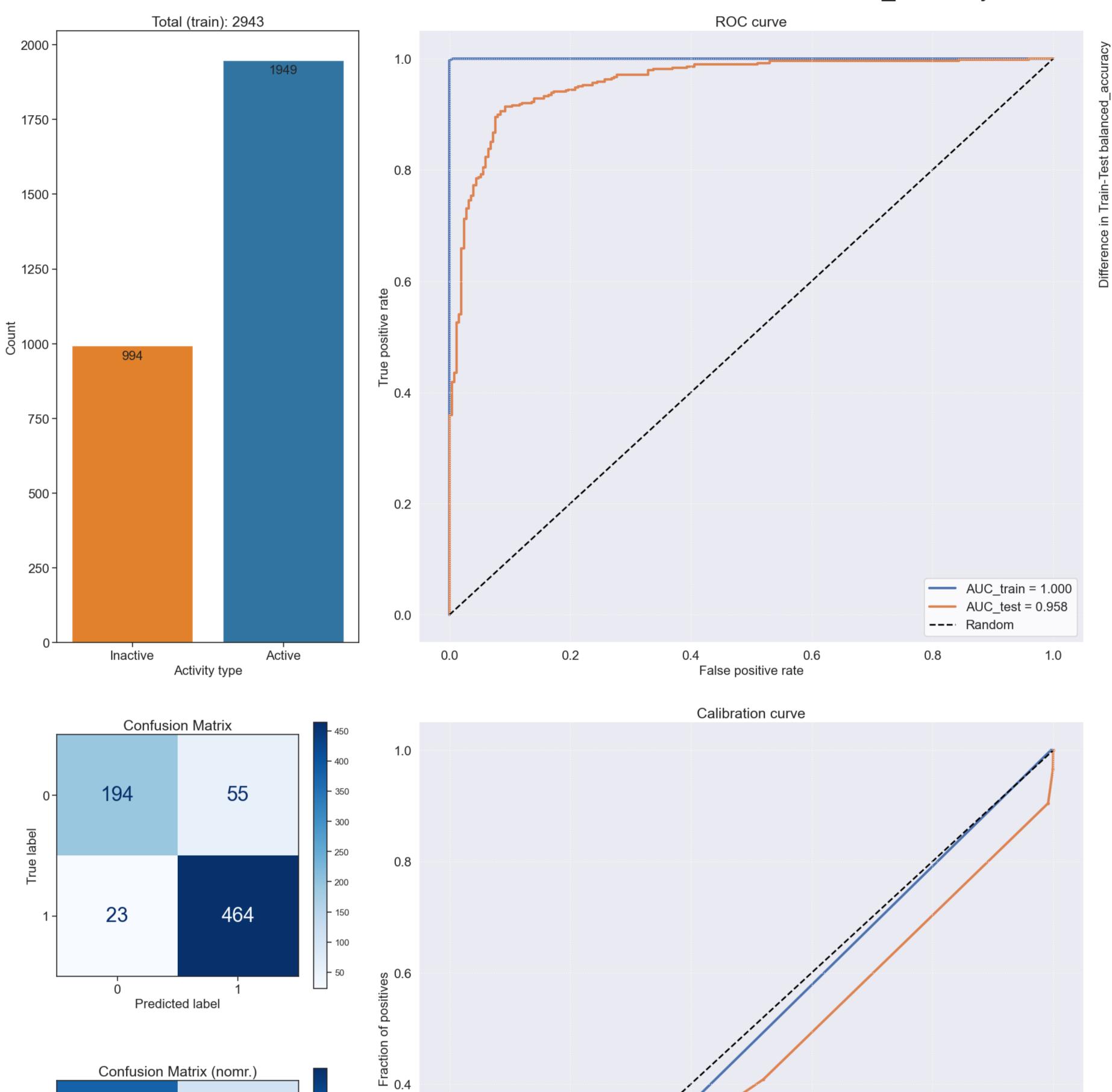
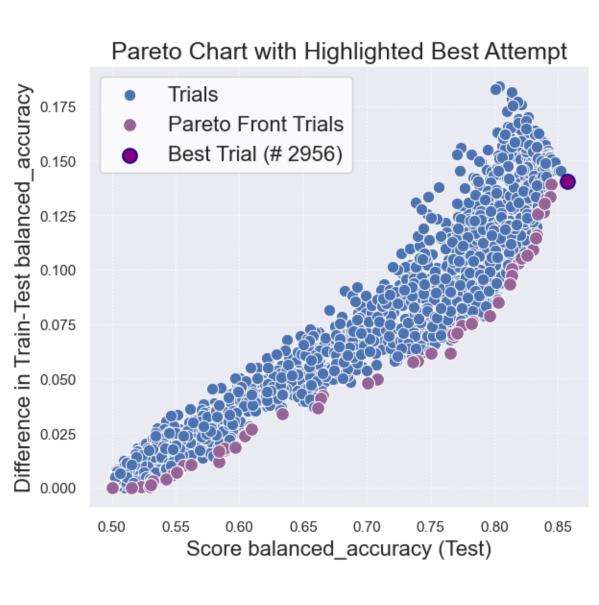
## Protein: P21554 - metric: balanced\_accuracy





XGBoost Parameters	Value
alpha	1.0245
colsample_bytree	0.646
gamma	0.0009
lambda	3.8996
learning_rate	0.1684
max_depth	10.0
min_child_weight	2.0
n_estimators	833.0
reg_alpha	0.3097
reg_lambda	0.1378
scale_pos_weight	4.0942
subsample	0.8821

		Active/Inactiv	e probability		
00	Inactive zon	е		Active zone	active (487)
70					decoy (733)
50 -					
40 -					
00					
20					
10 -	<b>-</b>				
0.0	0.2	0.4 Proba	0.6 ability	0.8	1.0

0.2

0.0

0.0

0.2

Confusion Matrix (nomr.)

Predicted label

0.22

0.95

0.78

0.047

0-

True label

Metric	Train	Test
AUC	1.0	0.9583
accuracy	0.998	0.894
balanced accuracy	0.997	0.8659
recall_(sens)	1.0	0.9528
specificity	0.994	0.7791
precision	0.9969	0.894
f1_score	0.998	0.8921

--- train: 0.0021

0.8

test: 0.084

--- perfectly\_calibrated

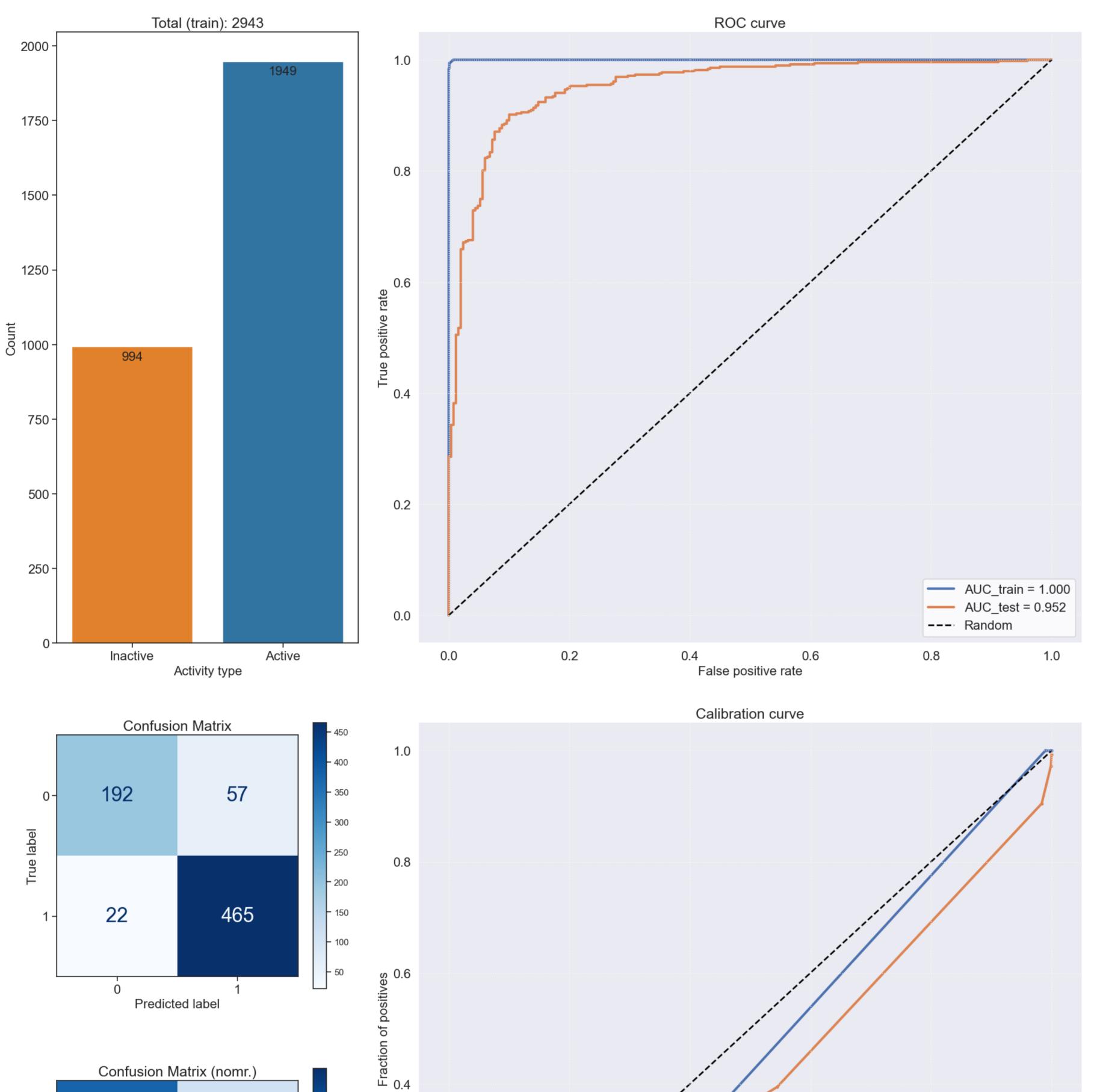
1.0

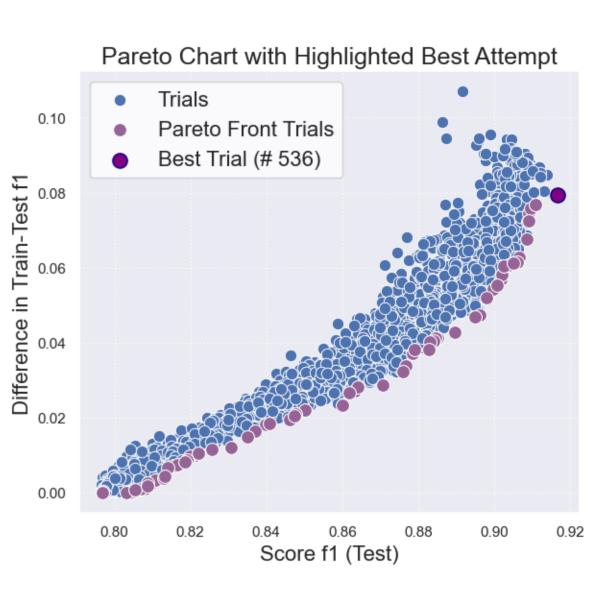
		Active/Inactive	e probability				
80	Inactive zone	İ		Active zone	type active (487)	Metric	Train
70					inactive (249) decoy (733)	AUC	•
60						accuracy	0.9
50						balanced accuracy	0.9
Percent 05						recall_(sens)	
30						specificity	0.9
20						precision	0.99
10						f1_score	0.9
0	0.0 0.2	0.4	0.6	0.8	1.0		

0.2

0.4 0.6 Mean predicted probability

## Protein: P21554 - metric: f1





XGBoost Parameters	Value
alpha	0.6166
colsample_bytree	0.9939
gamma	0.0022
lambda	2.3202
learning_rate	0.2396
max_depth	7.0
min_child_weight	6.0
n_estimators	784.0
reg_alpha	0.2377
reg_lambda	0.5569
scale_pos_weight	4.3505
subsample	0.6757

		Ac	tive/Inactive probab	ility		
00	In			Active	inac	ve (487) tive (249) by (733)
50						
Percent 30						
20						
0	0.0	1	.4 0 Probability	0.6	).8	1.0

0.2

0.0

0.0

0.2

0.2

Confusion Matrix (nomr.)

Predicted label

0.23

0.95

0.77

0.045

0-

True label

Metric	Train	Test
AUC	1.0	0.9523
accuracy	0.9949	0.8927
balanced accuracy	0.9925	0.863
recall_(sens)	1.0	0.9548
specificity	0.9849	0.7711
precision	0.9924	0.8908
f1_score	0.9949	0.8905

- train: 0.0075

0.8

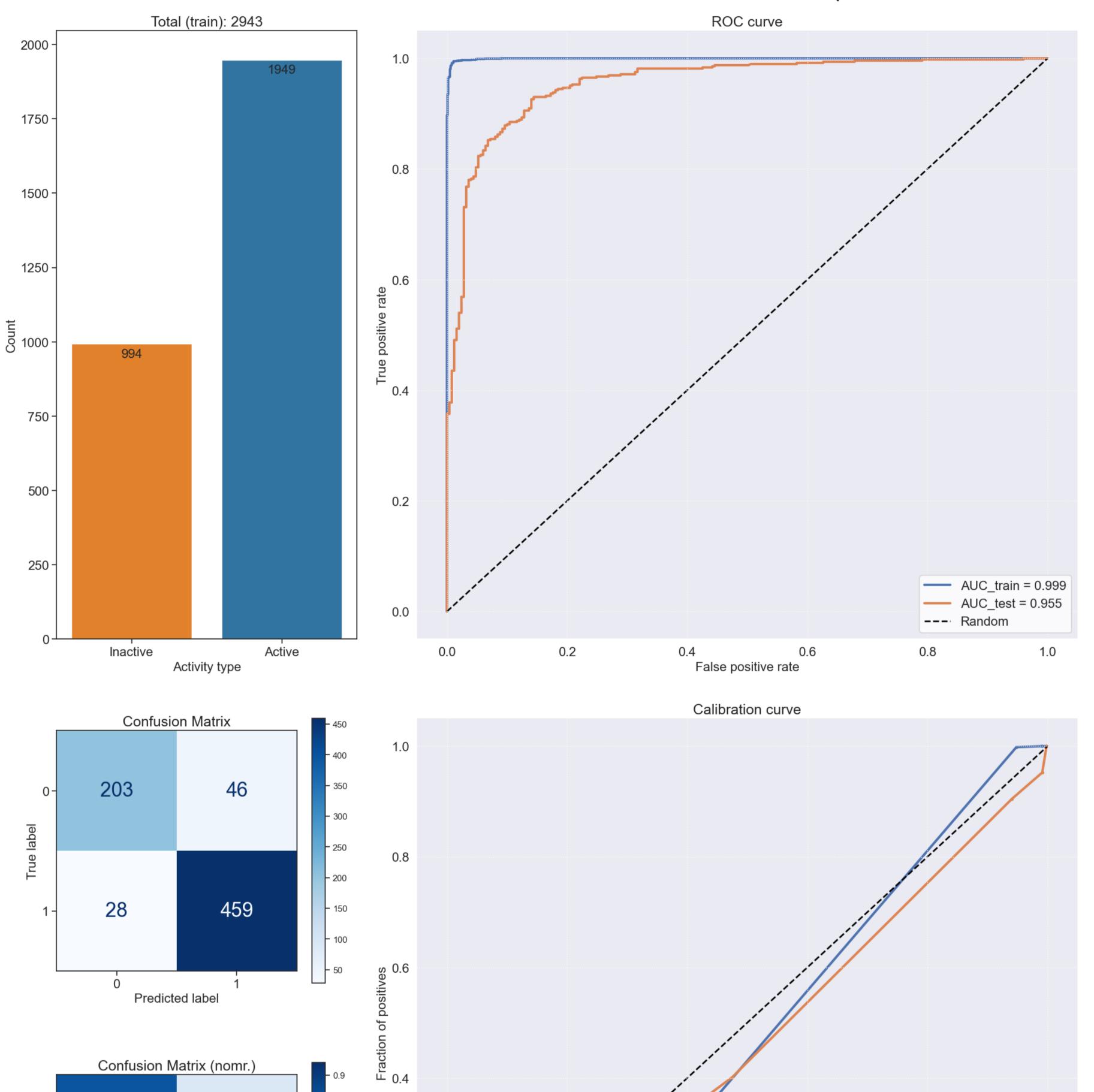
0.4 0.6 Mean predicted probability

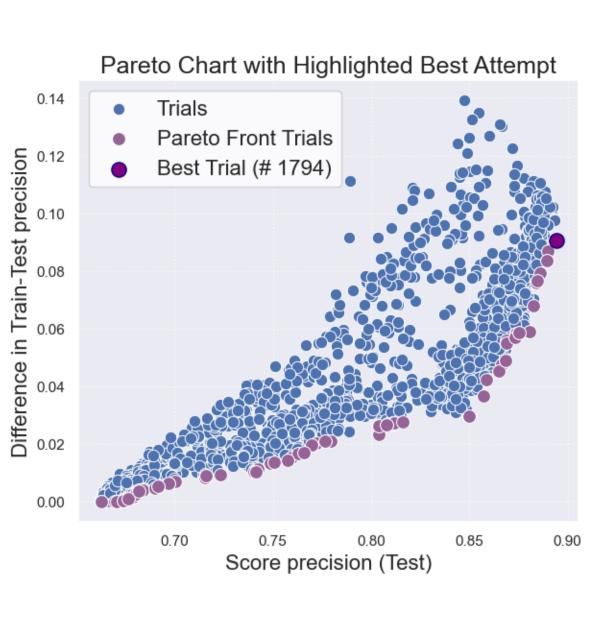
test: 0.0854

--- perfectly\_calibrated

1.0

## Protein: P21554 - metric: precision





XGBoost Parameters	Value
alpha	1.0329
colsample_bytree	0.7872
gamma	0.1934
lambda	0.1739
learning_rate	0.0834
max_depth	9.0
min_child_weight	4.0
n_estimators	803.0
reg_alpha	0.8518
reg_lambda	0.1466
scale_pos_weight	1.4653
subsample	0.5877

F0		Ad	ctive/Inactive probab	oility		
		active zone		Active	zone a	type active (487) nactive (249) lecoy (733)
30						
Percent 0 0						
10						
0	0.0	0.2	0.4 Probability	0.6	).8	1.0

Confusion Matrix (nomr.)

Predicted label

0.18

0.94

- 0.6

- 0.5

- 0.3

0.2

0.2

0.0

0.0

0.2

0.82

0.057

0-

True label

Metric	Train	Test
AUC	0.9995	0.955
accuracy	0.9878	0.8995
balanced accuracy	0.9834	0.8789
recall_(sens)	0.9969	0.9425
specificity	0.9698	0.8153
precision	0.9848	0.9089
f1_score	0.9877	0.8985

Metric	irain	iest
AUC	0.9995	0.955
accuracy	0.9878	0.8995
balanced accuracy	0.9834	0.8789
recall_(sens)	0.9969	0.9425
specificity	0.9698	0.8153
precision	0.9848	0.9089
f1 score	0.9877	0.8985

--- train: 0.016

0.8

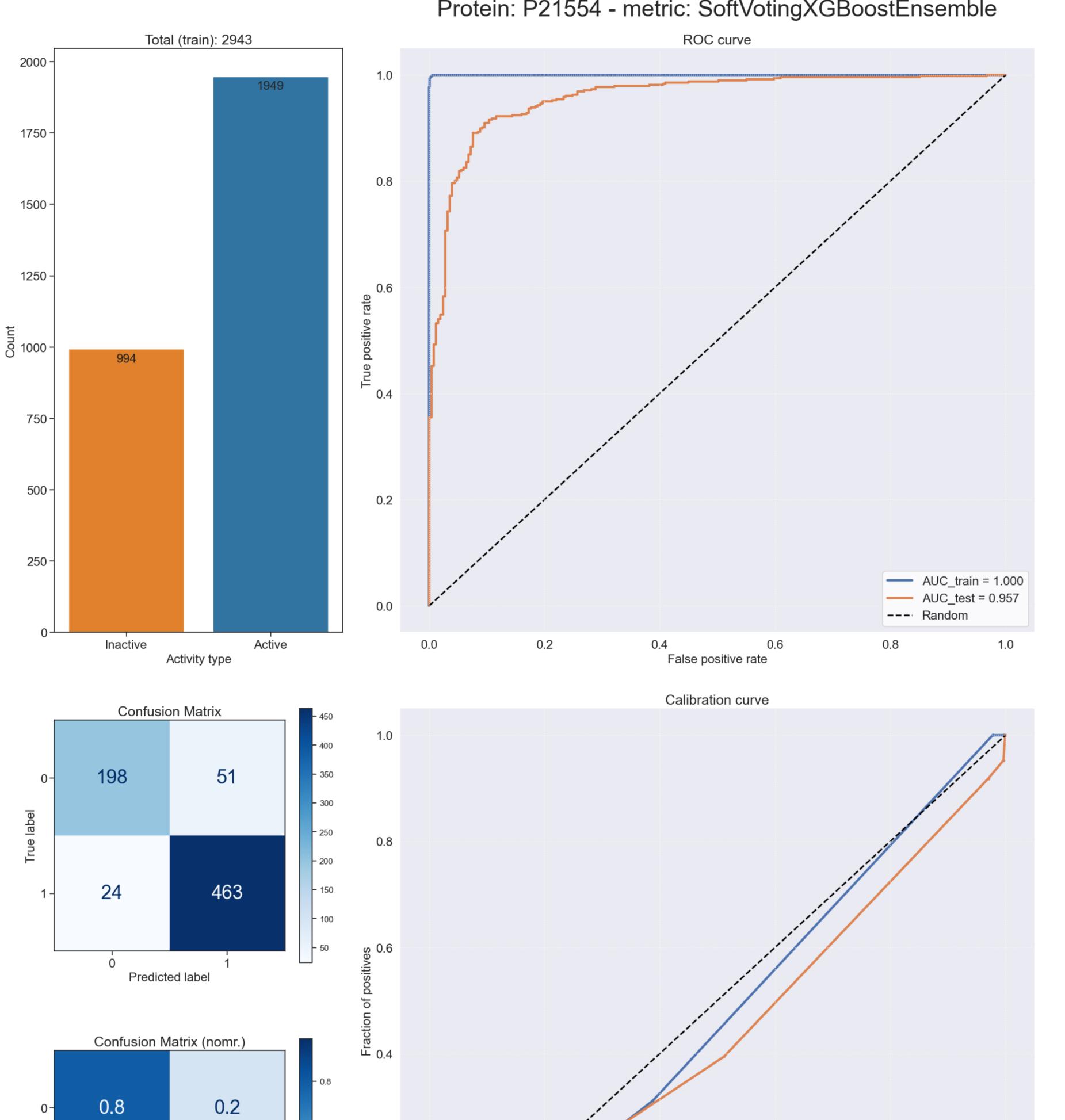
0.4 0.6 Mean predicted probability

- test: 0.0763

--- perfectly\_calibrated

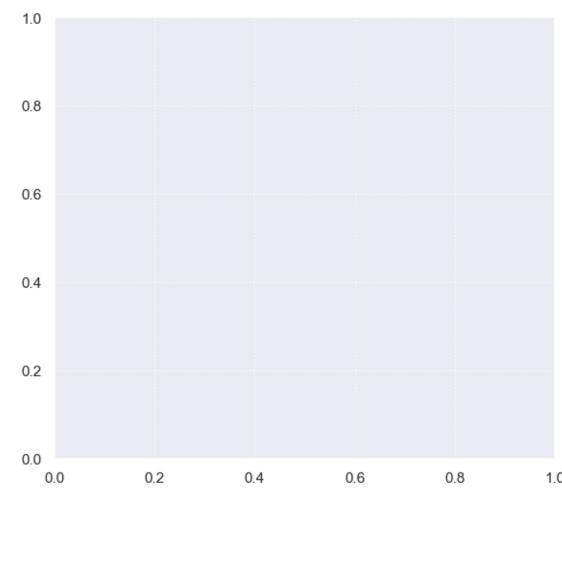
1.0

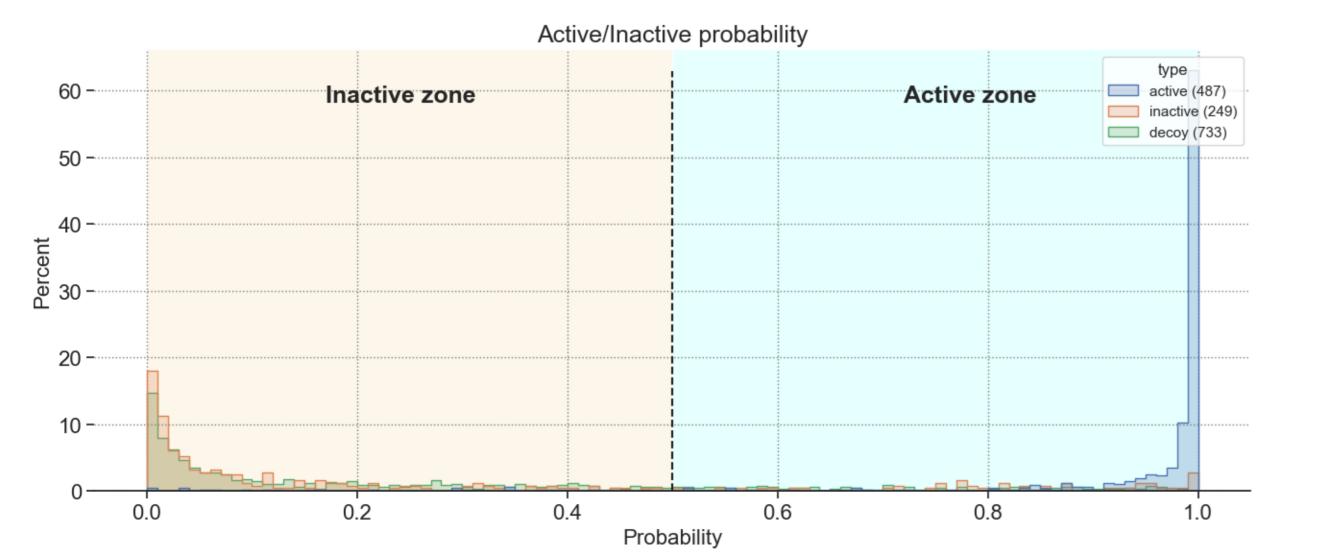
## Protein: P21554 - metric: SoftVotingXGBoostEnsemble



0.2

0.4 0.6 Mean predicted probability





- 0.6

- 0.4

0.2

0.95

0.2

0.0

0.0

True label

0.049

Predicted label

Metric	Train	Test
AUC	1.0	0.9575
accuracy	0.9966	0.8981
balanced accuracy	0.995	0.8729
recall_(sens)	1.0	0.9507
specificity	0.9899	0.7952
precision	0.9949	0.9008
f1_score	0.9966	0.8966

train: 0.0066

test: 0.0779

8.0

perfectly\_calibrated

1.0