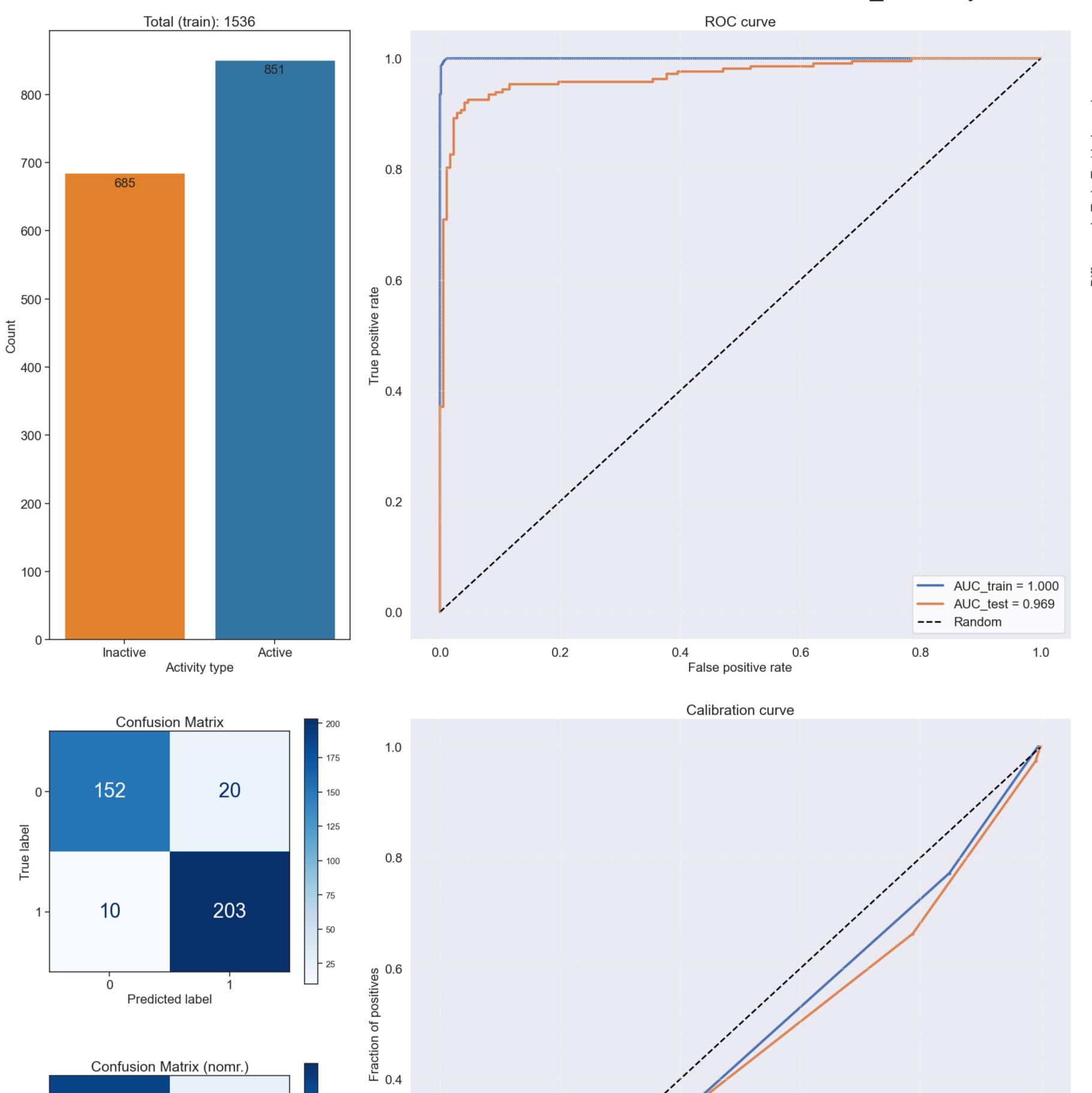
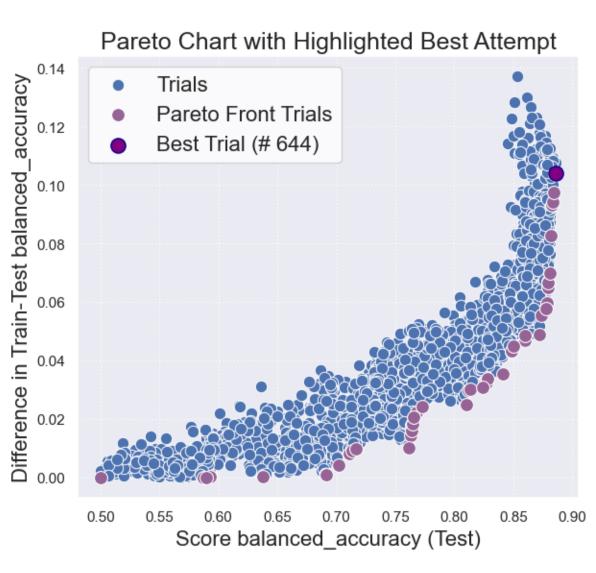
Protein: P45452 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	0.4176
colsample_bytree	0.514
gamma	0.0419
lambda	3.3773
learning_rate	0.0537
max_depth	7.0
min_child_weight	1.0
n_estimators	789.0
reg_alpha	0.0049
reg_lambda	0.883
scale_pos_weight	5.5136
subsample	0.9241

		Acti	ive/Inactive probat	oility		
60	Inac	tive zone		Active :	zone	type active (213) inactive (172) decoy (385)
40						
Percent 00						
20						
10						
	0.0 0.2	0.4	4 Probability	0.6 0	.8	1.0

Confusion Matrix (nomr.)

Predicted label

0.12

0.95

- 0.6

0.4

0.2

0.2

0.0

0.0

0.2

0.88

0.047

True label

Metric	Train	Test
AUC	0.9998	0.9685
accuracy	0.987	0.9221
balanced accuracy	0.9854	0.9184
recall_(sens)	1.0	0.9531
specificity	0.9708	0.8837
precision	0.977	0.9103
f1_score	0.987	0.9218

train: 0.012

0.8

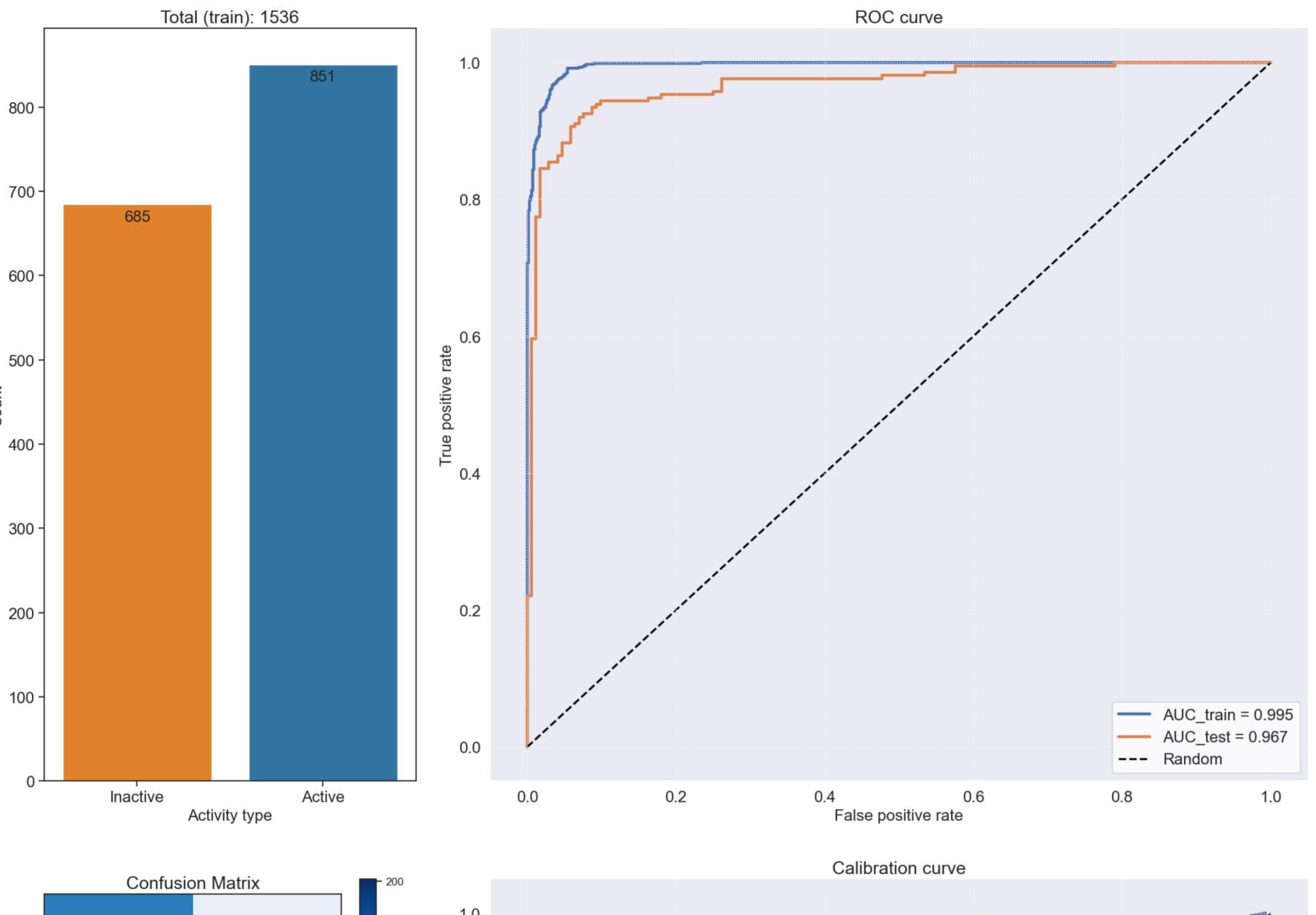
0.4 0.6 Mean predicted probability

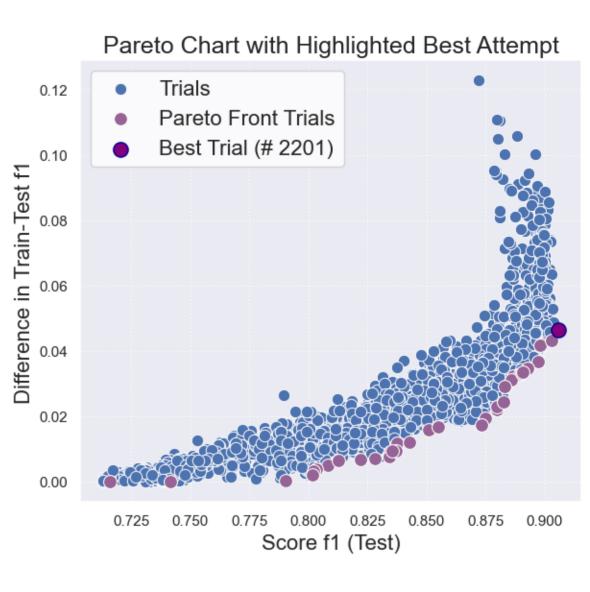
test: 0.0633

--- perfectly_calibrated

1.0

Protein: P45452 - metric: f1





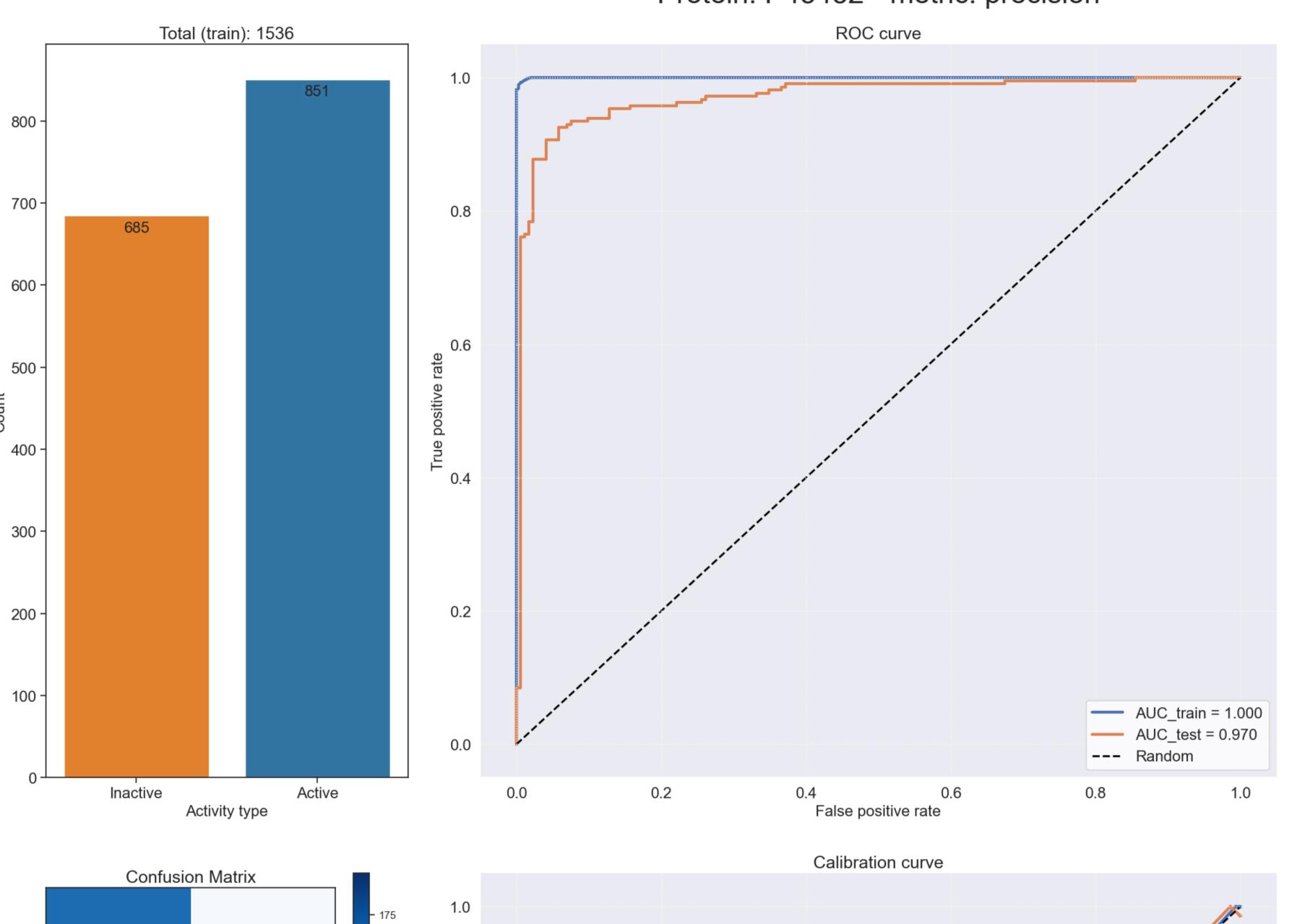
XGBoost Parameters	Value
alpha	2.5225
colsample_bytree	0.6923
gamma	0.1748
lambda	4.1516
learning_rate	0.0186
max_depth	10.0
min_child_weight	2.0
n_estimators	377.0
reg_alpha	0.8839
reg_lambda	0.1522
scale_pos_weight	3.5447
subsample	0.9389

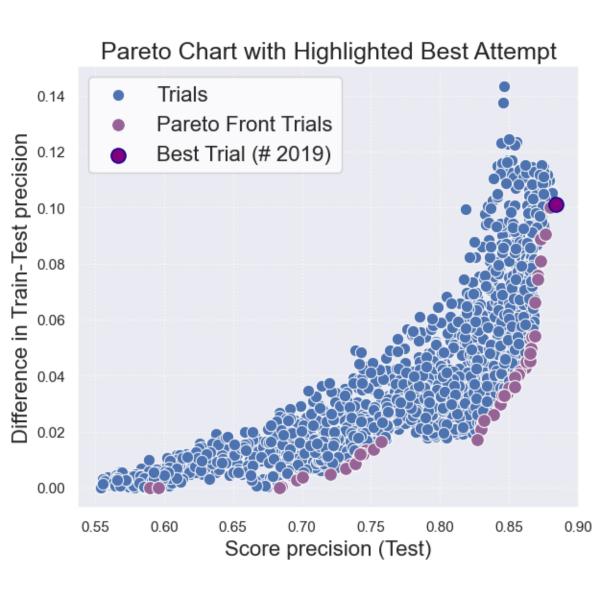
True label	145	27	- 175 - 150 - 125	1.0					, or	
True	12	201	- 100 - 75 - 50	0.8				, or		
	0 Predicte	1 ed label	- 25	Fraction of positives o			proces			
	Confusion M	latrix (nomr.)	- 0.9	Fract 7.0			proper			
label O	0.84	0.16	- 0.8 - 0.7 - 0.6 - 0.5	0.2		, or				
True	0.056	0.94	- 0.4 - 0.3 - 0.2 - 0.1	0.0	, co				train: 0	
	0 Predicte	1 ed label			0.0	0.2	0.4 Mean predicte	0.6 d probability	0.8	1.0

20			ve probability		
35	Inactive	e zone		Active zone	type active (213) inactive (172) decoy (385)
20					
Percent					
10					
5 0					
	0.0 0.2	0.4 Prob	0.6 ability	0.8	1.0

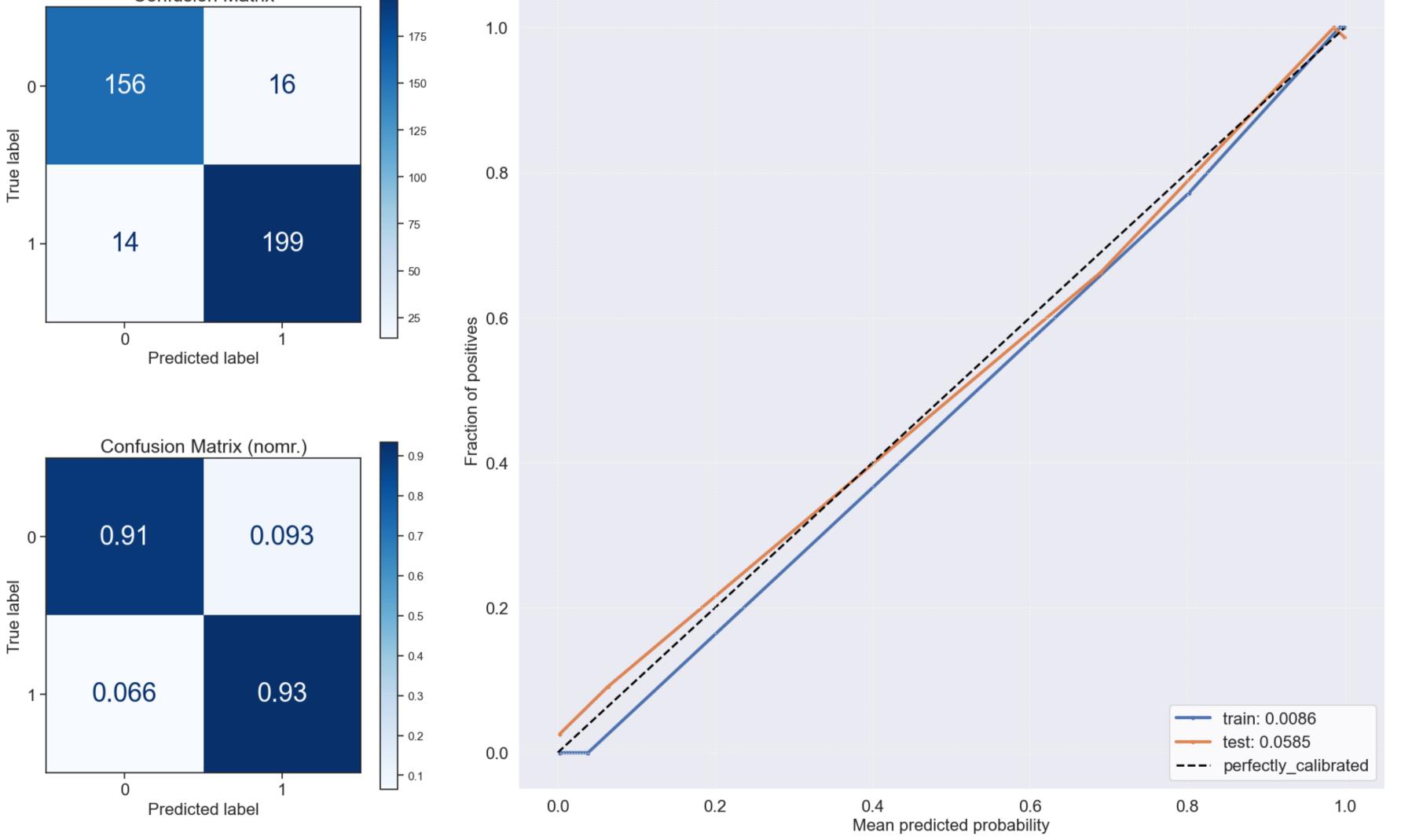
Metric	Train	Test
AUC	0.9954	0.967
accuracy	0.9408	0.8987
balanced accuracy	0.9337	0.8933
recall_(sens)	0.9988	0.9437
specificity	0.8686	0.843
precision	0.9043	0.8816
f1_score	0.9402	0.8981

Protein: P45452 - metric: precision





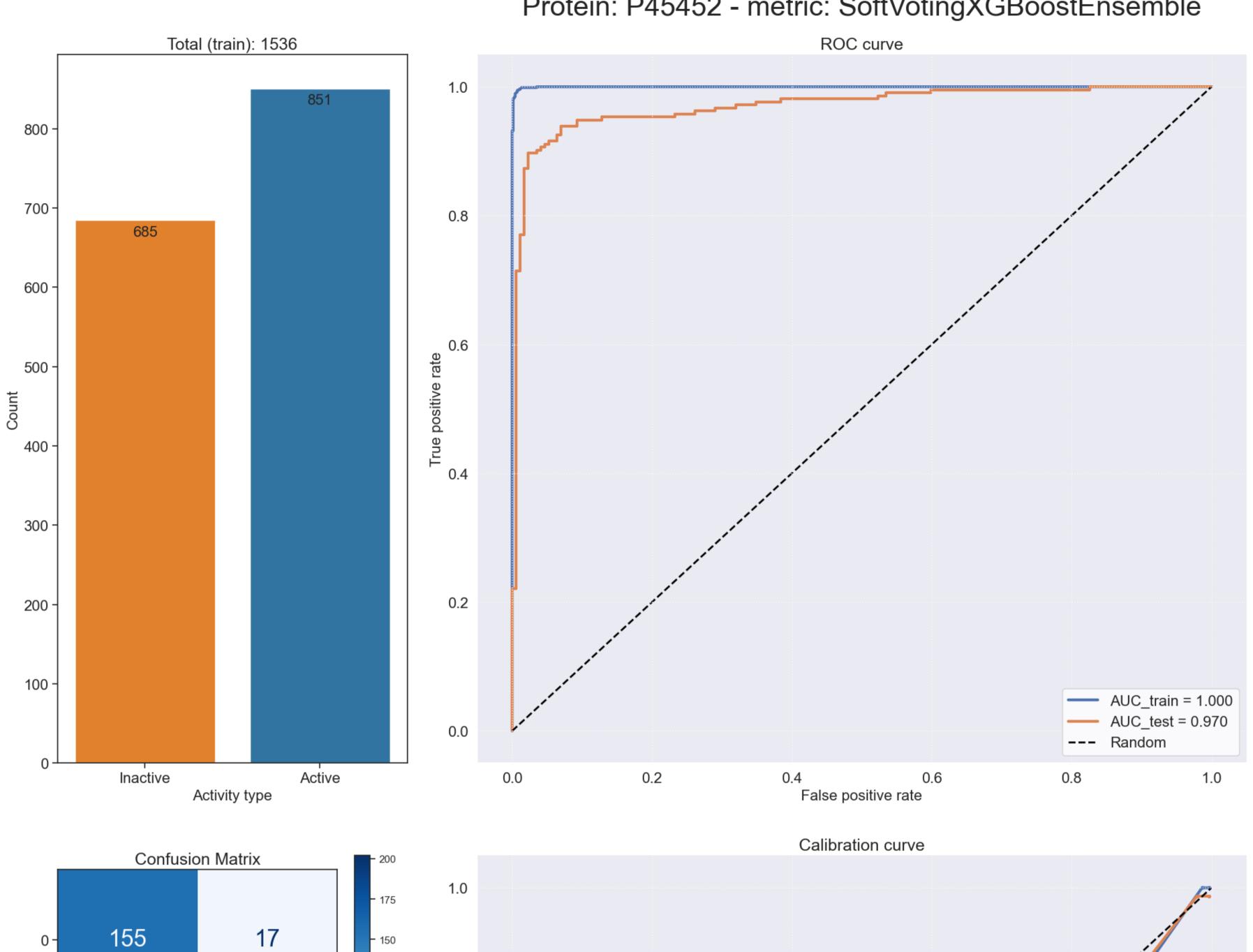
XGBoost Parameters	Value
alpha	3.6113
colsample_bytree	0.5958
gamma	0.0206
lambda	4.7026
learning_rate	0.113
max_depth	9.0
min_child_weight	3.0
n_estimators	516.0
reg_alpha	0.0278
reg_lambda	0.7203
scale_pos_weight	2.1743
subsample	0.9811

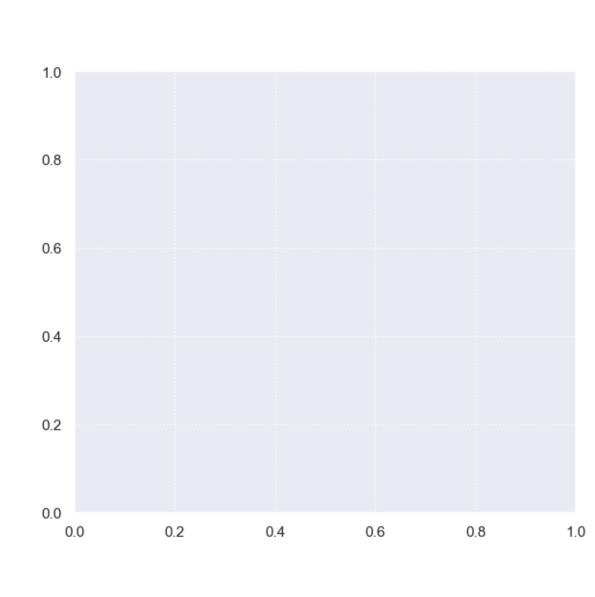


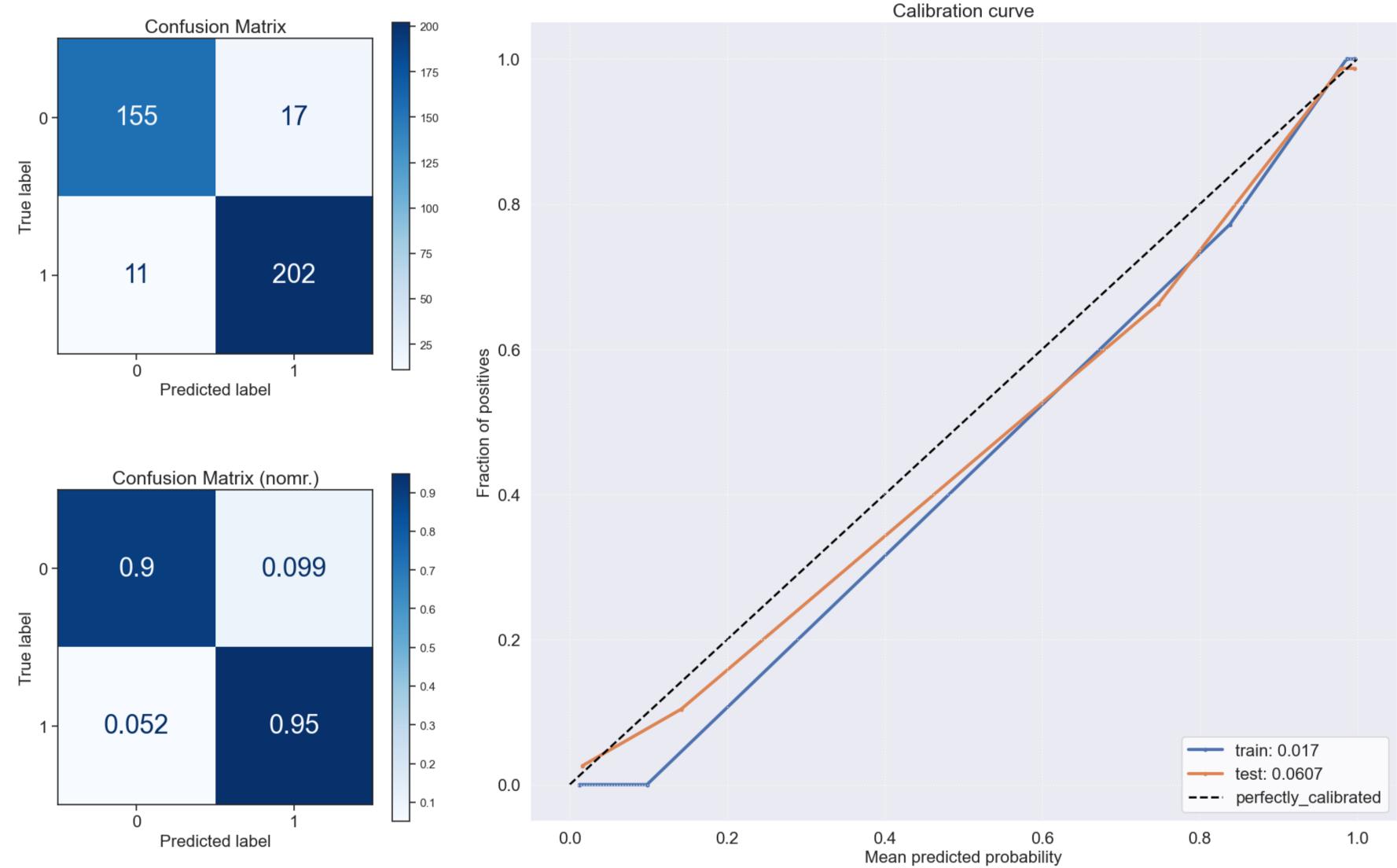
	Active/Inactive probability							
50	In	nactive zone			Active z	type active ((172)	
Percent 0.00								
10								
0	0.0	0.2).4 Probat	0.6 oility	8.0	3 1	.0	

Metric	Train	Test
AUC	0.9999	0.9703
accuracy	0.9909	0.9221
balanced accuracy	0.9898	0.9206
recall_(sens)	1.0	0.9343
specificity	0.9796	0.907
precision	0.9838	0.9256
f1_score	0.9909	0.922

Protein: P45452 - metric: SoftVotingXGBoostEnsemble







40	Ina	ictive zone		Active	zone type active ((172)
ercent						
40						
0		═ ┋ ═╸ <mark>┈┈╒╸┈╒╸</mark>				
O	0.0	.2 0.	.4 Probability	0.6).8 1.	.0

Active/Inactive probability

Metric	Train	Test
AUC	0.9998	0.97
accuracy	0.9805	0.9273
balanced accuracy	0.9781	0.9248
recall_(sens)	1.0	0.9484
specificity	0.9562	0.9012
precision	0.9659	0.9224
f1_score	0.9804	0.9271