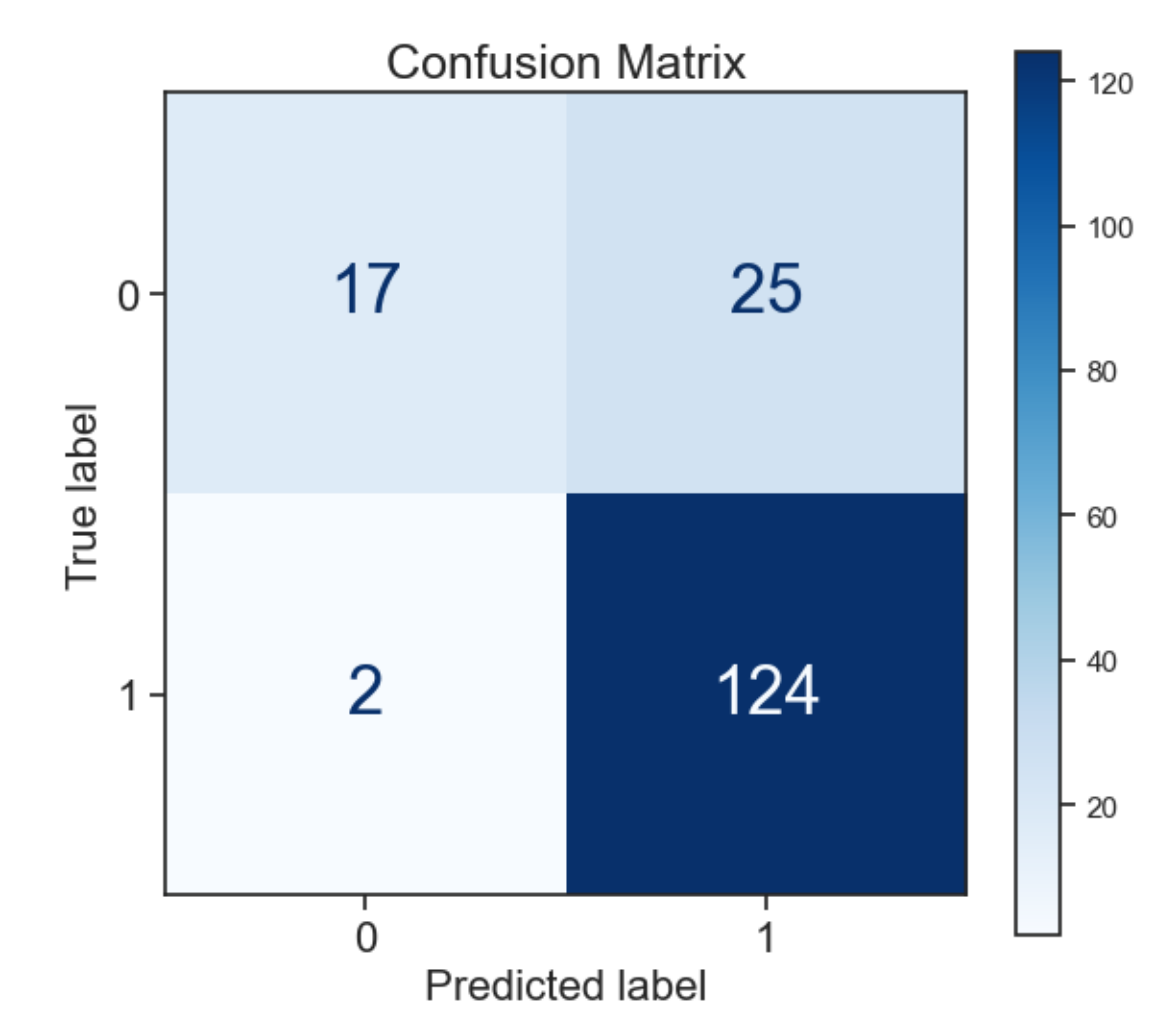
Metric	Train	Test
AUC	0.9989	0.8751
accuracy	0.9684	0.8393
balanced accuracy	0.9388	0.6944
recall_(sens)	0.998	0.9841
specificity	0.8795	0.4048
precision	0.9613	0.8322
f1_score	0.9677	0.8157

Protein: P35368 - metric: f1

Total (train): 664

ROC curve

AUC_train = 1.000
AUC_test = 0.867
Random

Pareto Chart with Highlighted Best Attempt

Trials
Pareto Front Trials
Best Trial (# 1244)
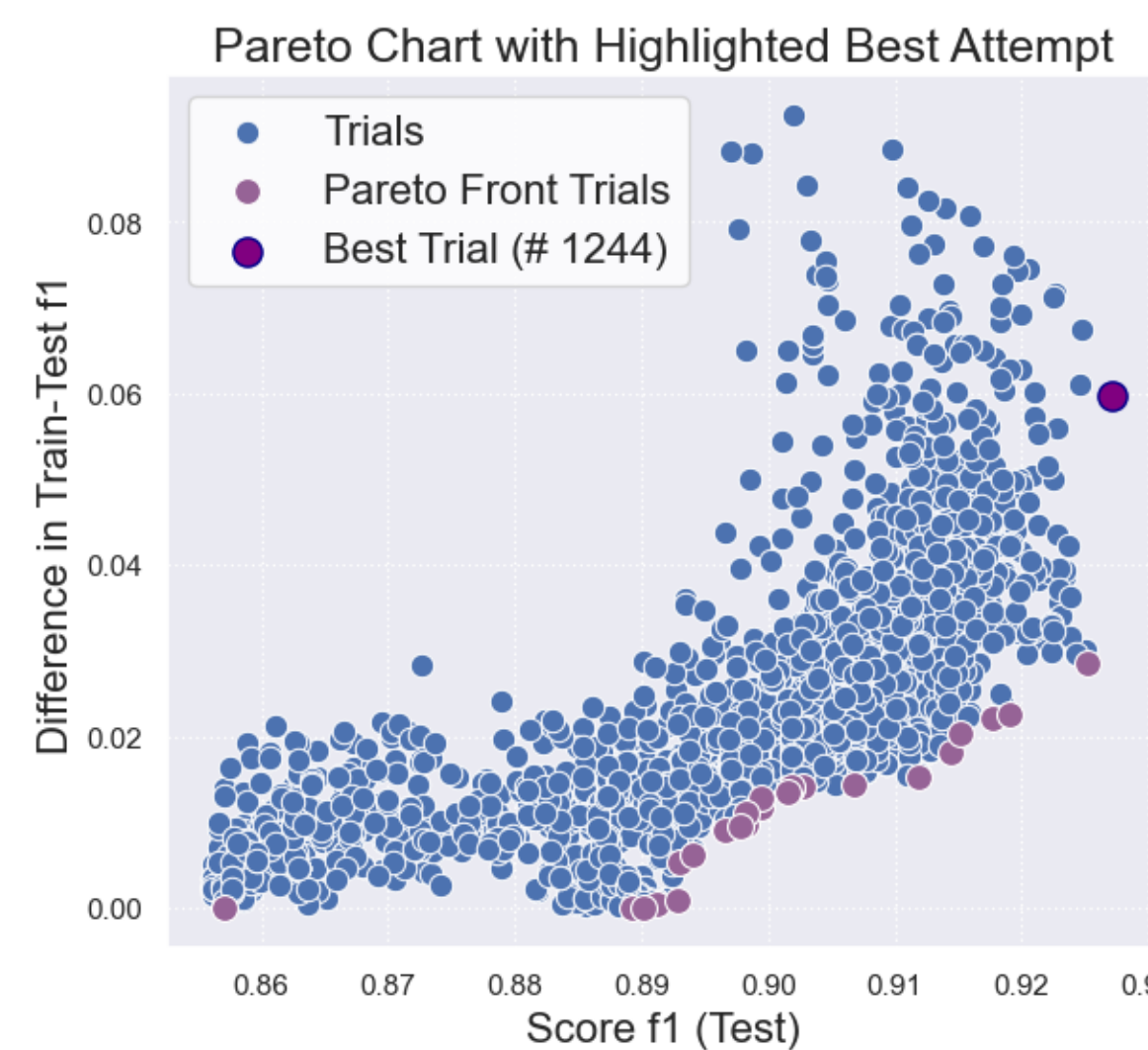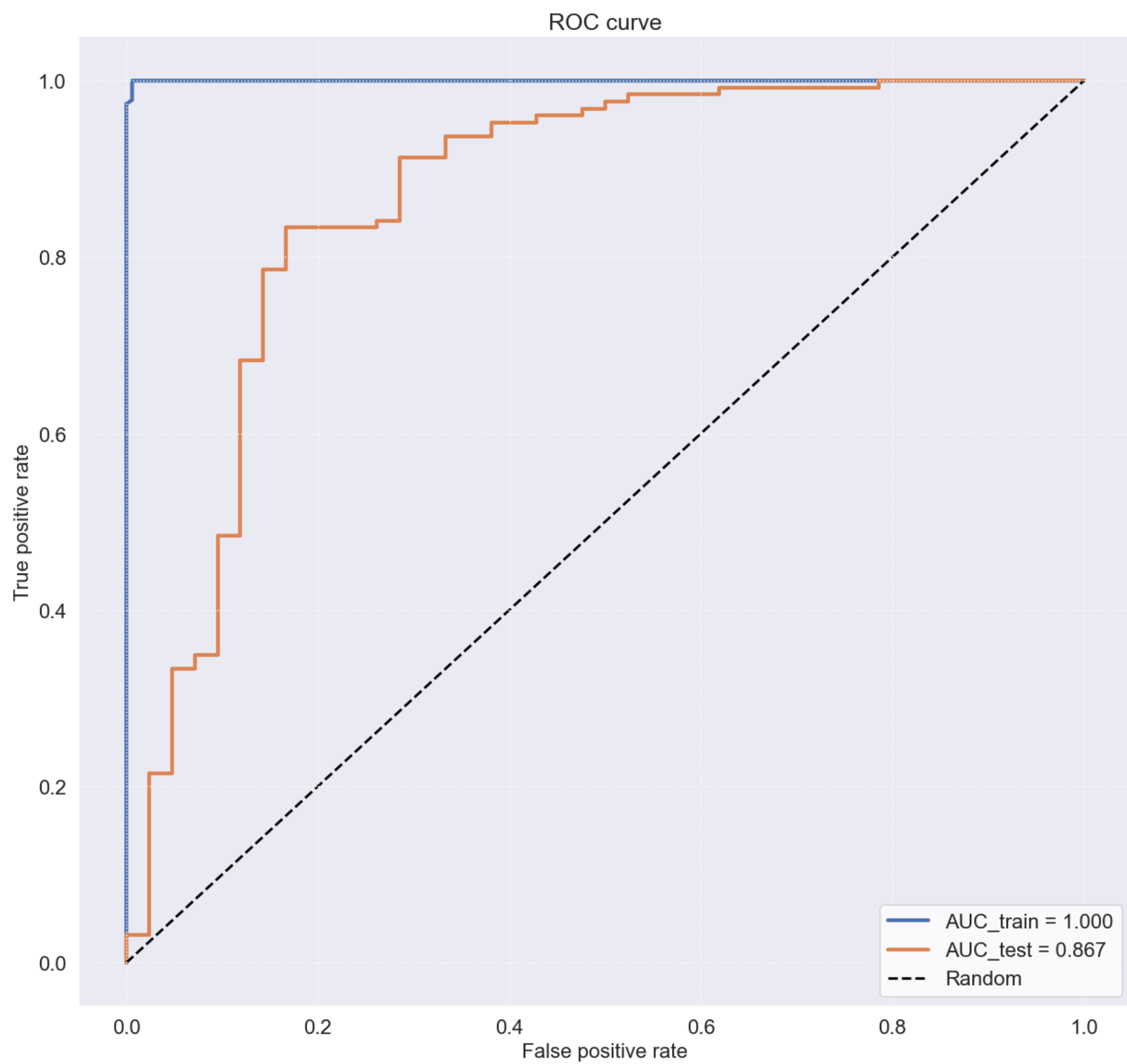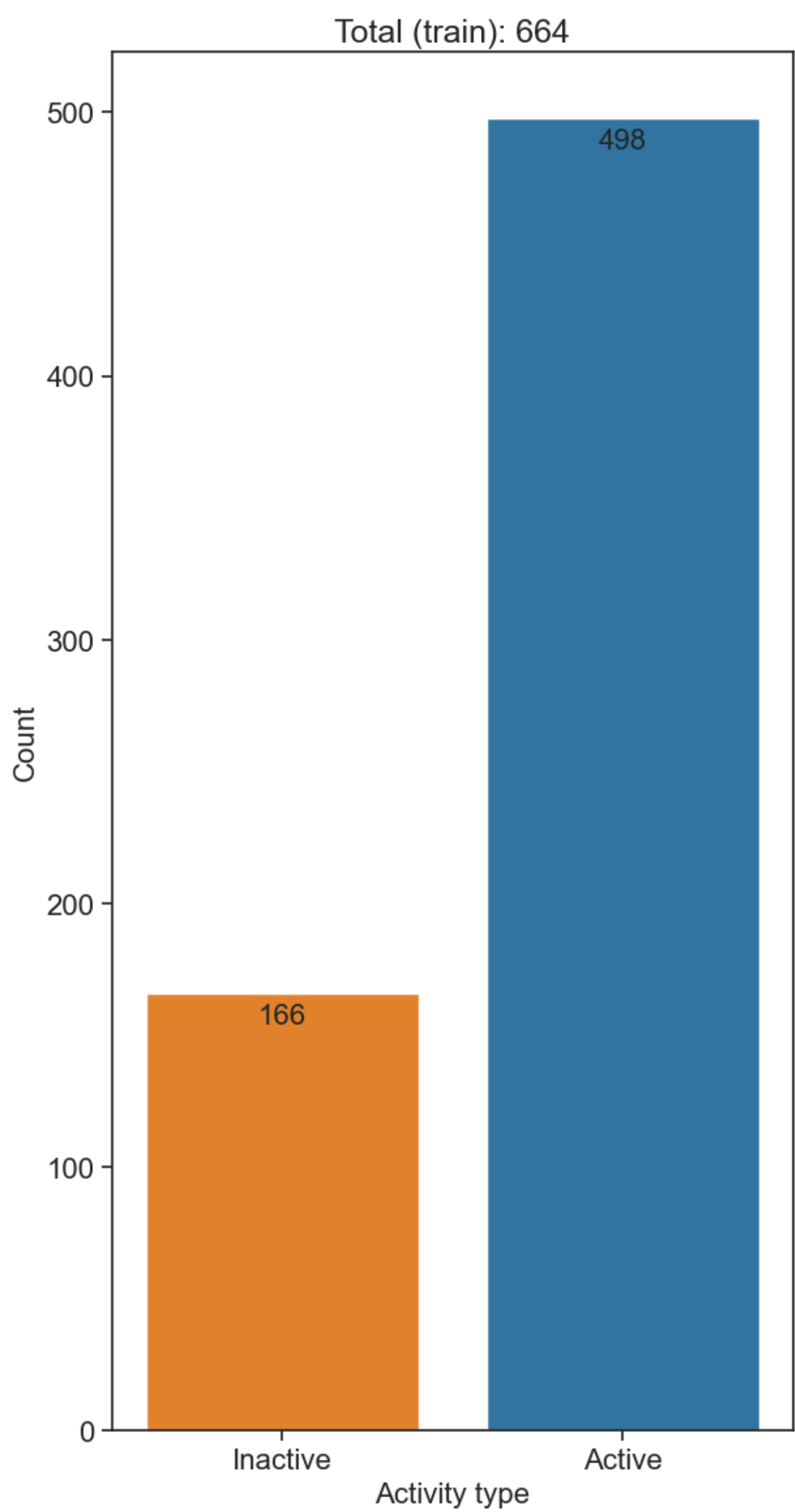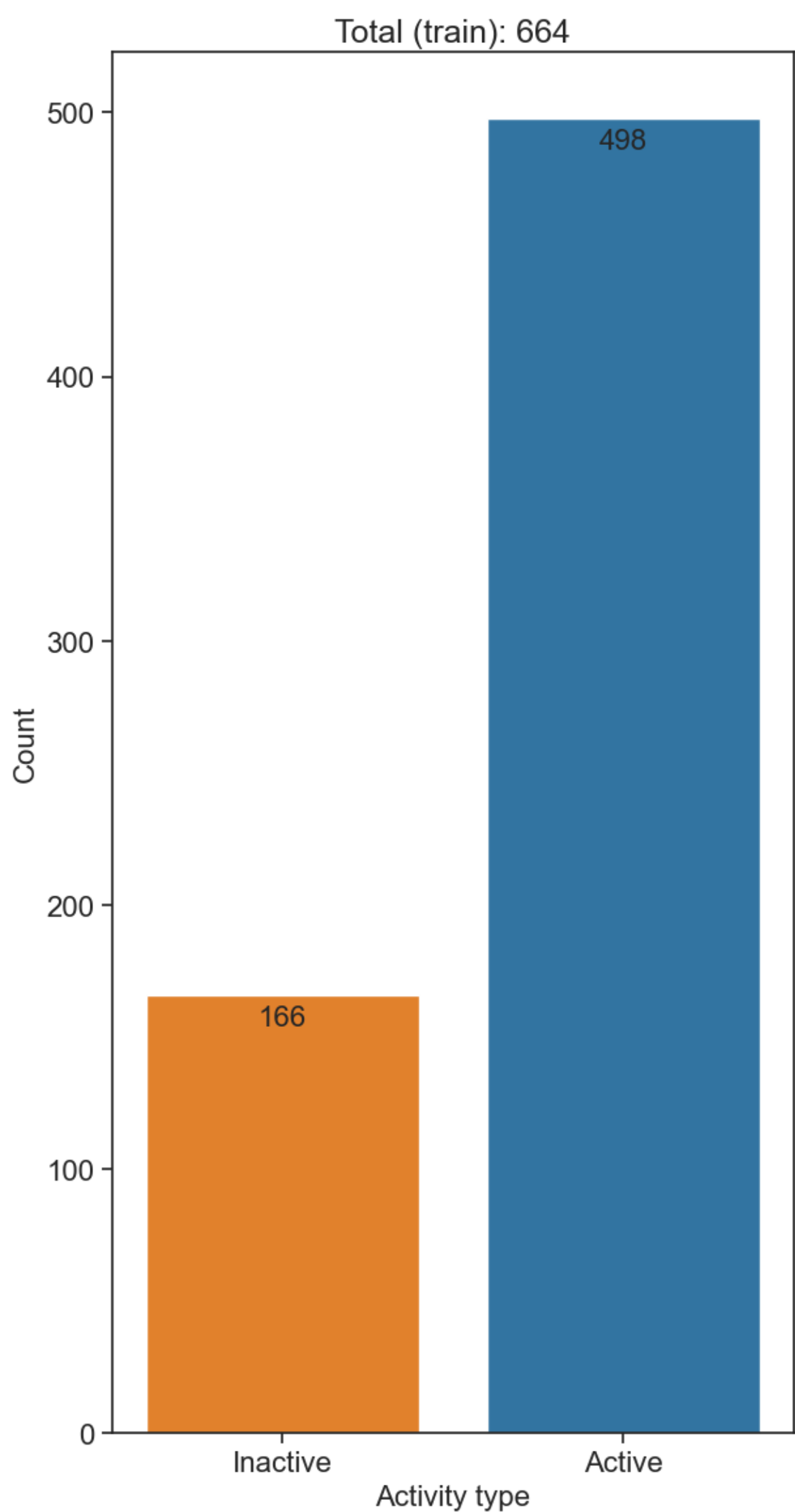
XGBoost Parameters	Value
alpha	1.0734
colsample_bytree	0.5018
gamma	0.3041
lambda	3.4809
learning_rate	0.0748
max_depth	9.0
min_child_weight	1.0
n_estimators	493.0
reg_alpha	0.0458
reg_lambda	0.6089
scale_pos_weight	76.6413
subsample	0.5228

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0164
test: 0.1415
perfectly_calibrated

Active/Inactive probability

type
active (126)
inactive (42)
decoy (167)
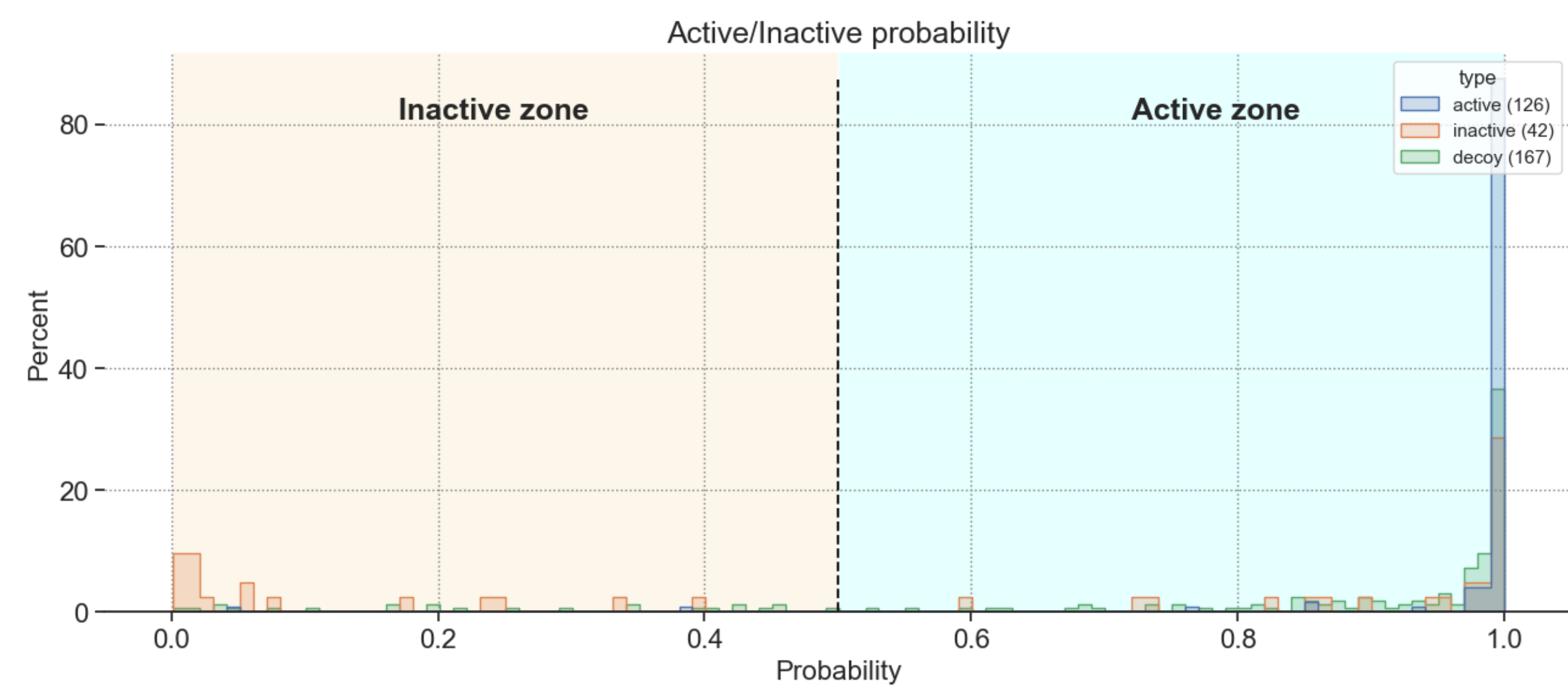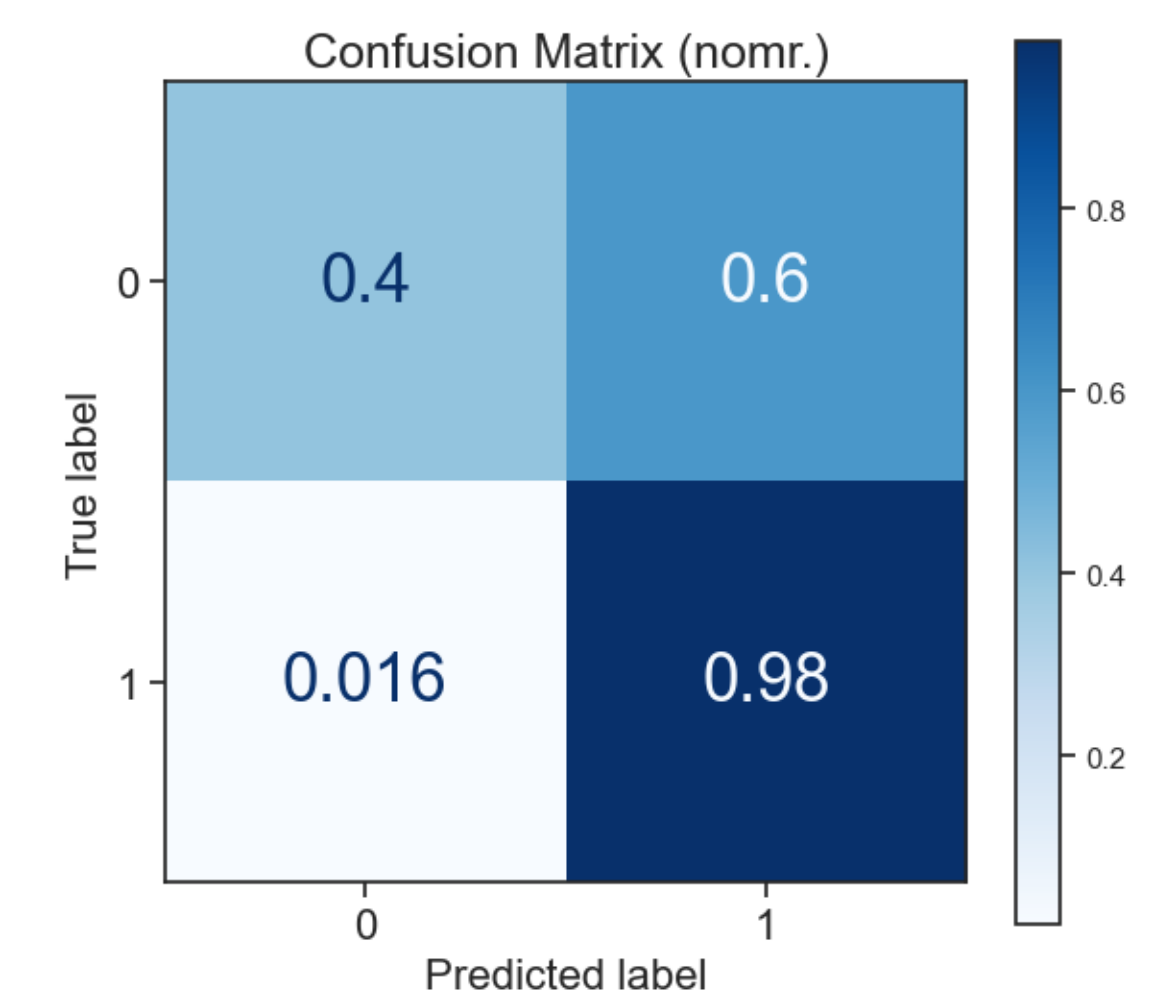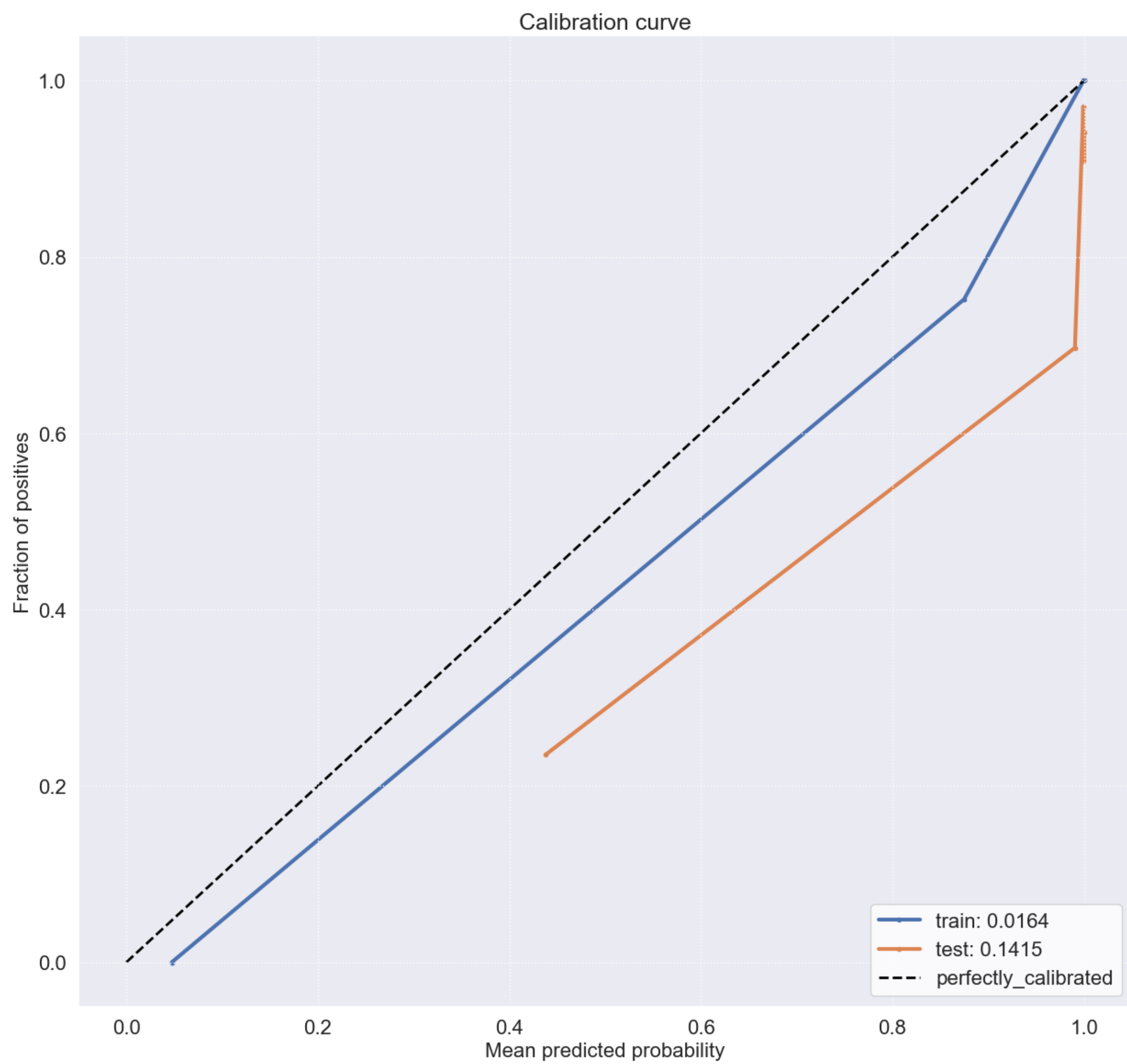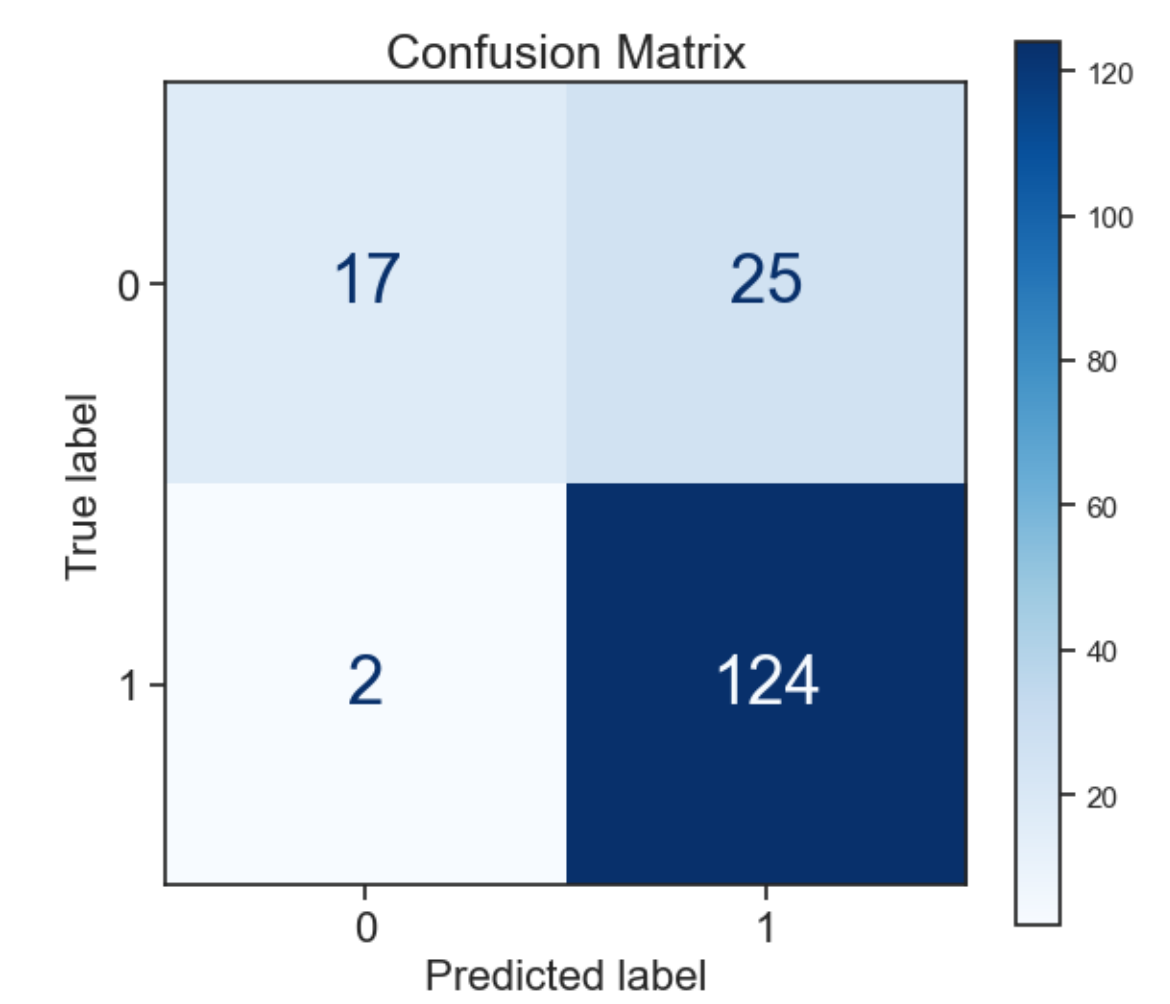
Metric	Train	Test
AUC	0.9999	0.8668
accuracy	0.9804	0.8393
balanced accuracy	0.9608	0.6944
recall_(sens)	1.0	0.9841
specificity	0.9217	0.4048
precision	0.9746	0.8322
f1_score	0.9801	0.8157

Protein: P35368 - metric: precision

Total (train): 664

ROC curve

Pareto Chart with Highlighted Best Attempt
- Trials
- Pareto Front Trials
- Best Trial (# 2236)
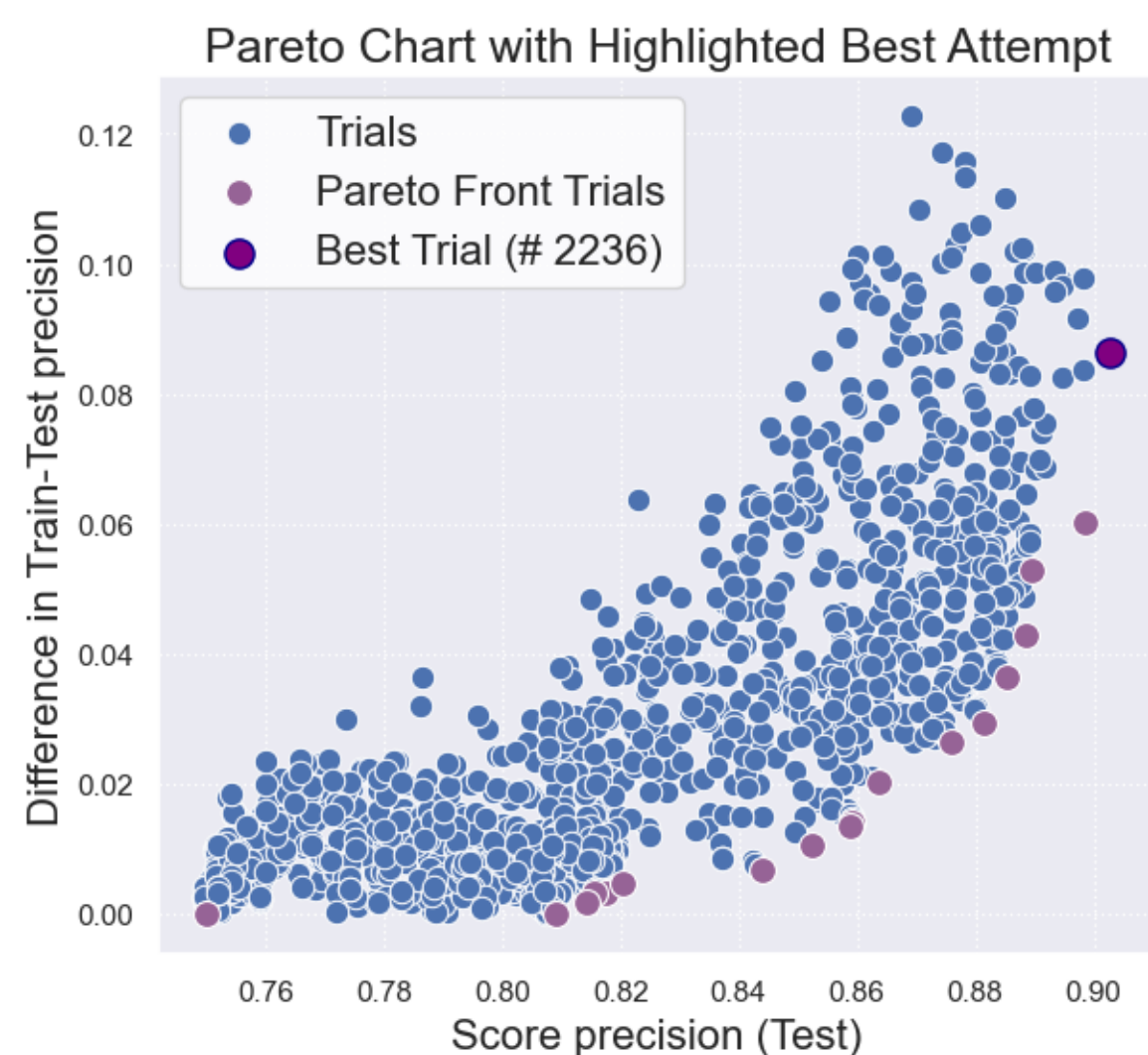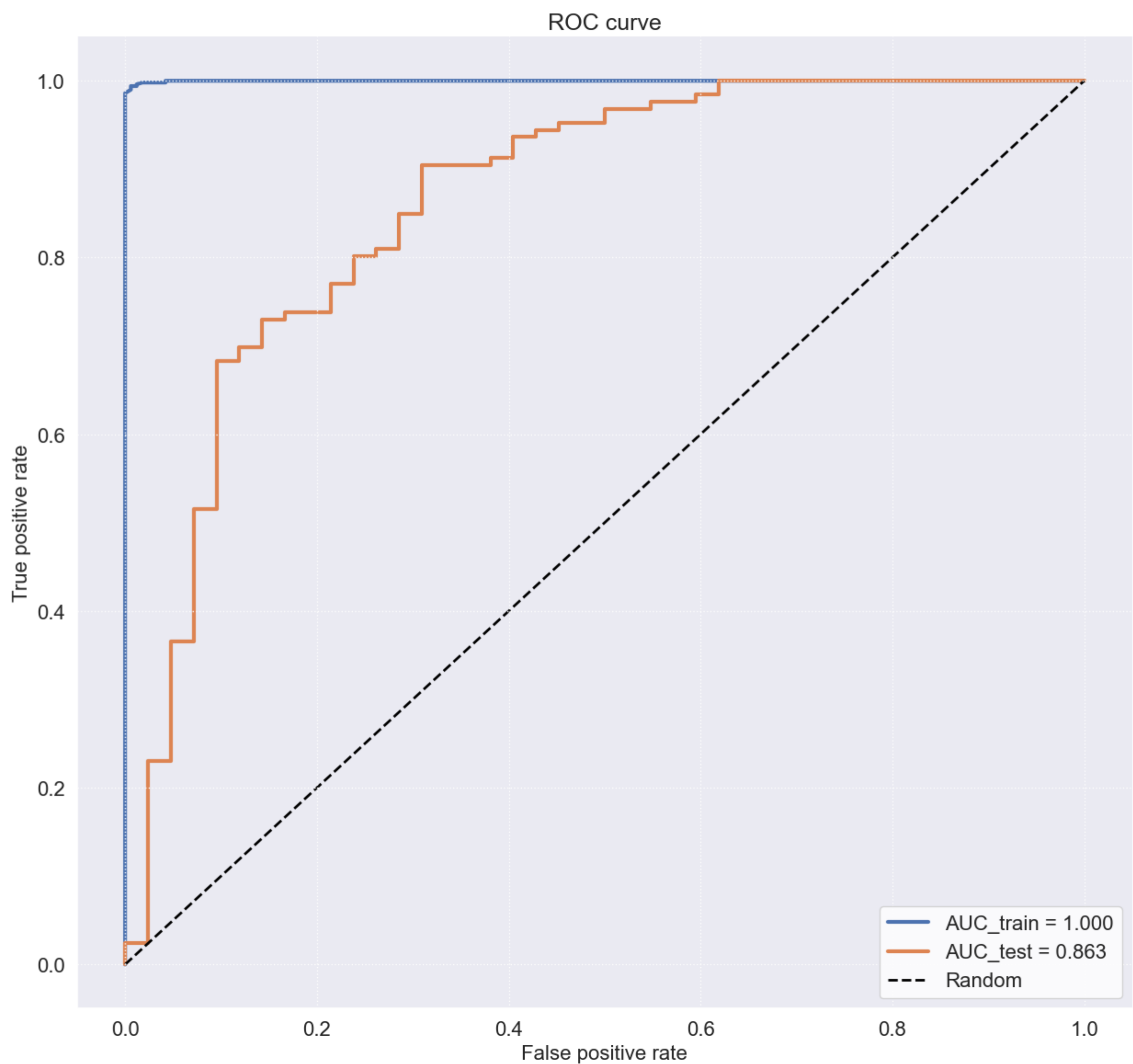
XGBoost Parameters	Value
alpha	1.307
colsample_bytree	0.9495
gamma	0.5528
lambda	2.2608
learning_rate	0.3126
max_depth	3.0
min_child_weight	1.0
n_estimators	652.0
reg_alpha	0.3944
reg_lambda	0.0982
scale_pos_weight	1.2743
subsample	0.6926

Confusion Matrix

23	19	
6	120	

Calibration curve

train: 0.0103
test: 0.1178
perfectly_calibrated

Confusion Matrix (nomr.)

0.55	0.45	
0.048	0.95	

Active/Inactive probability

Inactive zone — Active zone

type
- active (126)
- inactive (42)
- decoy (167)
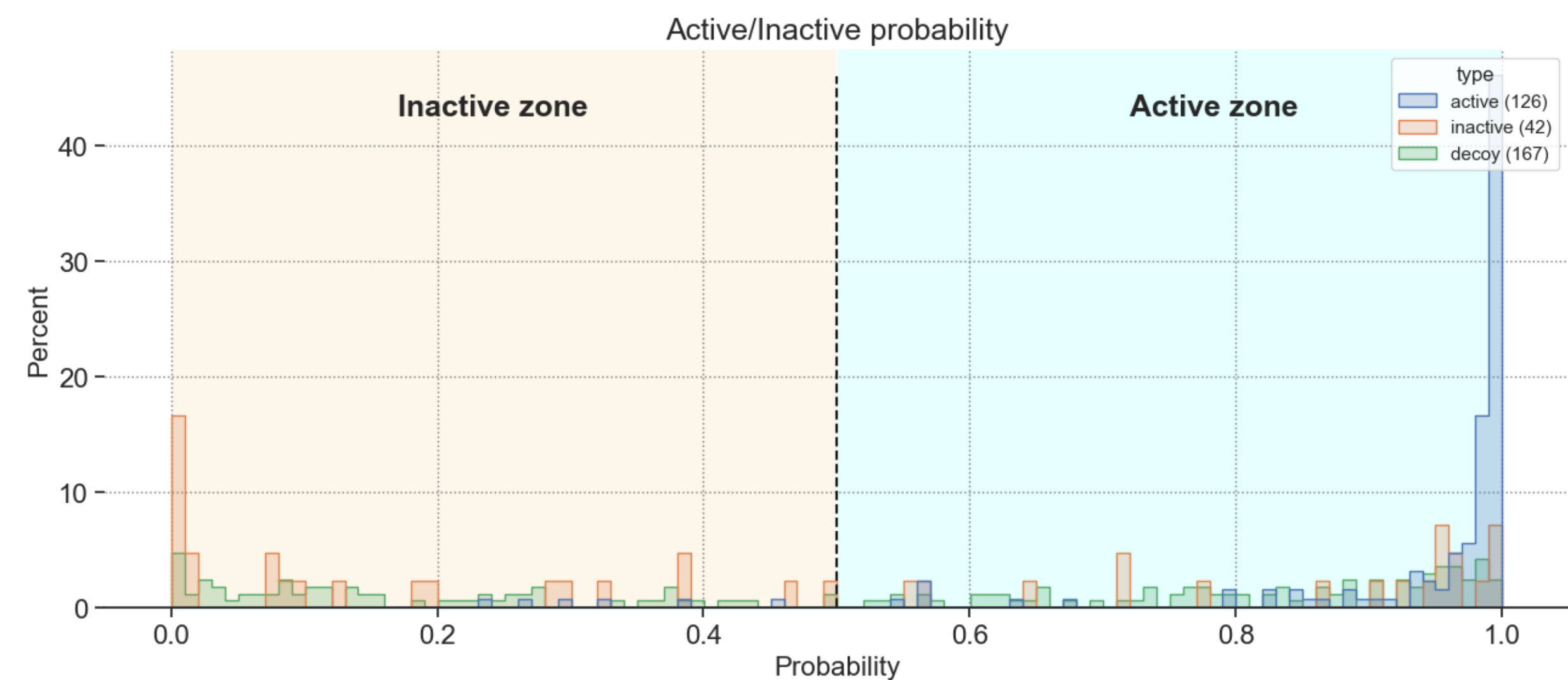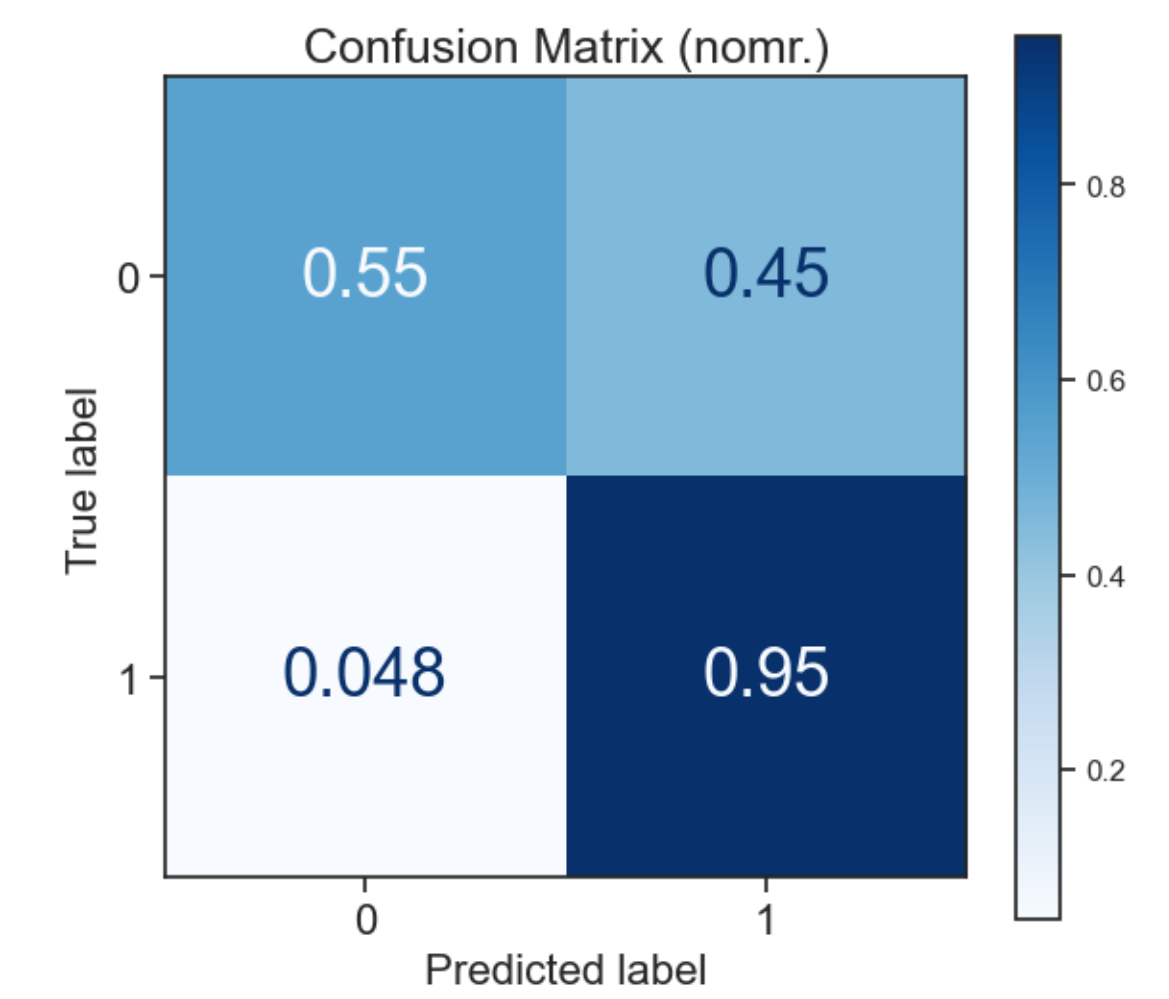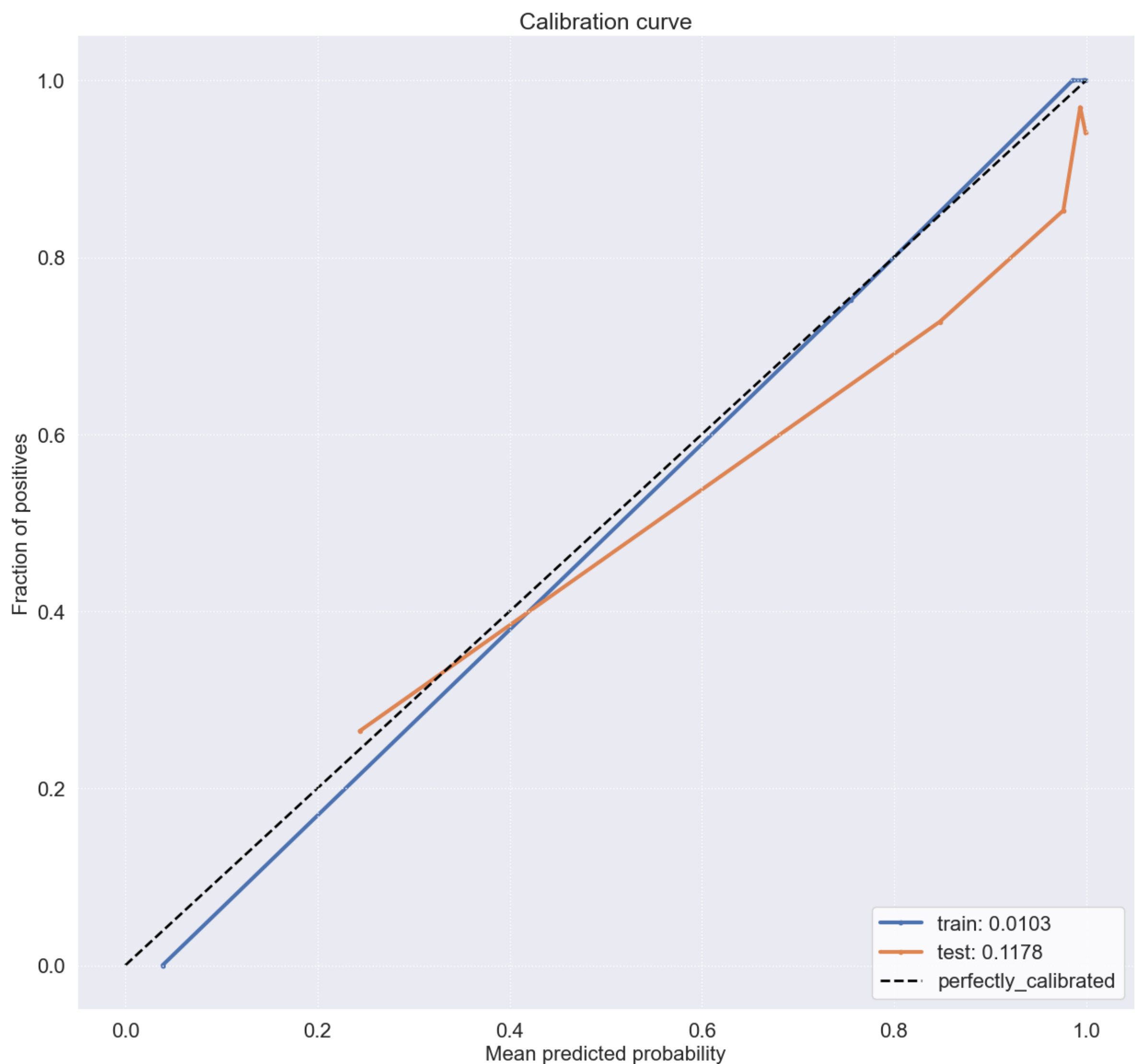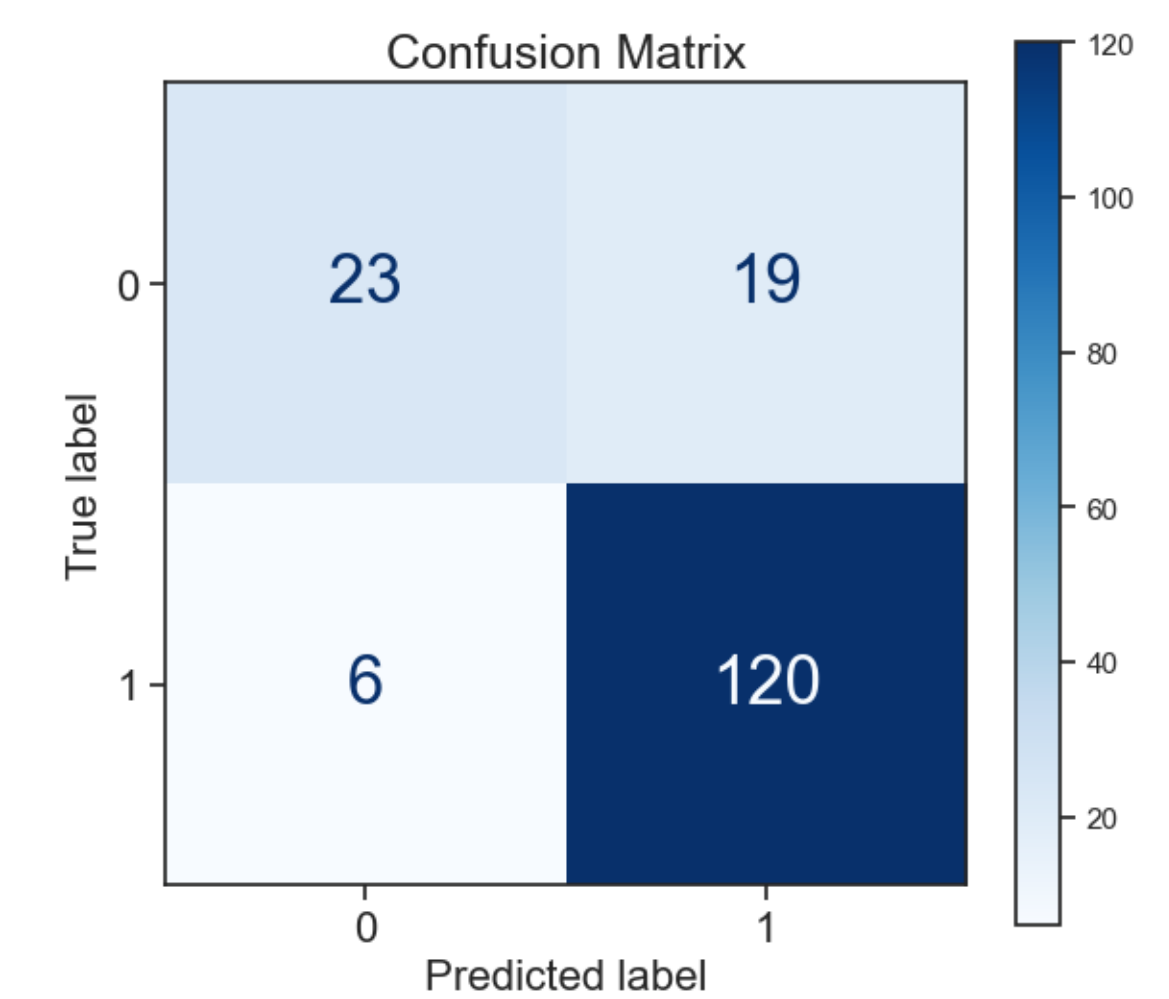
Metric	Train	Test
AUC	0.9998	0.8628
accuracy	0.991	0.8512
balanced accuracy	0.9839	0.75
recall_(sens)	0.998	0.9524
specificity	0.9699	0.5476
precision	0.99	0.8633
f1_score	0.9909	0.8412

Protein: P35368 - metric: SoftVotingXGBoostEnsemble
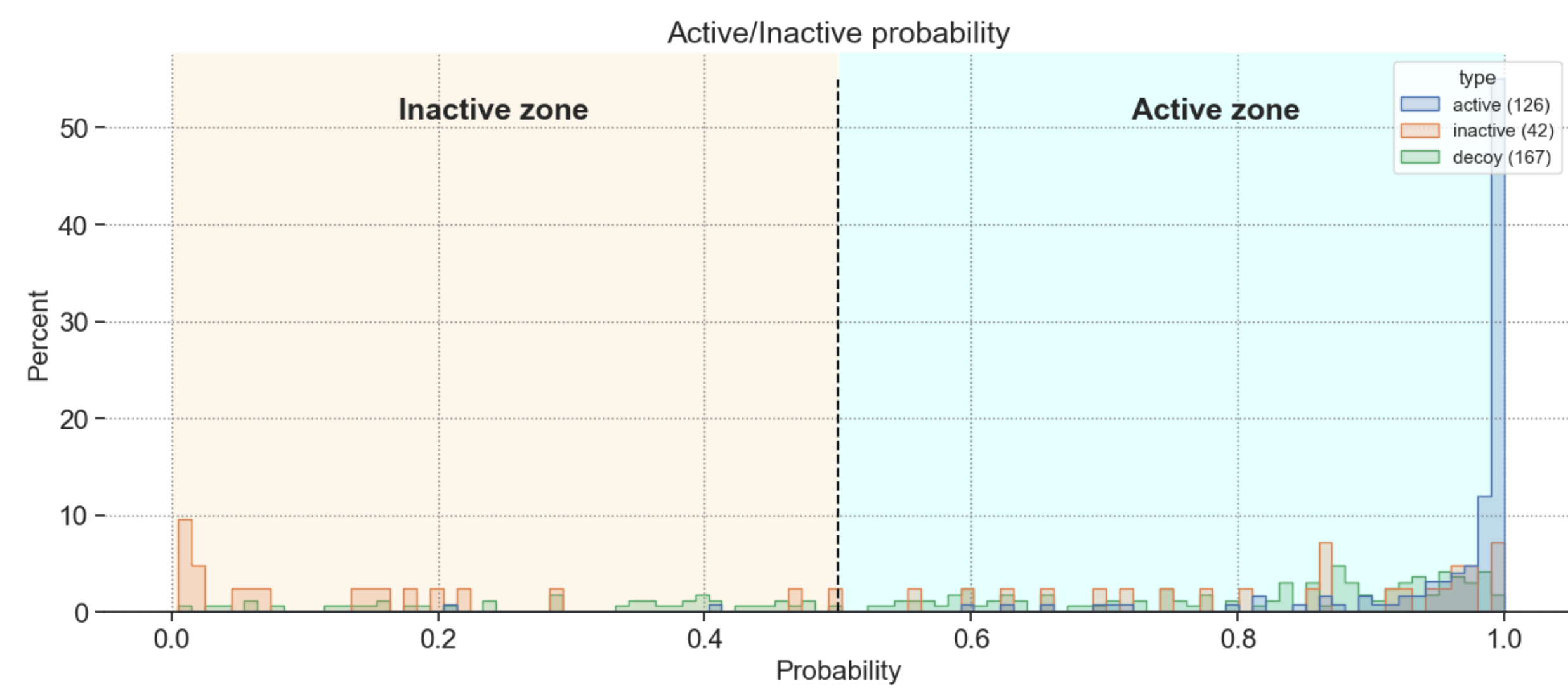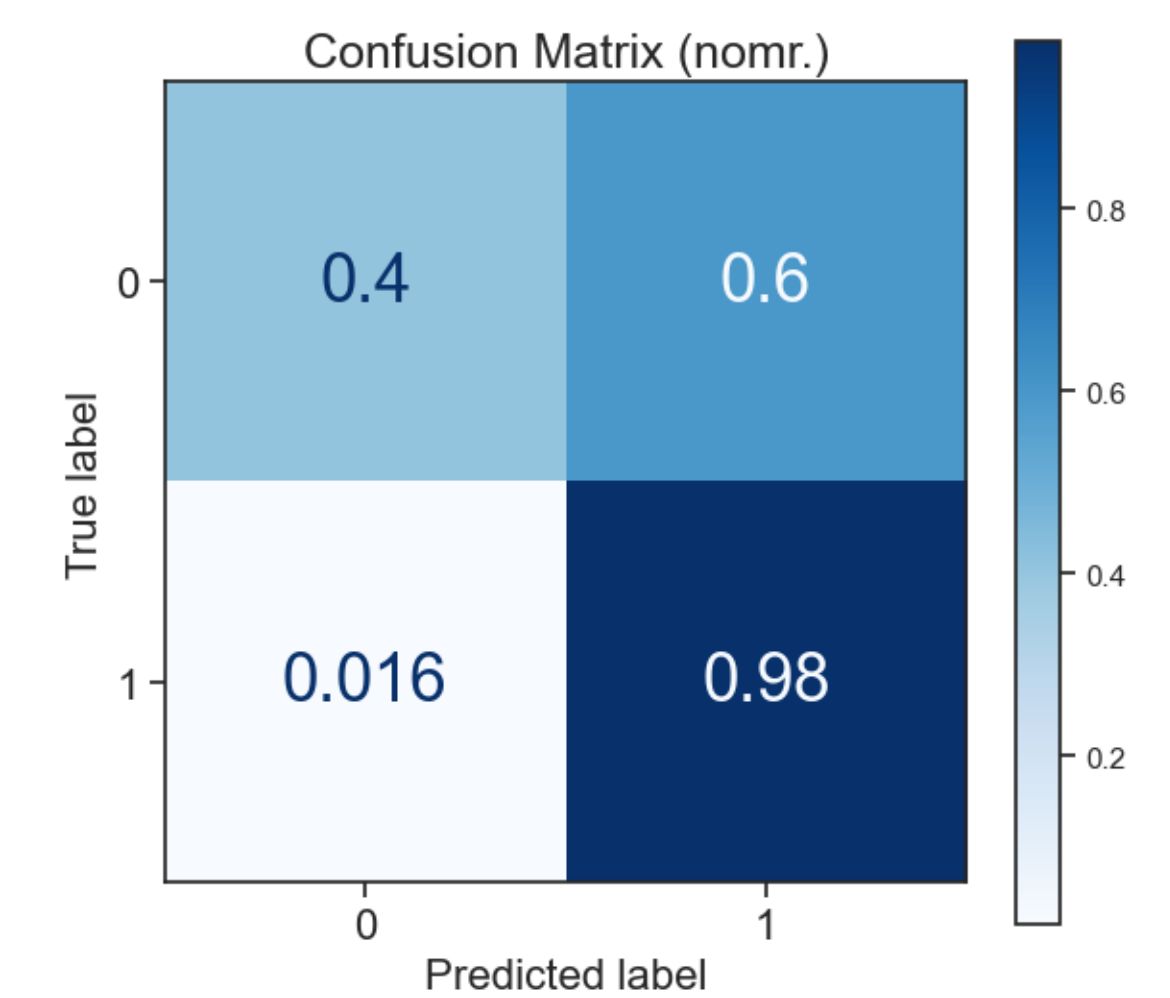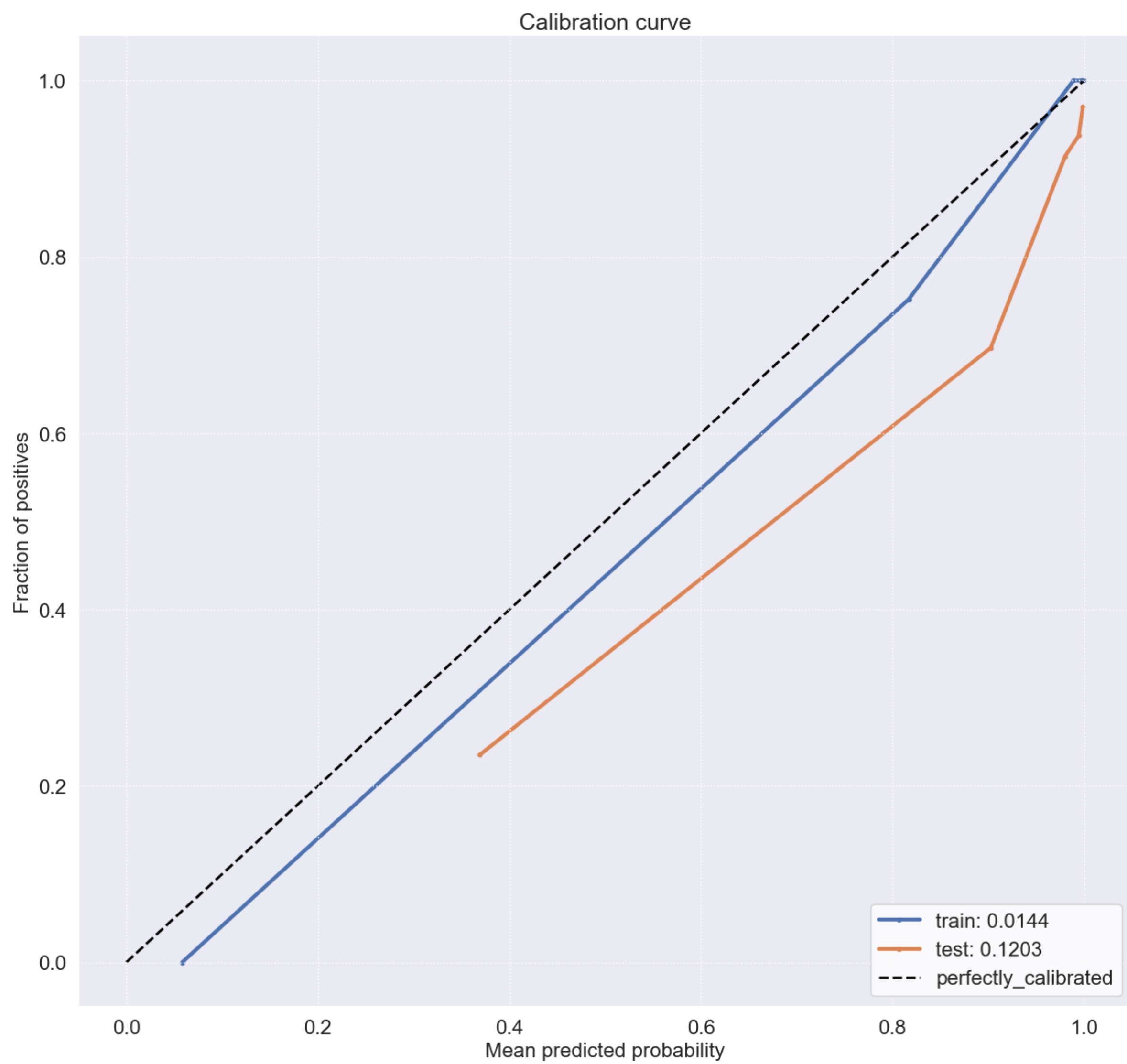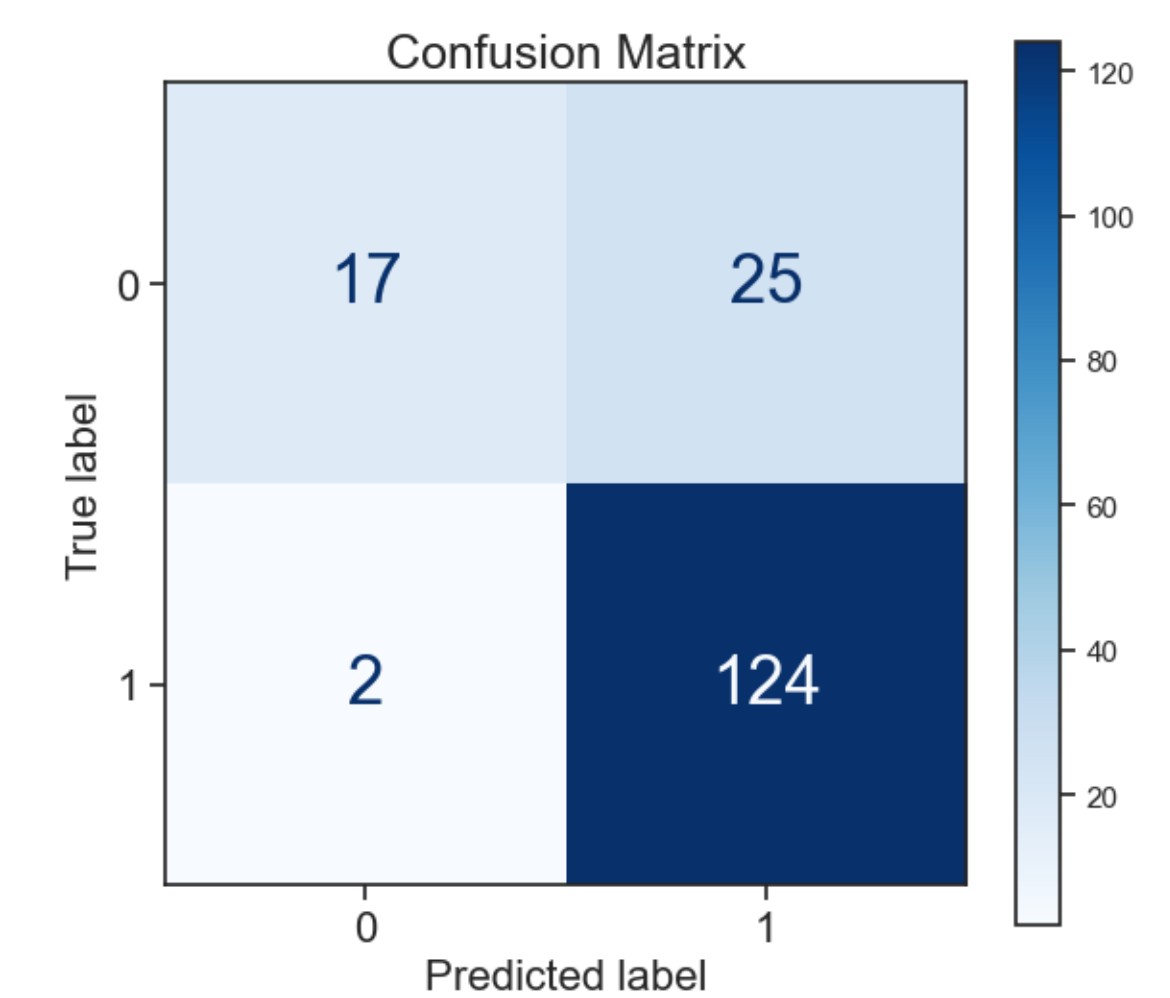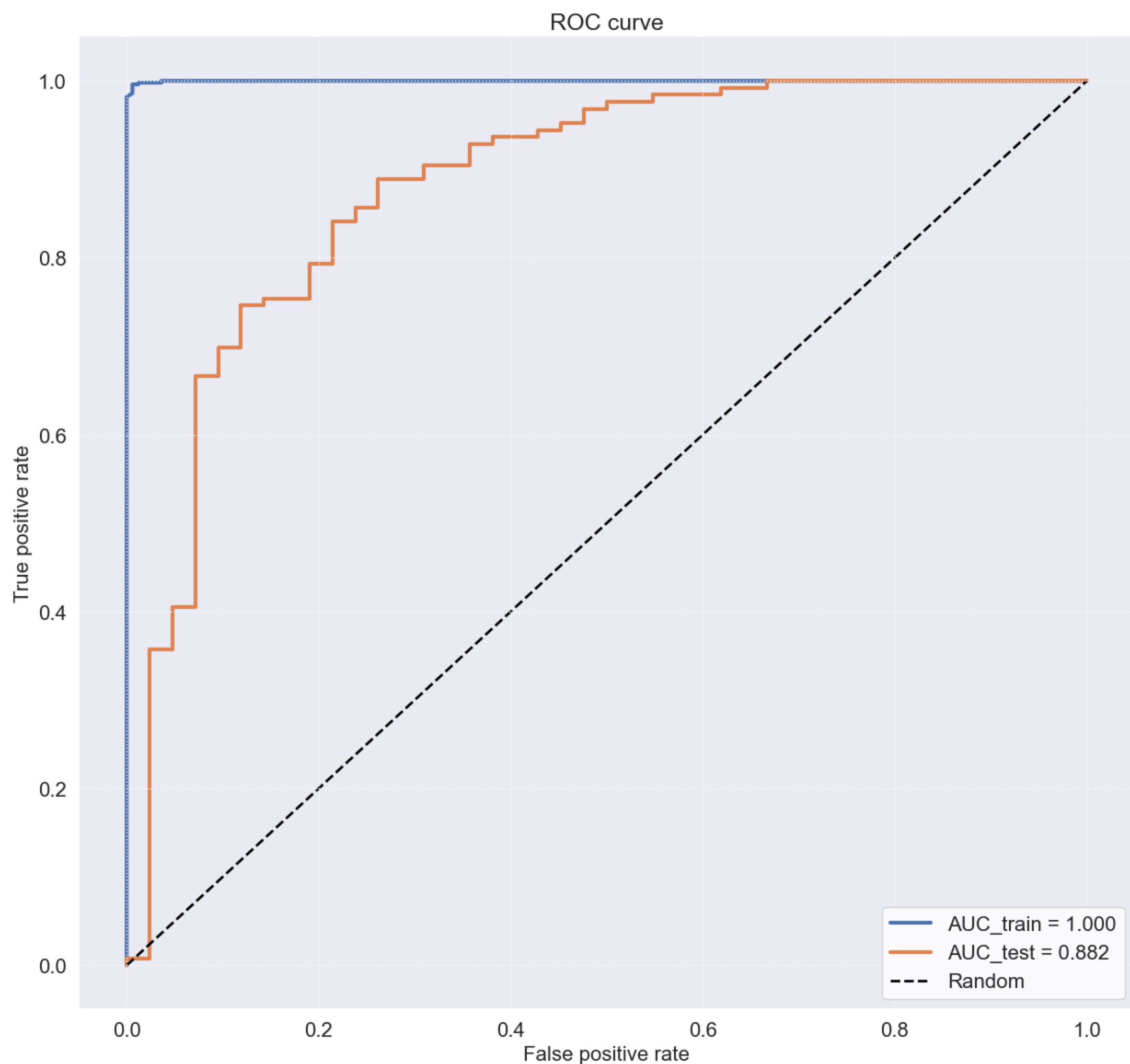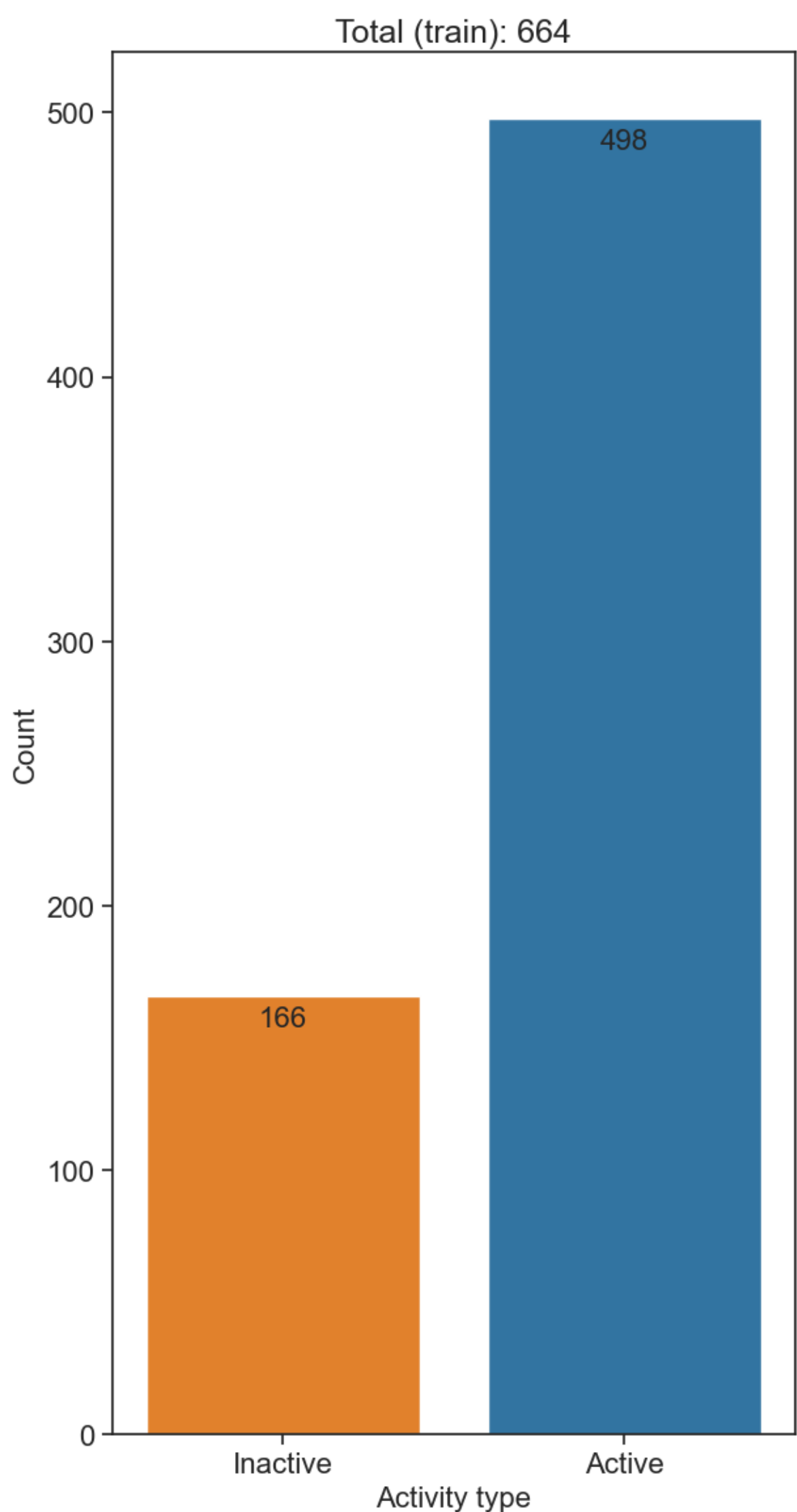
Metric	Train	Test
AUC	0.9998	0.8815
accuracy	0.9849	0.8393
balanced accuracy	0.9699	0.6944
recall_(sens)	1.0	0.9841
specificity	0.9398	0.4048
precision	0.9803	0.8322
f1_score	0.9848	0.8157