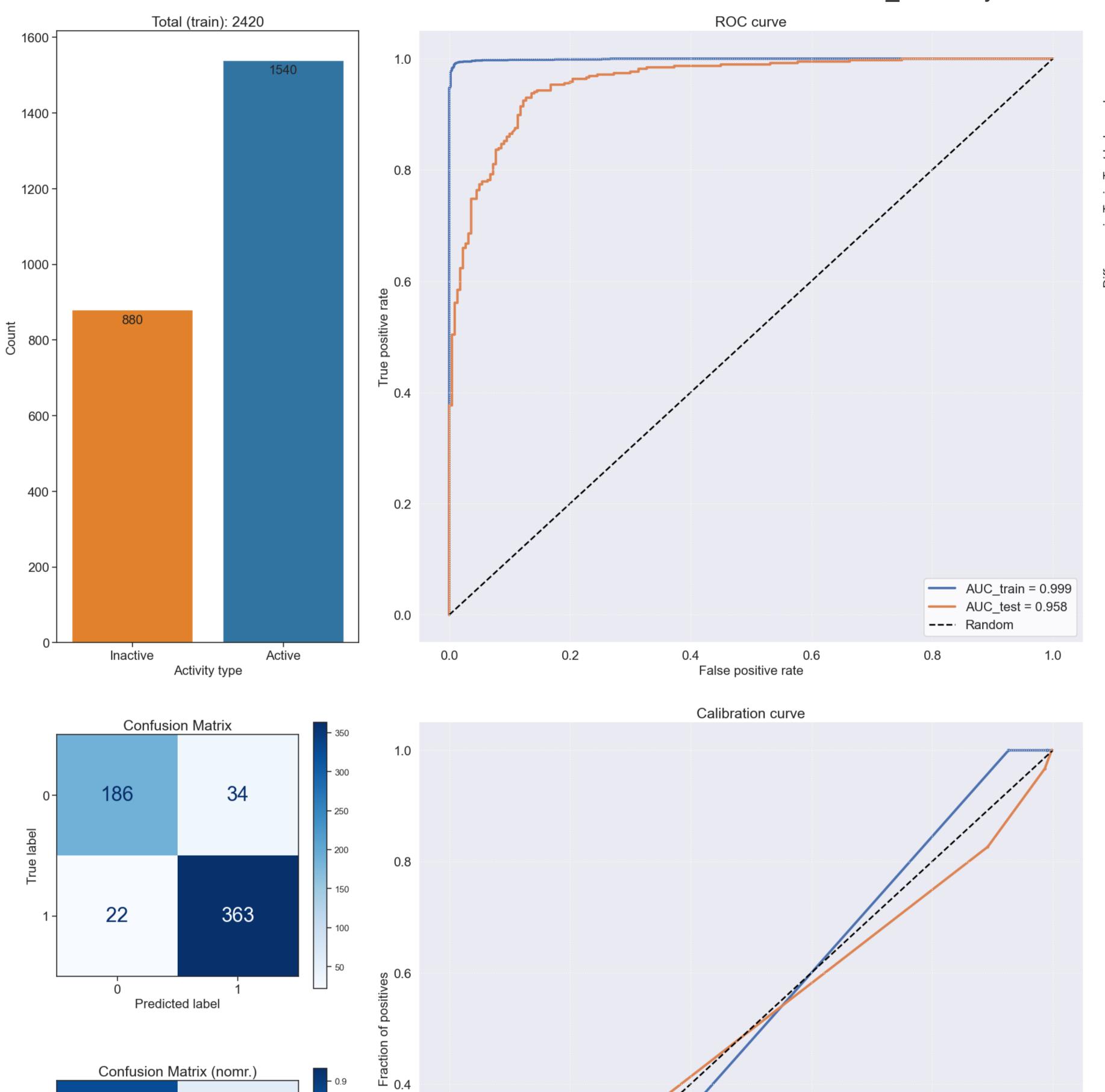
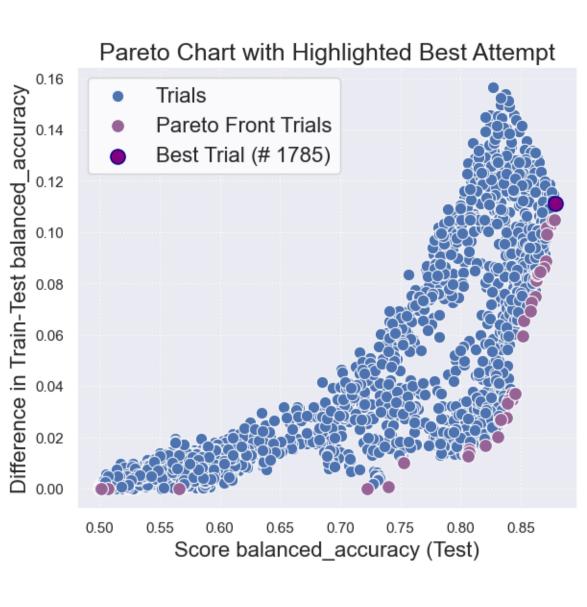
## Protein: Q9Y5N1 - metric: balanced\_accuracy





XGBoost Parameters	Value
alpha	2.0062
colsample_bytree	0.6769
gamma	0.2002
lambda	2.4677
learning_rate	0.0942
max_depth	6.0
min_child_weight	2.0
n_estimators	888.0
reg_alpha	0.6371
reg_lambda	0.7143
scale_pos_weight	1.017
subsample	0.7495

		Active/	Inactive probabi	lity		
40	Inact	ive zone		Active :	inac	oe ve (385) ctive (220) oy (604)
30						
Percent 00						
10						
0	.0 0.2	0.4	0.	6 0	.8	1.0
			Probability			

- 0.8

- 0.6

- 0.5

0.4

0.3

0.2

0.1

0.2

0.0

0.0

0.2

Confusion Matrix (nomr.)

Predicted label

0.15

0.94

0.85

0.057

0-

True label

Metric	Train	Test
AUC	0.9992	0.958
accuracy	0.9905	0.9074
balanced accuracy	0.9894	0.8942
recall_(sens)	0.9935	0.9429
specificity	0.9852	0.8455
precision	0.9916	0.9144
f1_score	0.9905	0.9069

Metric	Train	Test
AUC	0.9992	0.958
accuracy	0.9905	0.9074
balanced accuracy	0.9894	0.8942
recall_(sens)	0.9935	0.9429
specificity	0.9852	0.8455
precision	0.9916	0.9144
f1 score	0 9905	0 9069

--- train: 0.0144

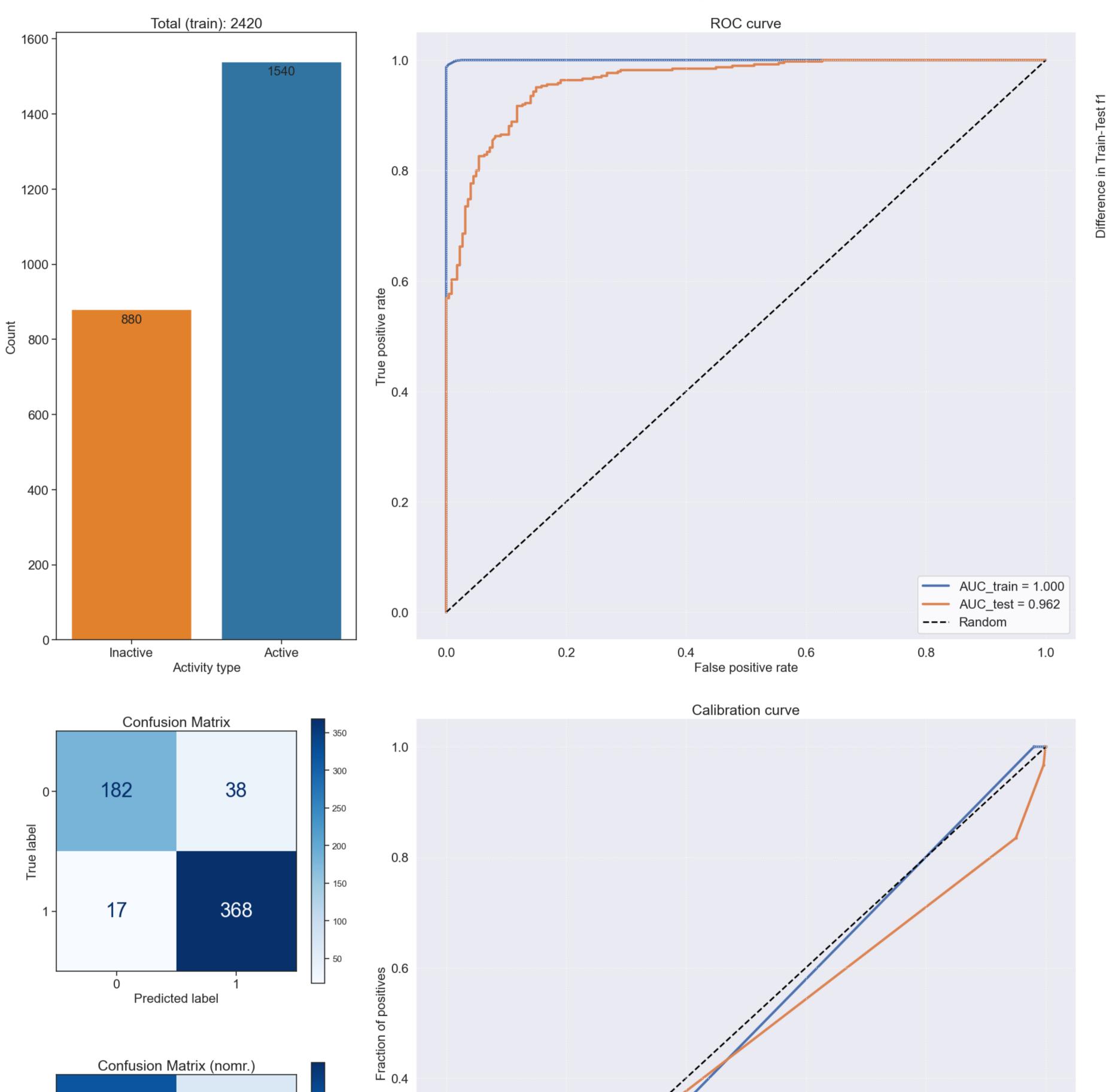
0.8

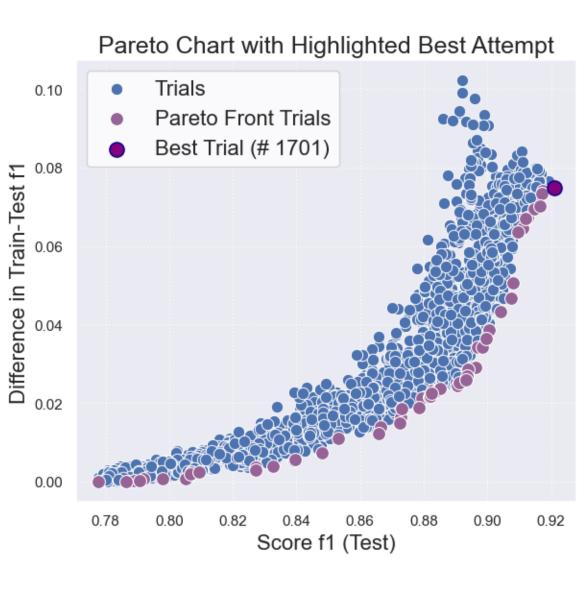
0.4 0.6 Mean predicted probability

test: 0.074

--- perfectly\_calibrated

## Protein: Q9Y5N1 - metric: f1





XGBoost Parameters	Value
alpha	4.0054
colsample_bytree	0.9797
gamma	0.0778
lambda	0.7014
learning_rate	0.0548
max_depth	9.0
min_child_weight	1.0
n_estimators	915.0
reg_alpha	0.0464
reg_lambda	0.607
scale_pos_weight	1.9013
subsample	0.7614

	Active/Inactive probability						
60	Ina	ctive zone			Active 2	zone	type active (385) inactive (220) decoy (604)
50							
Percent - 05							
20							
10							
0	0.0 0.	2 0	.4 Proba	0.6 ability	6 0.	.8	1.0

0.2

0.0

0.0

0.2

0.2

0.83

0.044

Predicted label

0-

True label

0.17

0.96

Metric	Train	Test
AUC	0.9999	0.9621
accuracy	0.9938	0.9091
balanced accuracy	0.9922	0.8916
recall_(sens)	0.9981	0.9558
specificity	0.9864	0.8273
precision	0.9923	0.9064
f1_score	0.9938	0.908

AUC	0.9999	0.9621
accuracy	0.9938	0.9091
balanced accuracy	0.9922	0.8916
recall_(sens)	0.9981	0.9558
specificity	0.9864	0.8273
precision	0.9923	0.9064
f1_score	0.9938	0.908

--- train: 0.006

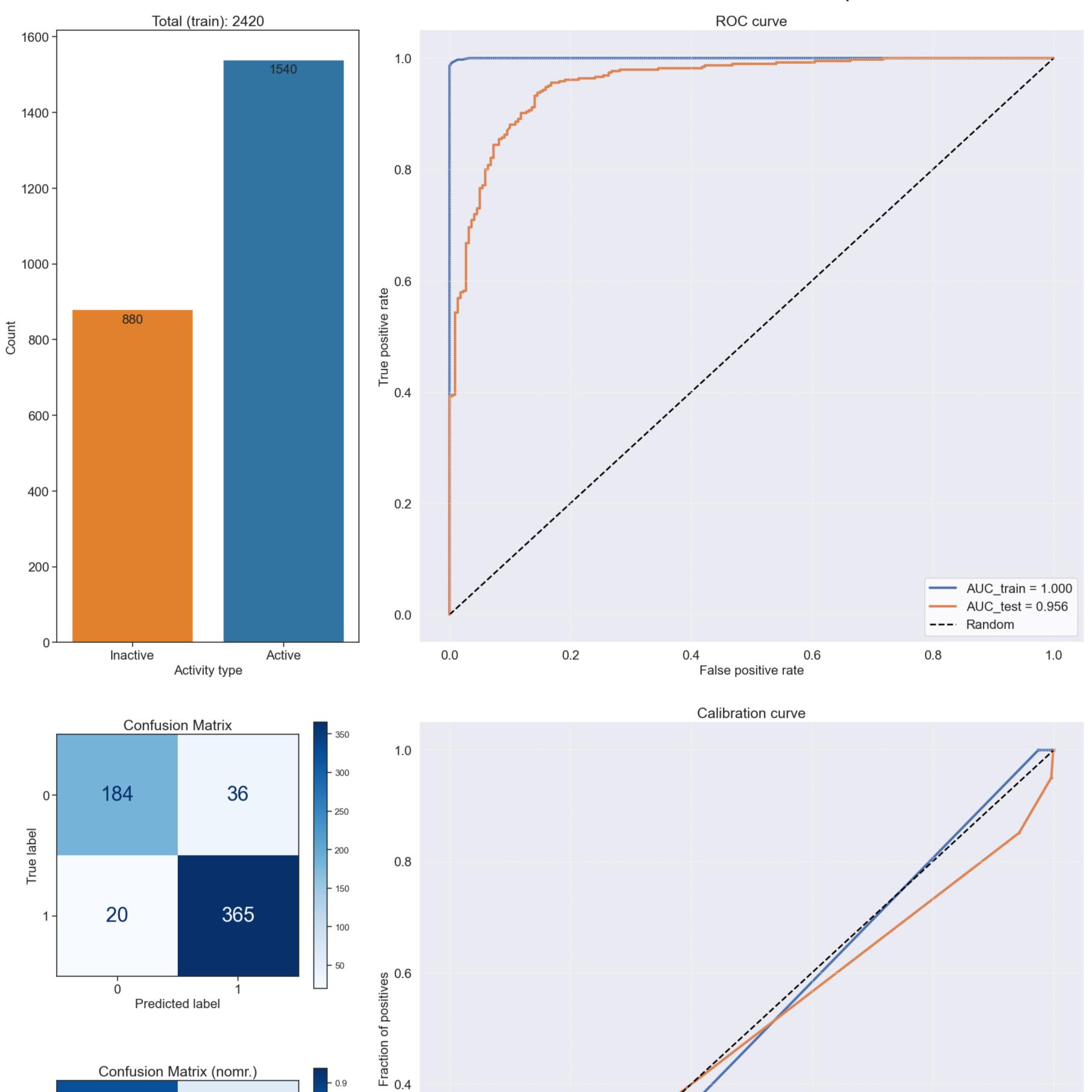
0.8

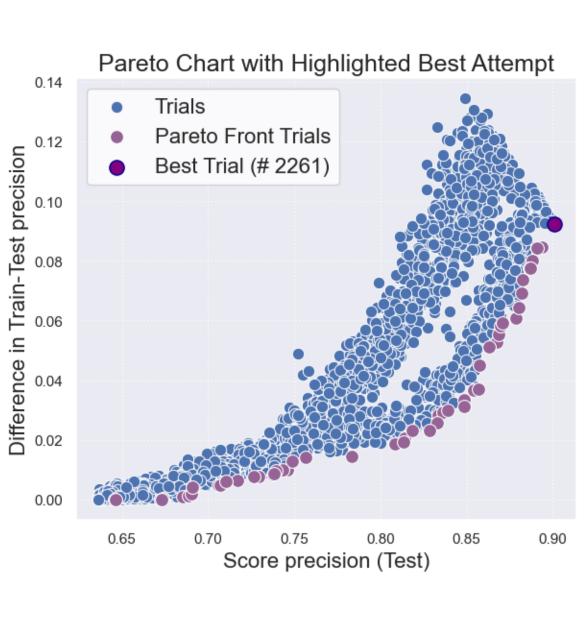
0.4 0.6 Mean predicted probability

test: 0.0744

--- perfectly\_calibrated

## Protein: Q9Y5N1 - metric: precision





XGBoost Parameters	Value
alpha	2.2589
colsample_bytree	0.8763
gamma	0.0703
lambda	0.8106
learning_rate	0.1665
max_depth	10.0
min_child_weight	3.0
n_estimators	559.0
reg_alpha	0.2258
reg_lambda	0.2166
scale_pos_weight	1.5554
subsample	0.9108

		Acti	ve/Inactive probab	ility		
60		ctive zone		Active	type active (	(385)
50					inactive decoy (	
40						
Percent						
20						
10		_				
0 —	0.0 0.	2 0.4	l 0 Probability	.6 0	.8 1	.0
			Trobability			

- 0.8

- 0.7

0.6

- 0.5

0.4

0.3

0.2

0.1

0.2

0.0

0.0

0.2

Confusion Matrix (nomr.)

Predicted label

0.16

0.95

0.84

0.052

0-

True label

Metric	Train	Test
AUC	0.9999	0.9564
accuracy	0.993	0.9074
balanced accuracy	0.9913	0.8922
recall_(sens)	0.9974	0.9481
specificity	0.9852	0.8364
precision	0.9916	0.9102
f1_score	0.993	0.9066

--- train: 0.0066

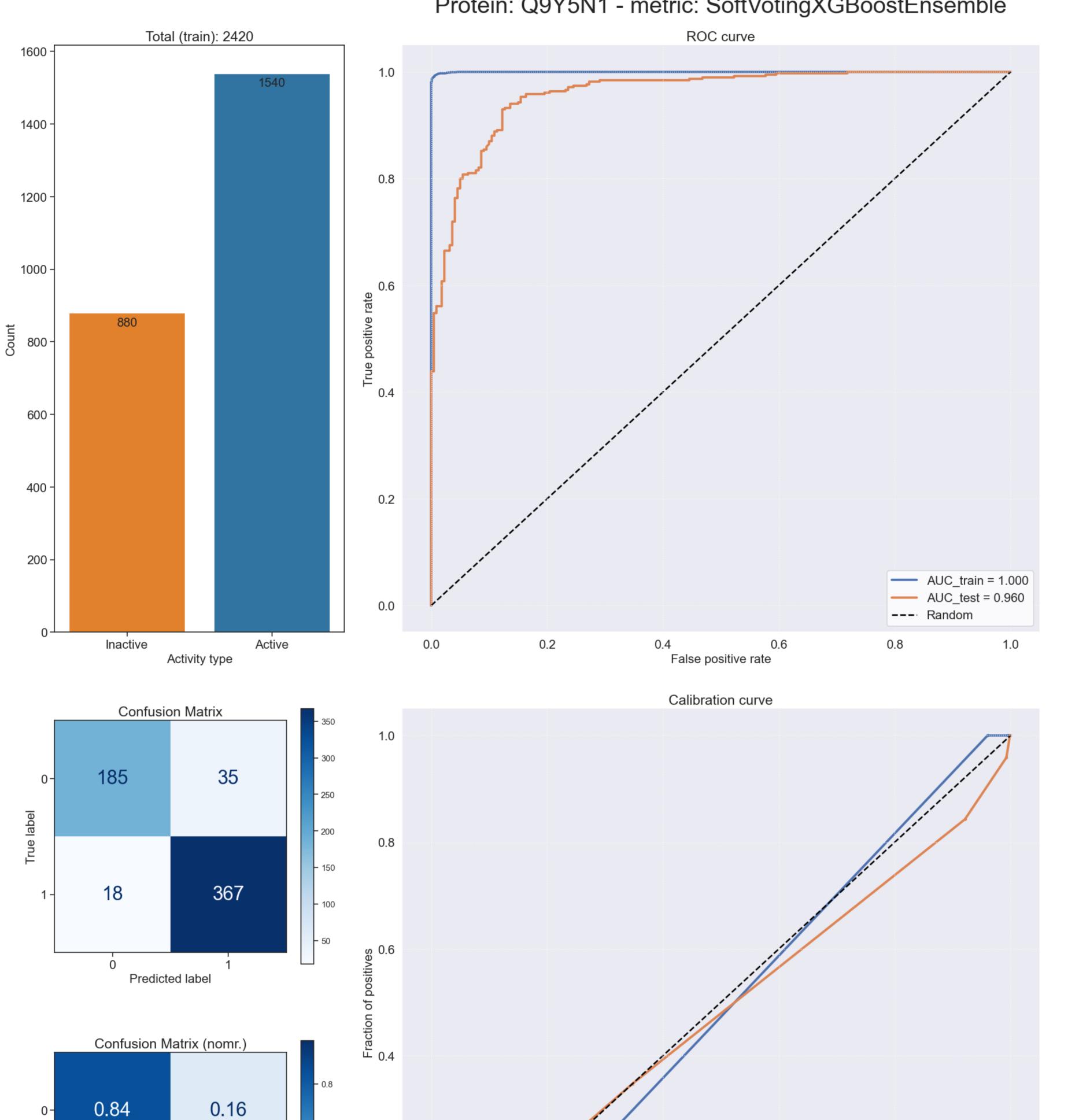
0.8

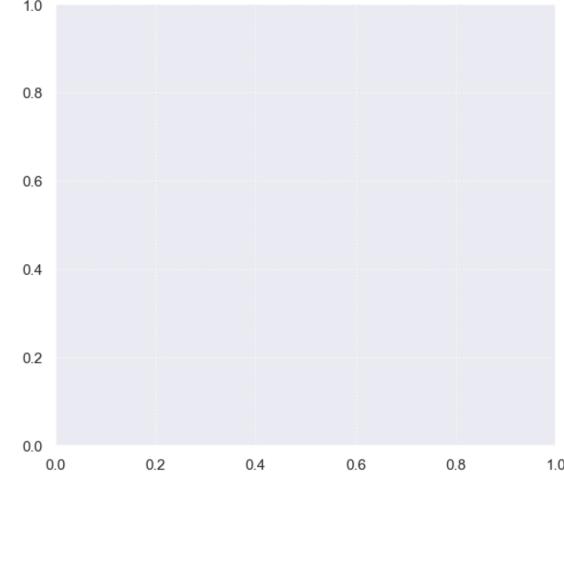
0.4 0.6 Mean predicted probability

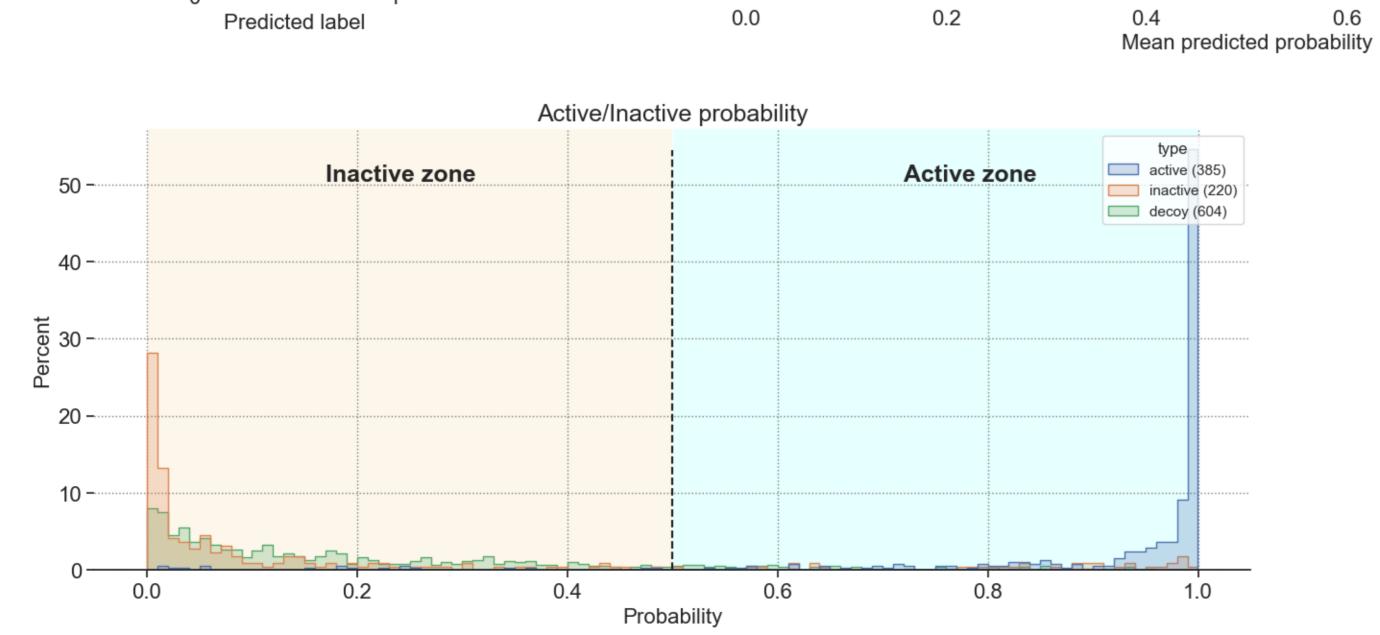
test: 0.0762

--- perfectly\_calibrated

## Protein: Q9Y5N1 - metric: SoftVotingXGBoostEnsemble







0.2

0.0

- 0.6

0.4

0.2

0.95

True label

0.047

Metric	Train	Test
AUC	0.9998	0.9596
accuracy	0.993	0.9124
balanced accuracy	0.9913	0.8971
recall_(sens)	0.9974	0.9532
specificity	0.9852	0.8409
precision	0.9916	0.9129
f1_score	0.993	0.9116

train: 0.0081

test: 0.0727

8.0

perfectly\_calibrated