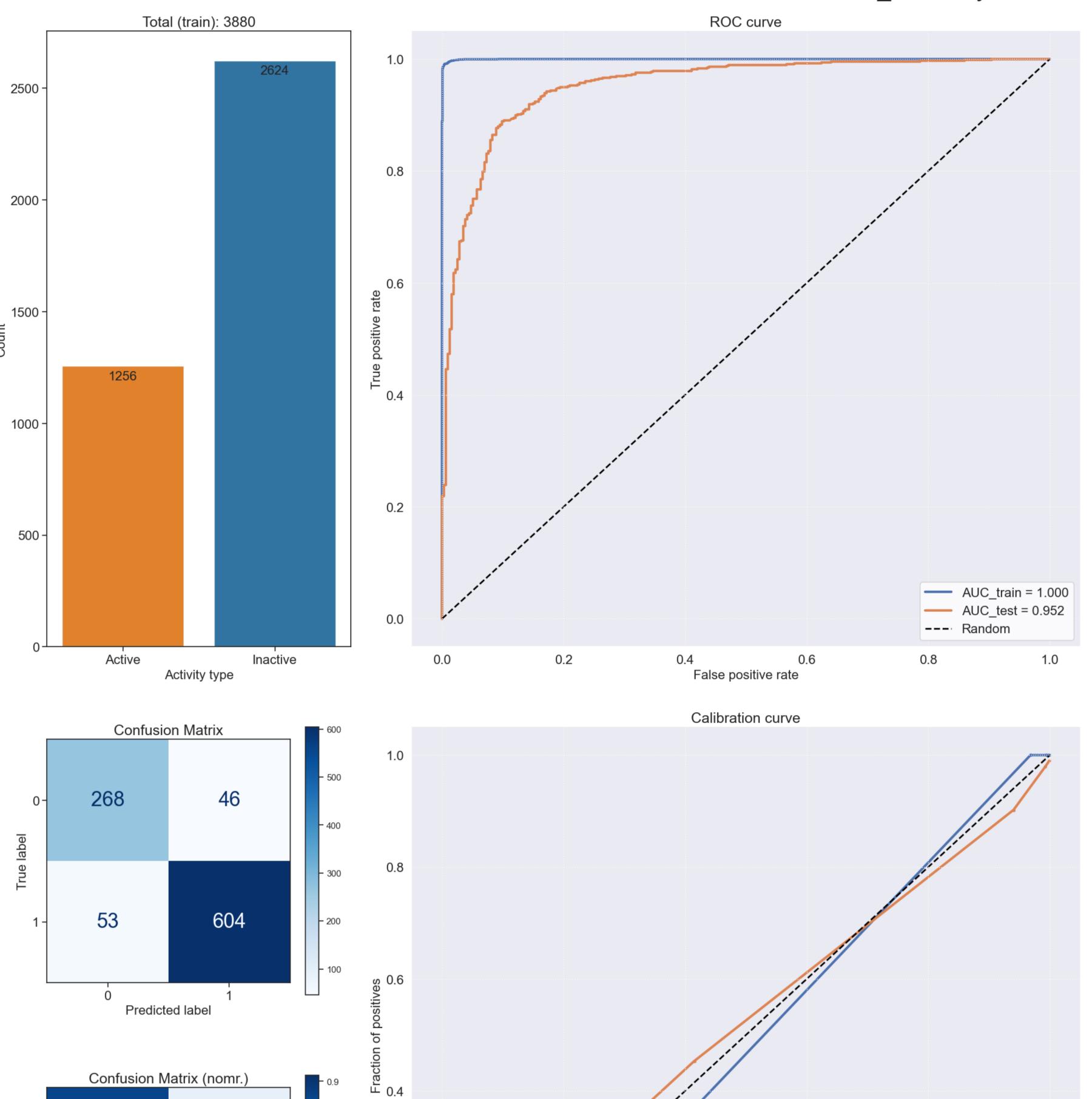
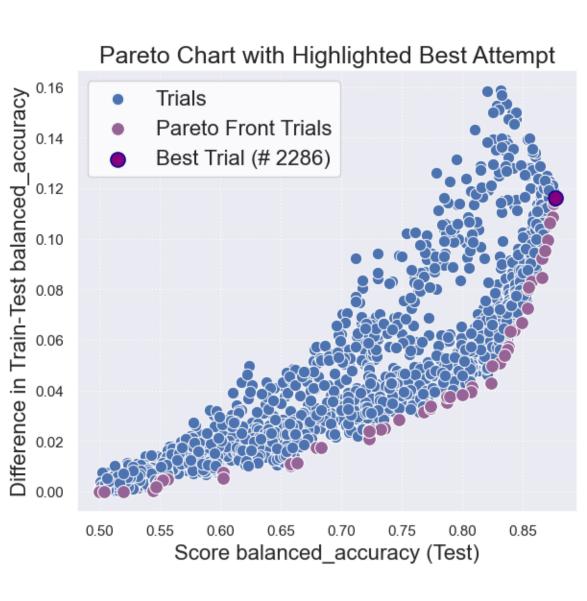
Protein: P14416 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	4.7966
colsample_bytree	0.5387
gamma	0.1068
lambda	2.1589
learning_rate	0.1118
max_depth	5.0
min_child_weight	1.0
n_estimators	917.0
reg_alpha	0.1172
reg_lambda	0.6887
scale_pos_weight	1.2314
subsample	0.7559

			Active/Inactive	e probability		
50		Inactive zone			Active zone	type active (657) inactive (314) decoy (968)
40						
Percent 20						
10						
0	0.0	0.2	0.4 Proba	0.6 bility	0.8	1.0

Confusion Matrix (nomr.)

Predicted label

0.15

0.92

- 0.6

- 0.5

0.4

0.3

0.2

0.2

0.0

0.0

0.2

0.85

0.081

0-

True label

Metric	Train	Test
AUC	0.9997	0.9521
accuracy	0.9928	0.898
balanced accuracy	0.9905	0.8864
recall_(sens)	0.997	0.9193
specificity	0.9841	0.8535
precision	0.9924	0.9292
f1_score	0.9928	0.8983

AUC	0.9997	0.9521
accuracy	0.9928	0.898
balanced accuracy	0.9905	0.8864
recall_(sens)	0.997	0.9193
specificity	0.9841	0.8535
precision	0.9924	0.9292
f1_score	0.9928	0.8983

--- train: 0.0101

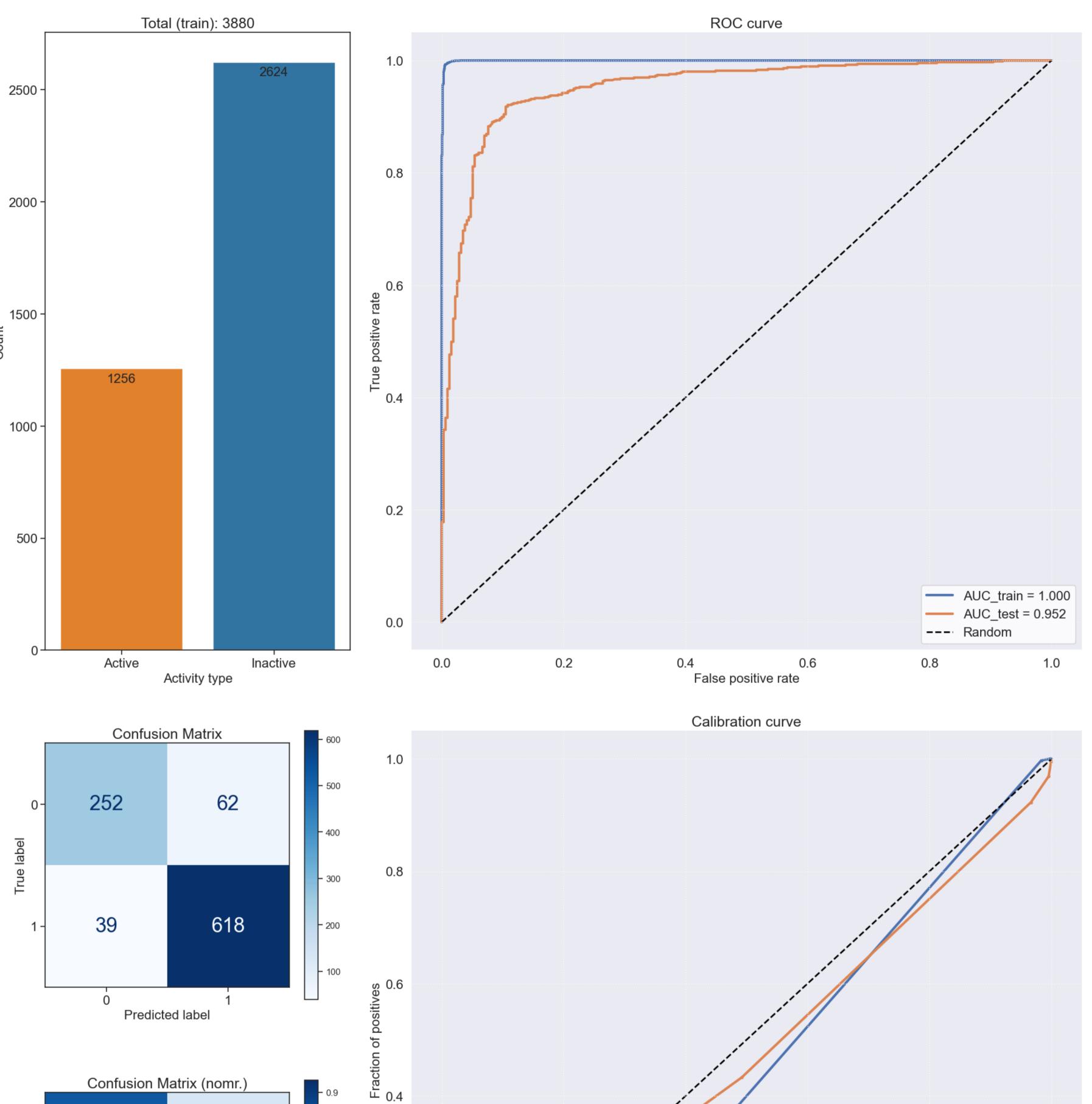
0.8

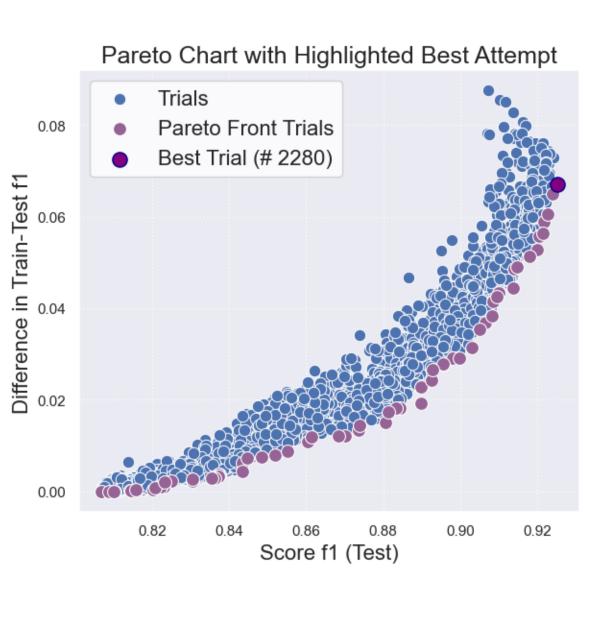
0.4 0.6 Mean predicted probability

test: 0.0764

--- perfectly_calibrated

Protein: P14416 - metric: f1





Value
4.1224
0.528
0.0785
1.471
0.0891
10.0
4.0
827.0
0.8145
0.294
2.9707
0.7229

		Act	tive/Inactive probab	ility		
60	Ina	ctive zone		Active		type active (657) inactive (314)
50						decoy (968)
10						
Percent						
20						
10						
0	0.0	.2 0.	4 0 Probability	0.6 0	0.8	1.0

Confusion Matrix (nomr.)

Predicted label

0.2

0.94

- 0.6

- 0.5

0.4

- 0.3

0.2

0.2

0.0

0.0

0.2

0.4 0.6 Mean predicted probability

8.0

0.059

0-

True label

Metric	Train	Test
AUC	0.9996	0.9516
accuracy	0.9881	0.896
balanced accuracy	0.9817	0.8716
recall_(sens)	1.0	0.9406
specificity	0.9634	0.8025
precision	0.9828	0.9088
f1_score	0.9881	0.8949

--- train: 0.0117

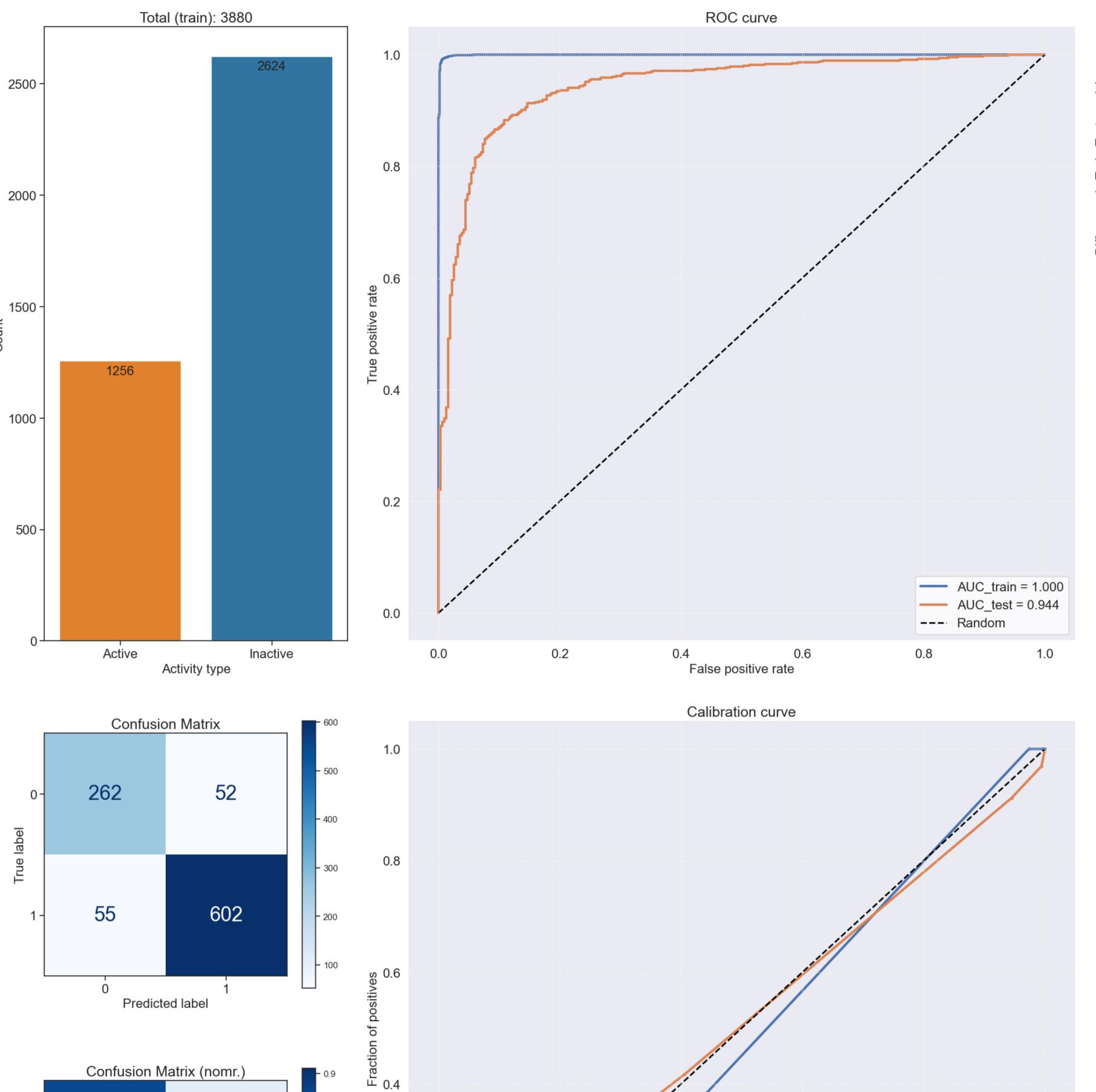
0.8

test: 0.0761

--- perfectly_calibrated

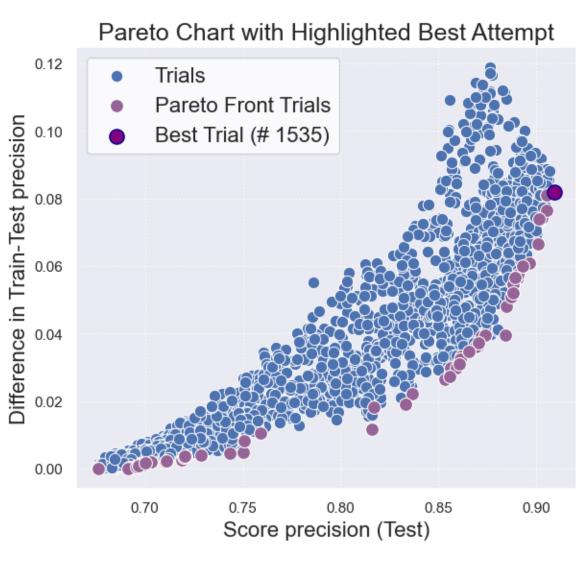
00	:	AC	ilive/iliactive pi	Obability	:	
60	Ina	active zone			Active zone	type active (657) inactive (314) decoy (968)
50						decoy (968)
40						
ent						
Percent 						
20						
10						
0 —	00	12	1	0.6	0.0	10
	0.0	J.Z C).4 Probabilit	0.6 /	0.8	1.0

Protein: P14416 - metric: precision



0.2

0.4 0.6 Mean predicted probability



XGBoost Parameters	Value
alpha	4.0187
colsample_bytree	0.884
gamma	0.2711
lambda	1.9282
learning_rate	0.2204
max_depth	6.0
min_child_weight	5.0
n_estimators	626.0
reg_alpha	0.0592
reg_lambda	0.3243
scale_pos_weight	1.5016
subsample	0.7435

	Active/Inactive probability									
50	lı	nactive zone			Active zone	type active (657) inactive (314) decoy (968)				
40										
Percent										
20										
10										
0—	0.0	0.2	0.4 Proba	0.6 ability	0.8	1.0				

- 0.8

- 0.6

- 0.5

0.4

0.3

0.2

0.2

0.0

0.0

Confusion Matrix (nomr.)

Predicted label

0.17

0.92

0.83

0.084

0-

True label

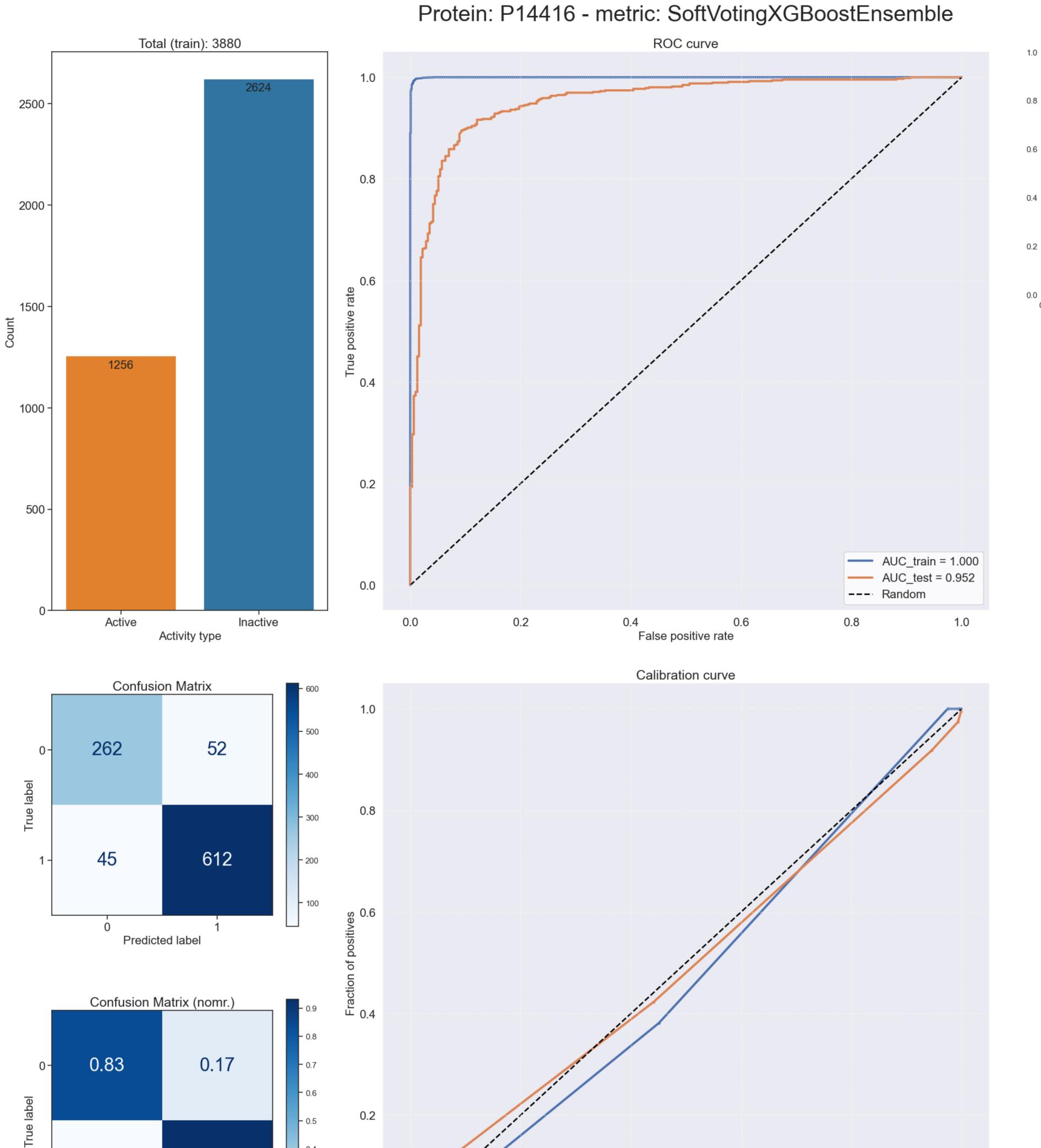
Metric	Train	Test
AUC	0.9997	0.9442
accuracy	0.9912	0.8898
balanced accuracy	0.9873	0.8753
recall_(sens)	0.9985	0.9163
specificity	0.9761	0.8344
precision	0.9887	0.9205
f1_score	0.9912	0.8899

--- train: 0.0103

0.8

test: 0.0835

--- perfectly_calibrated



- 0.5

- 0.4

0.3

0.2

0.1

0.0

0.0

0.93

0.068

Predicted label

	Active/	Inactive probability					
50	Inactive zone	Act	ive zone	type active (657)		Metric	Train
				inactive (314) decoy (968)		AUC	0.999
40						accuracy	0.991
30						balanced accuracy	0.987
Percent						recall_(sens)	0.998
20 -						specificity	0.976
10						precision	0.988
0					_	f1_score	0.991
0——(0.0 0.2 0.4	0.6 Probability	0.8	1.0			

0.4 0.6 Mean predicted probability

0.2

Metric	Train	Test
AUC	0.9998	0.9521
accuracy	0.9915	0.9001
balanced accuracy	0.9875	0.883
recall_(sens)	0.9989	0.9315
specificity	0.9761	0.8344
precision	0.9887	0.9217
f1_score	0.9915	0.8998

train: 0.0101

test: 0.0756

8.0

perfectly_calibrated

1.0

0.0

0.2