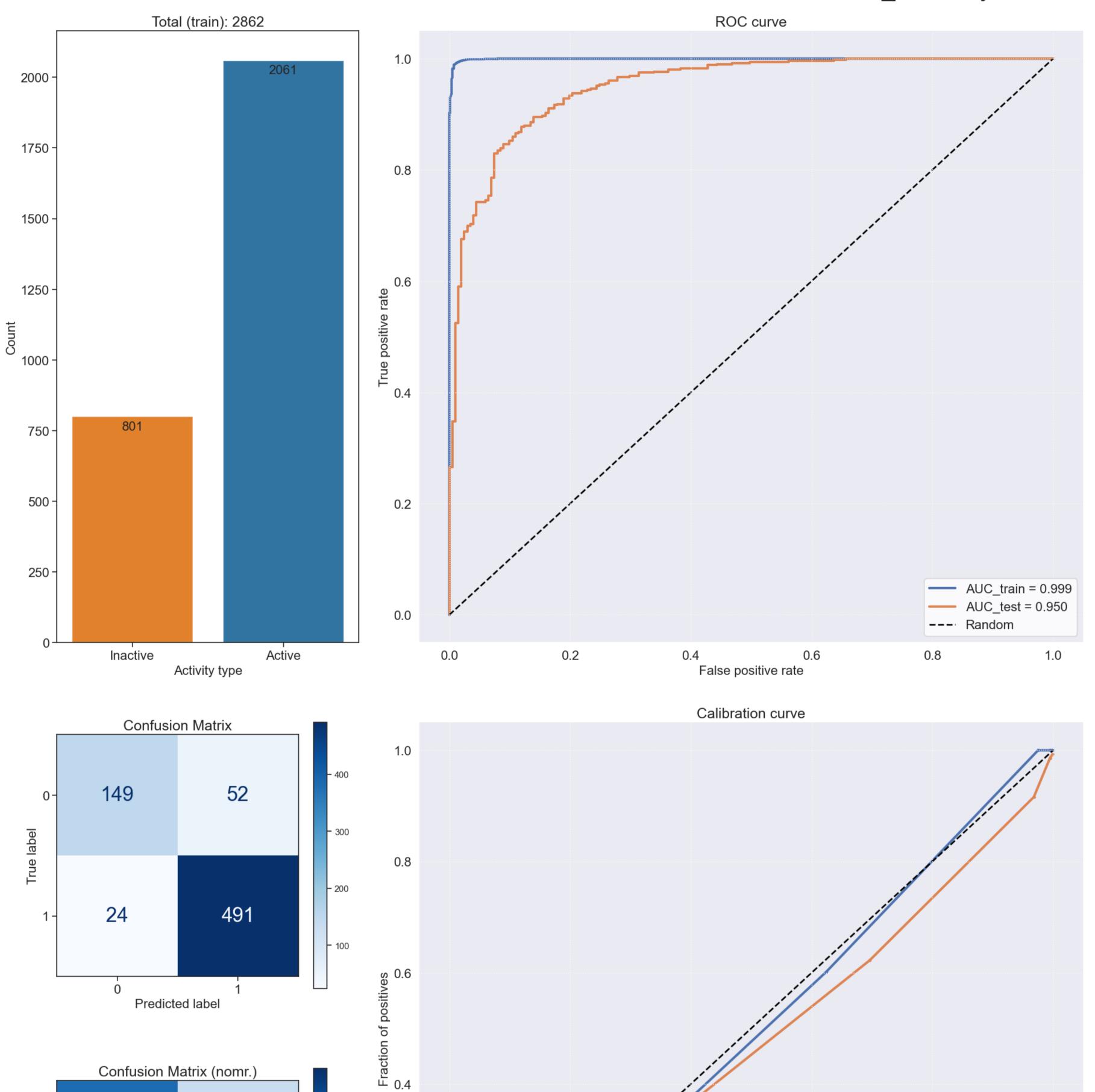
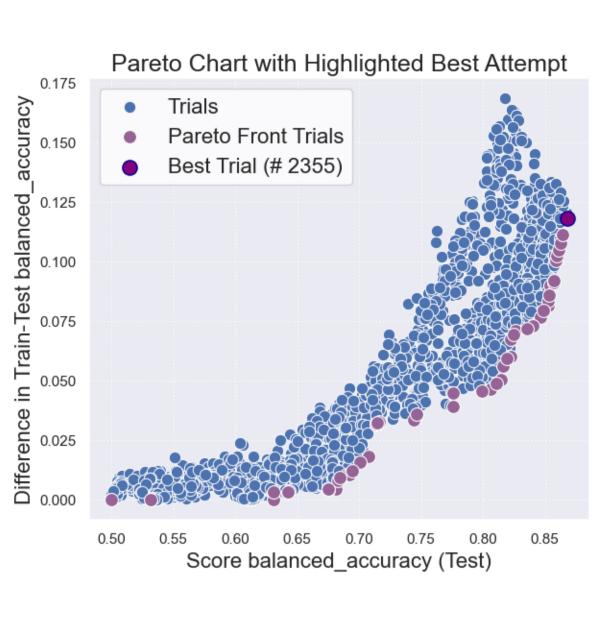
## Protein: P22303 - metric: balanced\_accuracy





XGBoost Parameters	Value
alpha	3.8053
colsample_bytree	0.8577
gamma	0.1053
lambda	3.6939
learning_rate	0.1257
max_depth	10.0
min_child_weight	4.0
n_estimators	819.0
reg_alpha	0.3784
reg_lambda	0.9479
scale_pos_weight	1.193
subsample	0.9574

	Active/Inactive probability						
50	Ina	ctive zone		Active z	type active (515) inactive (201) decoy (716)		
Percent 08 							
20							
0	0.0 0.	2 0.4	0.6 Probability	6 0.	8 1.0		

0.2

0.0

0.0

0.2

0.2

Confusion Matrix (nomr.)

Predicted label

0.26

0.95

0.74

0.047

0-

True label

Metric	Train	Test
AUC	0.9995	0.9502
accuracy	0.9906	0.8939
balanced accuracy	0.9843	0.8473
recall_(sens)	0.9985	0.9534
specificity	0.97	0.7413
precision	0.9885	0.9042
f1_score	0.9905	0.8913

MELLIC	IIaiii	1631
AUC	0.9995	0.9502
accuracy	0.9906	0.8939
balanced accuracy	0.9843	0.8473
recall_(sens)	0.9985	0.9534
specificity	0.97	0.7413
precision	0.9885	0.9042
f1 score	0.9905	0.8913

--- train: 0.0113

0.8

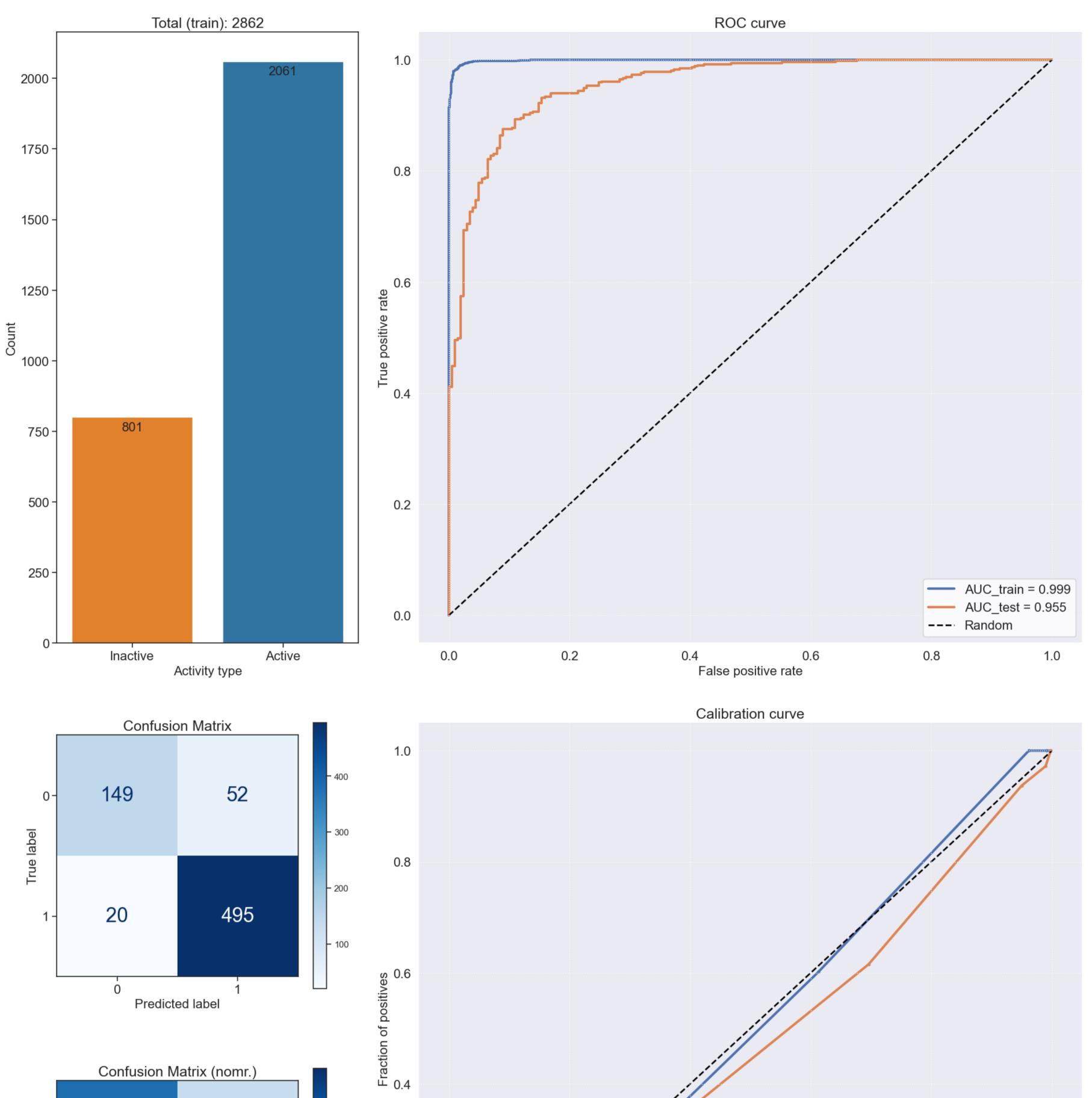
0.4 0.6 Mean predicted probability

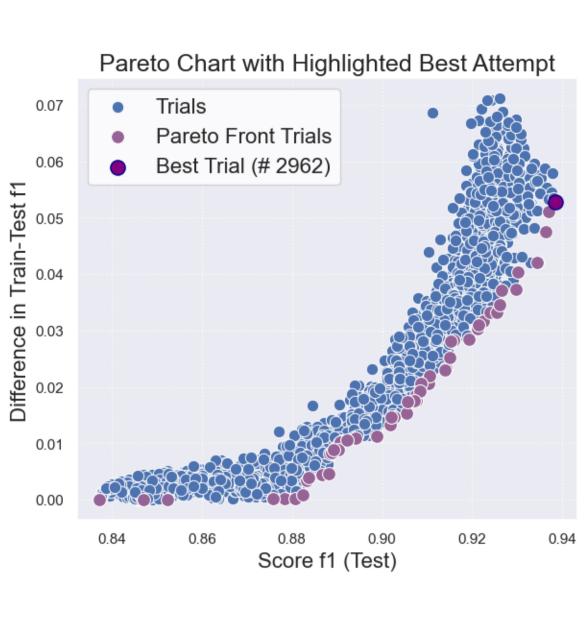
test: 0.0777

--- perfectly\_calibrated

1.0

## Protein: P22303 - metric: f1





Value
4.2569
0.7697
0.6113
2.9558
0.0457
10.0
1.0
968.0
0.3906
0.689
1.0114
0.6931

Active/Inactive probability						
40	Ina	ctive zone		Active	zone	type ctive (515) active (201)
35					de	ecoy (716)
30						
ti 25						
Percent 50						
15						
10						
5						
0						
	0.0	.2 0	.4 Probability	0.6	0.8	1.0

0.2

0.0

0.0

0.2

0.2

Confusion Matrix (nomr.)

Predicted label

0.26

0.96

0.74

0.039

0-

True label

Metric	Train	Test
AUC	0.9991	0.9552
accuracy	0.9878	0.8994
balanced accuracy	0.9816	0.8512
recall_(sens)	0.9956	0.9612
specificity	0.9675	0.7413
precision	0.9875	0.9049
f1_score	0.9877	0.8966

--- train: 0.0154

0.8

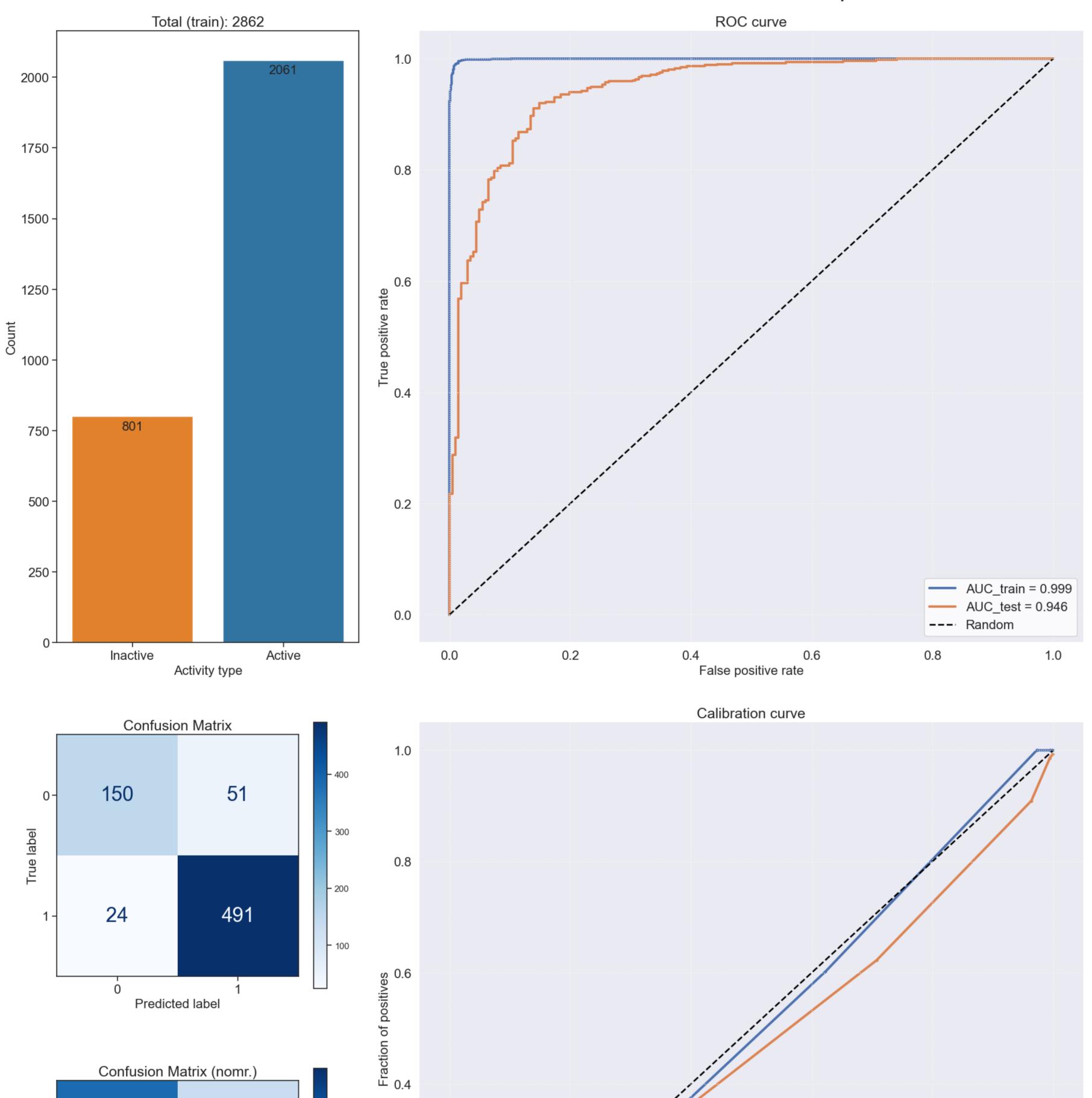
0.4 0.6 Mean predicted probability

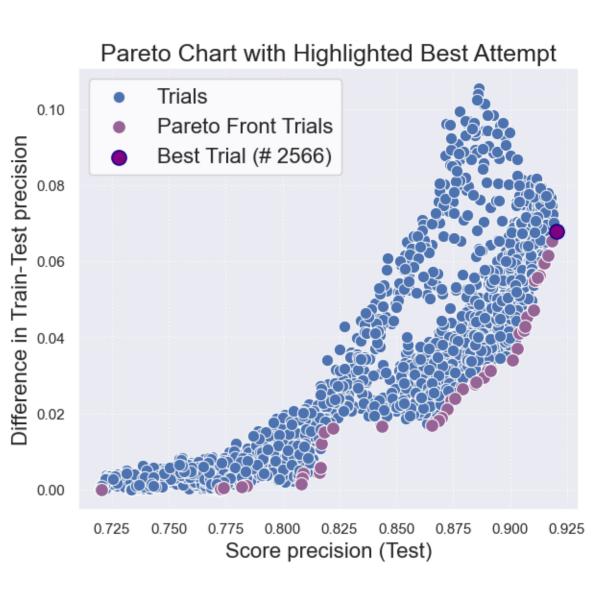
test: 0.073

--- perfectly\_calibrated

1.0

## Protein: P22303 - metric: precision





Value
3.3512
0.948
0.3442
2.5896
0.0741
10.0
2.0
559.0
0.4796
0.0089
1.1565
0.5113

	Active/Inactive probability					
50		Inactive zone			Active zone	type active (515) inactive (201) decoy (716)
± 00						
00						
10						
O	0.0	0.2	0.4 Probabi	0.6 ility	0.8	1.0

Confusion Matrix (nomr.)

Predicted label

0.25

0.95

- 0.6

0.4

0.2

0.2

0.0

0.0

0.2

0.75

0.047

0-

True label

Metric	Train	Test
AUC	0.9995	0.9456
accuracy	0.9913	0.8953
balanced accuracy	0.9855	0.8498
recall_(sens)	0.9985	0.9534
specificity	0.9725	0.7463
precision	0.9894	0.9059
f1_score	0.9912	0.8928

Metric	Train	Test
AUC	0.9995	0.9456
accuracy	0.9913	0.8953
balanced accuracy	0.9855	0.8498
recall_(sens)	0.9985	0.9534
specificity	0.9725	0.7463
precision	0.9894	0.9059
f1_score	0.9912	0.8928
	AUC accuracy balanced accuracy recall_(sens) specificity precision	AUC         0.9995           accuracy         0.9913           balanced accuracy         0.9855           recall_(sens)         0.9985           specificity         0.9725           precision         0.9894

1.0

- train: 0.0121

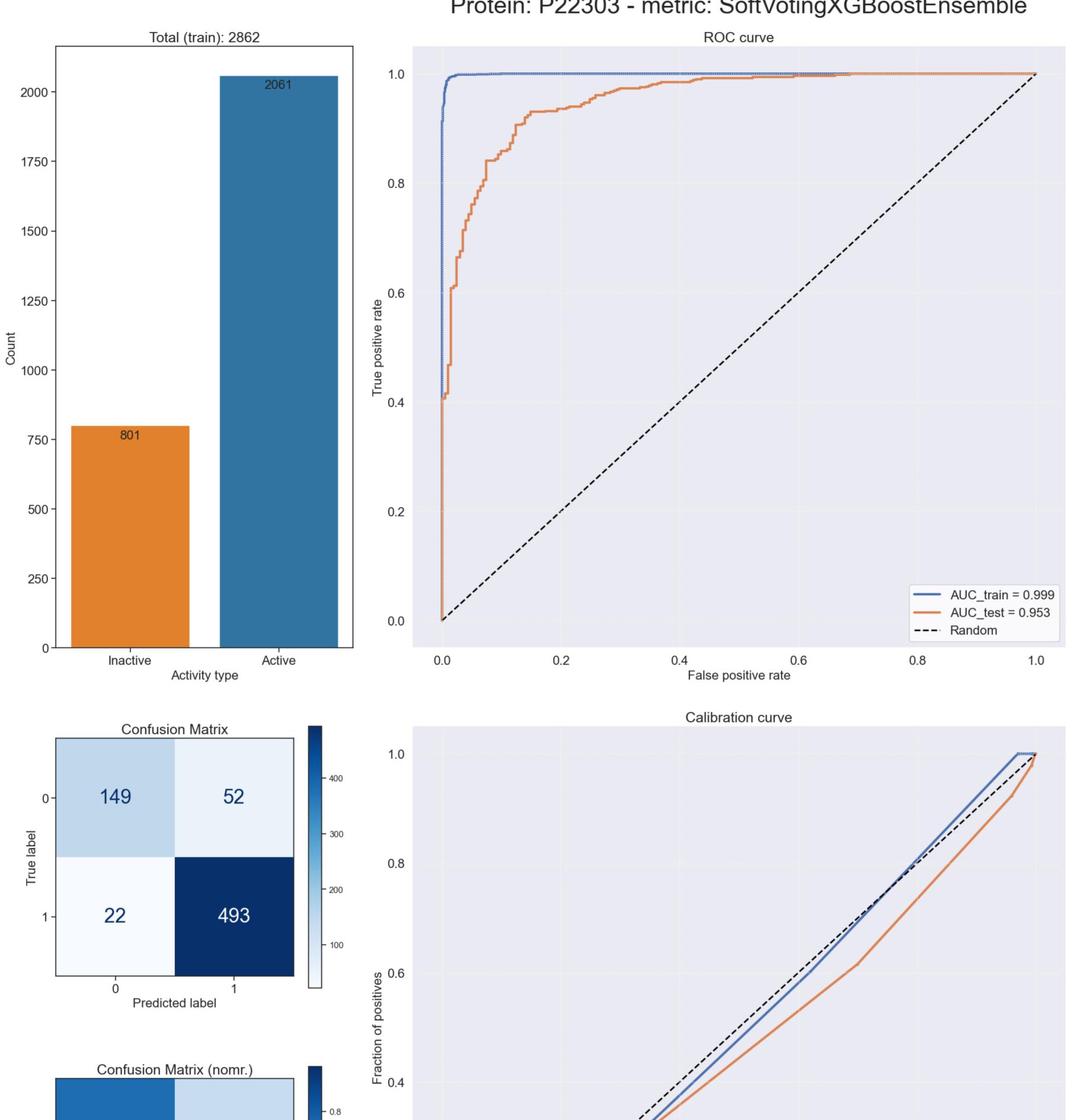
8.0

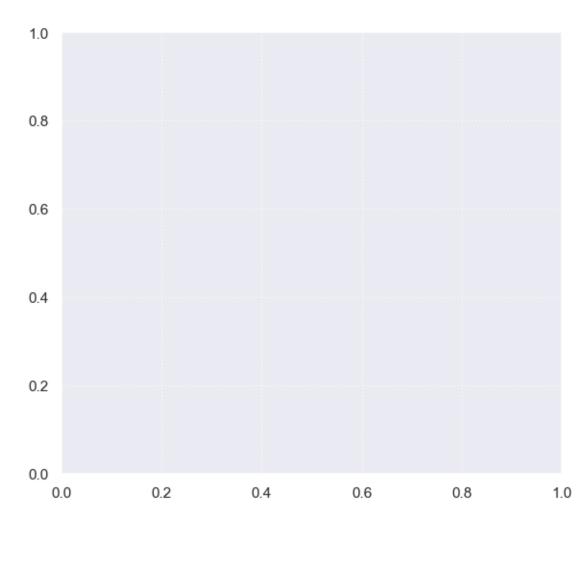
0.4 0.6 Mean predicted probability

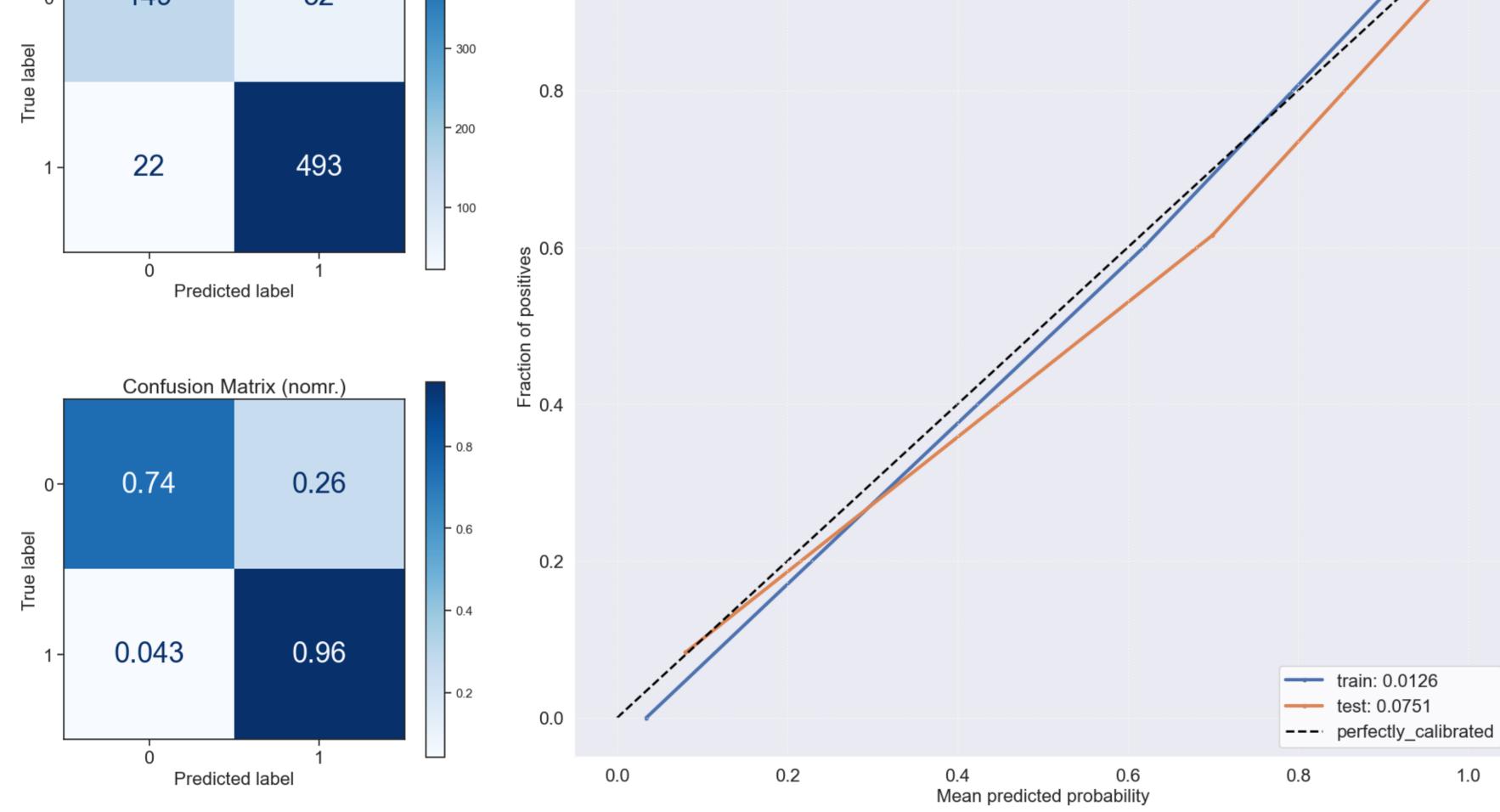
test: 0.0783

--- perfectly\_calibrated

## Protein: P22303 - metric: SoftVotingXGBoostEnsemble







<b>5</b> 0 -	Active/Inactive probability							
40		Inactive zone			Active zone	type active (515) inactive (201) decoy (716)		
ent 30								
Percent								
10								
0-	0.0	0.2	0.4 Proba	0.6	0.8	1.0		
			11000	ibility				

Metric	Train	Test
AUC	0.9995	0.9533
accuracy	0.9913	0.8966
balanced accuracy	0.9855	0.8493
recall_(sens)	0.9985	0.9573
specificity	0.9725	0.7413
precision	0.9894	0.9046
f1_score	0.9912	0.8939