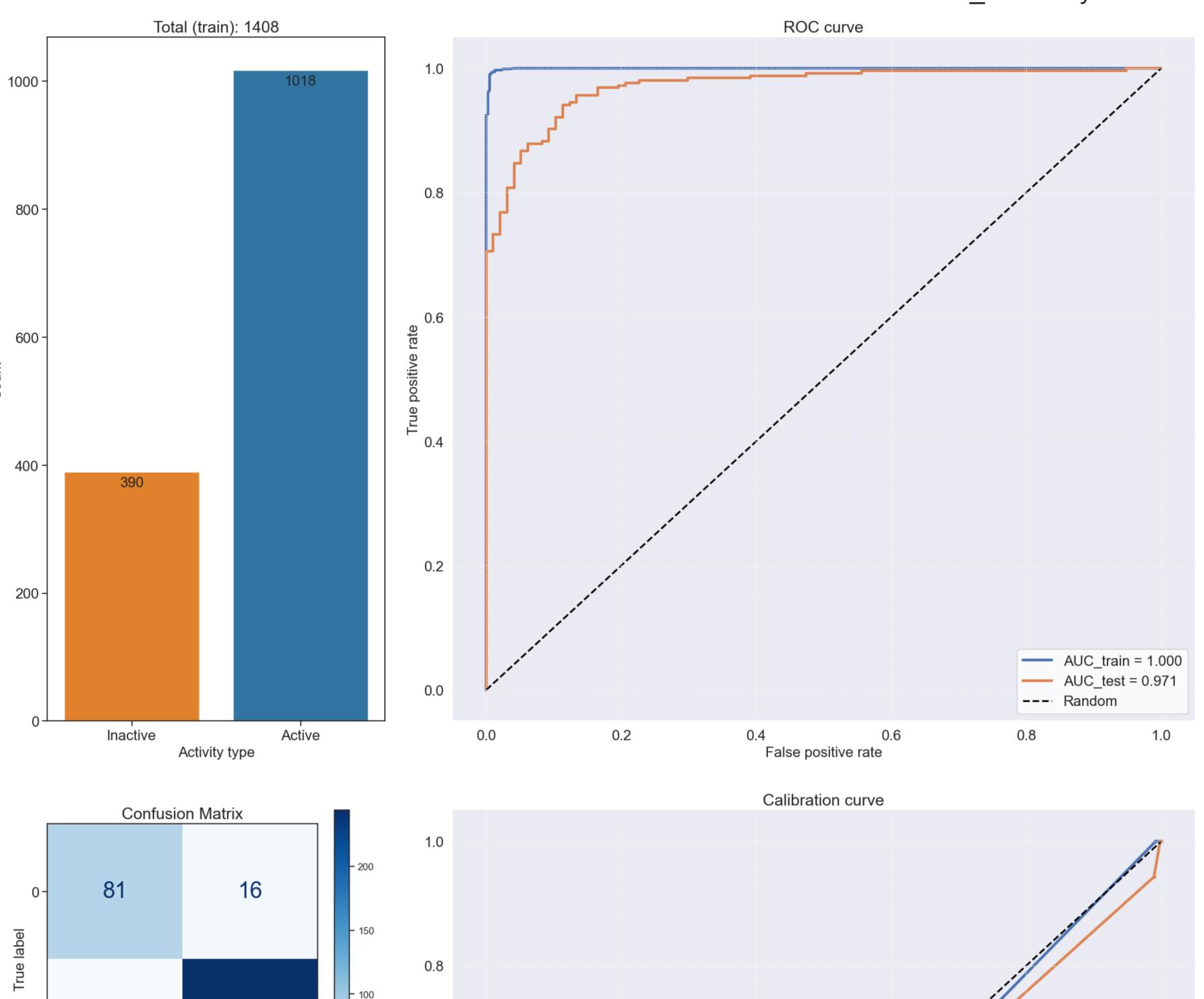
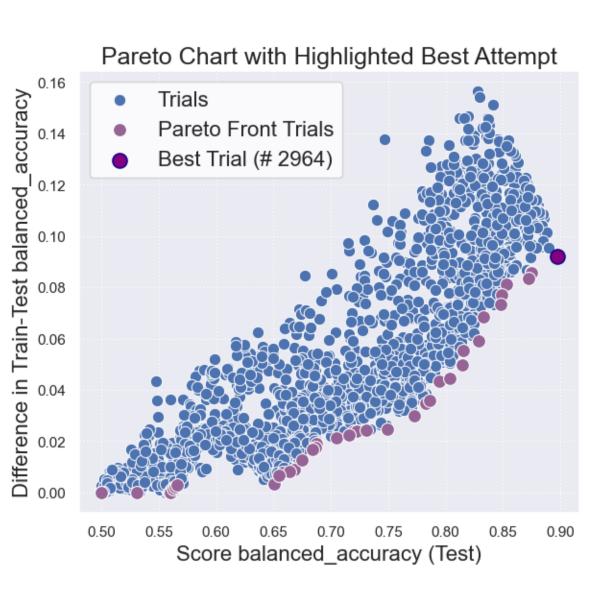
Protein: P20309 - metric: balanced_accuracy





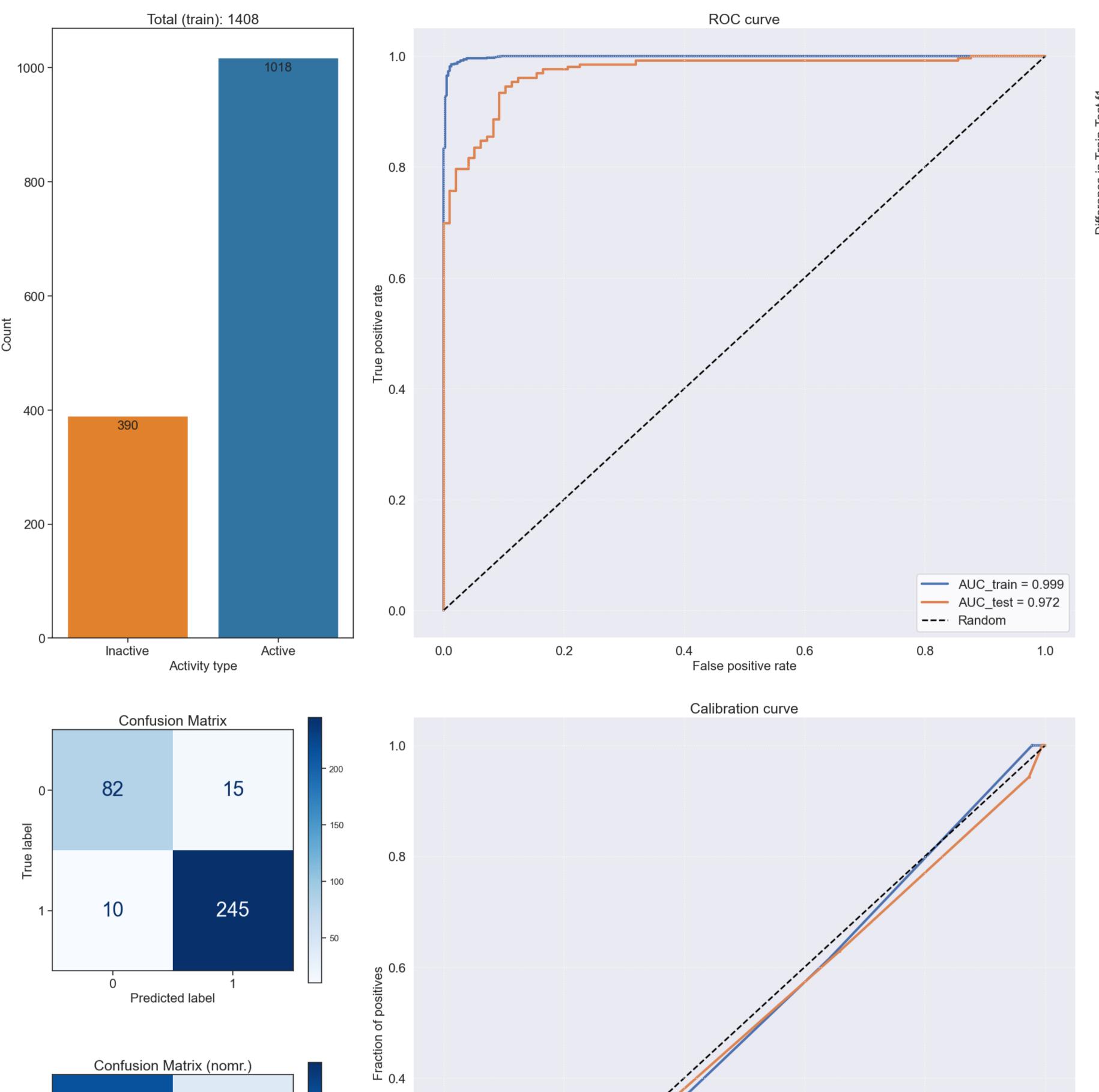
XGBoost Parameters	Value
alpha	3.9115
colsample_bytree	0.6372
gamma	0.248
lambda	3.6878
learning_rate	0.3996
max_depth	5.0
min_child_weight	1.0
n_estimators	324.0
reg_alpha	0.1848
reg_lambda	0.252
scale_pos_weight	1.6569
subsample	0.9733

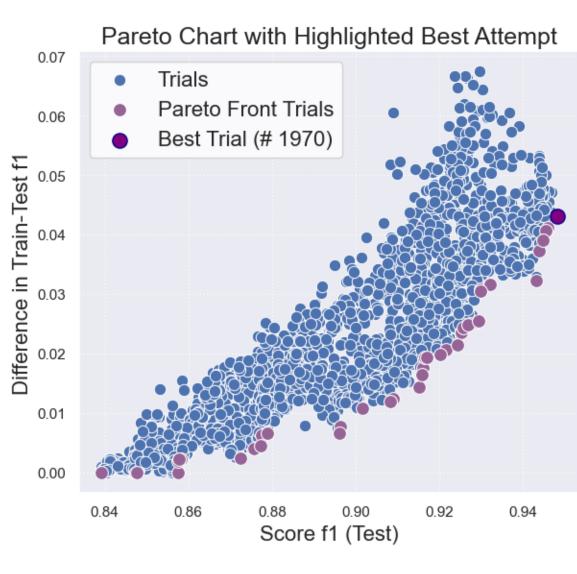
True label			- 150 O.	.8					
⊢ 1-	11	244	- 100 - 50						
	0 Predicte	1 ed label	on of positives						
	Confusion M	atrix (nomr.)	Fraction	.4					
abel - 0	0.84	0.16	- 0.8 - 0.6	2					
True label	0.043	0.96	- 0.4 - 0.2 0.					train: 0.0076 test: 0.059 perfectly_cal	
	0 Predicte	1 ed label		0.0	0.2	0.4 Mean predicted pro	0.6 bability	0.8	1.0

		Ac	tive/Inactive pro	obability		
70 60	lna	active zone			Active zone	active (255) inactive (97)
50						
Percent 30						
20						
10		2		0.6	0.8	10
	0.0		.4 Probability	0.0	0.0	1.0

Metric	Train	Test
AUC	0.9996	0.9711
accuracy	0.9922	0.9233
balanced accuracy	0.9867	0.896
recall_(sens)	0.999	0.9569
specificity	0.9744	0.8351
precision	0.9903	0.9385
f1_score	0.9922	0.9227

Protein: P20309 - metric: f1





XGBoost Parameters	Value
alpha	4.142
colsample_bytree	0.5632
gamma	0.2078
lambda	0.4349
learning_rate	0.05
max_depth	5.0
min_child_weight	1.0
n_estimators	578.0
reg_alpha	0.0692
reg_lambda	0.6113
scale_pos_weight	1.3826
subsample	0.9081

	Active/Inactive probability									
	Ina				Active 2	inacti deco	e (255) ve (97)			
ır ç										
20										
0	0.0 0.	2 0	.4 Probab	0.6 oility	6 0.	.8 1	.0			

0.2

0.0

0.0

0.2

0.2

Confusion Matrix (nomr.)

Predicted label

0.15

0.96

0.85

0.039

0-

True label

Metric	Train	Test
AUC	0.9987	0.9719
accuracy	0.9858	0.929
balanced accuracy	0.9791	0.9031
recall_(sens)	0.9941	0.9608
specificity	0.9641	0.8454
precision	0.9864	0.9423
f1_score	0.9857	0.9284

--- train: 0.0147

0.8

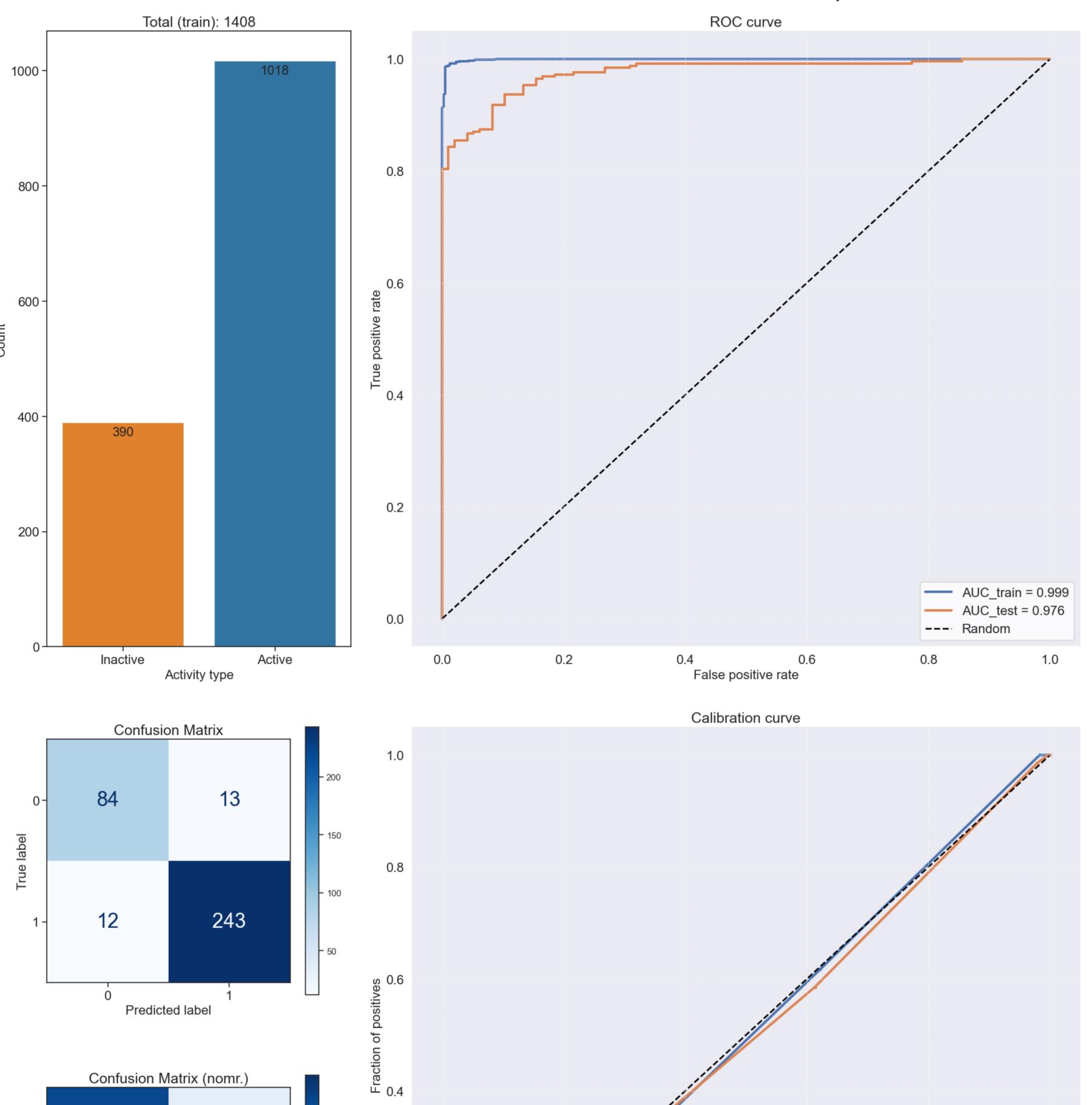
0.4 0.6 Mean predicted probability

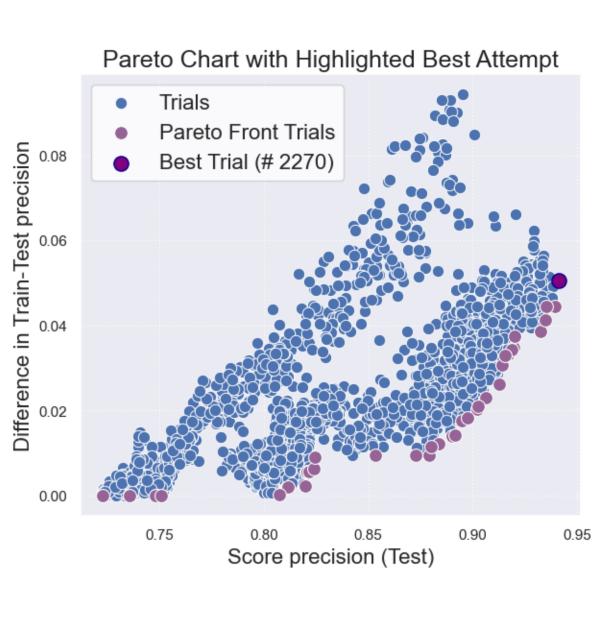
test: 0.0535

--- perfectly_calibrated

1.0

Protein: P20309 - metric: precision





XGBoost Parameters	Value
alpha	1.6289
colsample_bytree	0.6236
gamma	0.4449
lambda	1.2536
learning_rate	0.1007
max_depth	6.0
min_child_weight	1.0
n_estimators	356.0
reg_alpha	0.2072
reg_lambda	0.2327
scale_pos_weight	1.1623
subsample	0.7283

	Active/Inactive probability									
60 -····		Inactive zone			Active zone	type active (255) inactive (97) decoy (351)				
40										
Percent										
20										
0 —				0.6		10				
	0.0	0.2	0.4 Prob	0.6 ability	0.8	1.0				

Confusion Matrix (nomr.)

Predicted label

0.13

0.95

- 0.6

0.4

0.2

0.2

0.0

0.0

0.2

0.87

0.047

0-

True label

Metric	Train	Test
AUC	0.9993	0.9757
accuracy	0.9893	0.929
balanced accuracy	0.9847	0.9095
recall_(sens)	0.9951	0.9529
specificity	0.9744	0.866
precision	0.9902	0.9492
f1_score	0.9893	0.9289

Metric	Train	Test
AUC	0.9993	0.9757
 accuracy	0.9893	0.929
balanced accuracy	0.9847	0.9095
 recall_(sens)	0.9951	0.9529
 specificity	0.9744	0.866
 precision	0.9902	0.9492
f1_score	0.9893	0.9289

train: 0.011

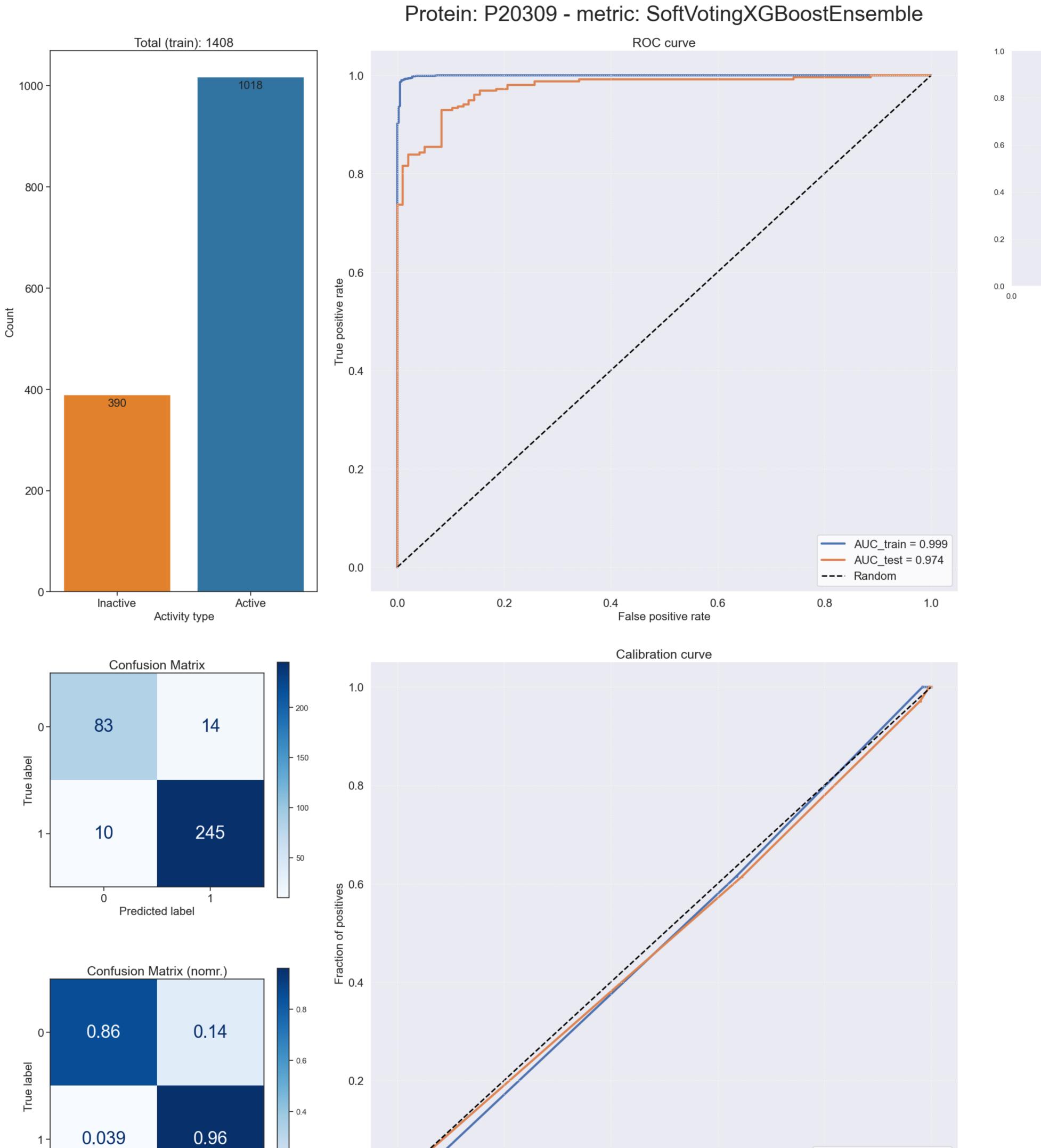
8.0

0.4 0.6 Mean predicted probability

test: 0.0556

perfectly_calibrated

1.0



0.2

0.4 0.6 Mean predicted probability

		А	ctive/Inactiv	e probability				
50	lna	ctive zone			Active z		type active inactiv decoy	(255) e (97) (351)
40								
Sercent								
- 00								
10								
0	0.0 0	.2	0.4 Proba	0.6 bility	0.	8	1.	0

0.2

Predicted label

0.0

0.0

Metric	Train	Test
AUC	0.9994	0.9744
accuracy	0.9886	0.9318
balanced accuracy	0.9834	0.9082
recall_(sens)	0.9951	0.9608
specificity	0.9718	0.8557
precision	0.9893	0.9459
f1_score	0.9886	0.9314

train: 0.0106

test: 0.0547

8.0

perfectly_calibrated

1.0

0.2

0.6