Protein: P25100 - metric: balanced_accuracy
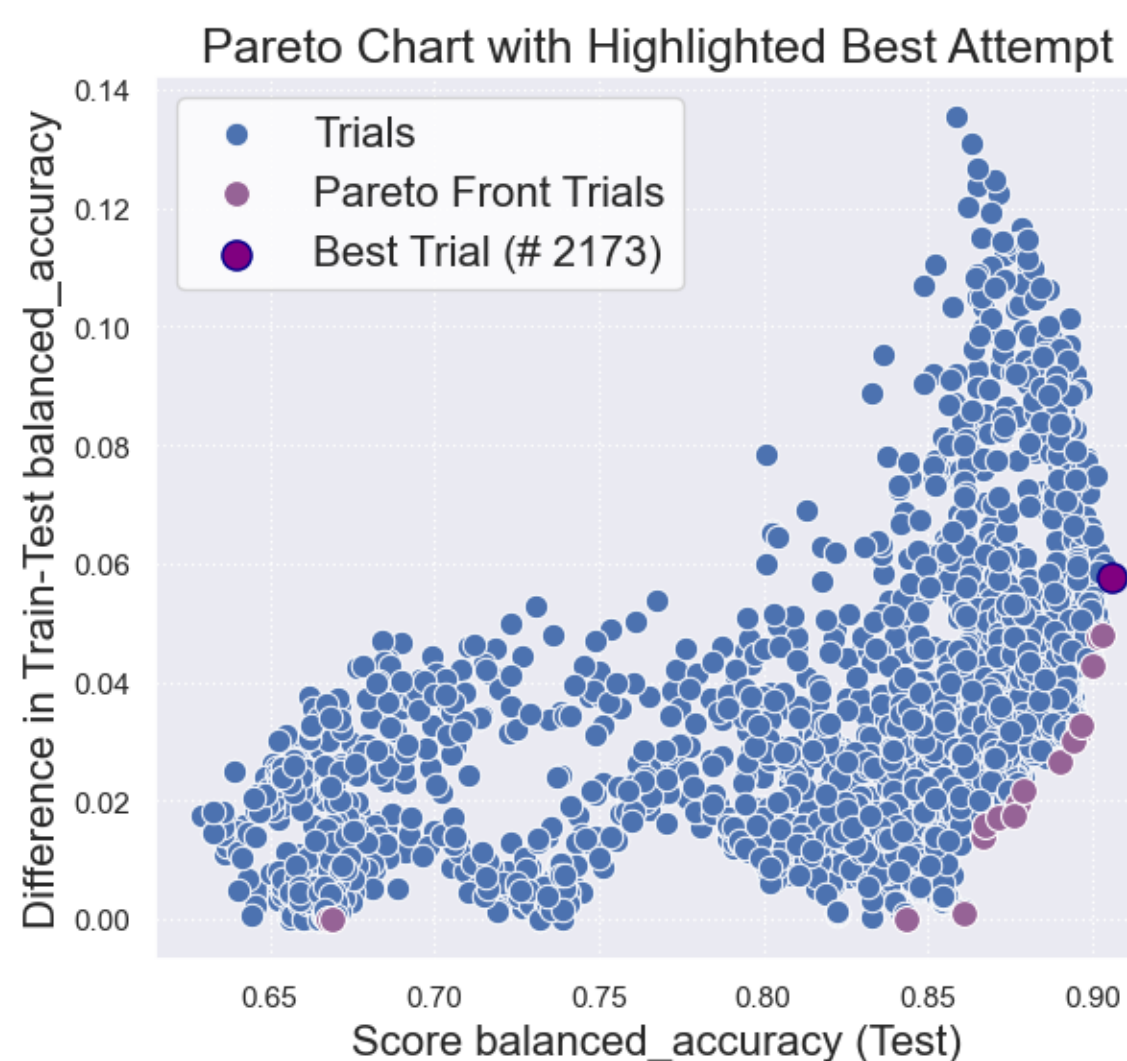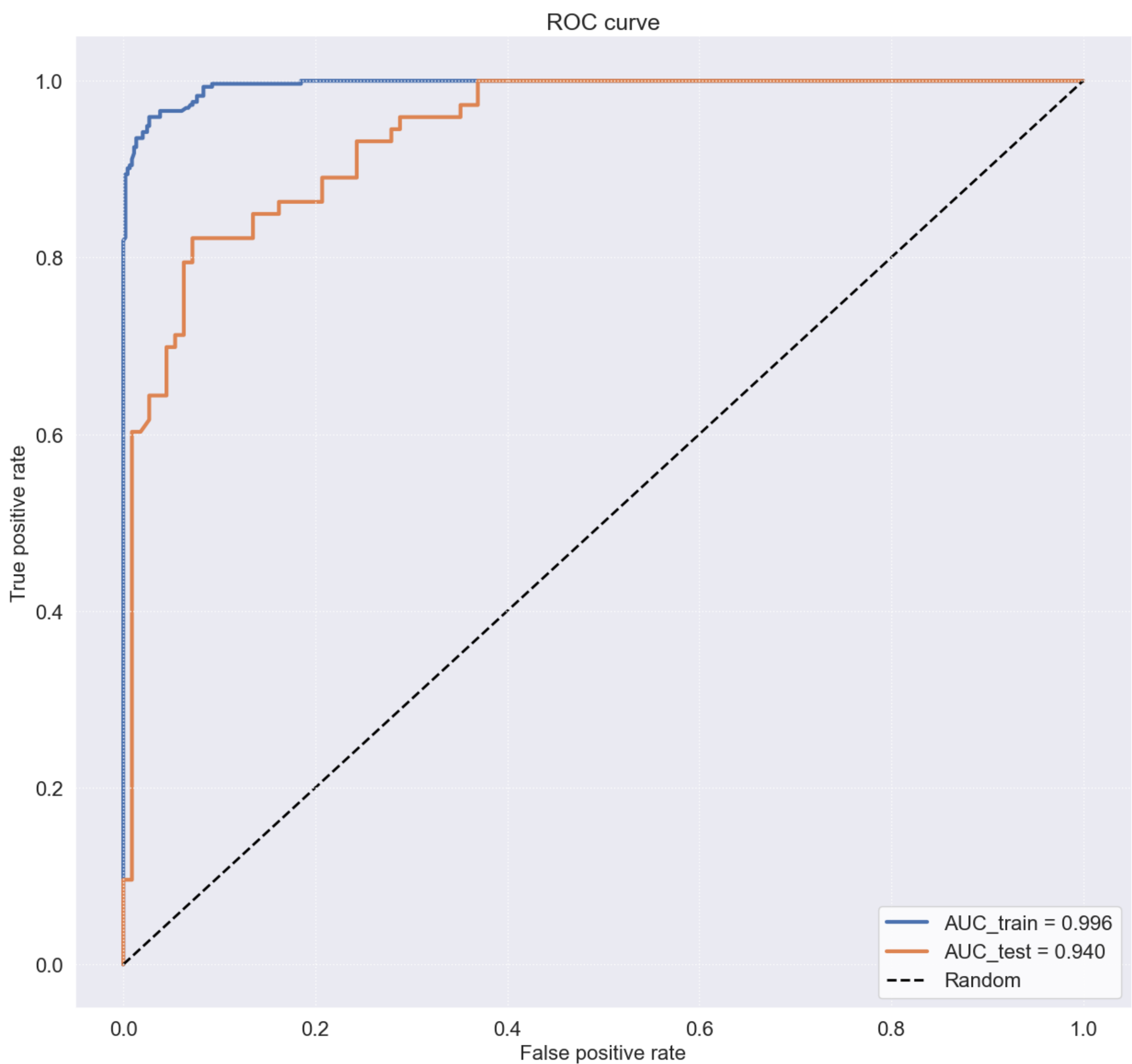
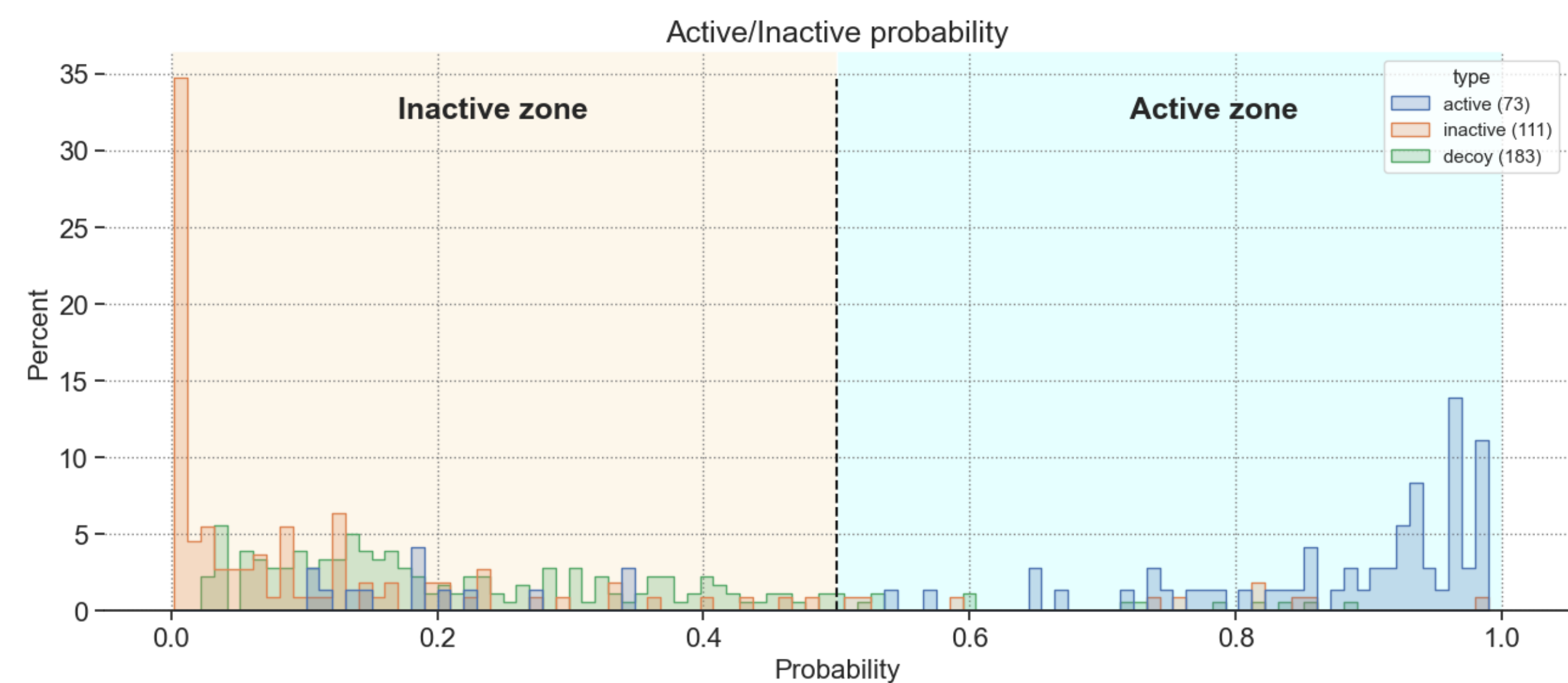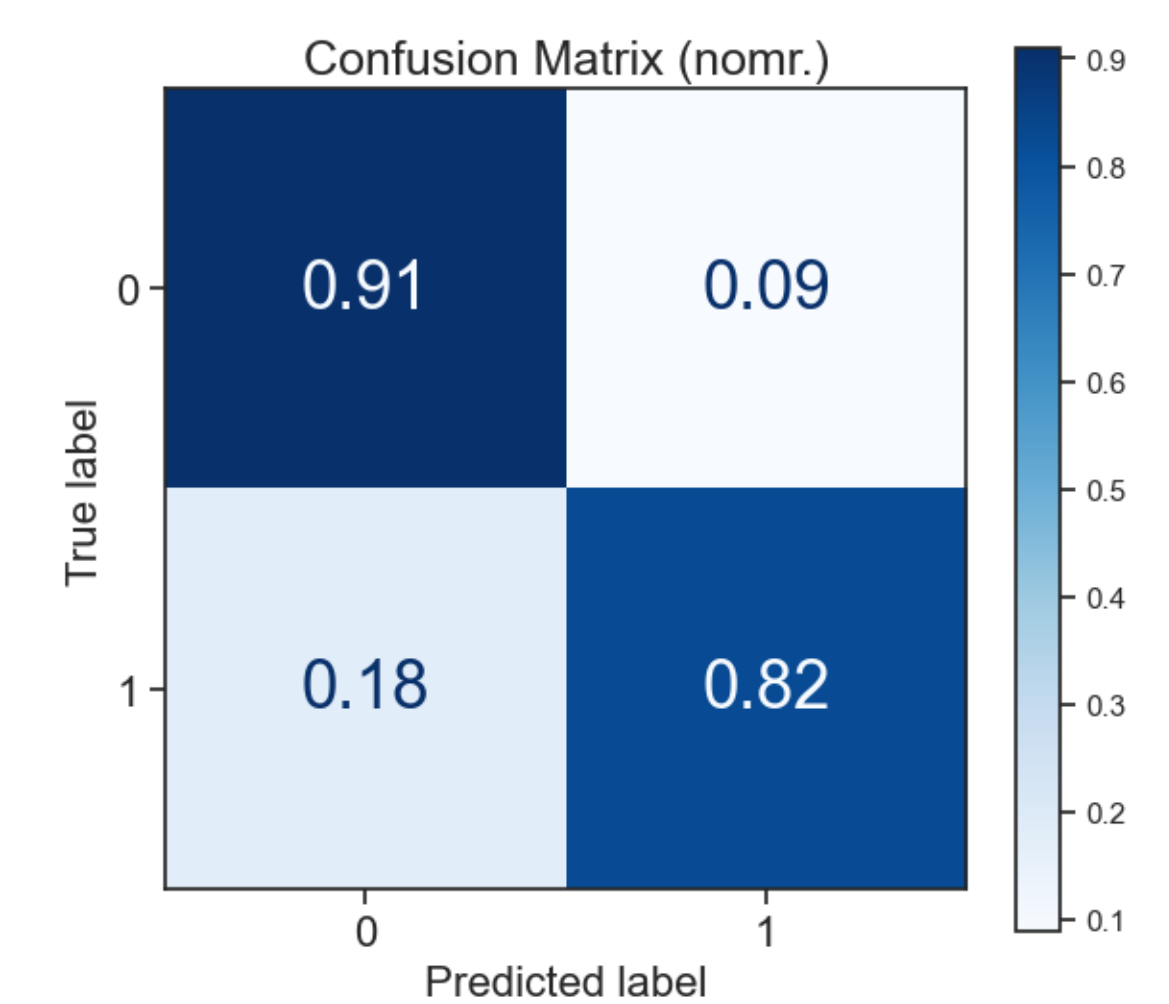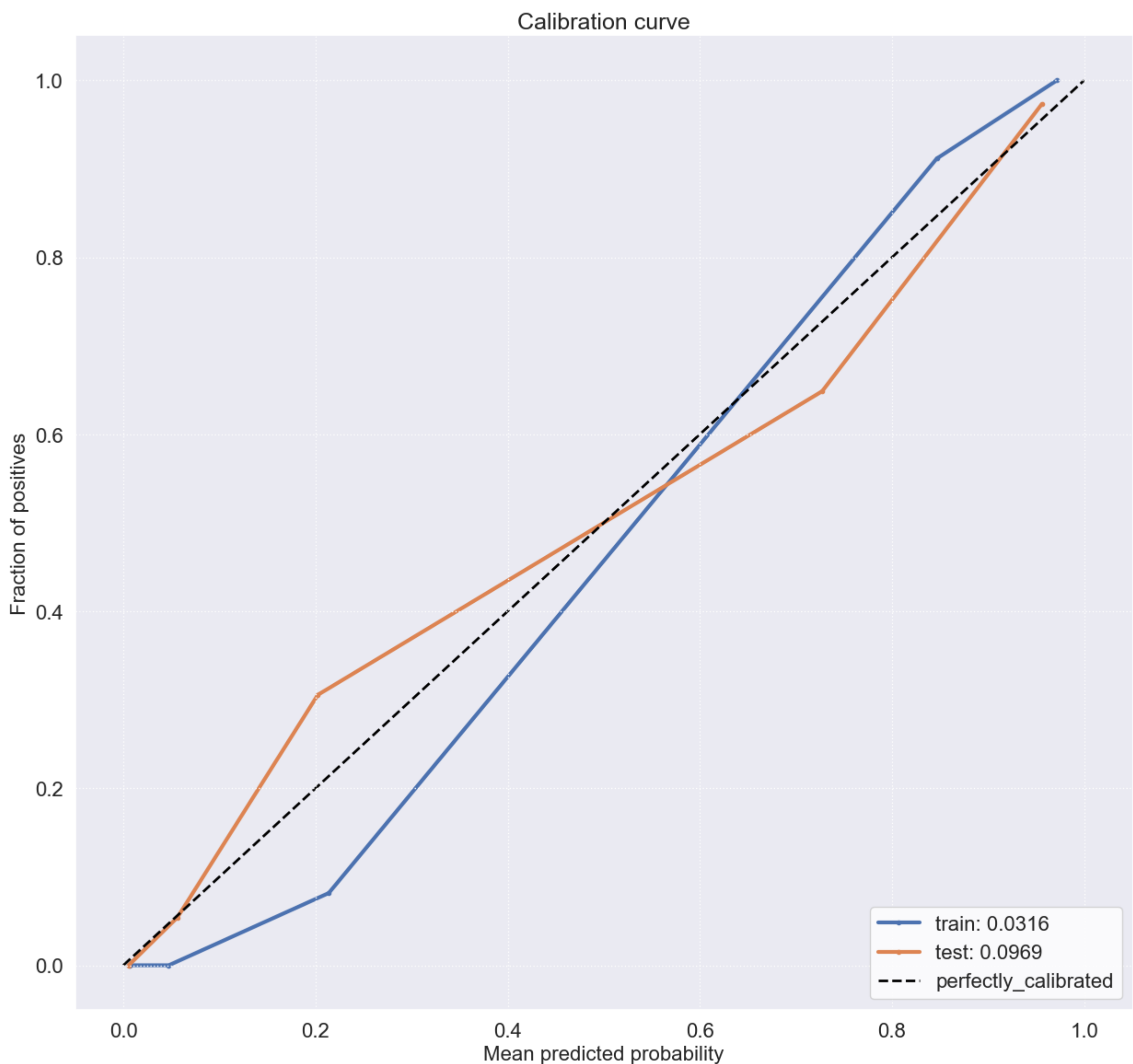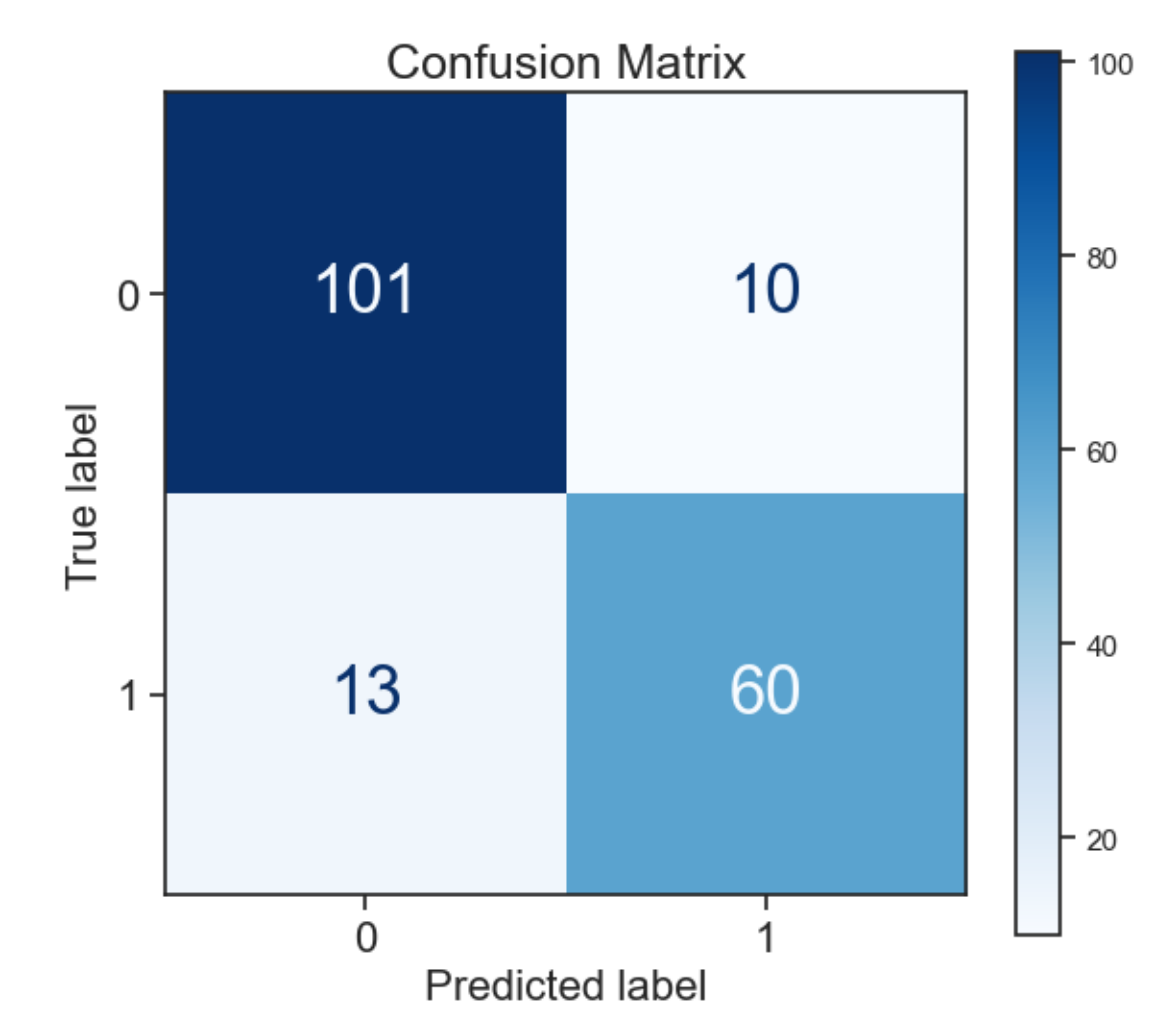| XGBoost Parameters	Value
alpha	3.6326
colsample_bytree	0.7895
gamma	0.9776
lambda	3.9318
learning_rate	0.0317
max_depth	7.0
min_child_weight	1.0
n_estimators	474.0
reg_alpha	0.4099
reg_lambda	0.0001
scale_pos_weight	1.4569
subsample	0.9314

Metric	Train	Test
AUC	0.9956	0.9397
accuracy	0.9606	0.875
balanced accuracy	0.9603	0.8659
recall_(sens)	0.959	0.8219
specificity	0.9616	0.9099
precision	0.943	0.8571
f1_score	0.9607	0.8745

Protein: P25100 - metric: f1

Total (train): 736

ROC curve

AUC_train = 0.996
AUC_test = 0.946
Random

Pareto Chart with Highlighted Best Attempt

Trials
Pareto Front Trials
Best Trial (# 2077)
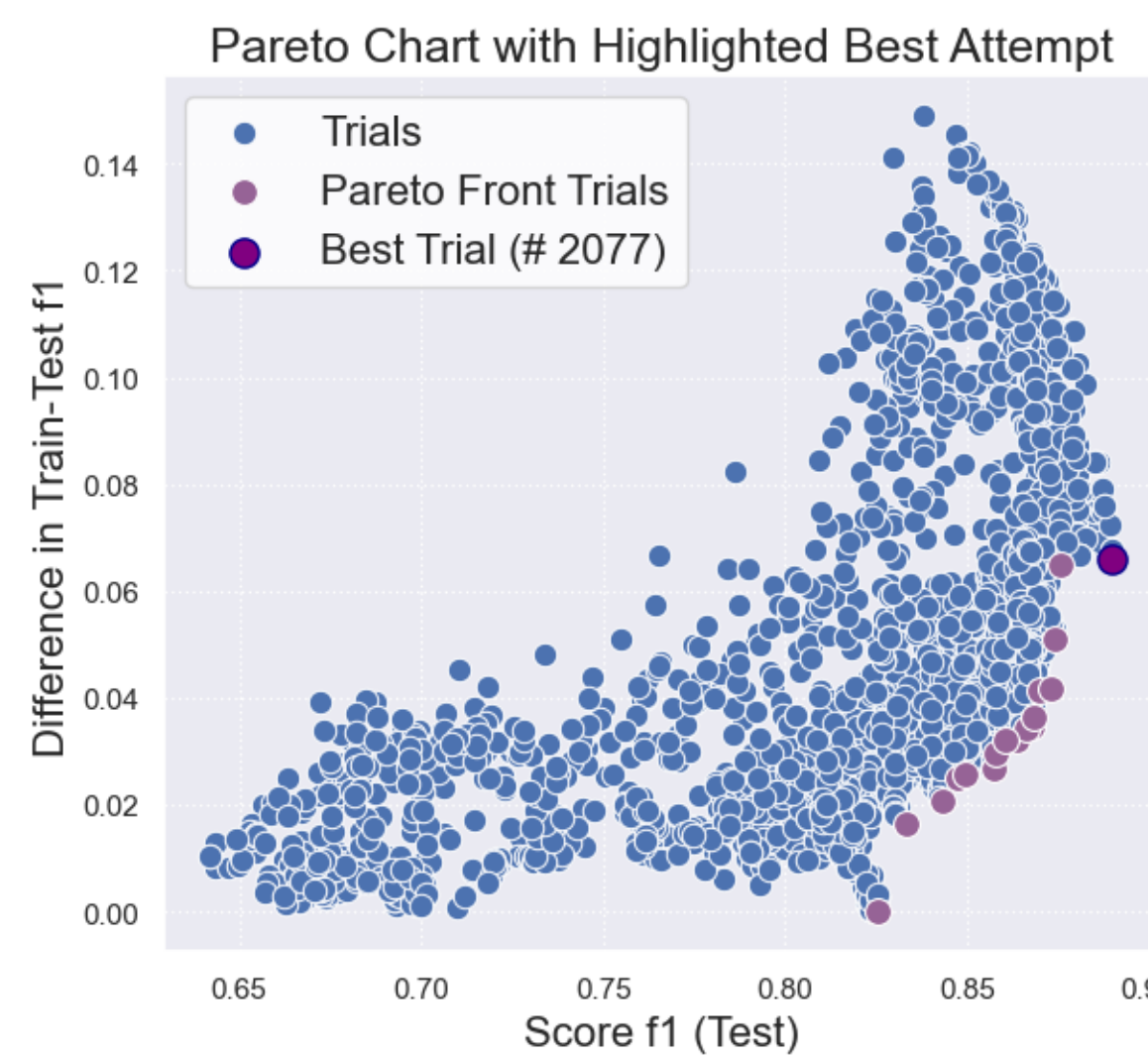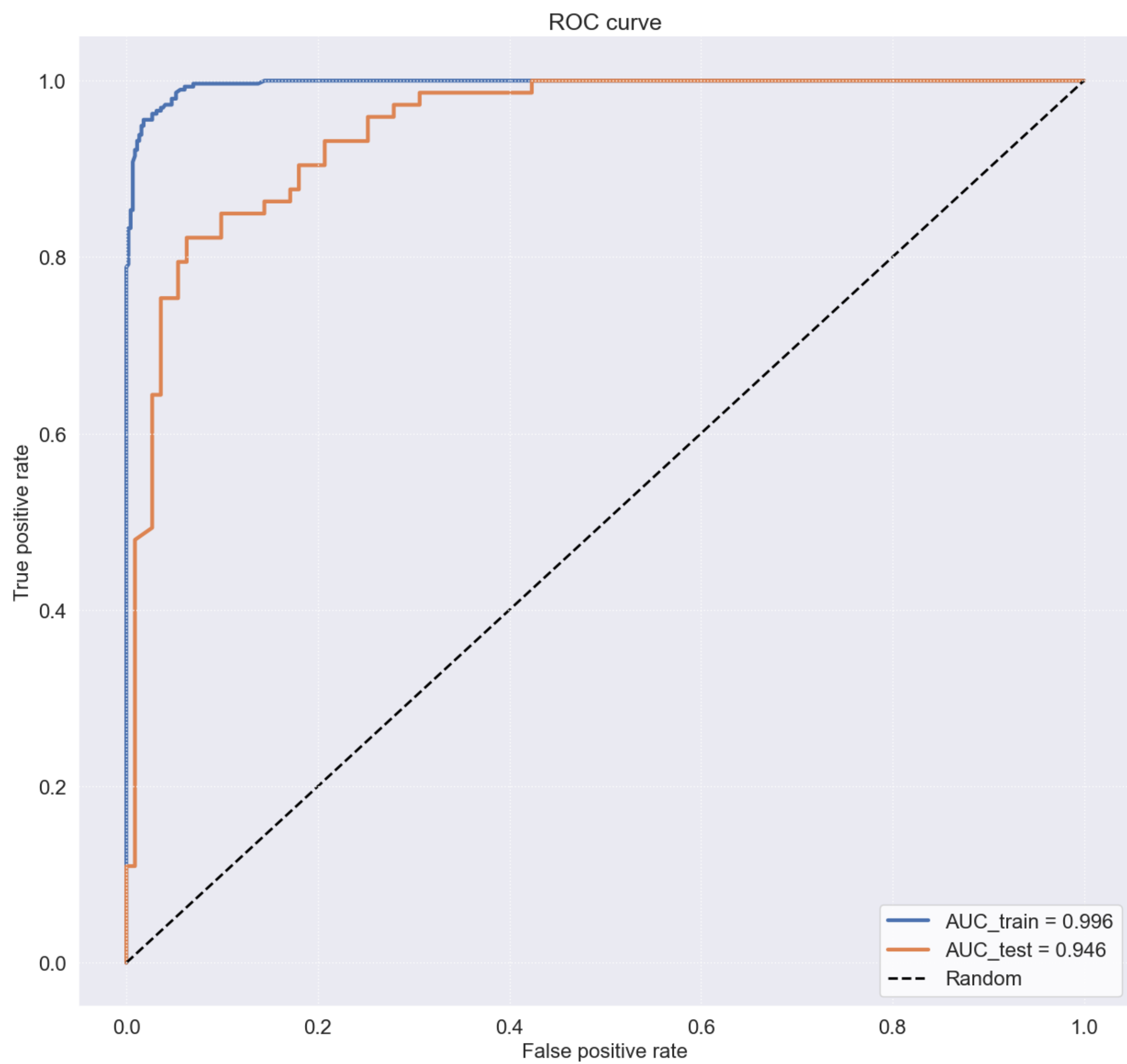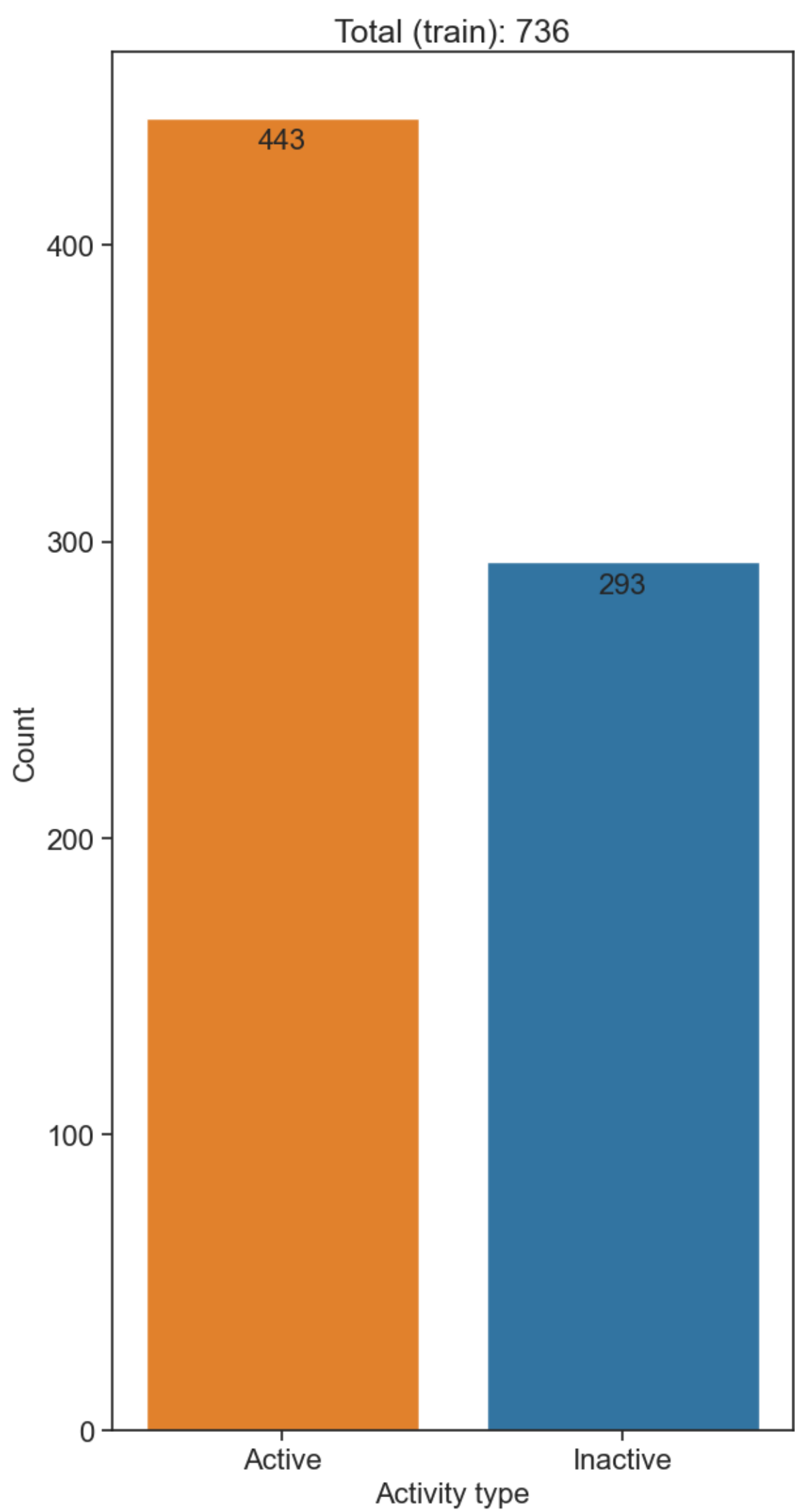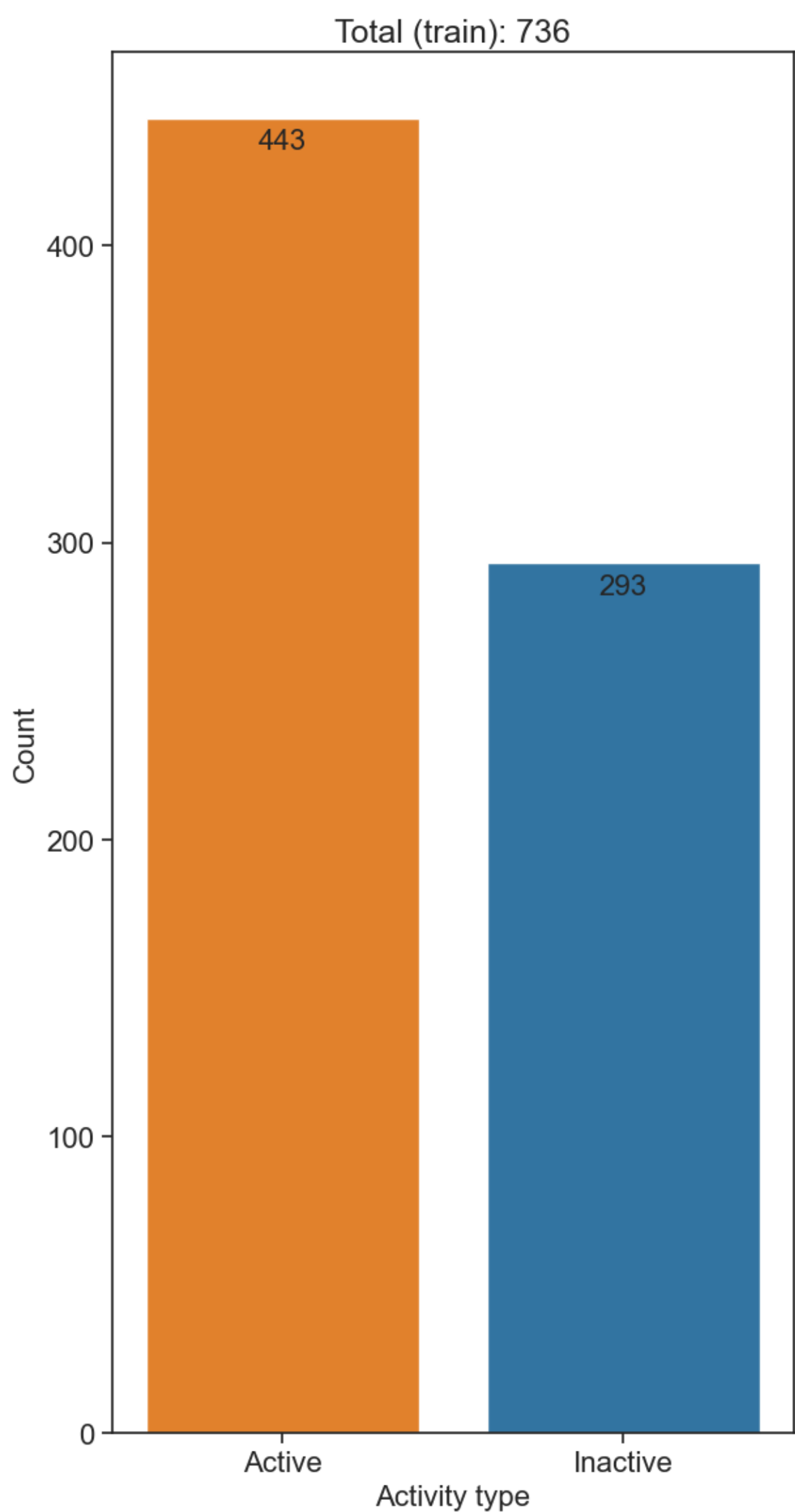
XGBoost Parameters	Value
alpha	3.0111
colsample_bytree	0.6001
gamma	0.5059
lambda	4.8668
learning_rate	0.0931
max_depth	9.0
min_child_weight	1.0
n_estimators	169.0
reg_alpha	0.6445
reg_lambda	0.9401
scale_pos_weight	1.8423
subsample	0.9915

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0303
test: 0.0885
perfectly_calibrated

Active/Inactive probability

Inactive zone   Active zone

type
active (73)
inactive (111)
decoy (183)
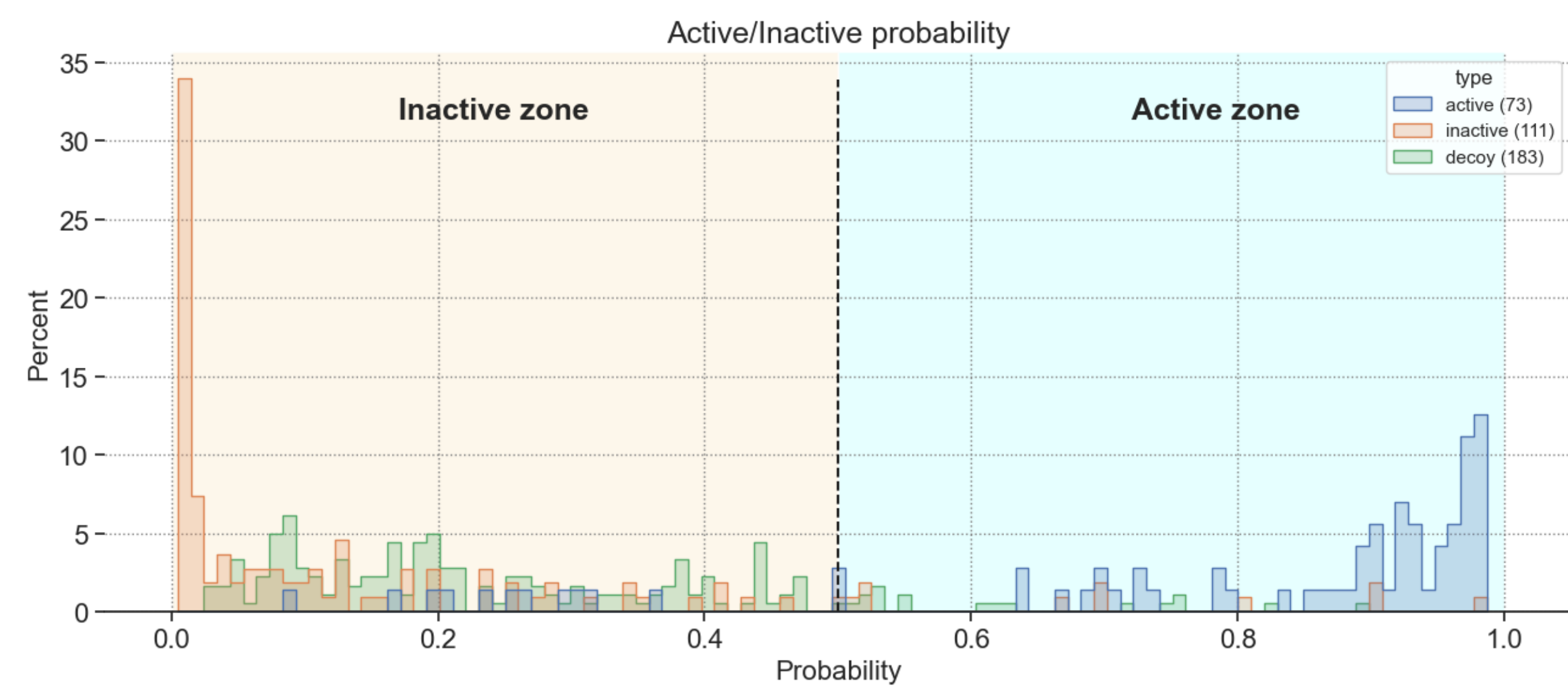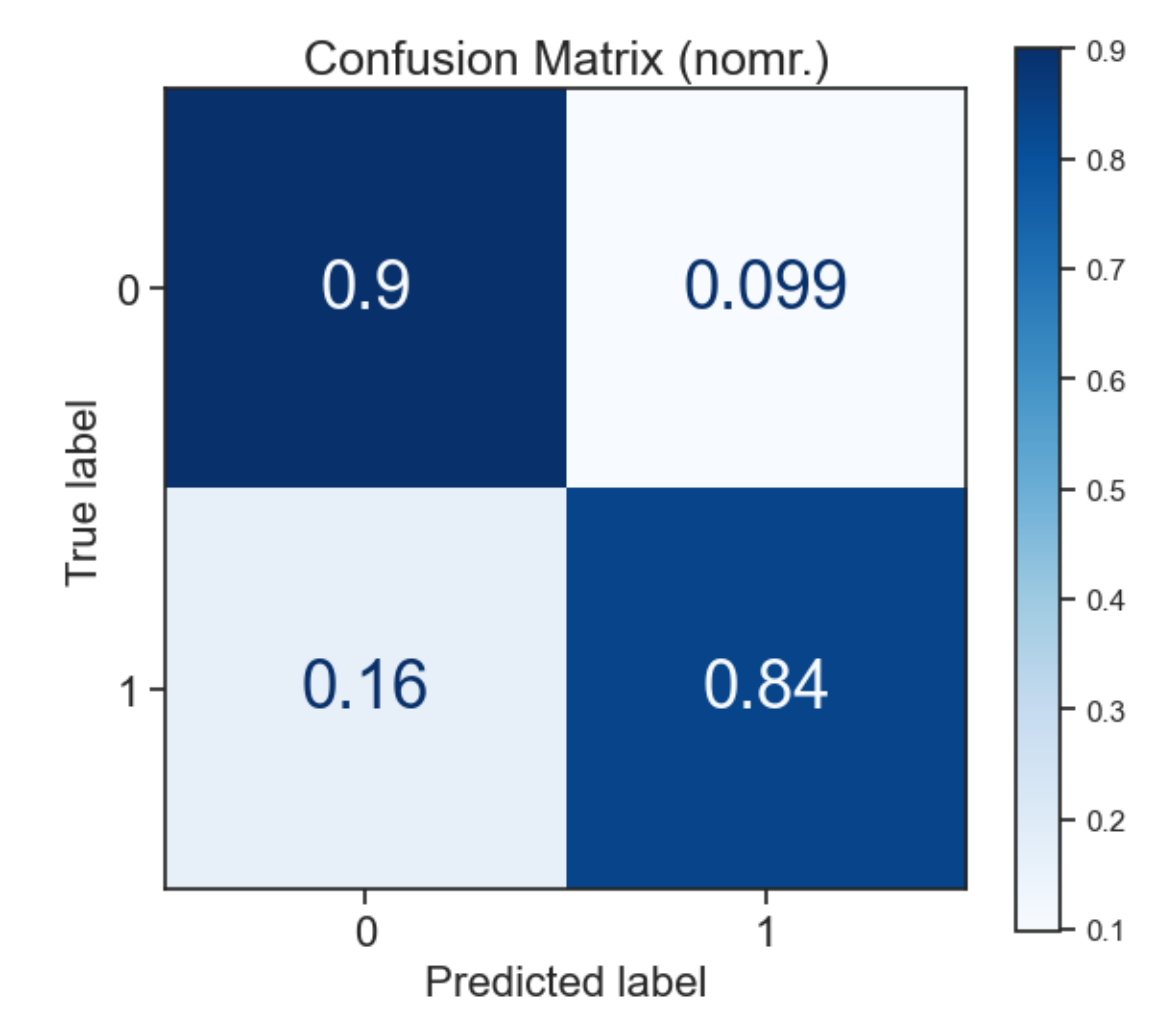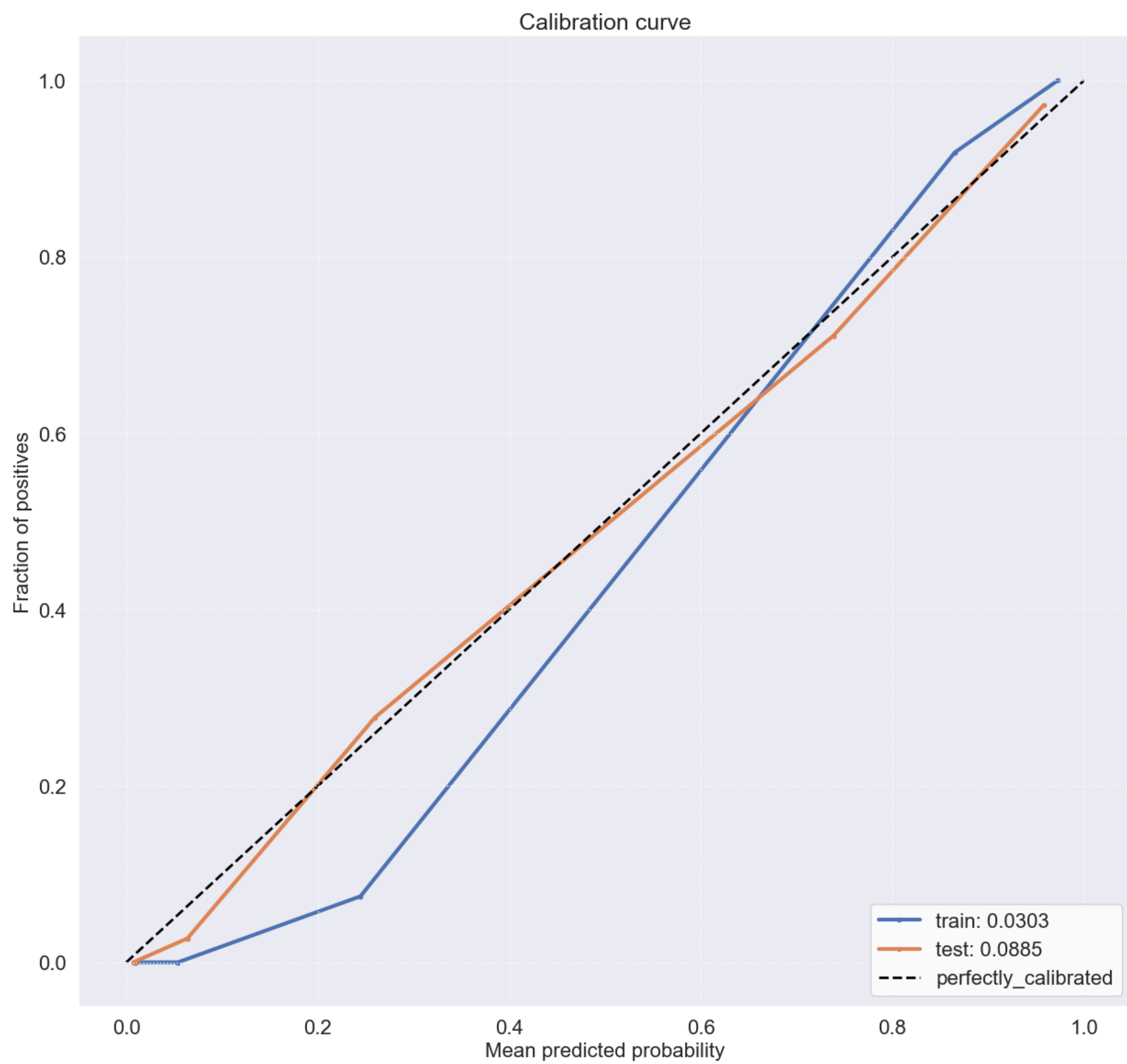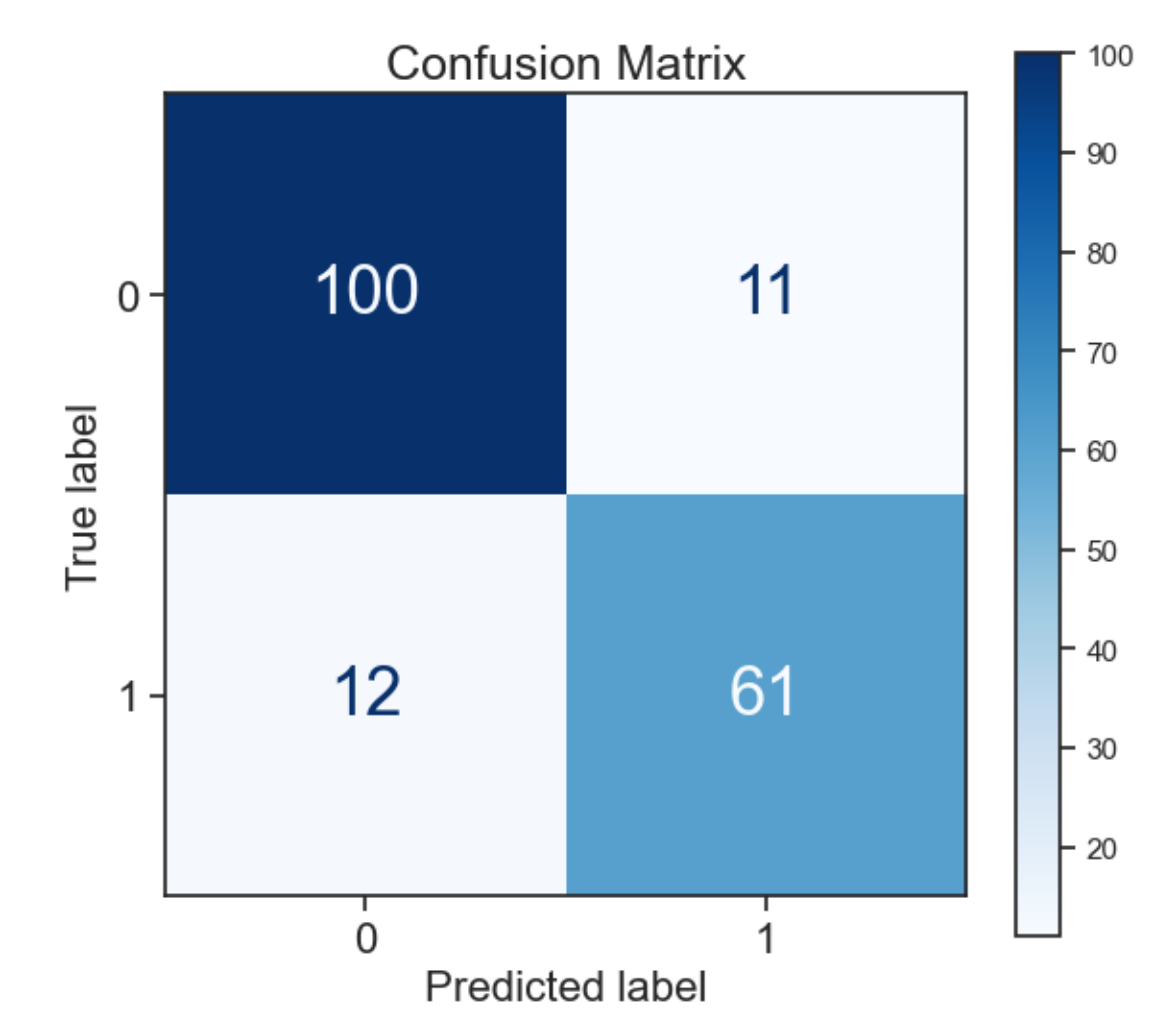
Metric	Train	Test
AUC	0.9965	0.9461
accuracy	0.9647	0.875
balanced accuracy	0.966	0.8683
recall_(sens)	0.9727	0.8356
specificity	0.9594	0.9009
precision	0.9406	0.8472
f1_score	0.9648	0.8748

Protein: P25100 - metric: precision

Total (train): 736

ROC curve

- AUC_train = 0.998
- AUC_test = 0.955
- Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 2836)
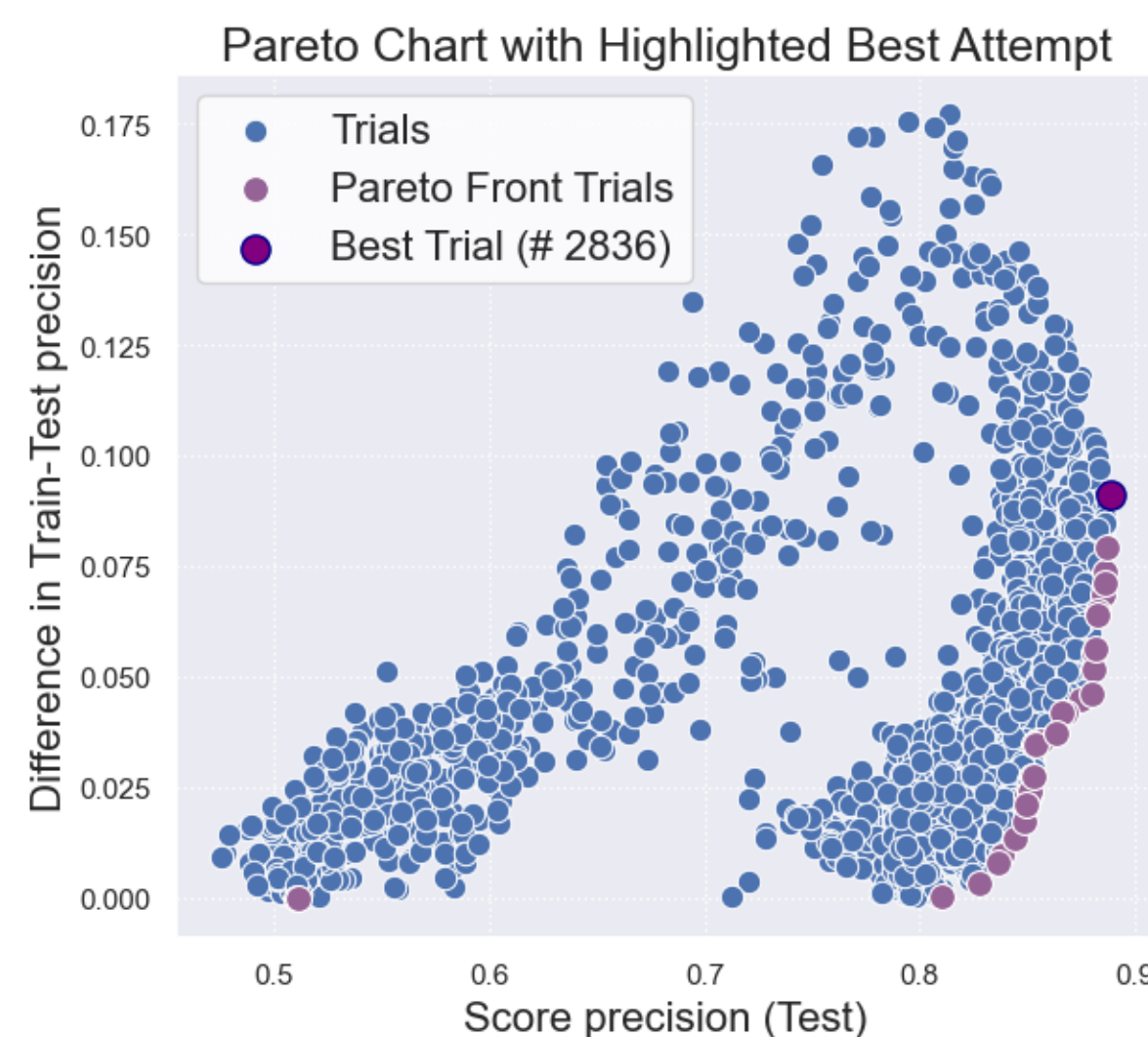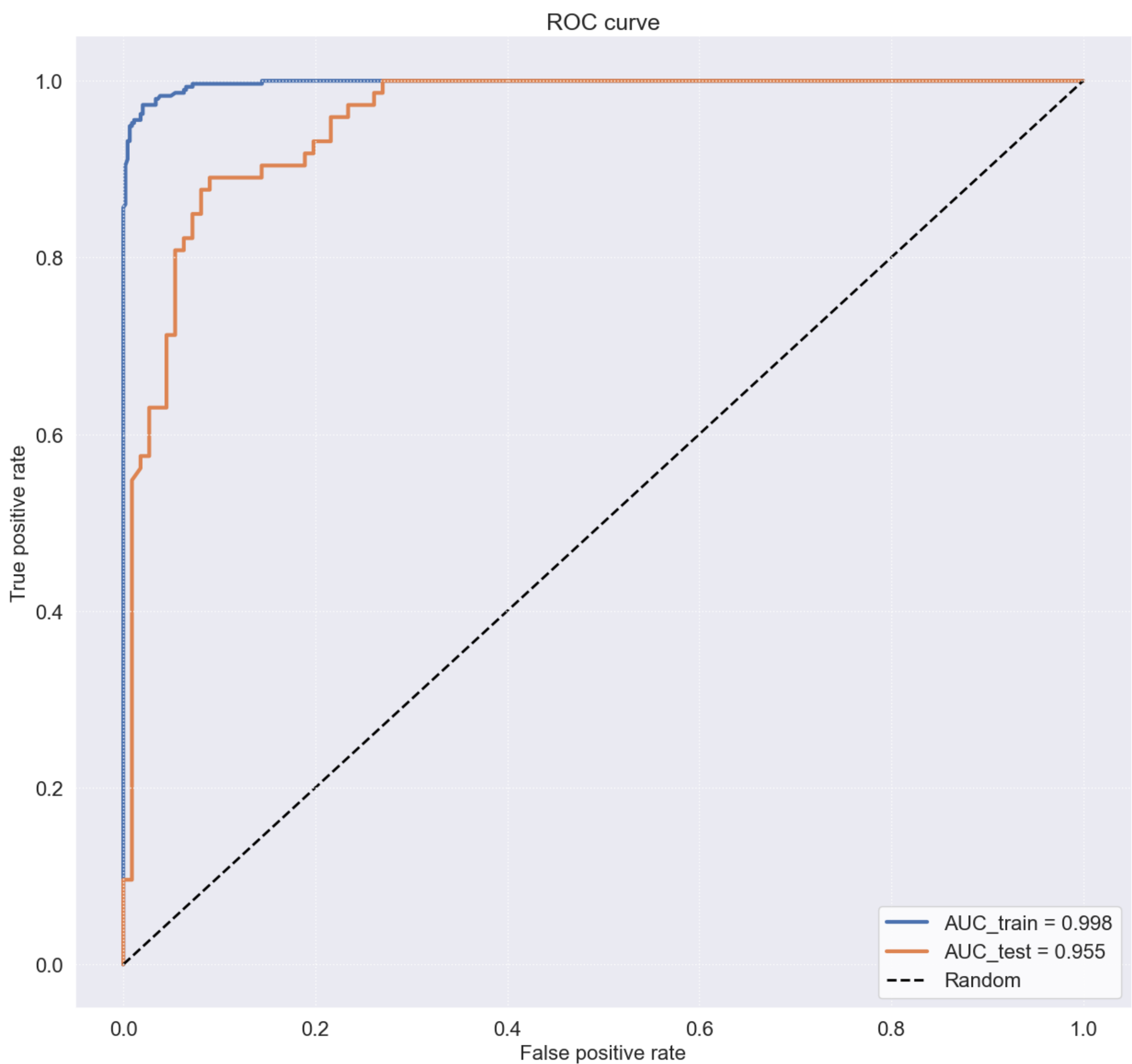
XGBoost Parameters	Value
alpha	3.3533
colsample_bytree	0.6315
gamma	0.2572
lambda	2.6961
learning_rate	0.0058
max_depth	10.0
min_child_weight	1.0
n_estimators	977.0
reg_alpha	0.1646
reg_lambda	0.1431
scale_pos_weight	1.0473
subsample	0.683

Confusion Matrix

Calibration curve

- train: 0.0246
- test: 0.0872
- perfectly_calibrated

Confusion Matrix (nomr.)

Active/Inactive probability

type
- active (73)
- inactive (111)
- decoy (183)
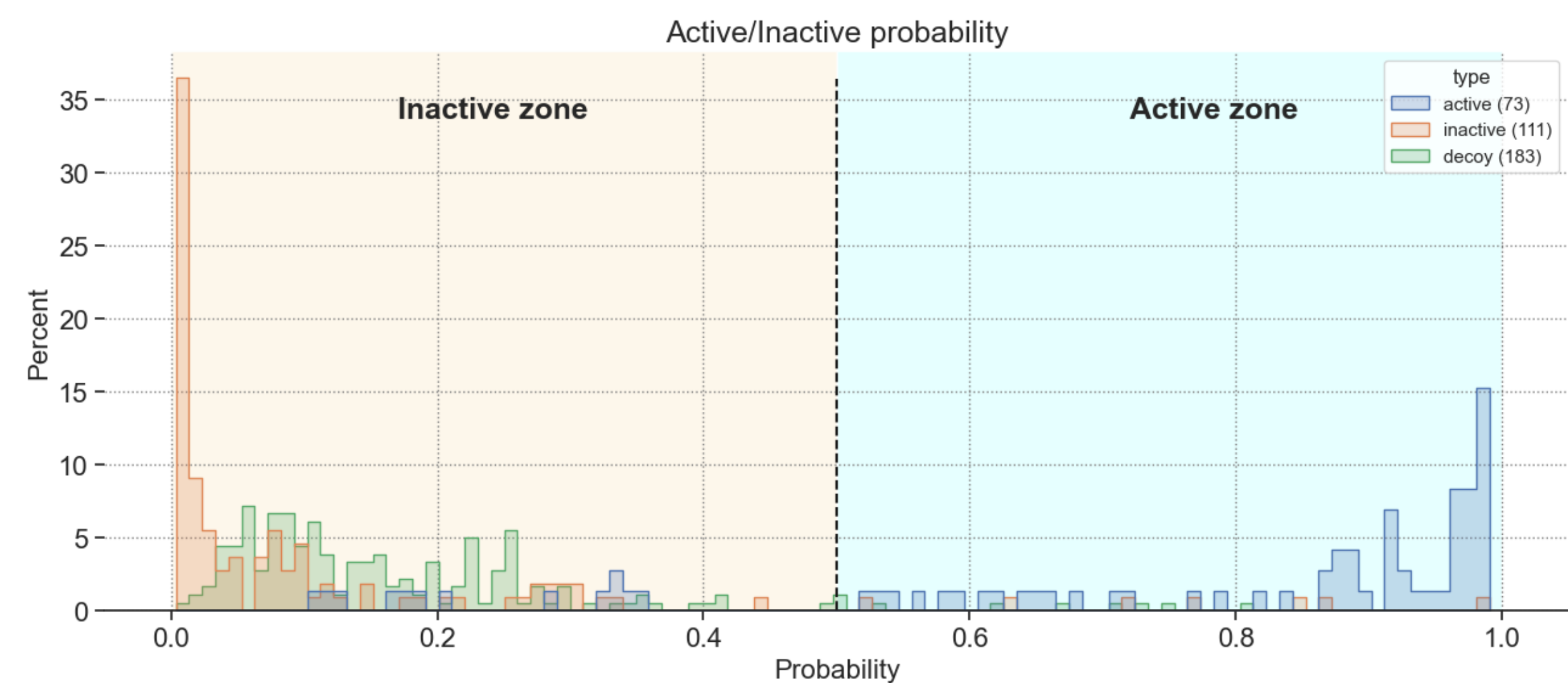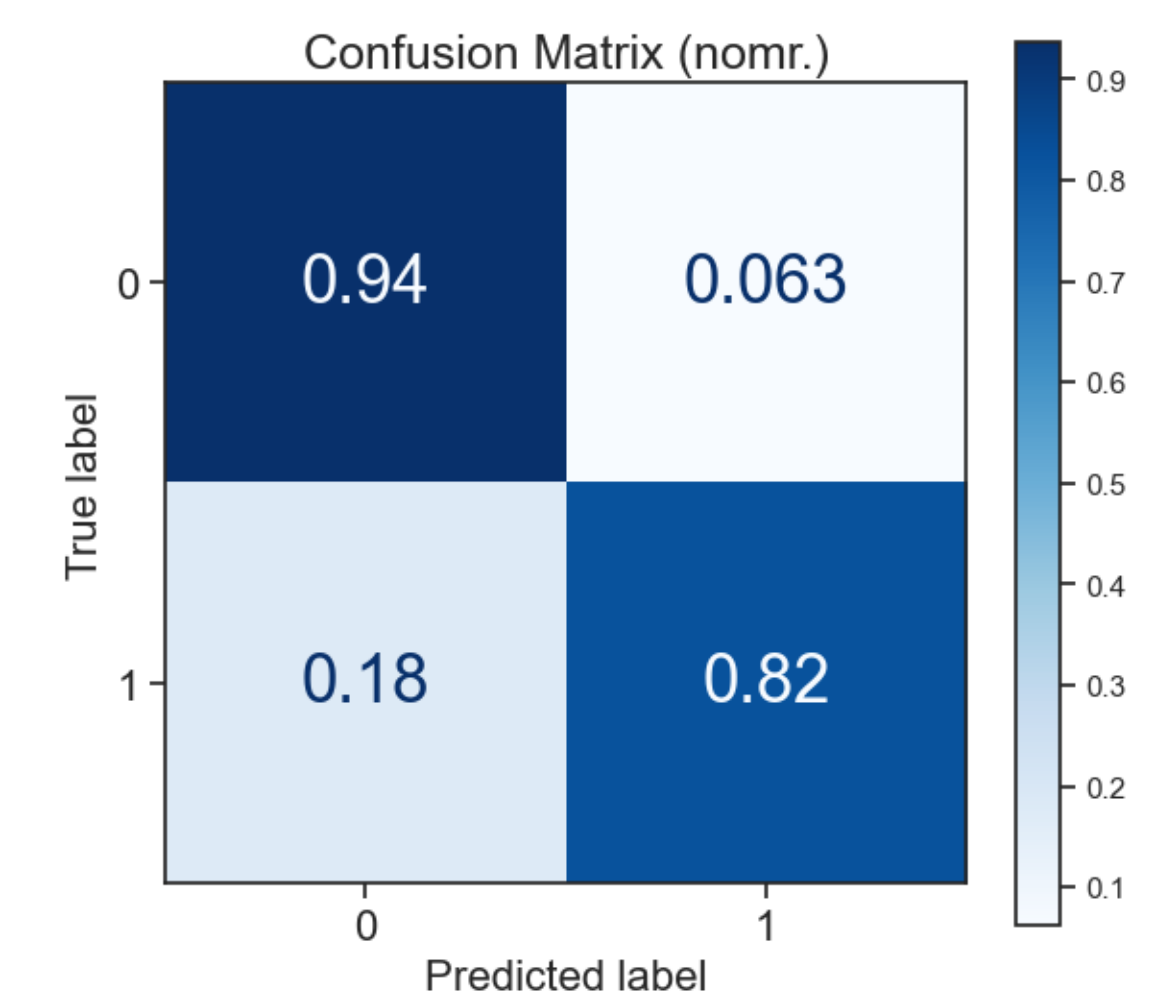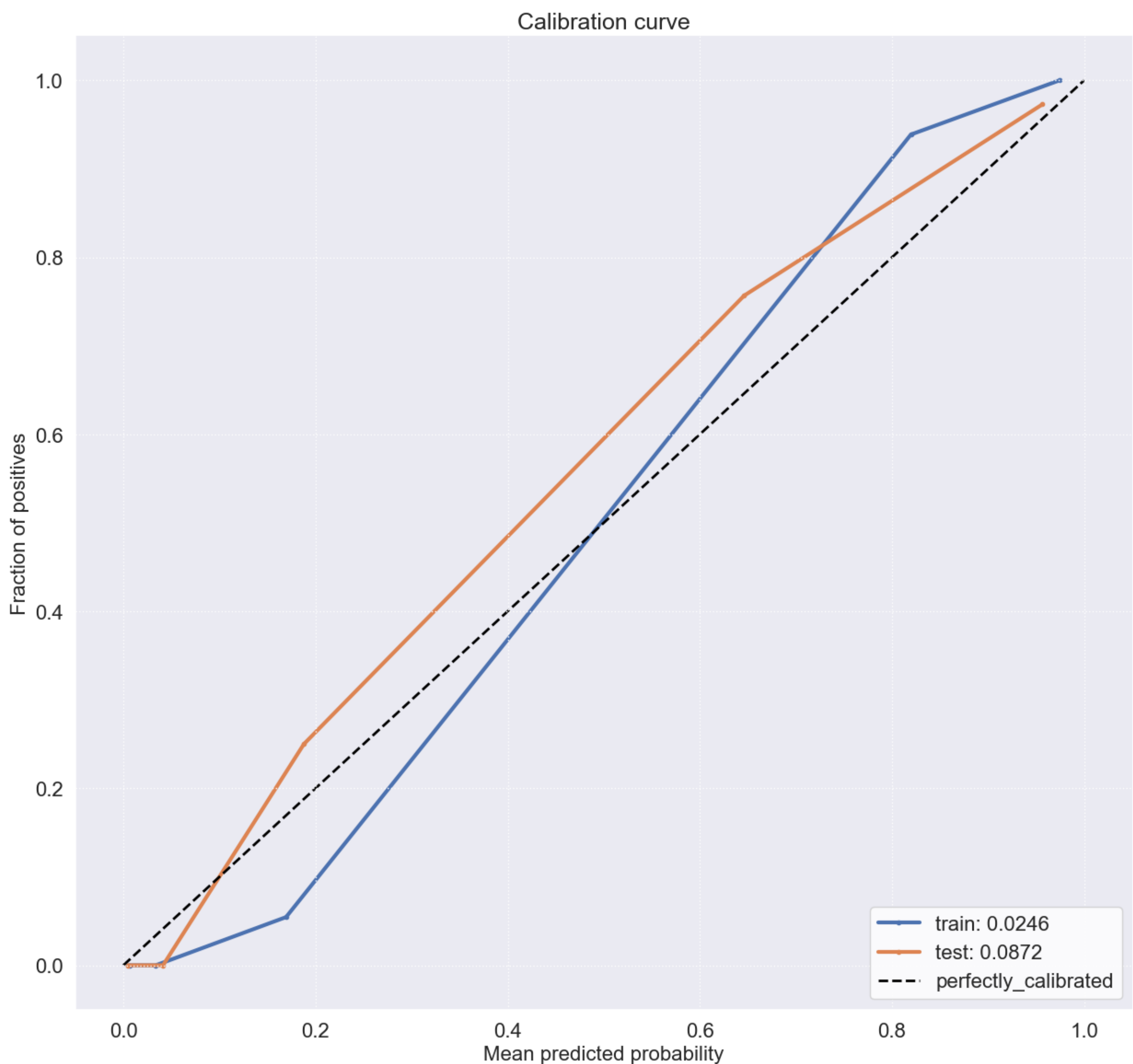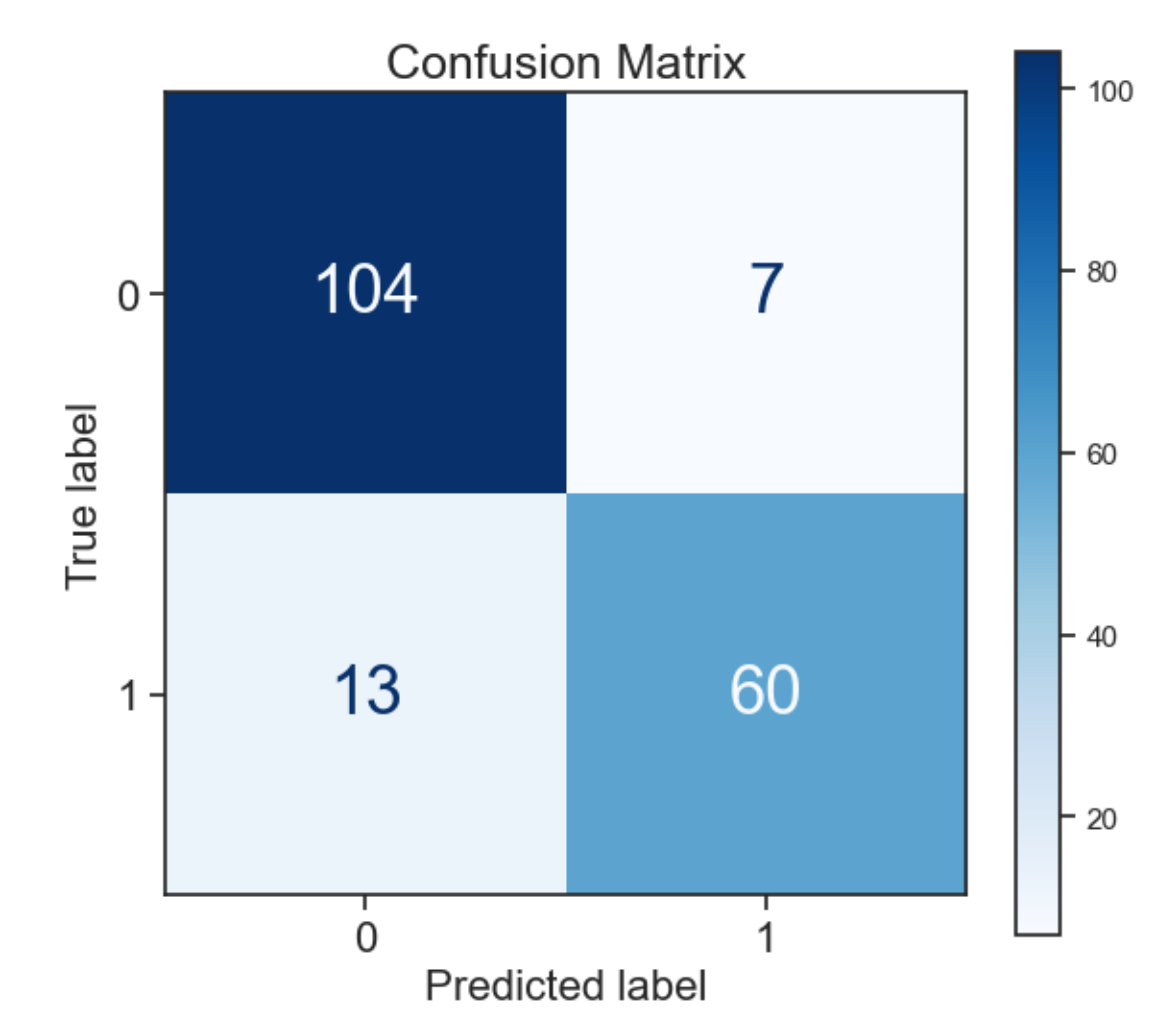
Metric	Train	Test
AUC	0.9976	0.9551
accuracy	0.9728	0.8913
balanced accuracy	0.9699	0.8794
recall_(sens)	0.9556	0.8219
specificity	0.9842	0.9369
precision	0.9756	0.8955
f1_score	0.9728	0.8904

Protein: P25100 - metric: SoftVotingXGBoostEnsemble
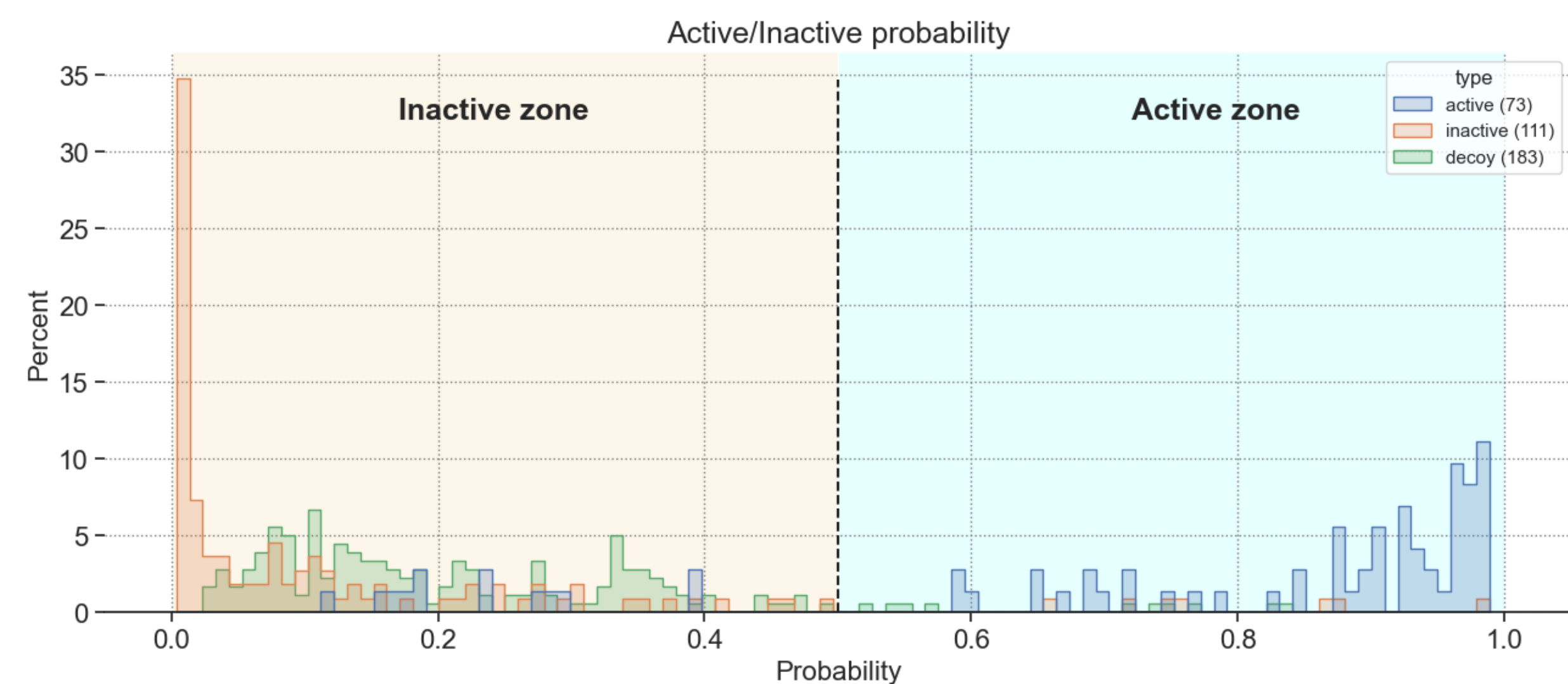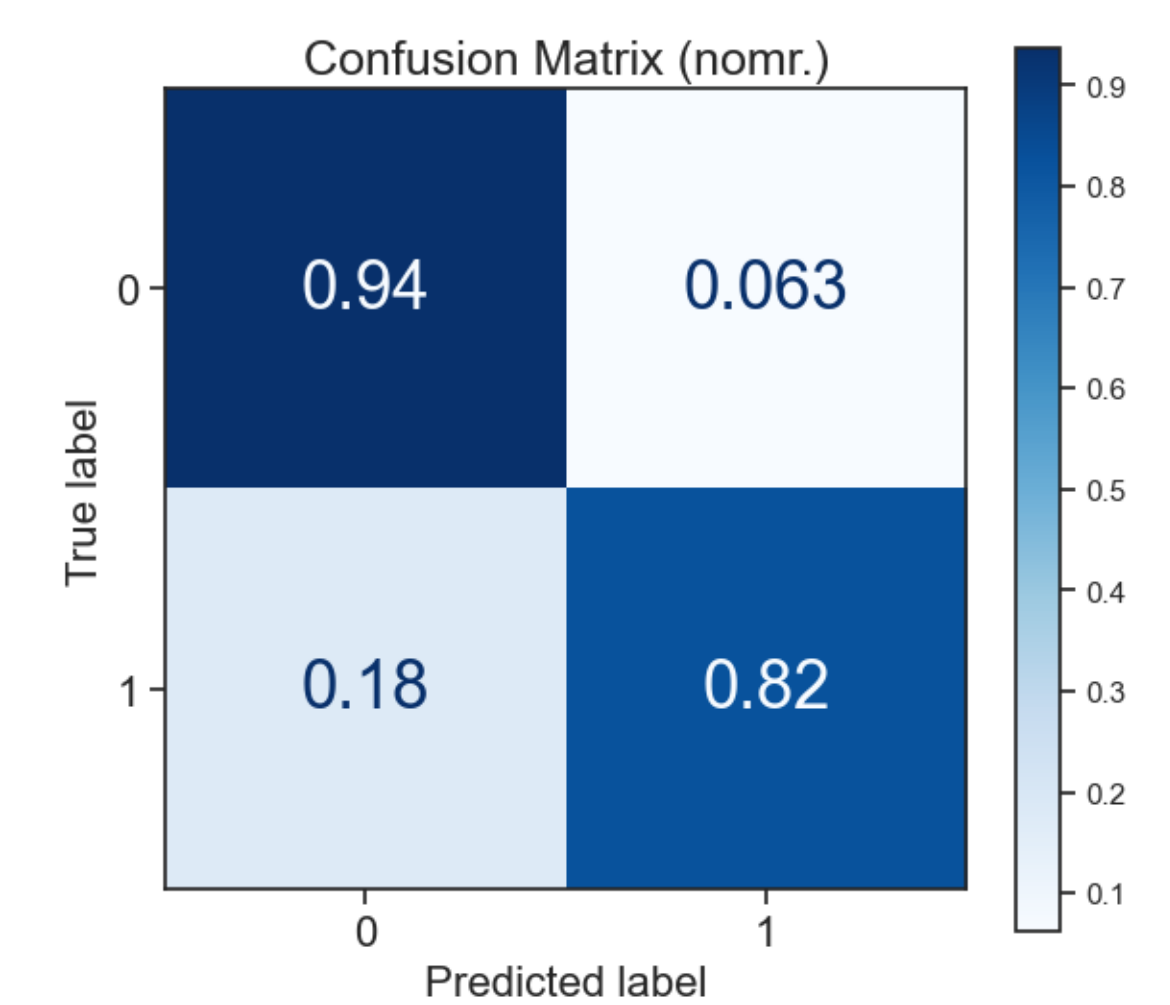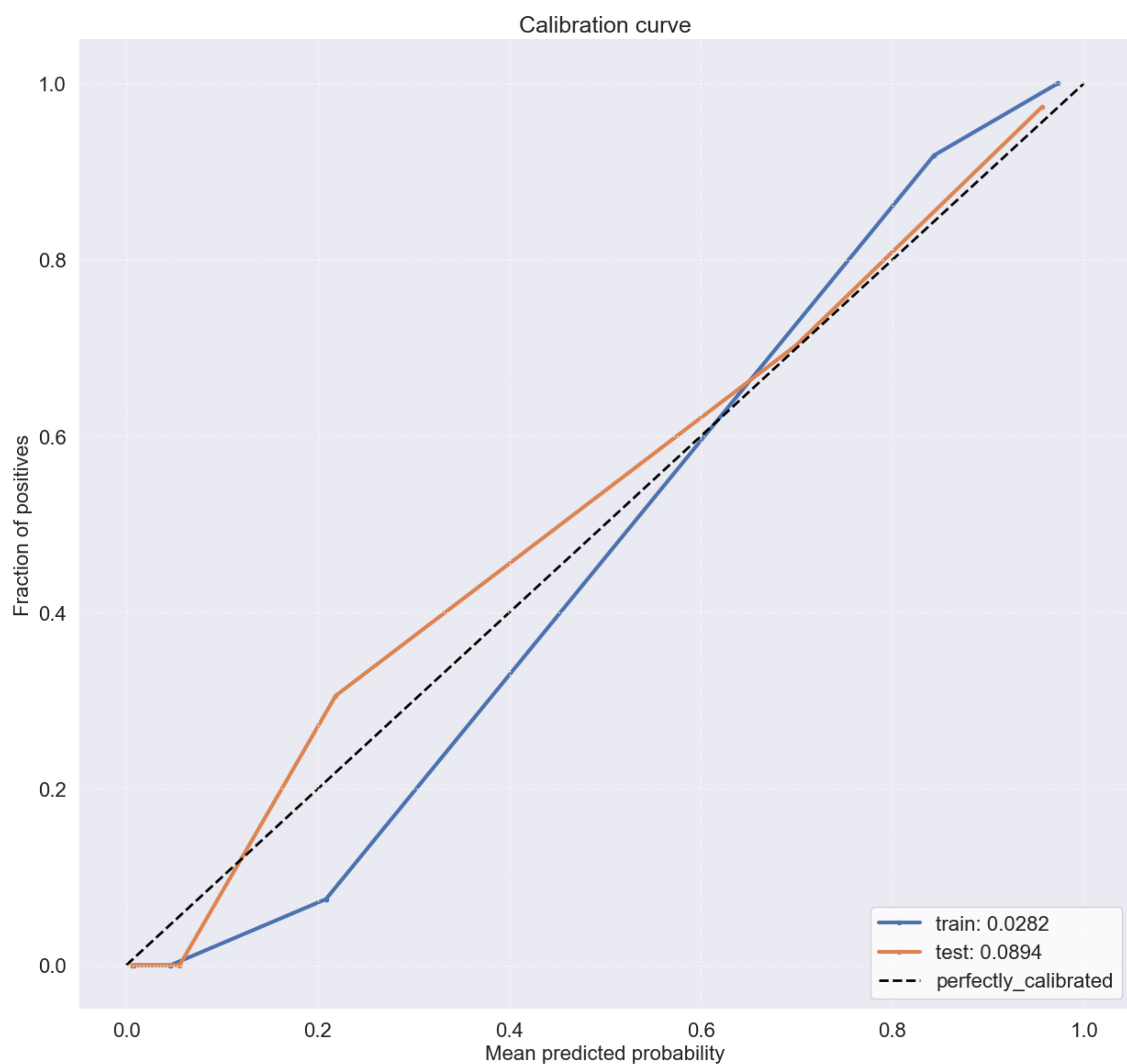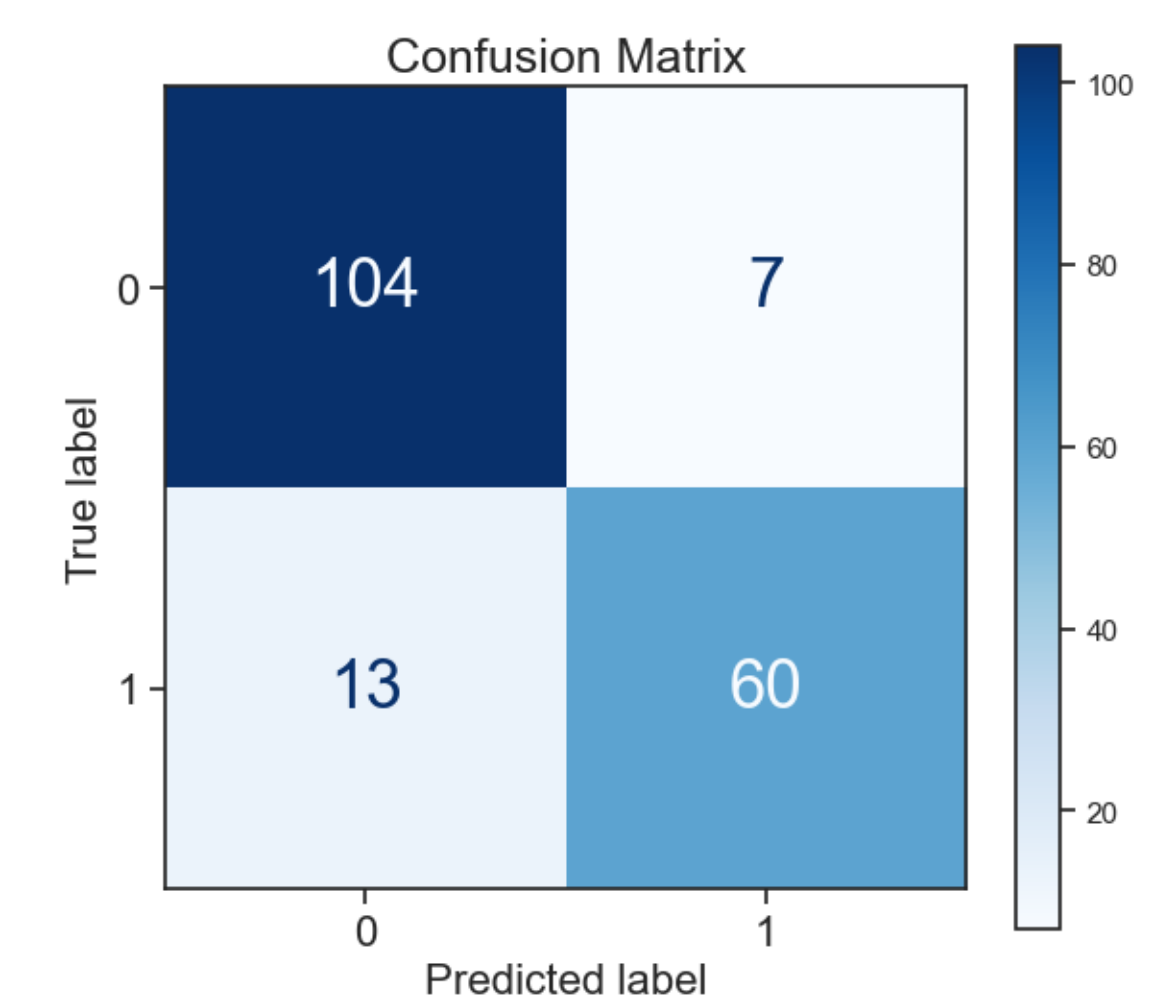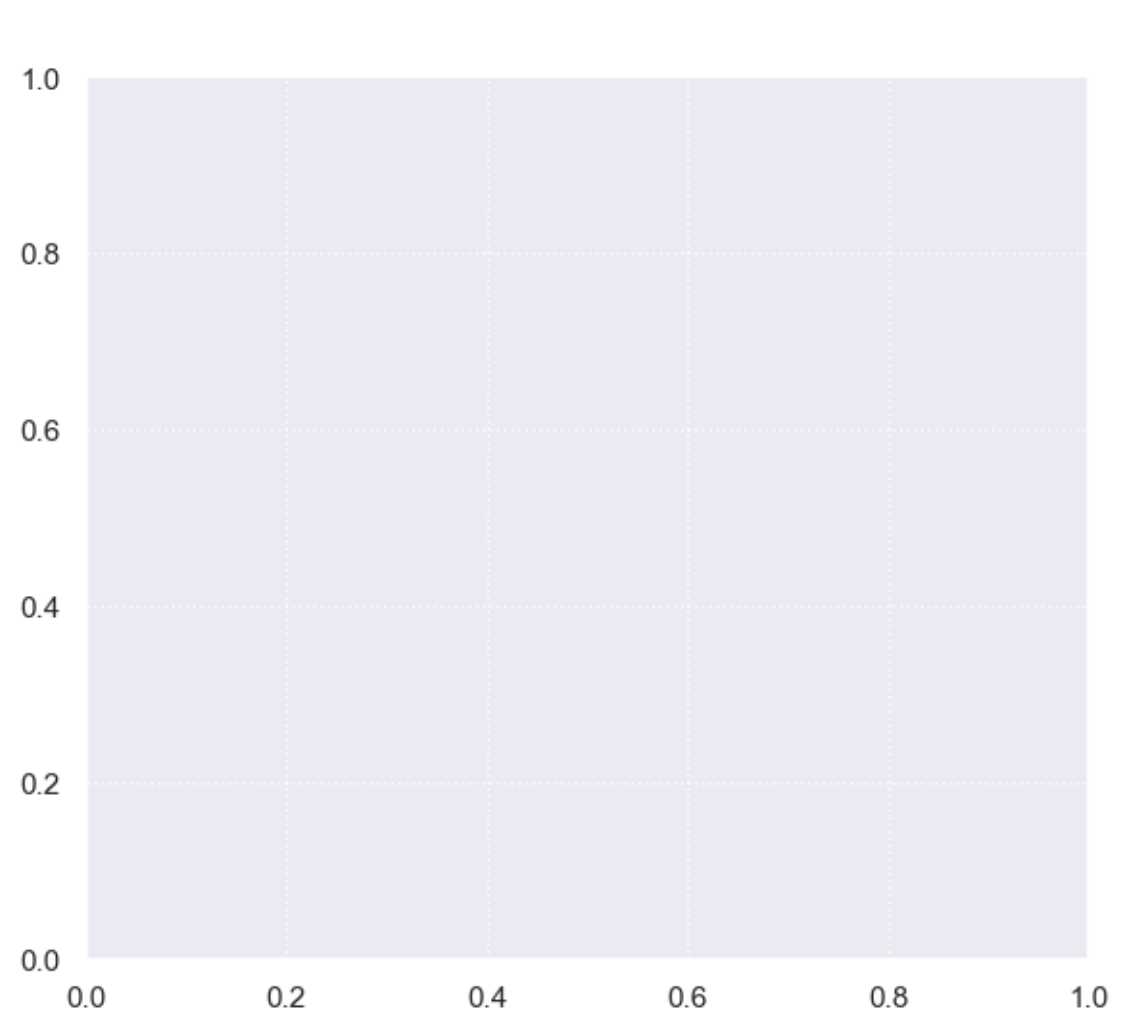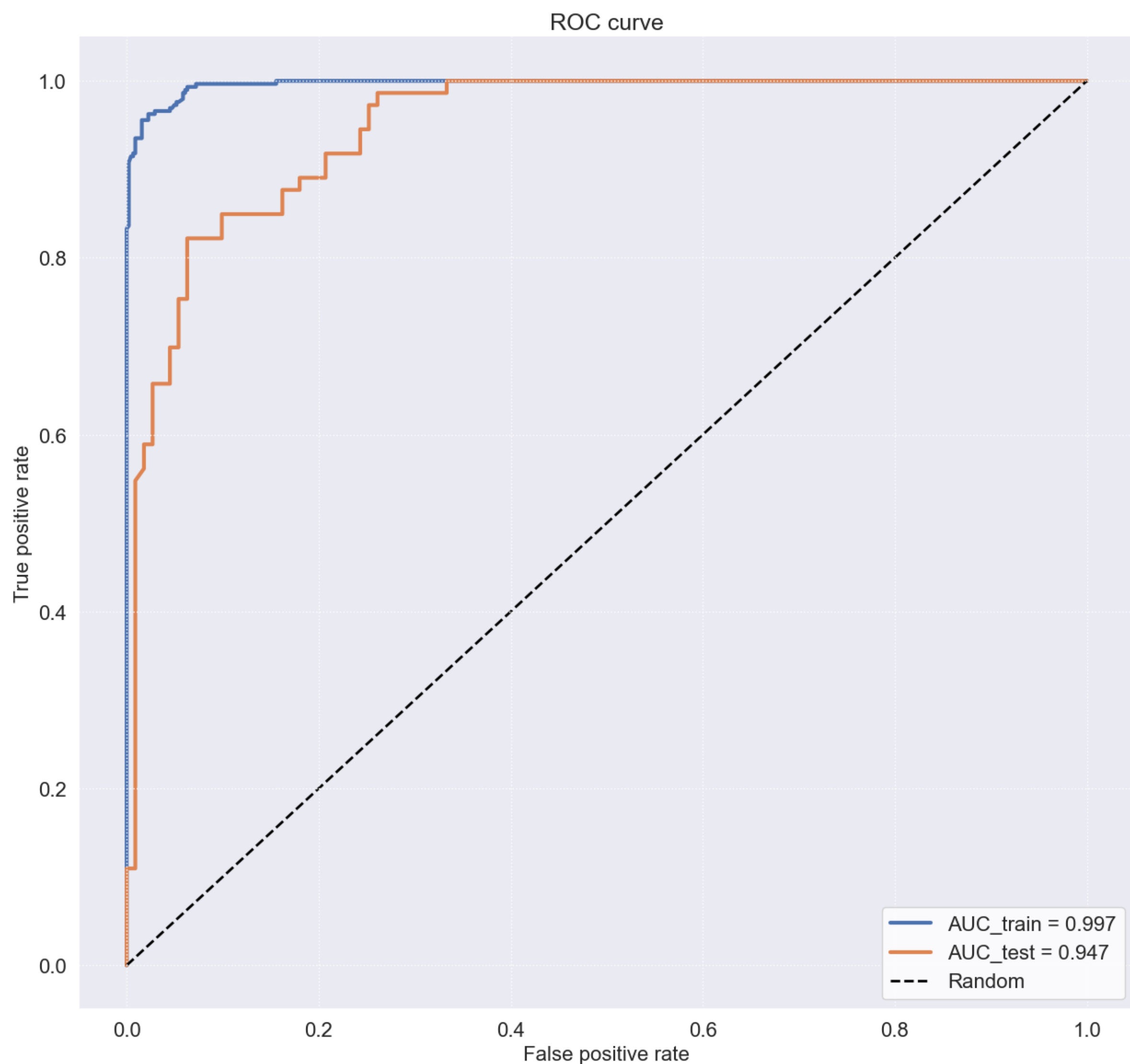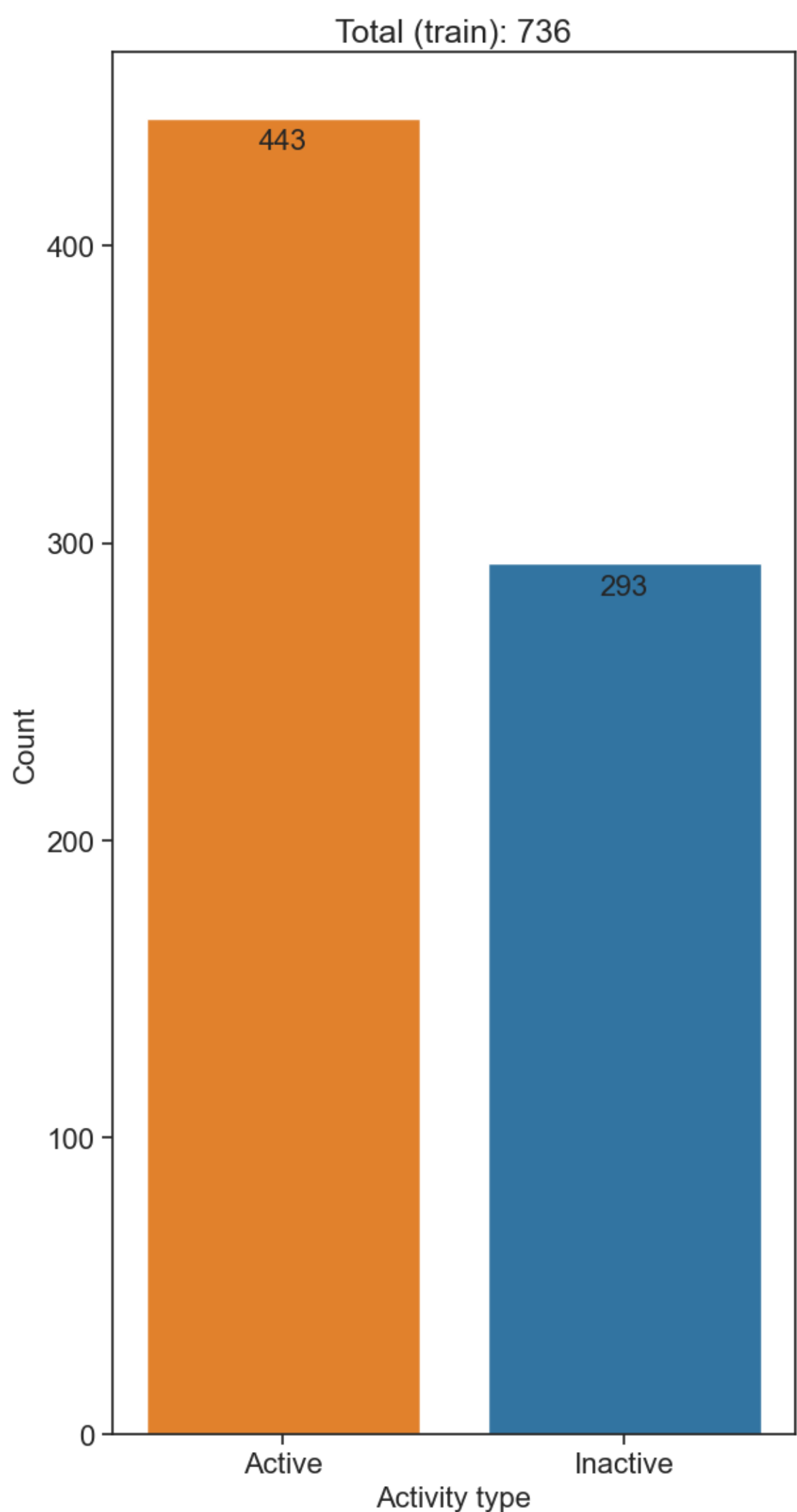
Metric	Train	Test
AUC	0.9968	0.9474
accuracy	0.9688	0.8913
balanced accuracy	0.9677	0.8794
recall_(sens)	0.9625	0.8219
specificity	0.9729	0.9369
precision	0.9592	0.8955
f1_score	0.9688	0.8904