Protein: P35348 - metric: balanced_accuracy

Total (train): 992

ROC curve

AUC_train = 0.996
AUC_test = 0.941
Random

Pareto Chart with Highlighted Best Attempt

Trials
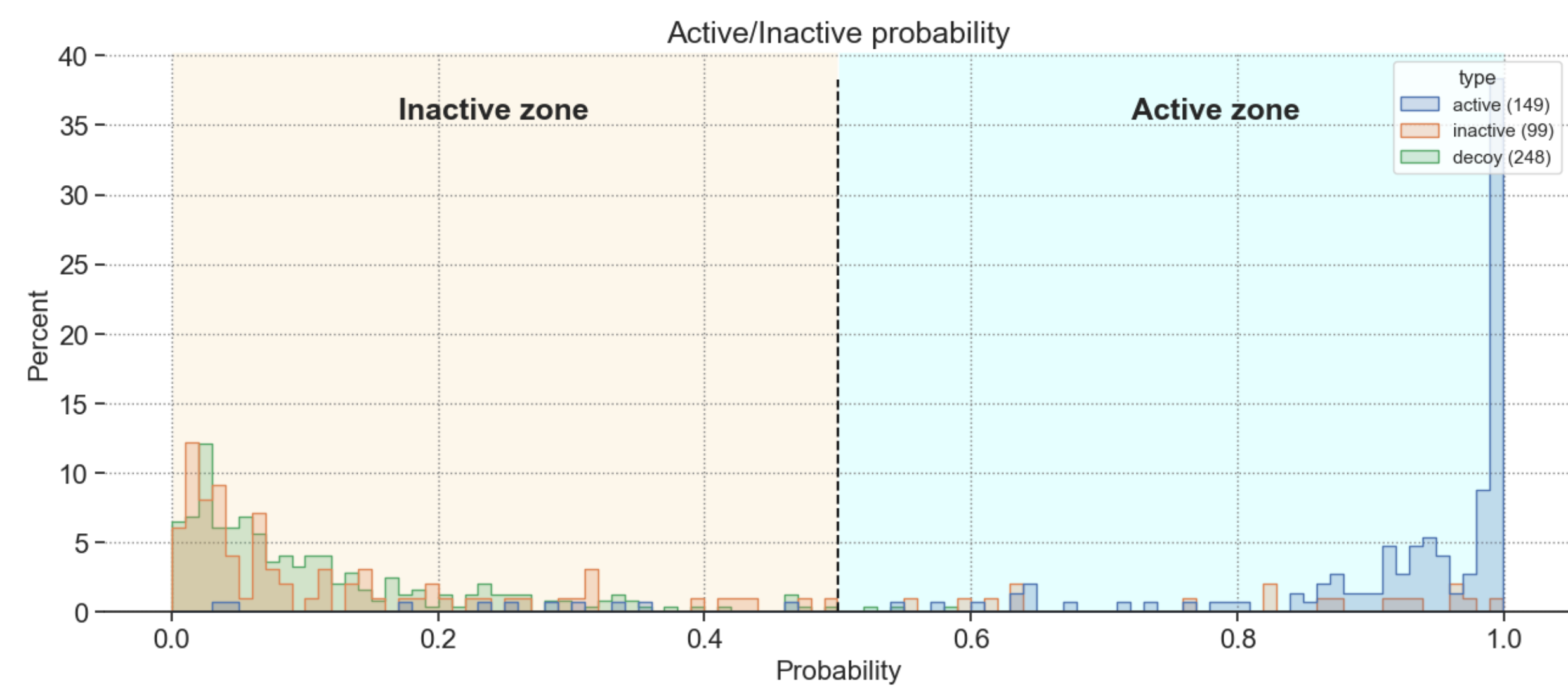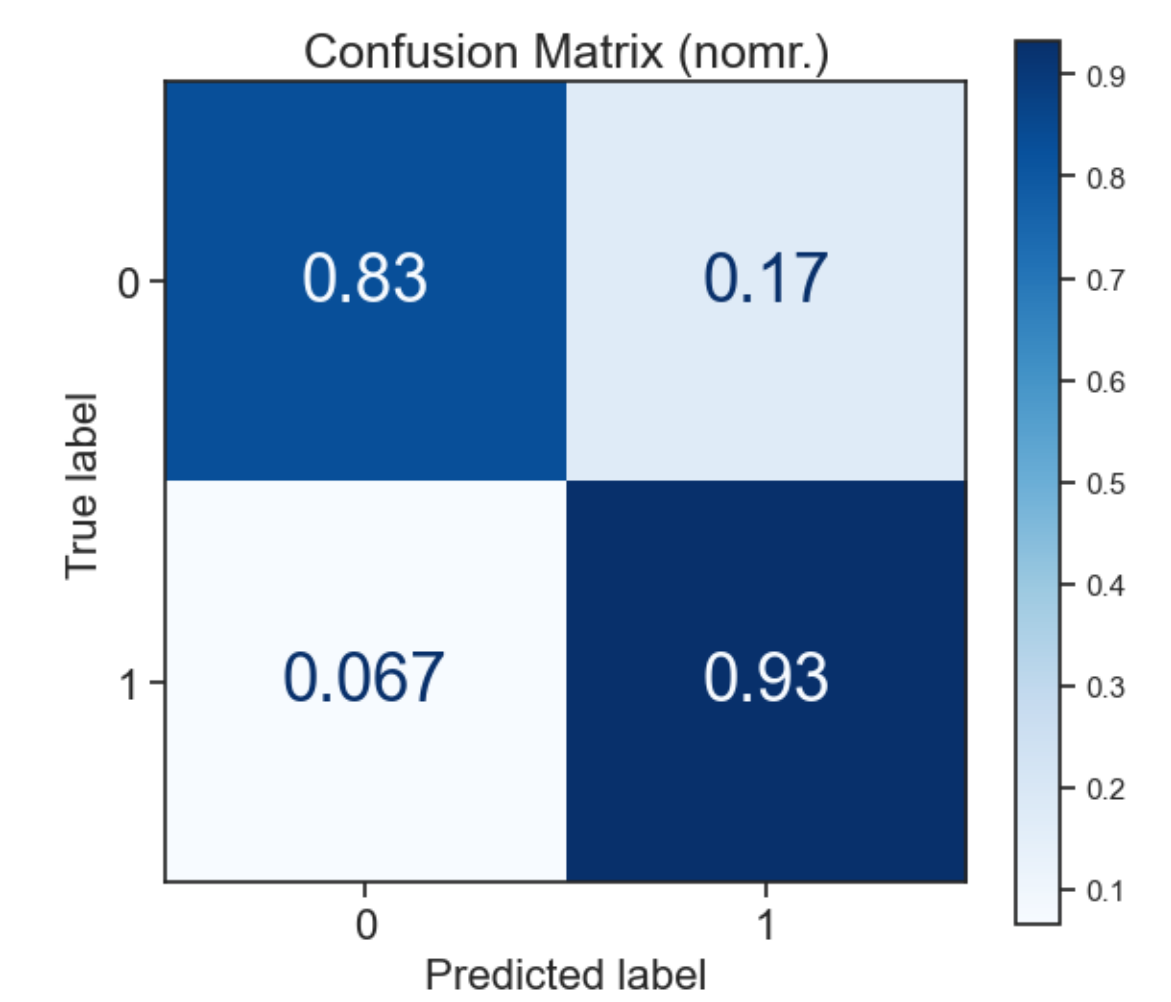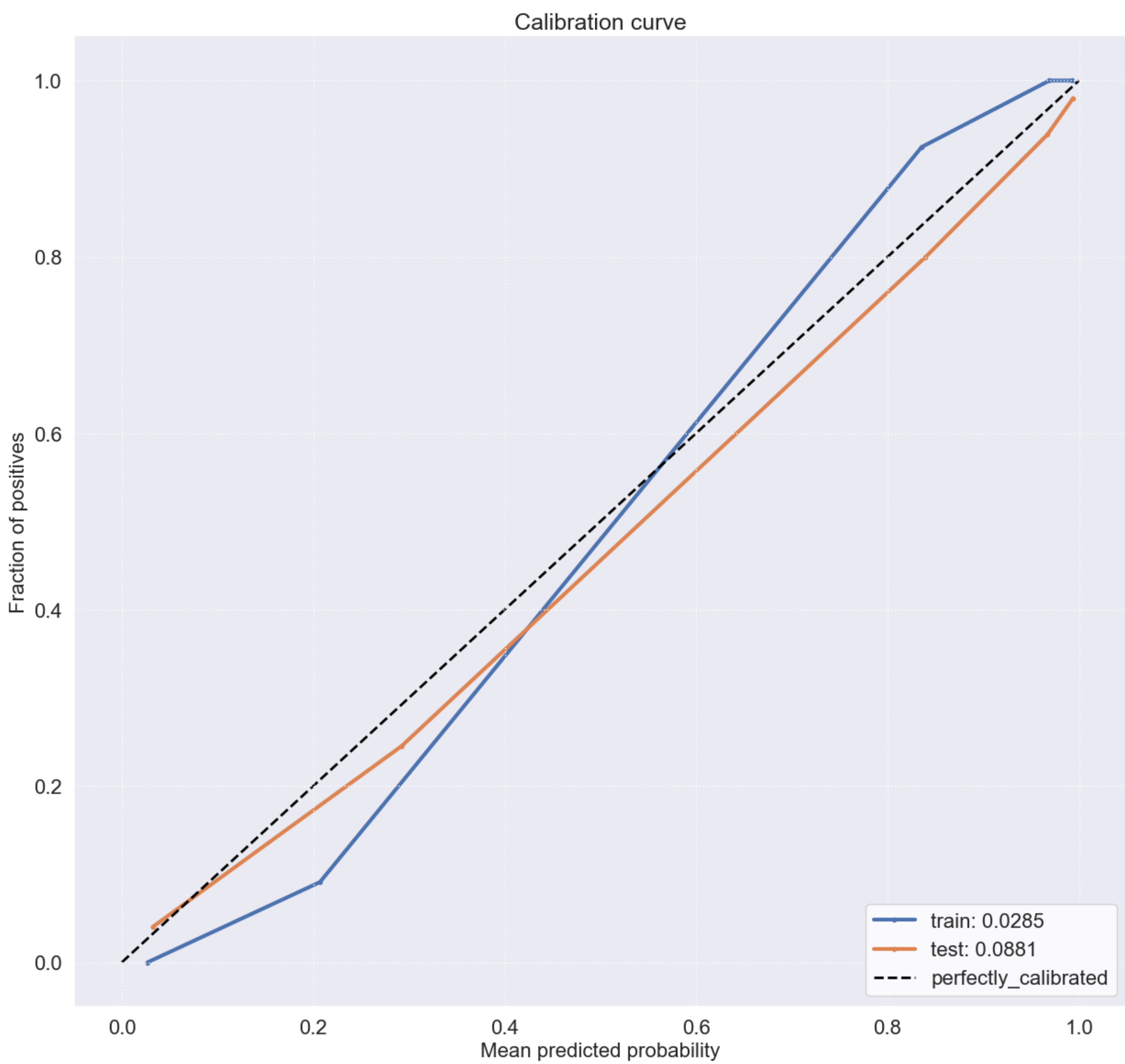Pareto Front Trials
Best Trial (# 2973)
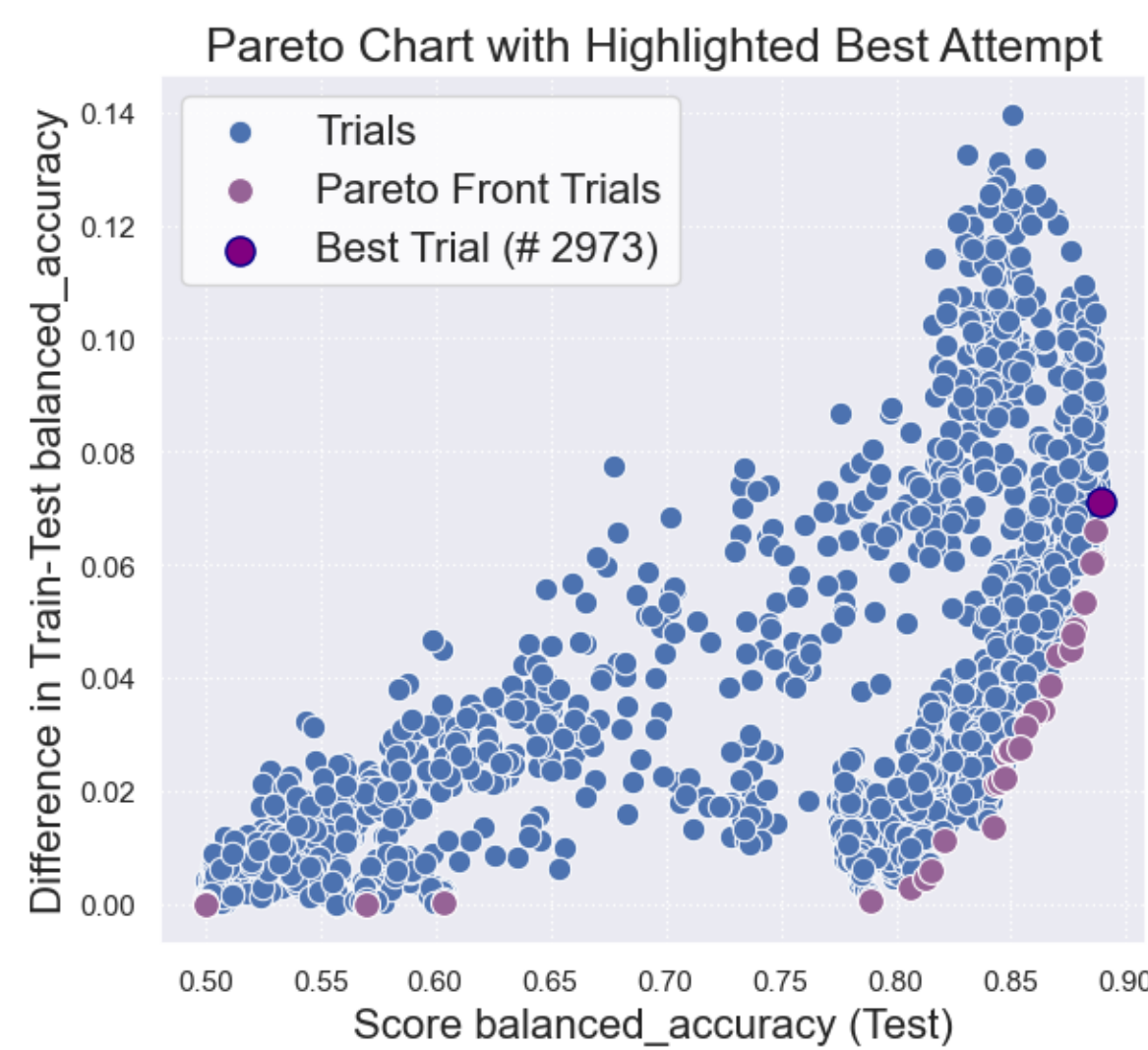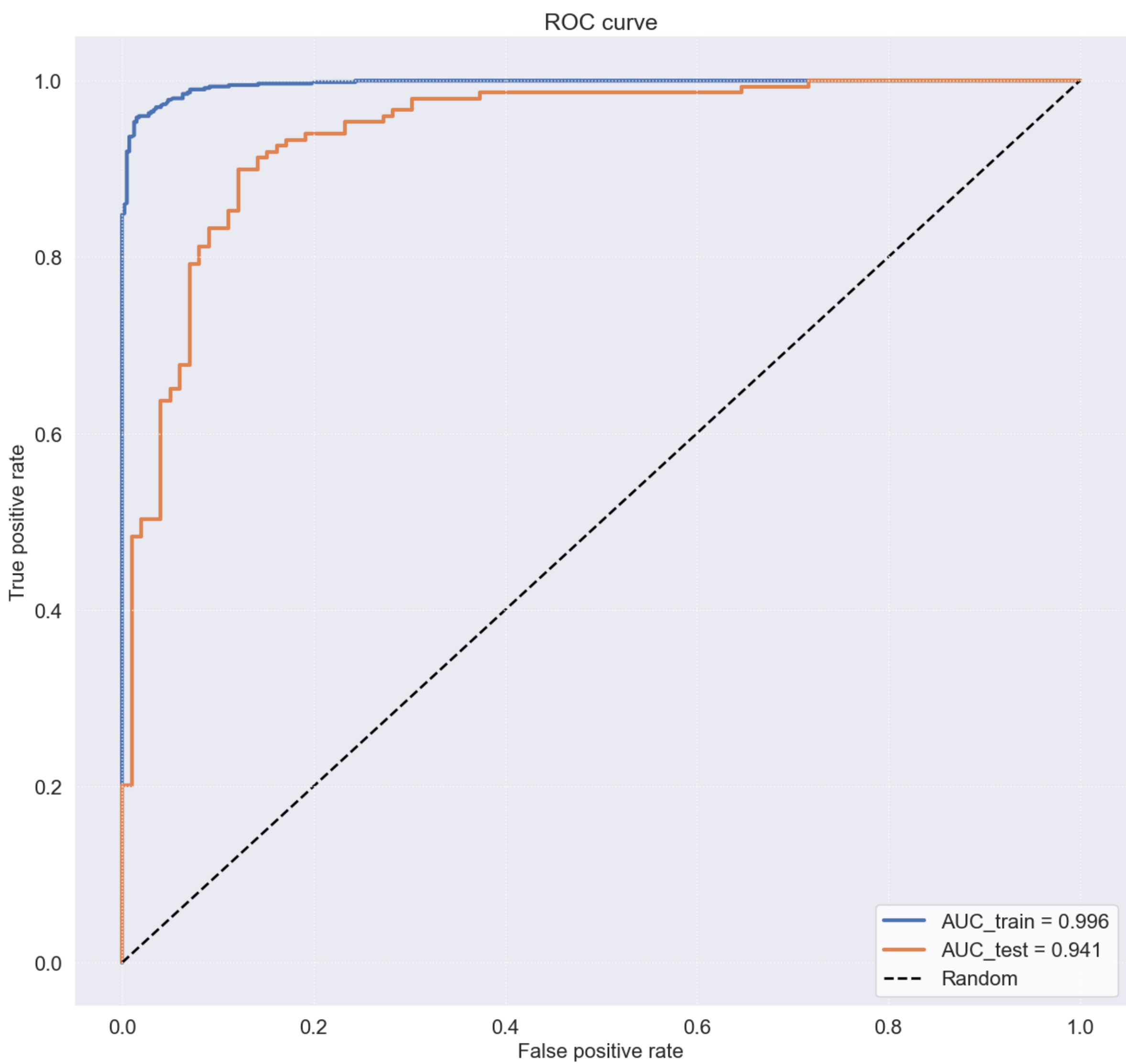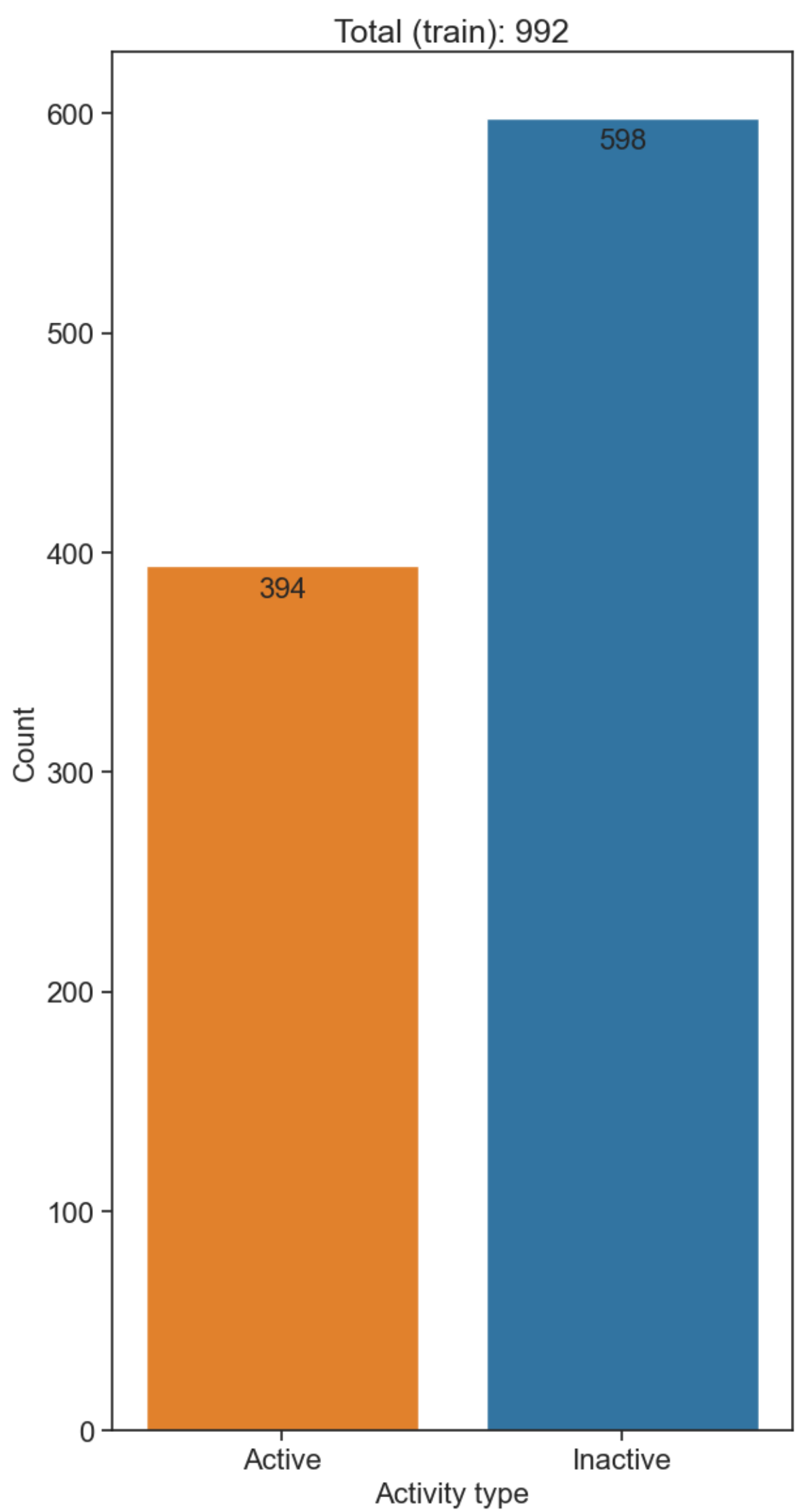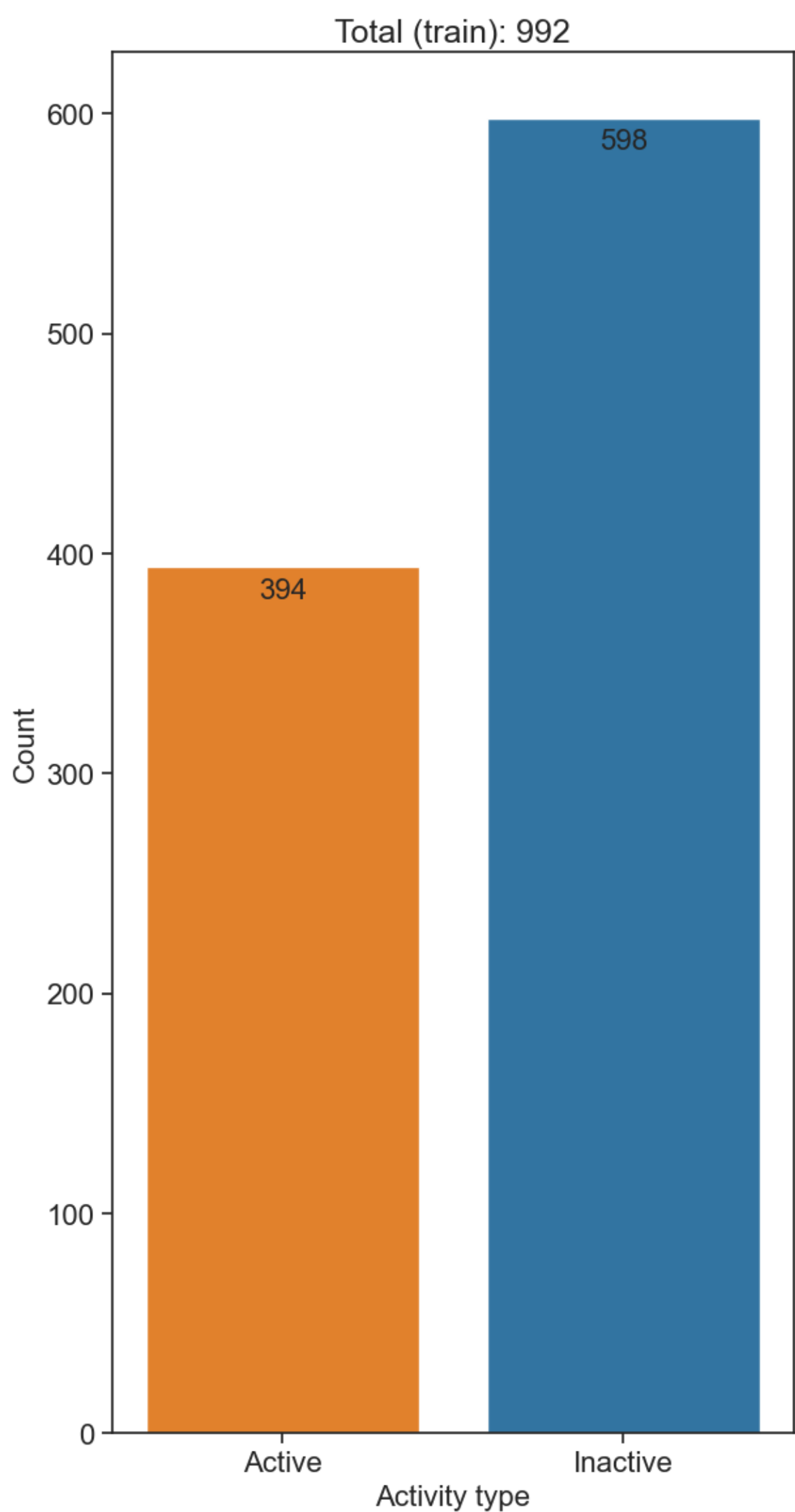
XGBoost Parameters	Value
alpha	4.6648
colsample_bytree	0.8727
gamma	0.5344
lambda	0.6546
learning_rate	0.0553
max_depth	9.0
min_child_weight	2.0
n_estimators	643.0
reg_alpha	0.9244
reg_lambda	0.2911
scale_pos_weight	1.0705
subsample	0.72

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0285
test: 0.0881
perfectly_calibrated

Active/Inactive probability

type
active (149)
inactive (99)
decoy (248)

Inactive zone
Active zone
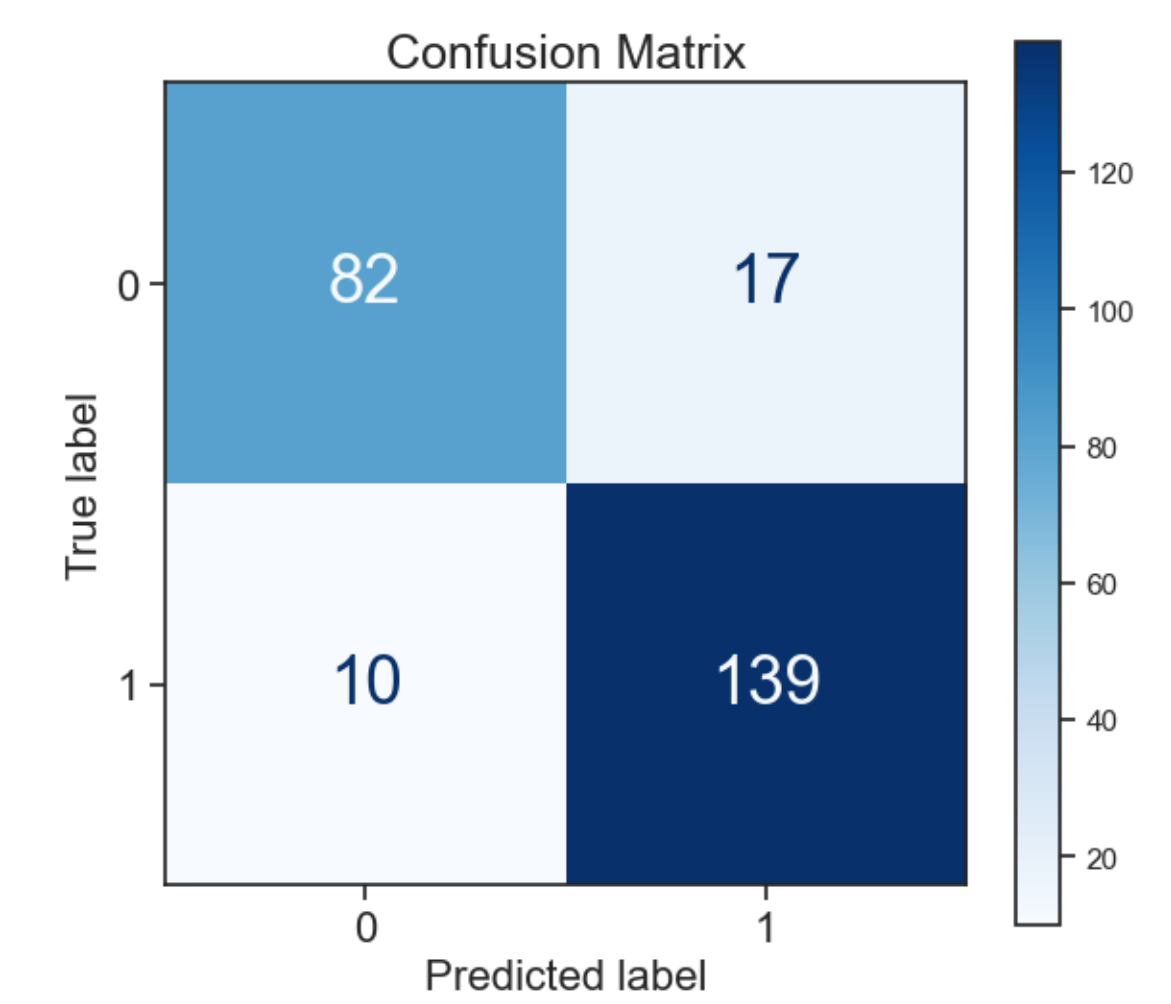
Metric	Train	Test
AUC	0.9963	0.9405
accuracy	0.9647	0.8911
balanced accuracy	0.9608	0.8806
recall_(sens)	0.9799	0.9329
specificity	0.9416	0.8283
precision	0.9622	0.891
f1_score	0.9646	0.8904

Protein: P35348 - metric: f1

Total (train): 992

ROC curve

	AUC_train = 0.998
	AUC_test = 0.941
	Random

Pareto Chart with Highlighted Best Attempt

- Trials
- Pareto Front Trials
- Best Trial (# 2296)
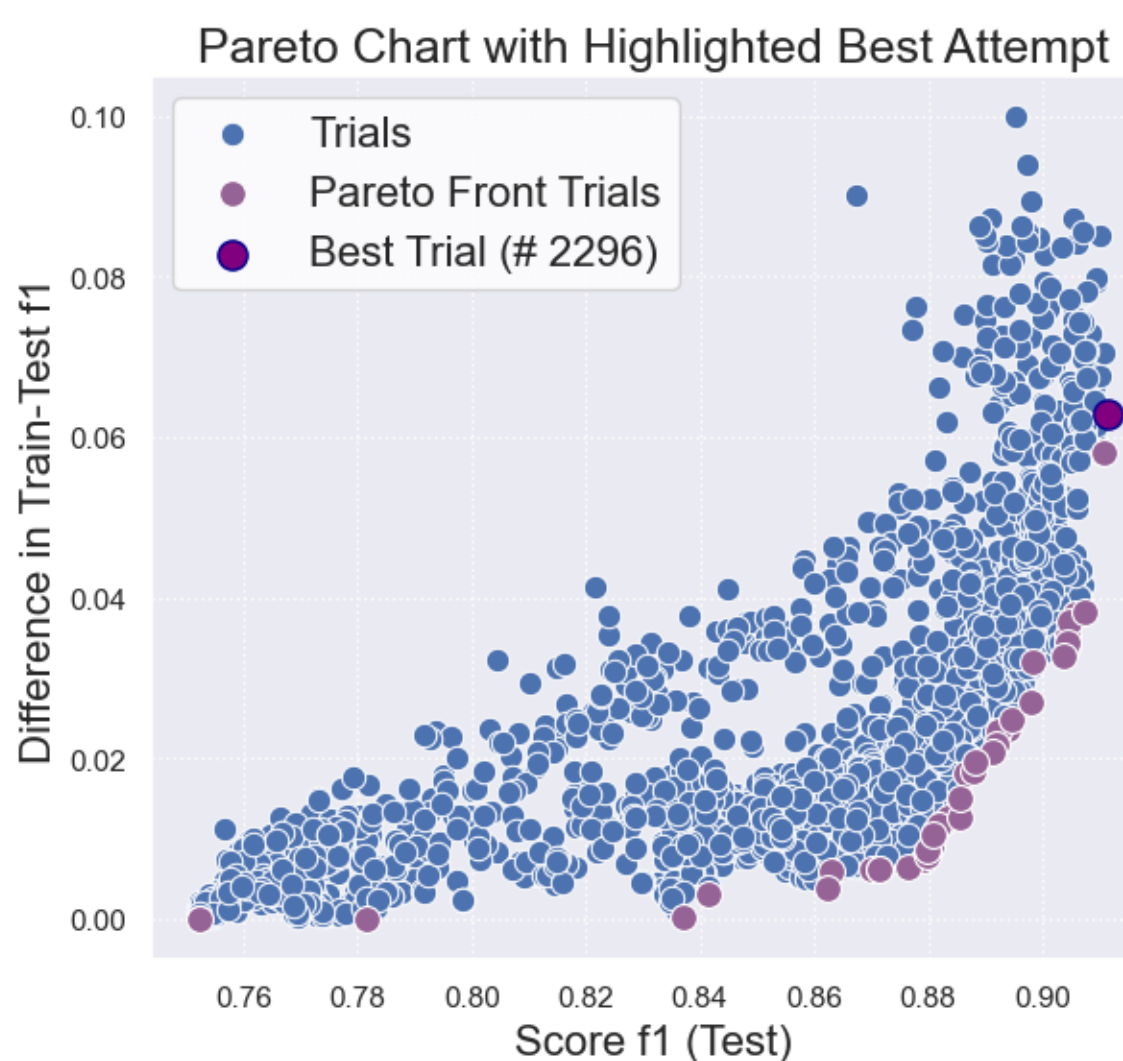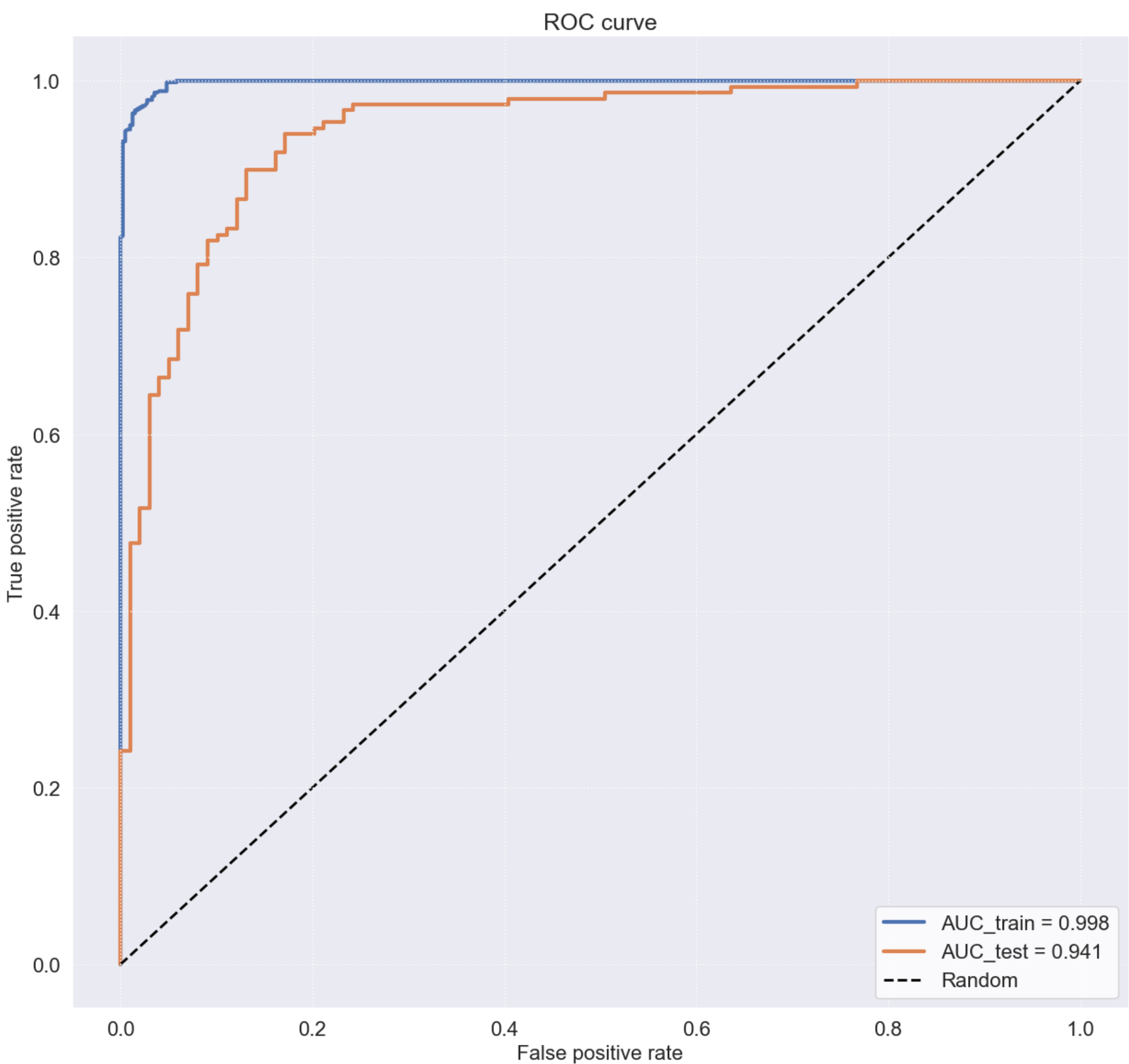
XGBoost Parameters	Value
alpha	4.367
colsample_bytree	0.5469
gamma	0.9849
lambda	3.2189
learning_rate	0.0137
max_depth	10.0
min_child_weight	1.0
n_estimators	505.0
reg_alpha	0.049
reg_lambda	0.3468
scale_pos_weight	2.5066
subsample	0.6759

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

	train: 0.0286
	test: 0.0939
	perfectly_calibrated

Active/Inactive probability

Inactive zone — **Active zone**

type
- active (149)
- inactive (99)
- decoy (248)
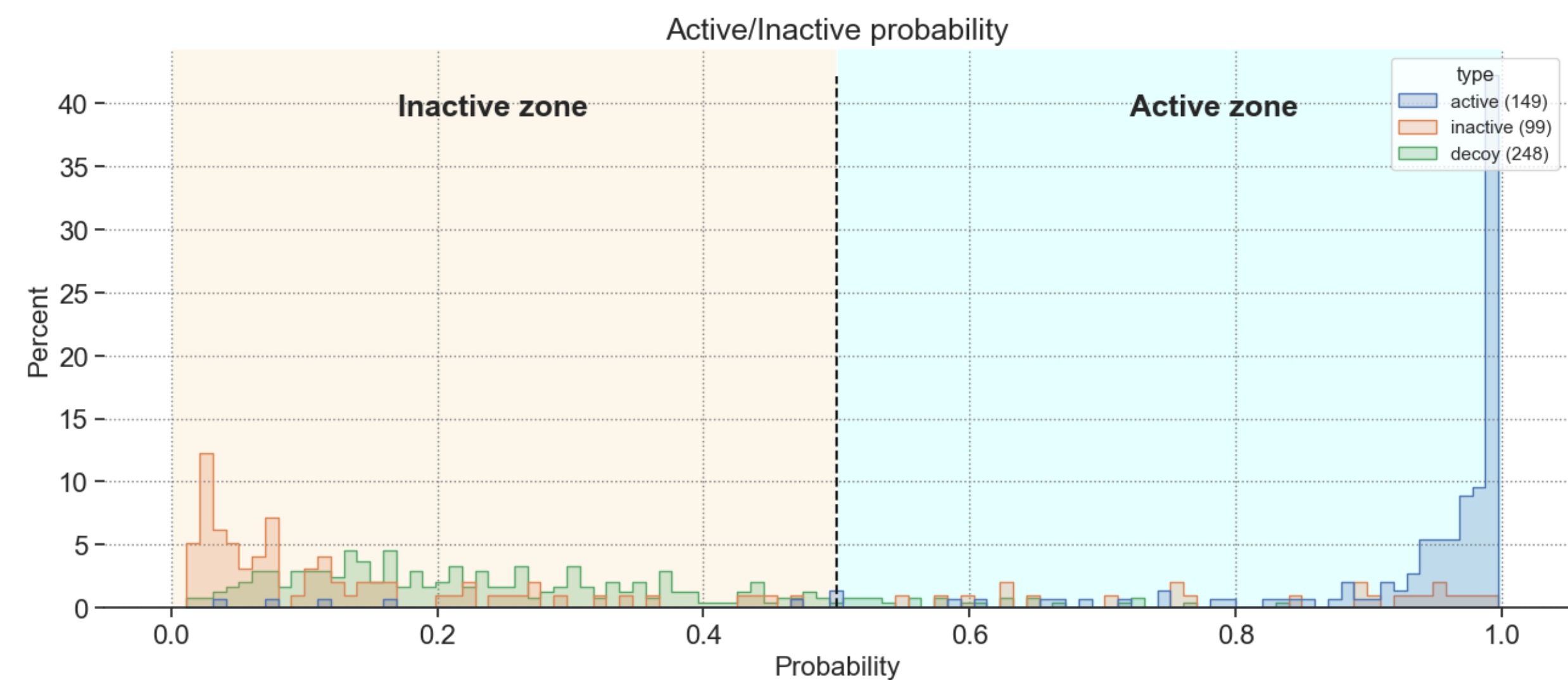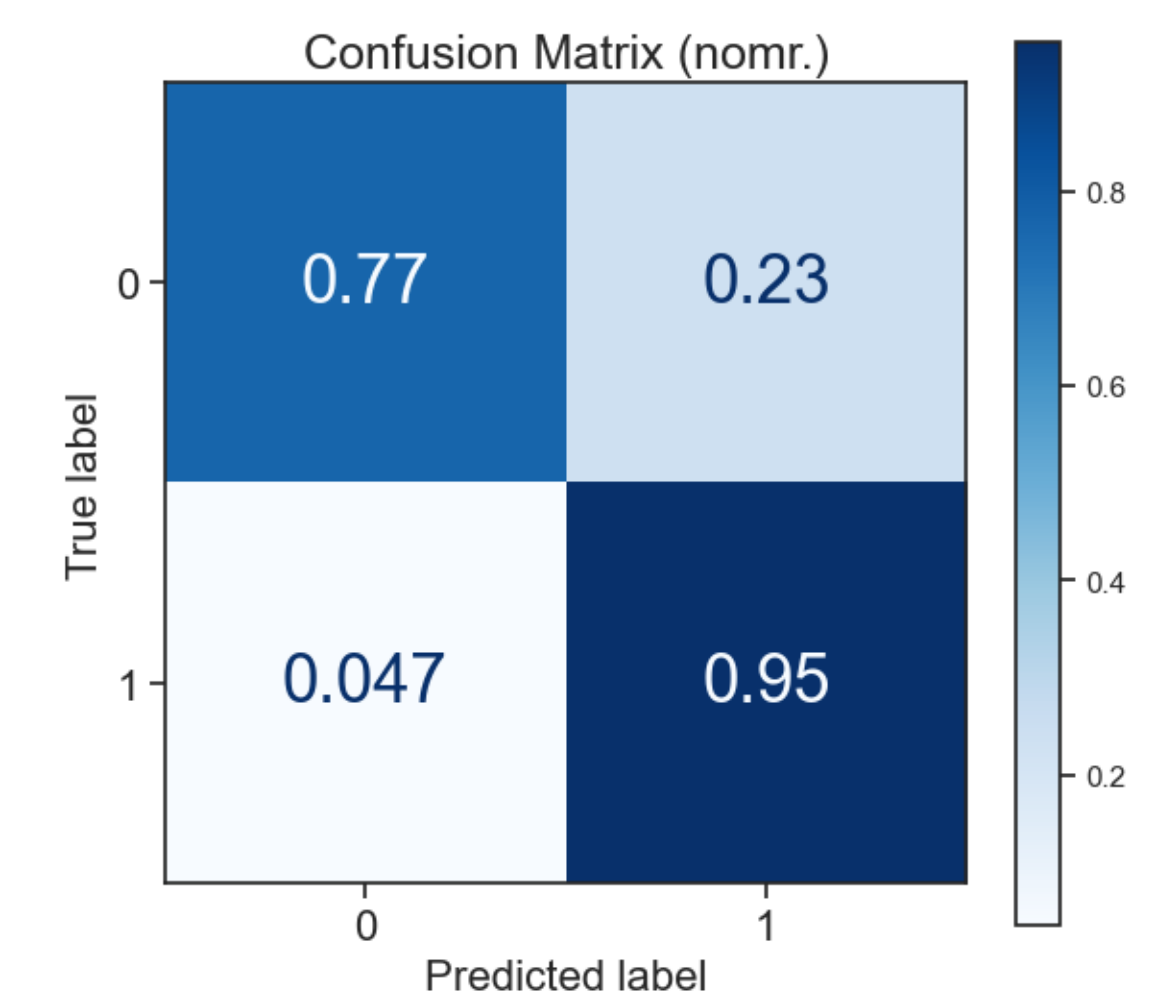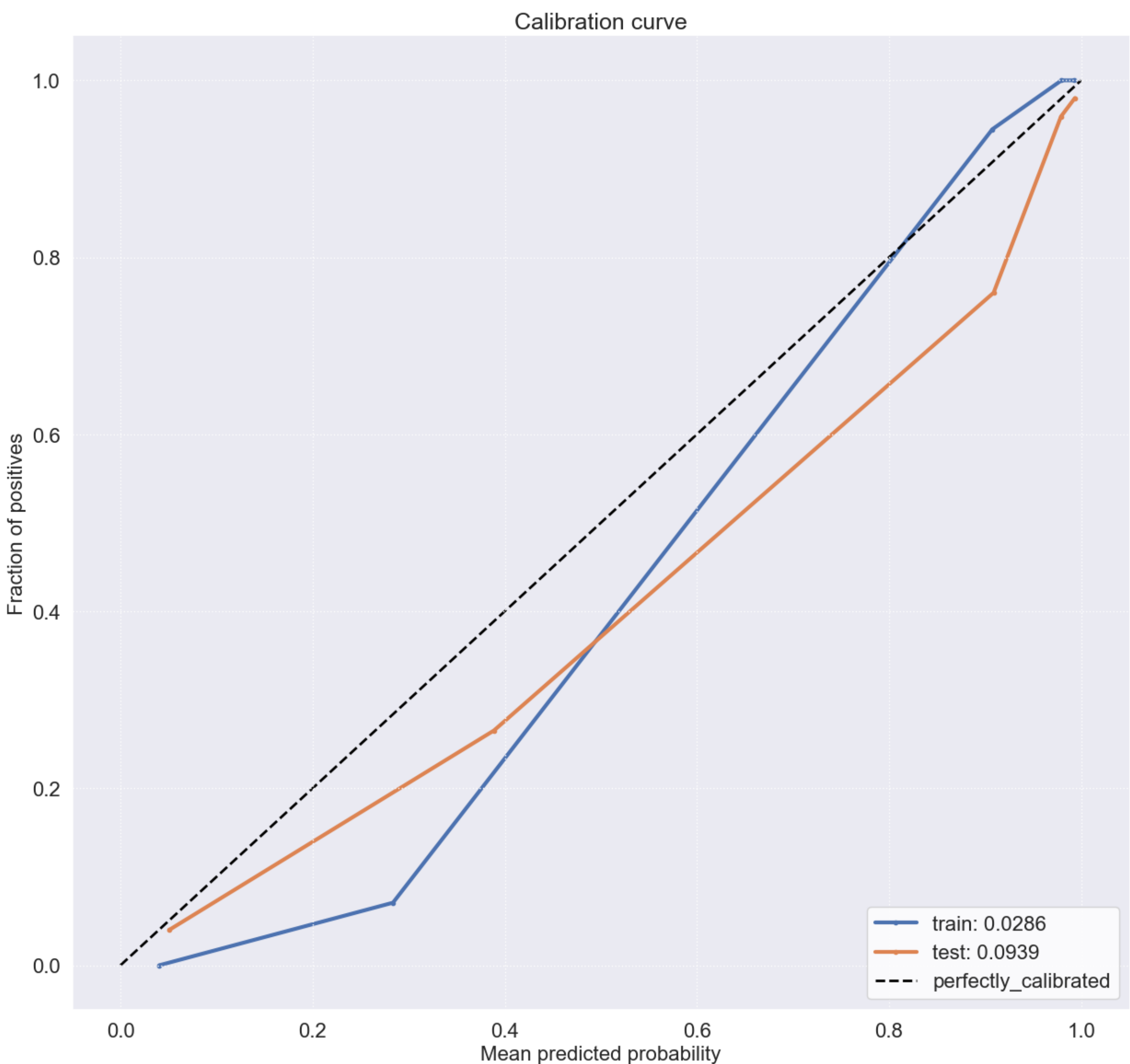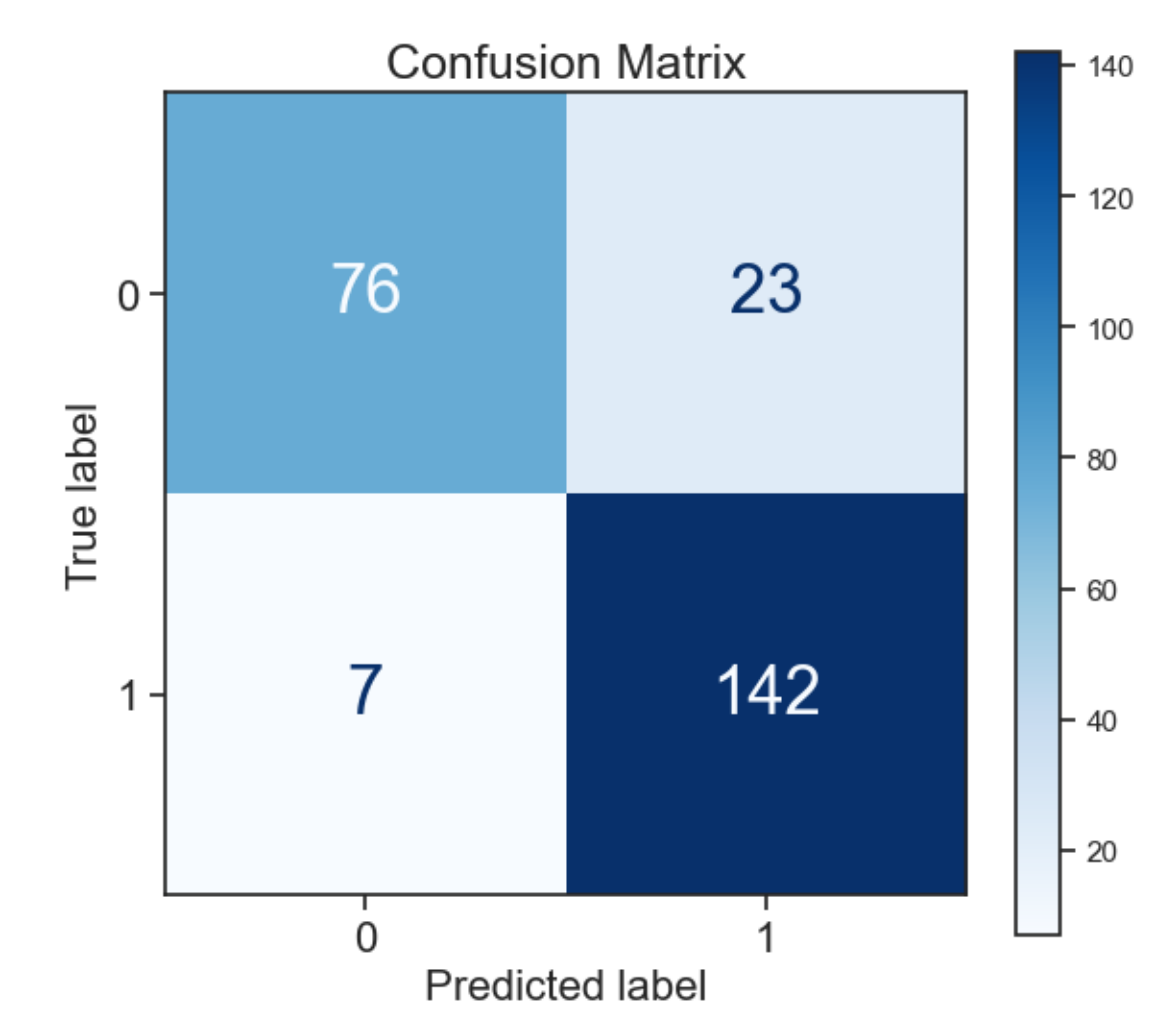
Metric	Train	Test
AUC	0.9982	0.9414
accuracy	0.9667	0.879
balanced accuracy	0.9581	0.8603
recall_(sens)	1.0	0.953
specificity	0.9162	0.7677
precision	0.9477	0.8606
f1_score	0.9665	0.8768

Protein: P35348 - metric: precision

Total (train): 992

ROC curve

AUC_train = 0.999
AUC_test = 0.936
Random

Confusion Matrix

Confusion Matrix (nomr.)

Calibration curve

train: 0.0196
test: 0.0915
perfectly_calibrated

Pareto Chart with Highlighted Best Attempt

Trials
Pareto Front Trials
Best Trial (# 2961)
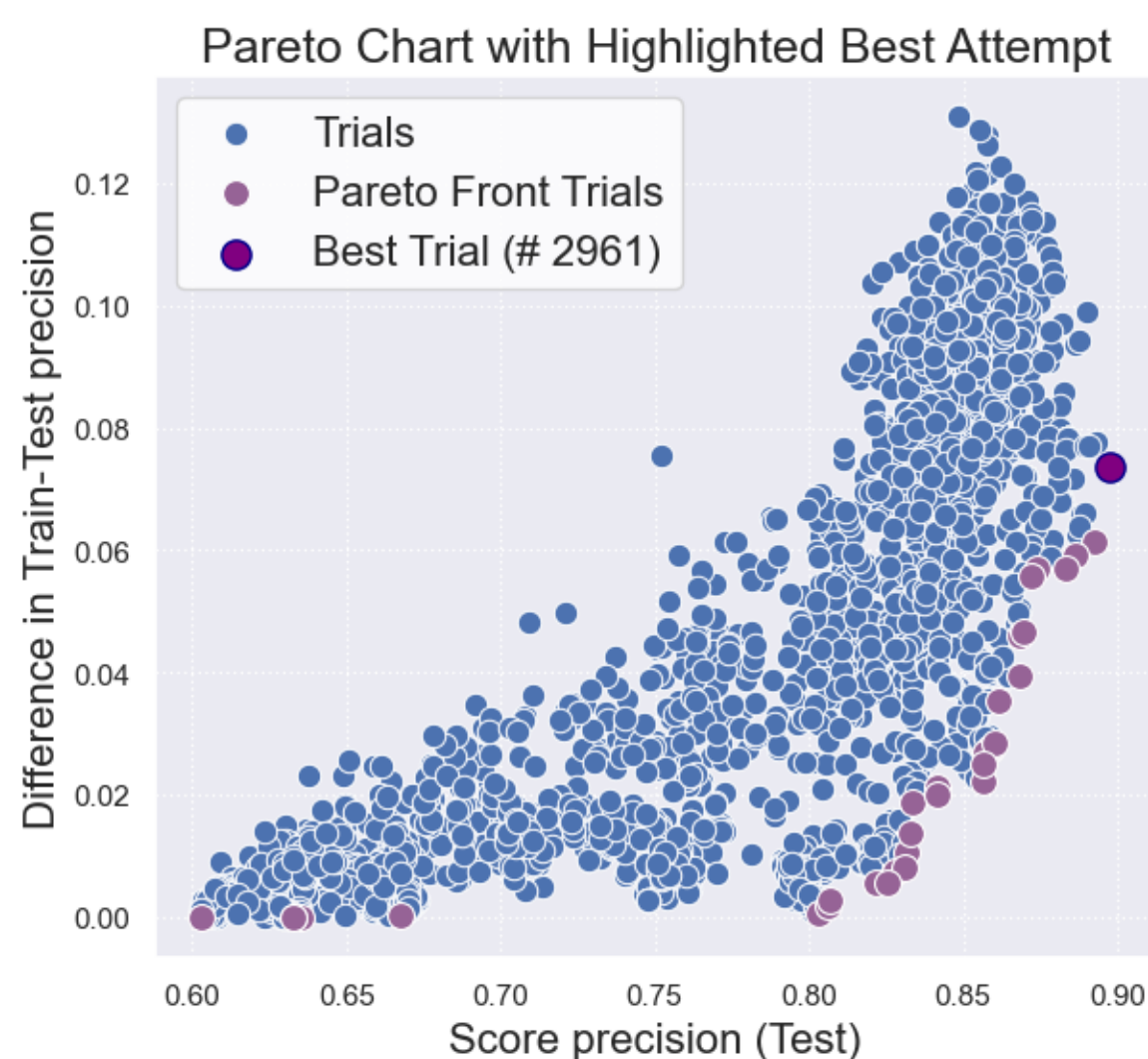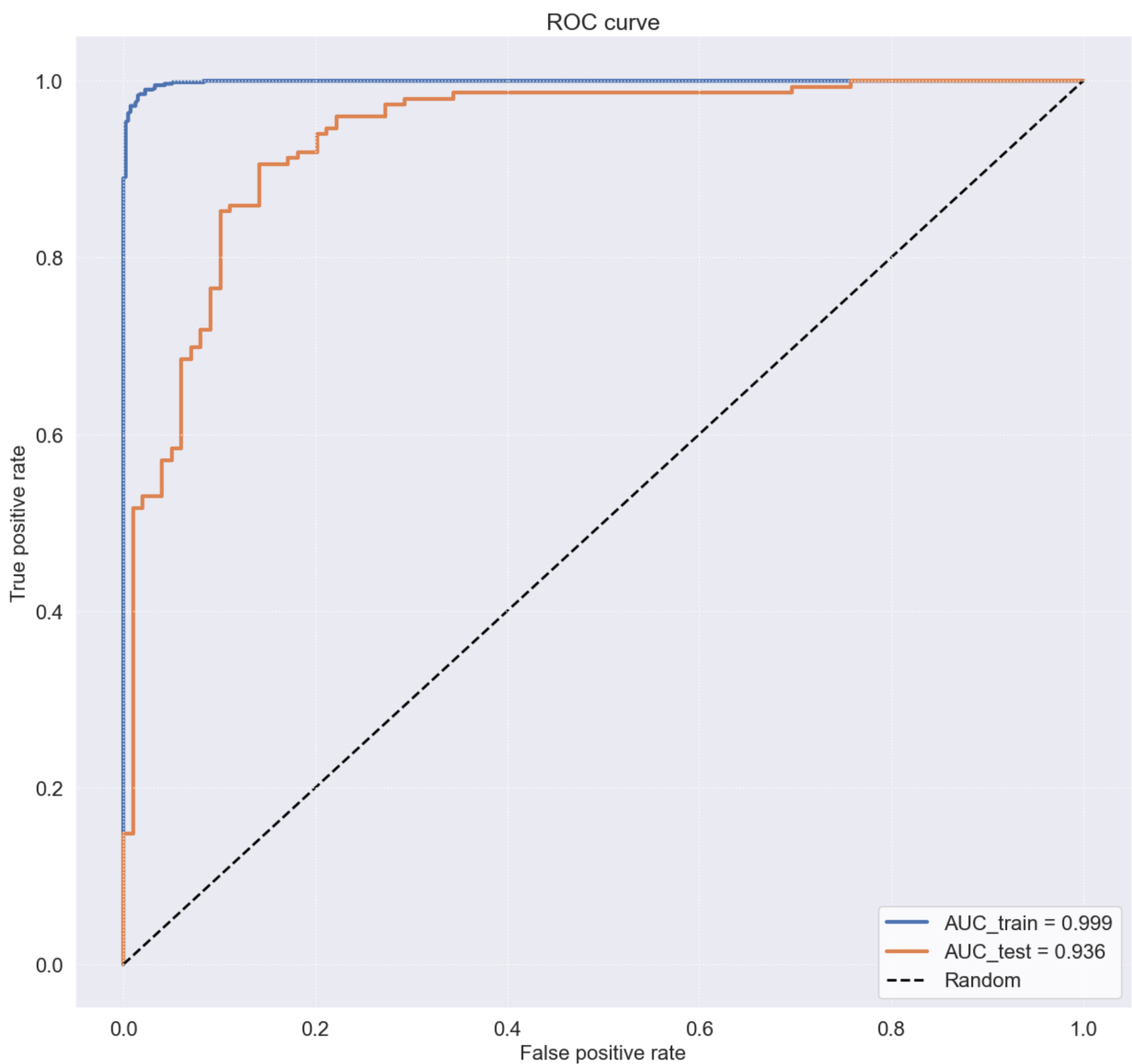
XGBoost Parameters	Value
alpha	2.304
colsample_bytree	0.7191
gamma	0.1072
lambda	0.9881
learning_rate	0.0477
max_depth	8.0
min_child_weight	2.0
n_estimators	410.0
reg_alpha	0.4066
reg_lambda	0.5145
scale_pos_weight	1.4876
subsample	0.6372

Active/Inactive probability

type
active (149)
inactive (99)
decoy (248)
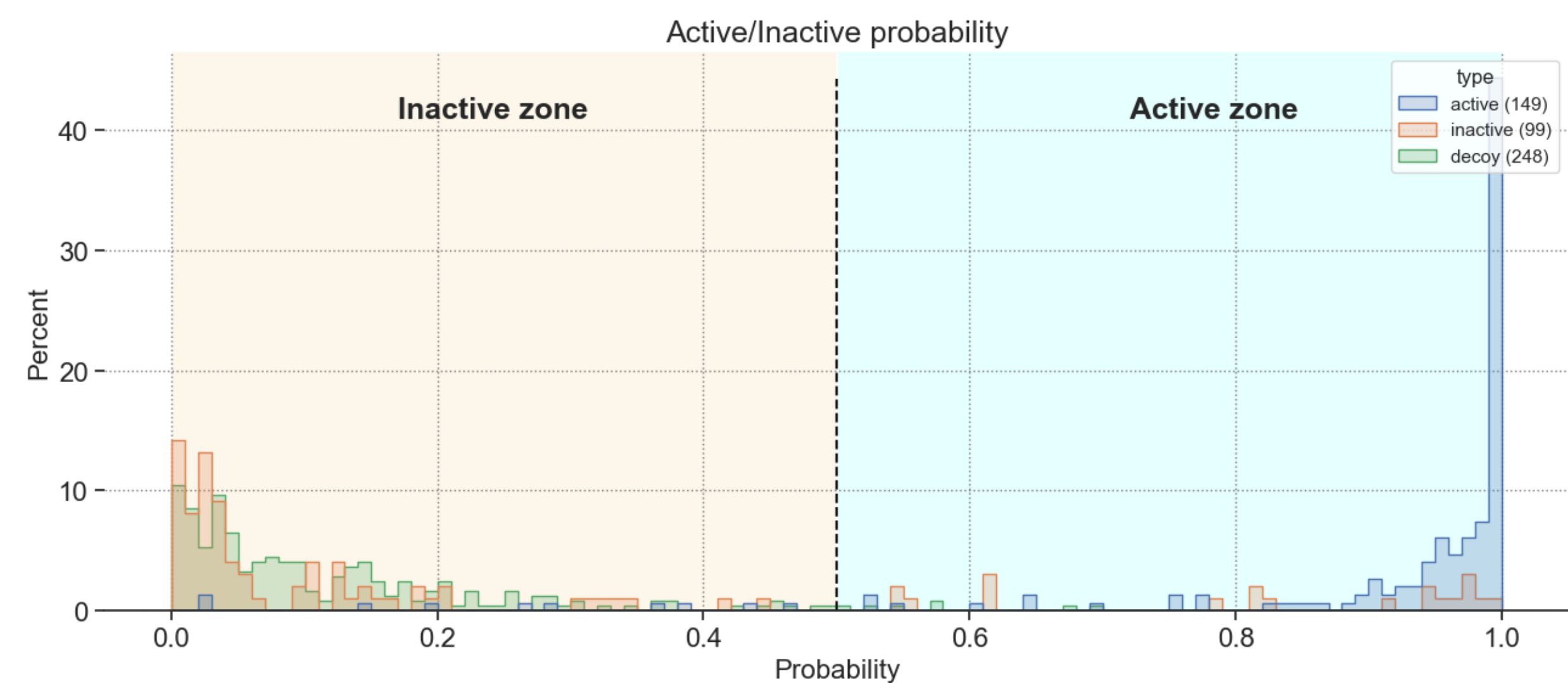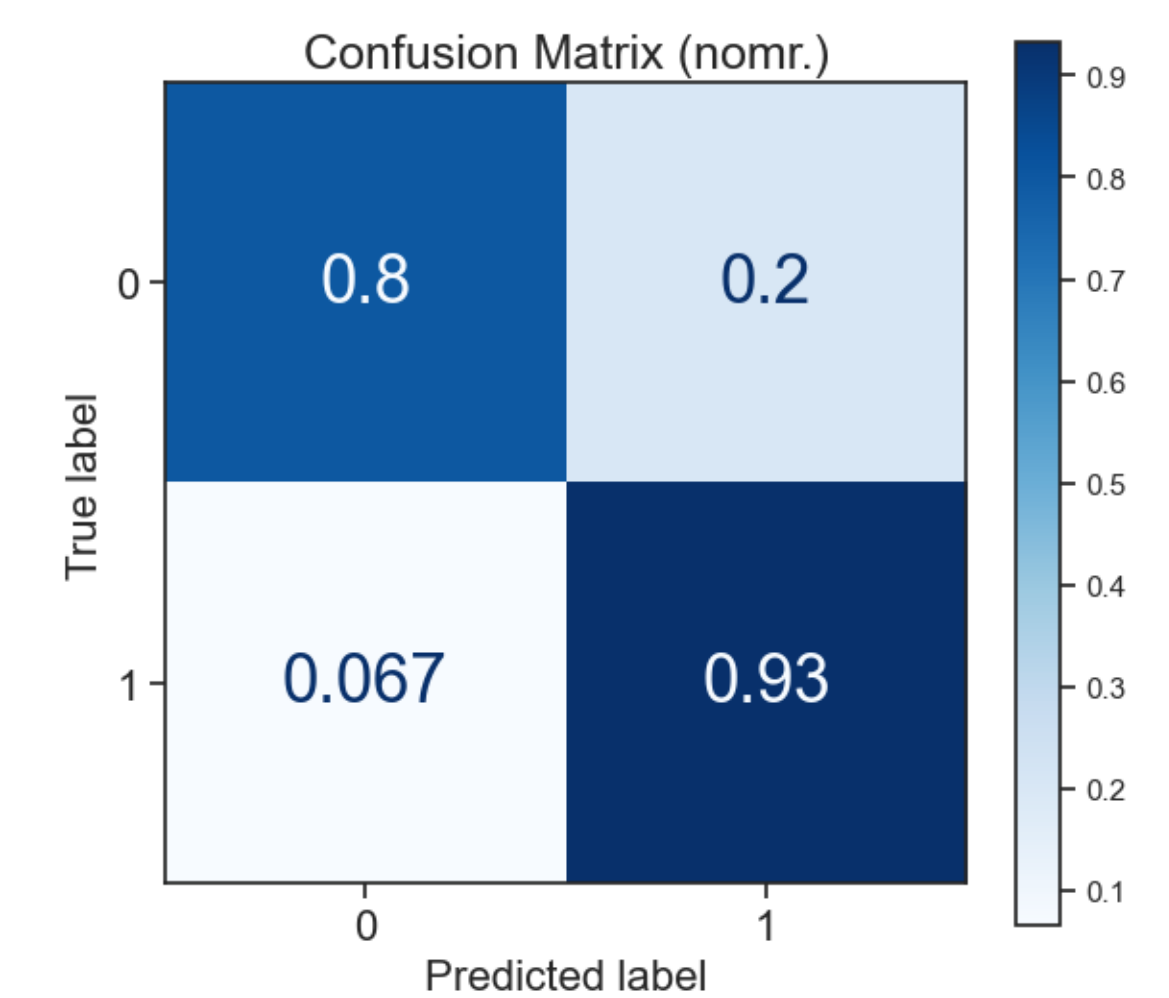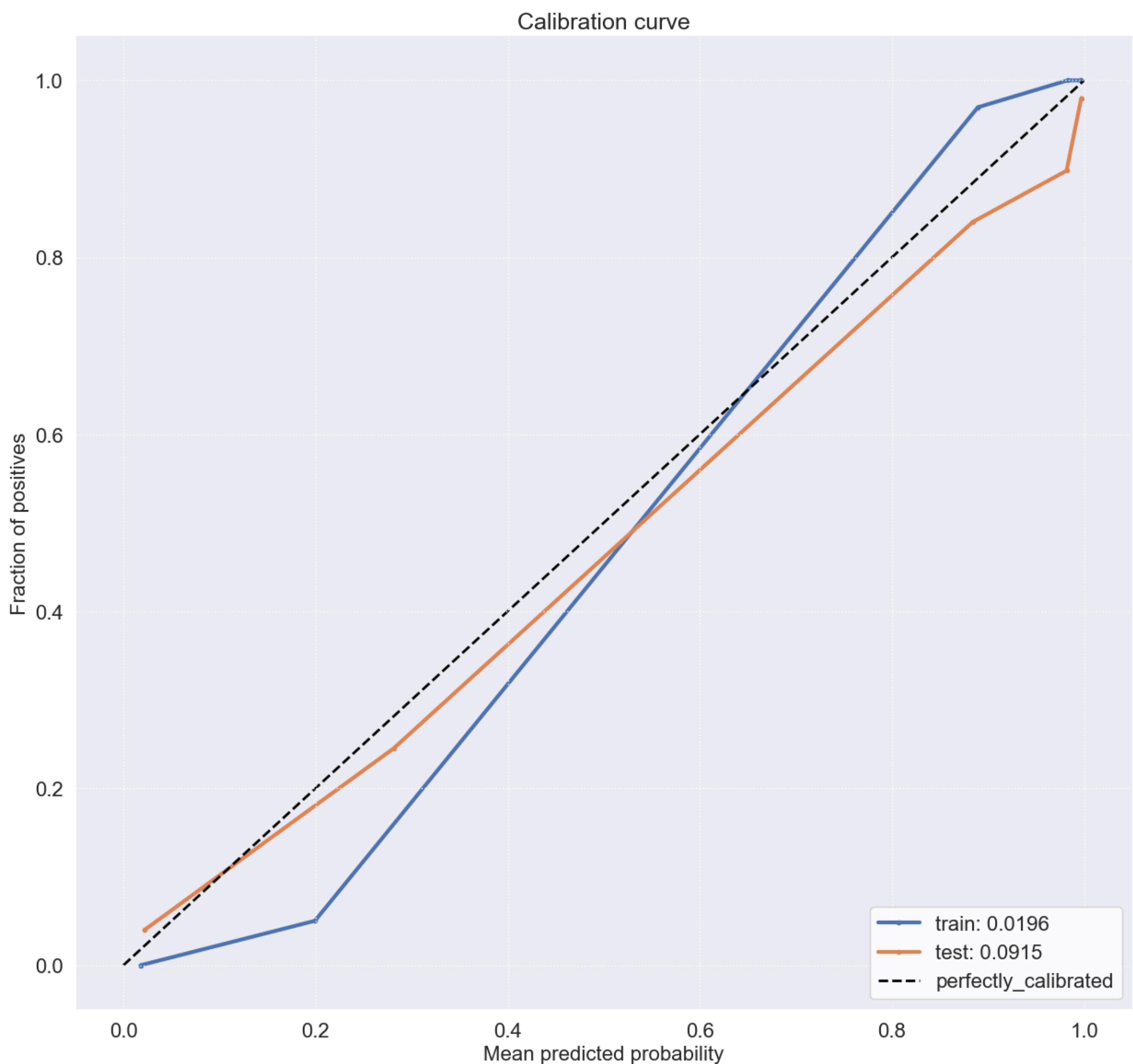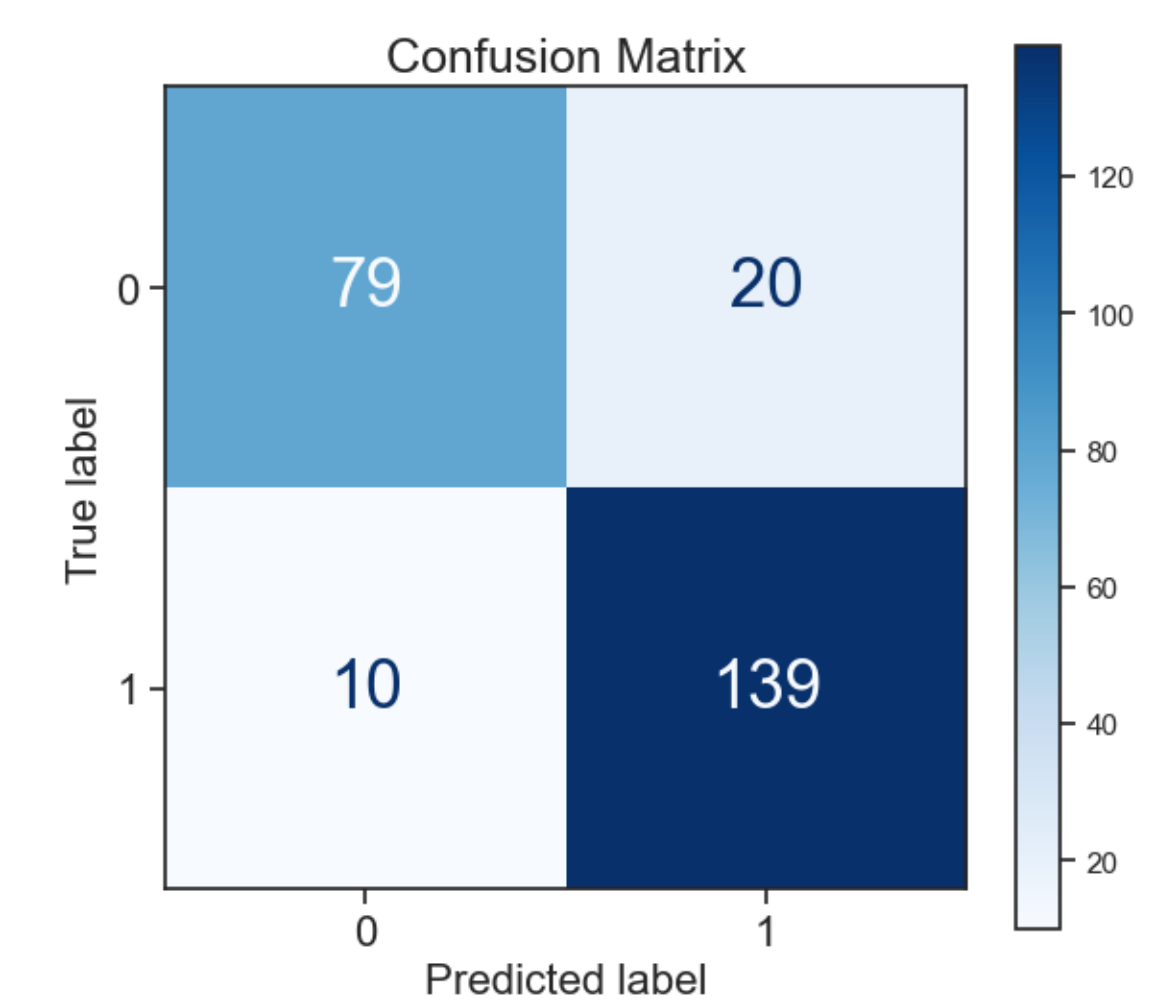
Metric	Train	Test
AUC	0.999	0.936
accuracy	0.9788	0.879
balanced accuracy	0.9738	0.8654
recall_(sens)	0.9983	0.9329
specificity	0.9492	0.798
precision	0.9676	0.8742
f1_score	0.9787	0.8778

Protein: P35348 - metric: SoftVotingXGBoostEnsemble
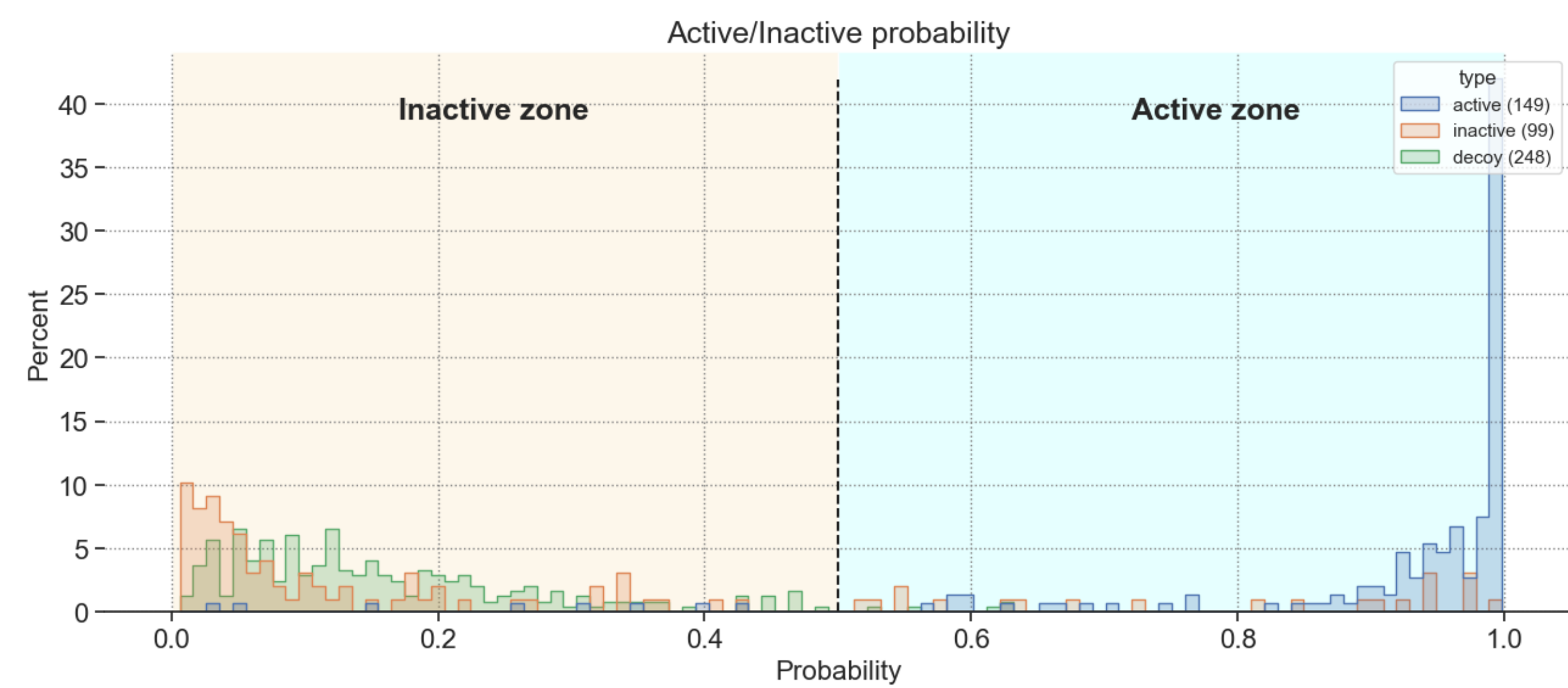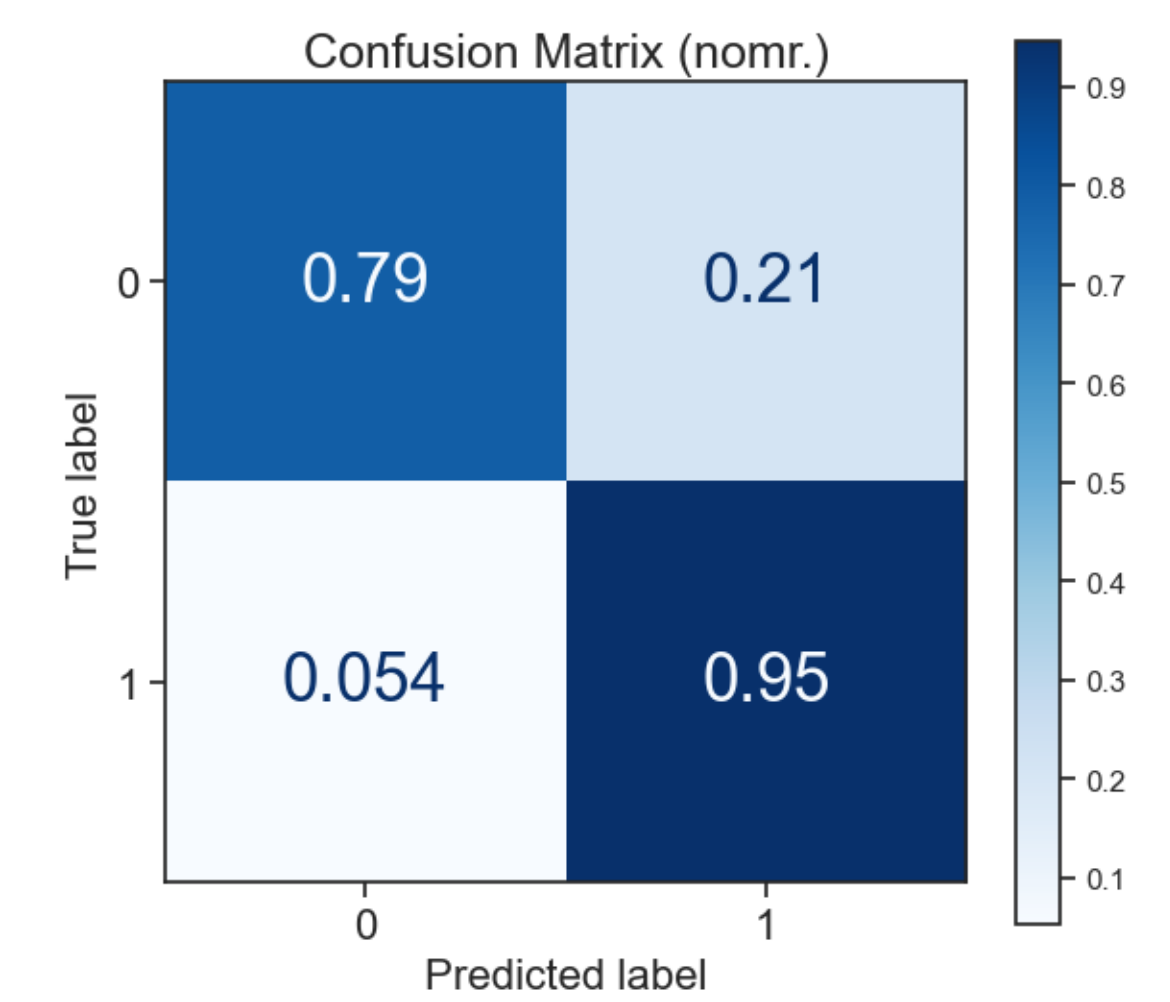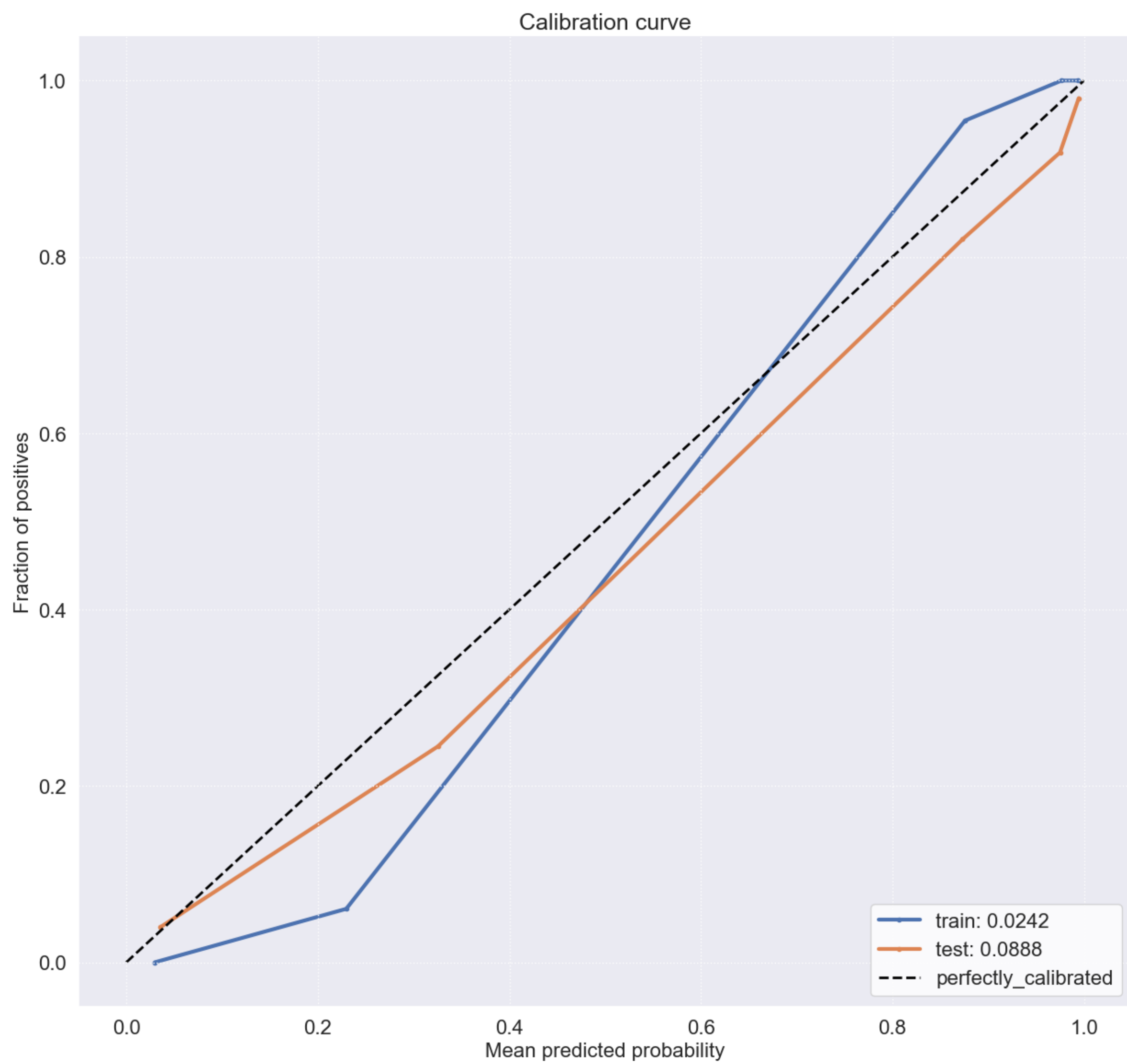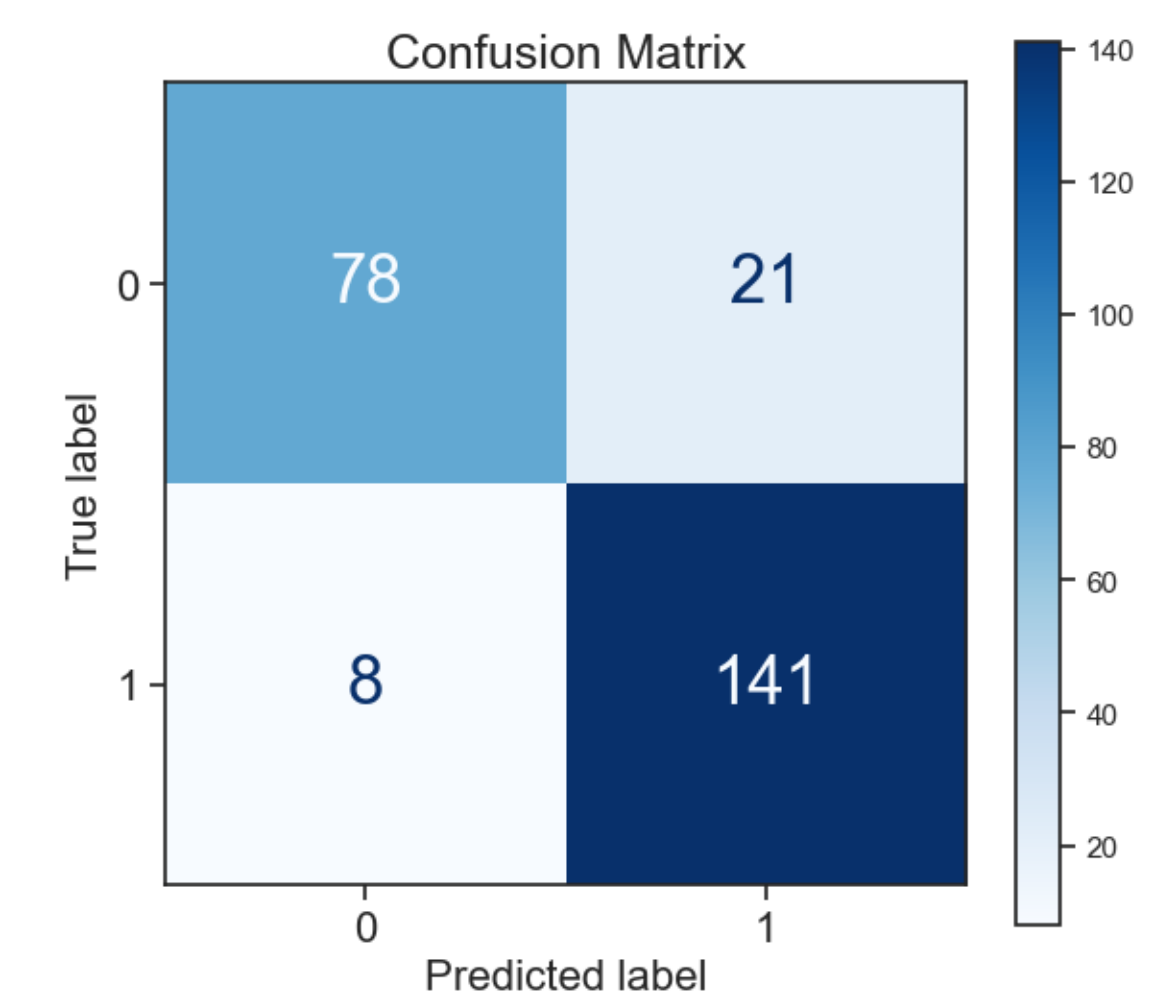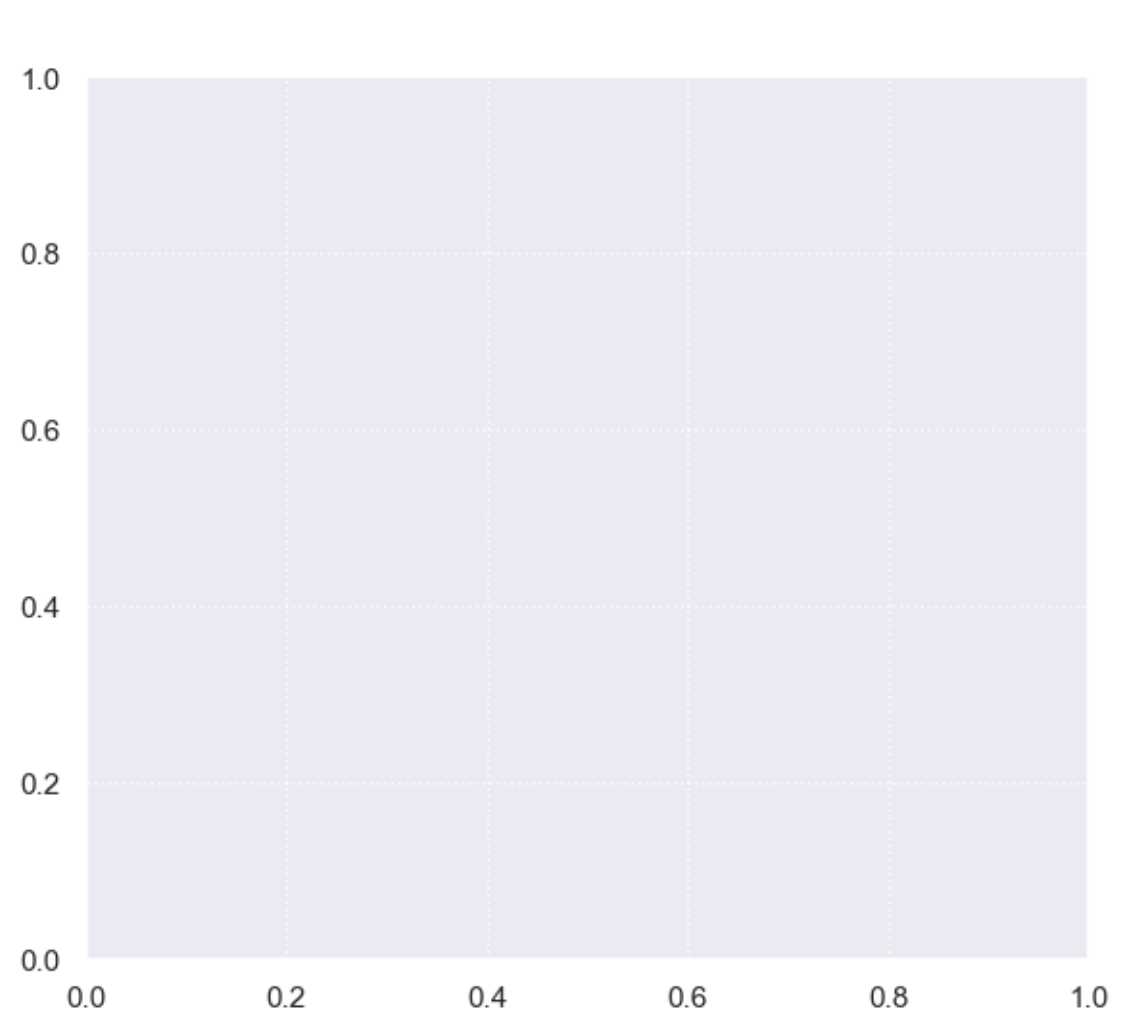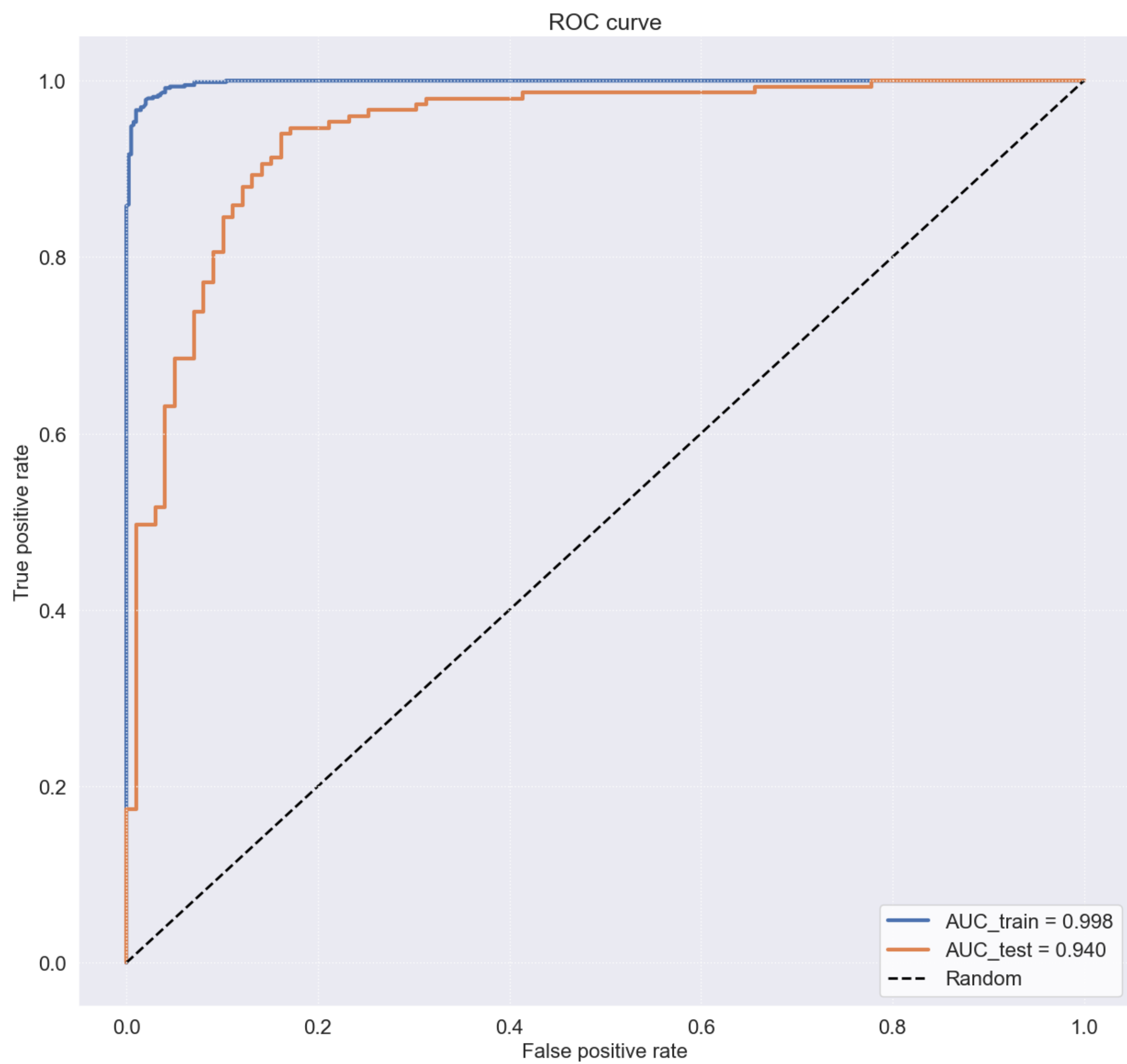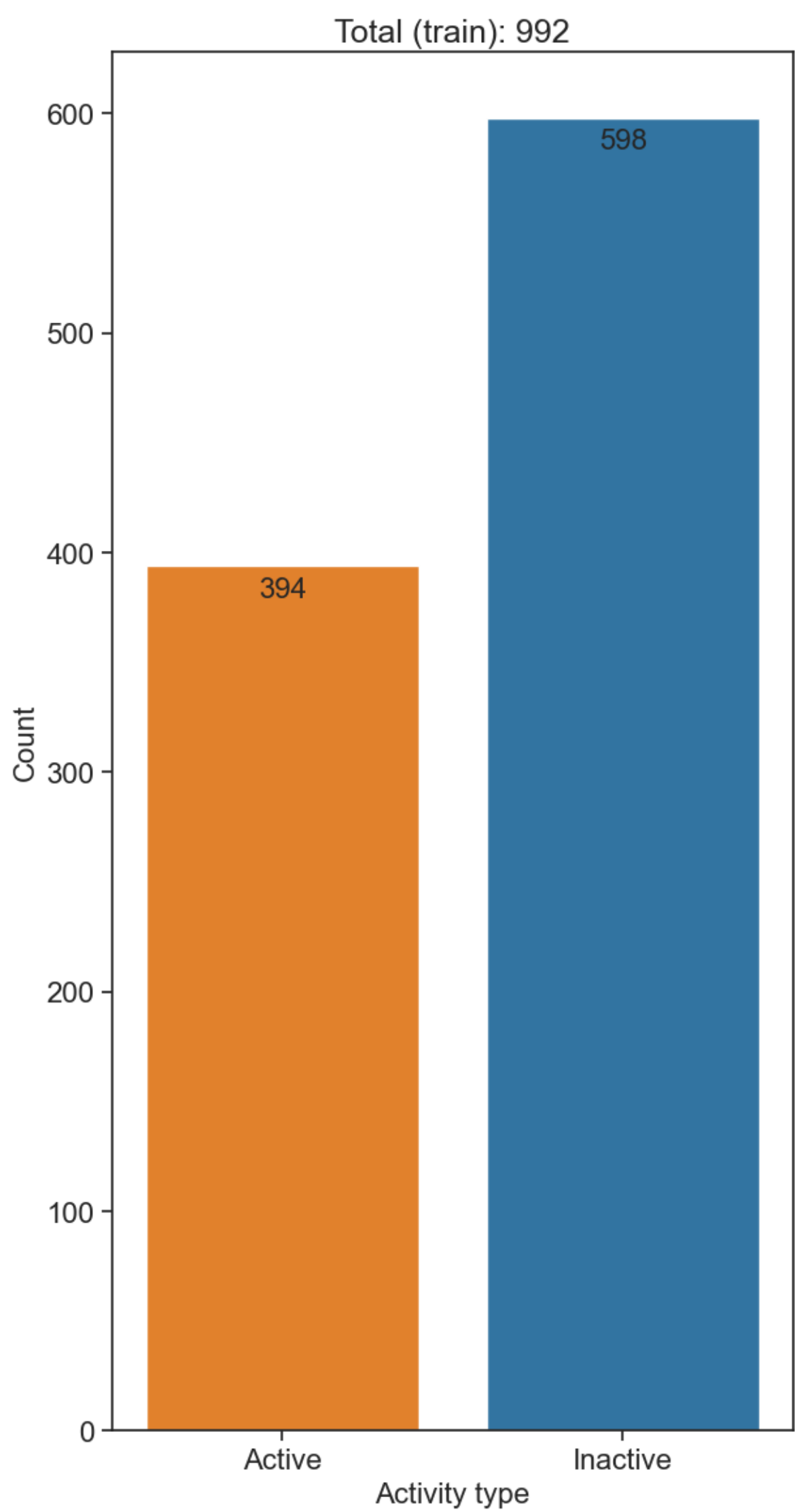
Metric	Train	Test
AUC	0.9983	0.9399
accuracy	0.9718	0.8831
balanced accuracy	0.9658	0.8671
recall_(sens)	0.995	0.9463
specificity	0.9365	0.7879
precision	0.9597	0.8704
f1_score	0.9716	0.8814