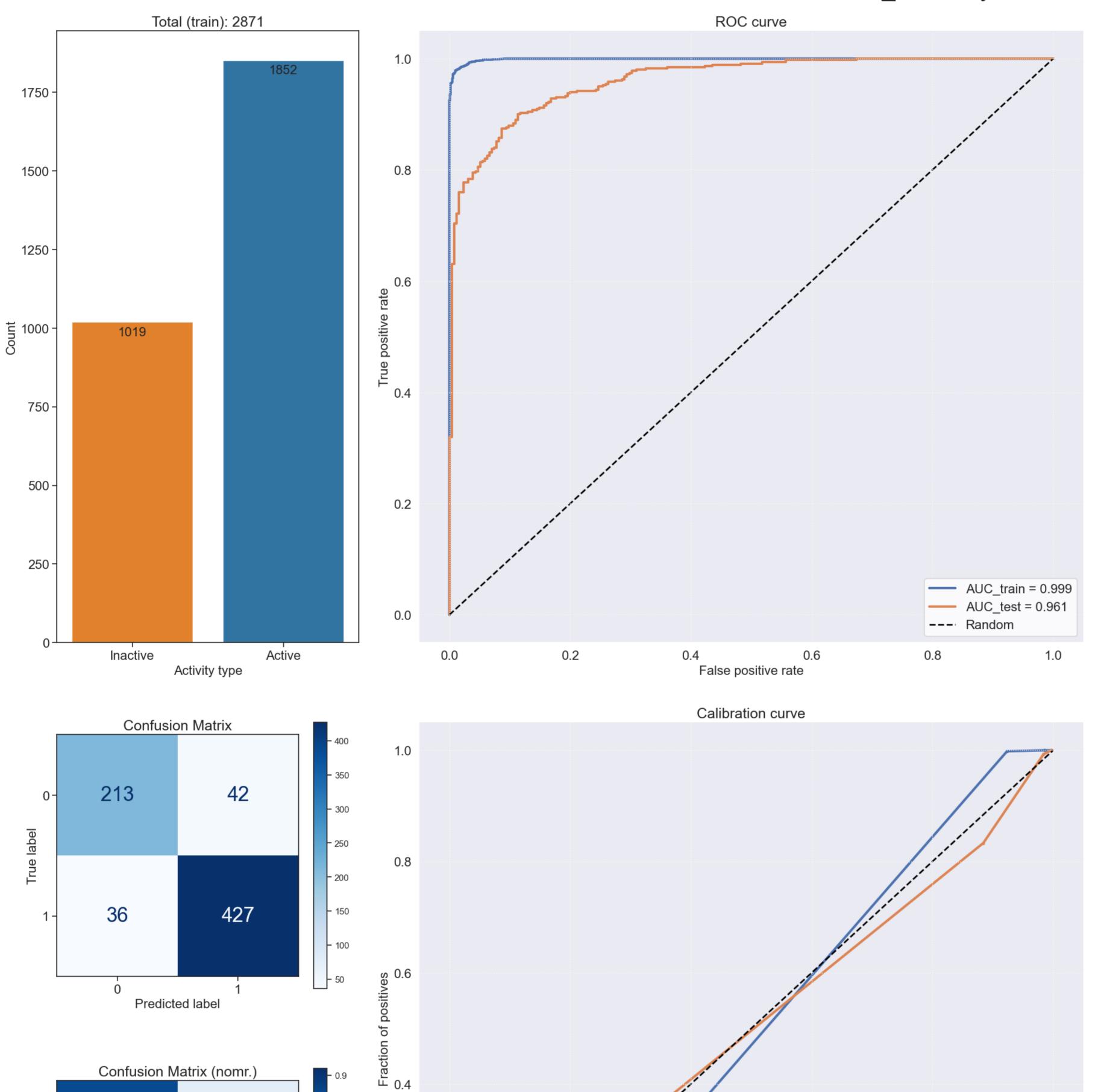
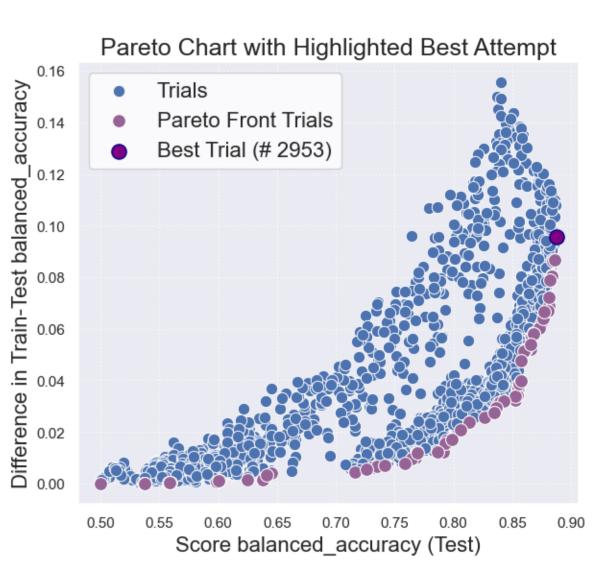
Protein: P29274 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	4.6264
colsample_bytree	0.7128
gamma	0.104
lambda	1.6636
learning_rate	0.1247
max_depth	9.0
min_child_weight	2.0
n_estimators	215.0
reg_alpha	0.1564
reg_lambda	0.9924
scale_pos_weight	1.0323
subsample	0.807

	Active/Inactive probability								
35	ln	active zone			Active zo	ne type active inactive decoy	e (255)		
30									
20									
Dercent 20									
10									
5									
0	0.0	0.2	0.4	0.6	0.8		.0		
			Probab	oility					

- 0.8

0.6

- 0.5

0.4

0.3

0.2

0.2

0.0

0.0

0.2

Confusion Matrix (nomr.)

Predicted label

0.16

0.92

0.84

0.078

0-

True label

Metric	Train	Test
AUC	0.9991	0.9611
accuracy	0.9829	0.8914
balanced accuracy	0.9795	0.8788
recall_(sens)	0.9914	0.9222
specificity	0.9676	0.8353
precision	0.9823	0.9104
f1_score	0.9829	0.8911

train: 0.0178

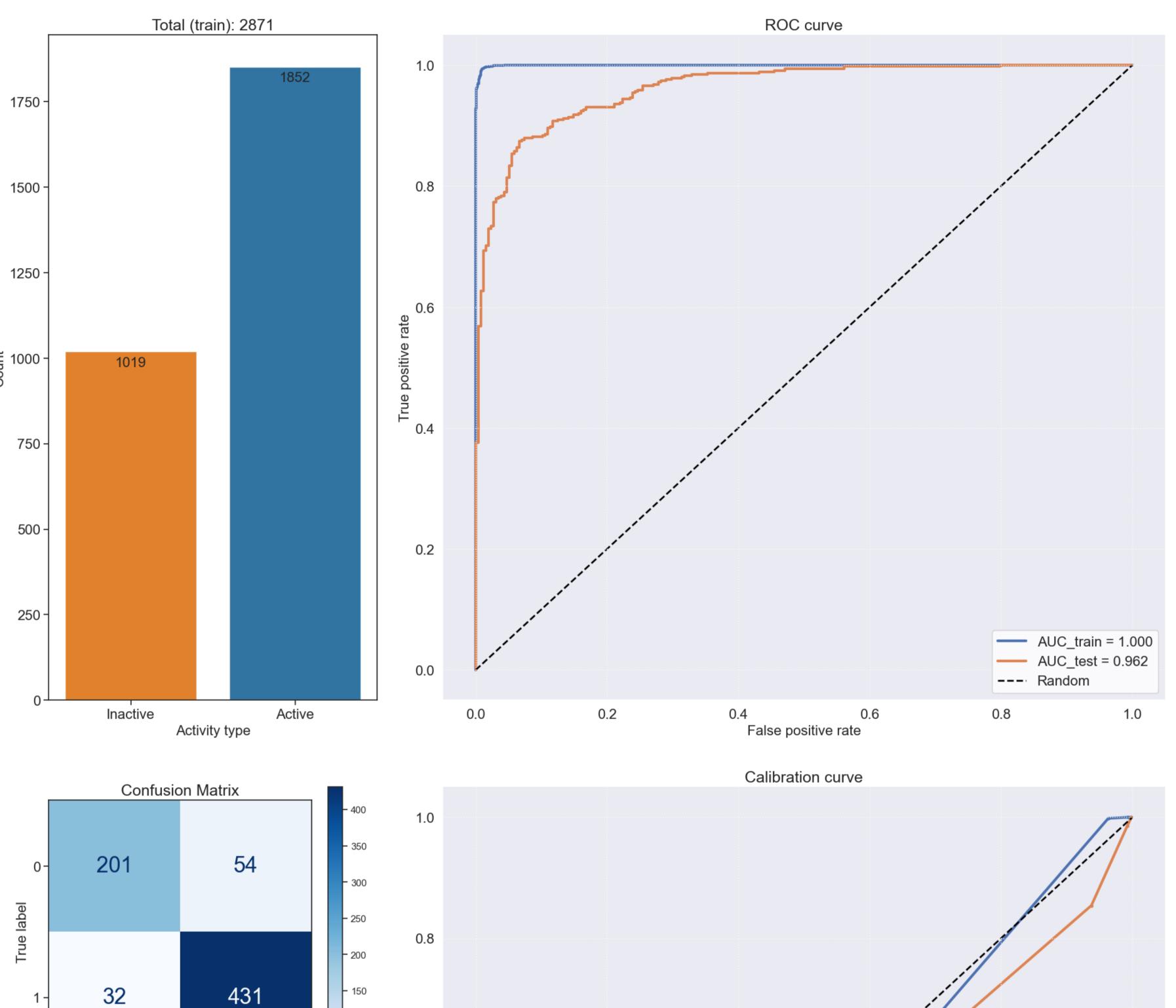
8.0

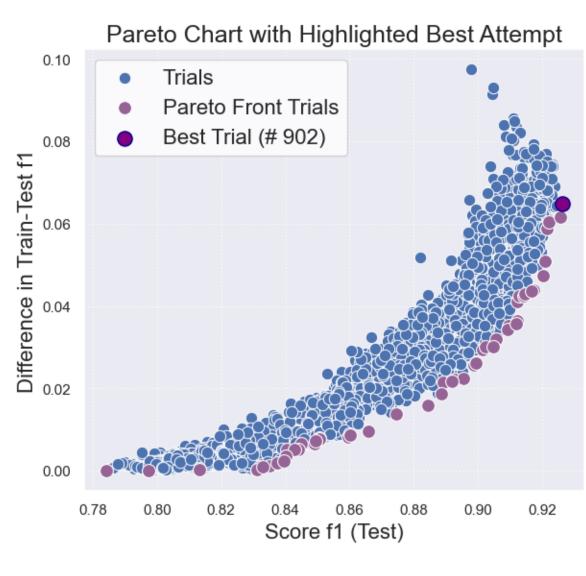
0.4 0.6 Mean predicted probability

test: 0.0772

--- perfectly_calibrated

Protein: P29274 - metric: f1





XGBoost Parameters	Value
alpha	4.5954
colsample_bytree	0.6981
gamma	0.0337
lambda	0.5454
learning_rate	0.1151
max_depth	9.0
min_child_weight	1.0
n_estimators	221.0
reg_alpha	0.8969
reg_lambda	0.1285
scale_pos_weight	2.1214
subsample	0.6179

l	0 Dradiated k	1	- 0.1	0.0	0.2	0.4	0.6
	Predicted la	apei		0.0	0.2	Mean predicted	
			Active/Inactiv	e probability			
50 -		Inactive zone			Active zone	type active (463) inactive (255) decoy (717)	
40 -						decoy (111)	
Percent 08							
20 -							
10 -							
0-	0.0	0.2	0.4	0.6	0.8	1.0	

- 100

- 0.8

- 0.6

- 0.5

0.4

0.3

0.2

Predicted label

Confusion Matrix (nomr.)

0.21

0.93

0.79

0.069

0-

True label

Fraction of positives
0.0
7.0

0.2

0.0

Metric	Train	Test
AUC	0.9997	0.9622
accuracy	0.9861	0.8802
balanced accuracy	0.9806	0.8596
recall_(sens)	0.9995	0.9309
specificity	0.9617	0.7882
precision	0.9794	0.8887
f1_score	0.986	0.8789

--- train: 0.0144

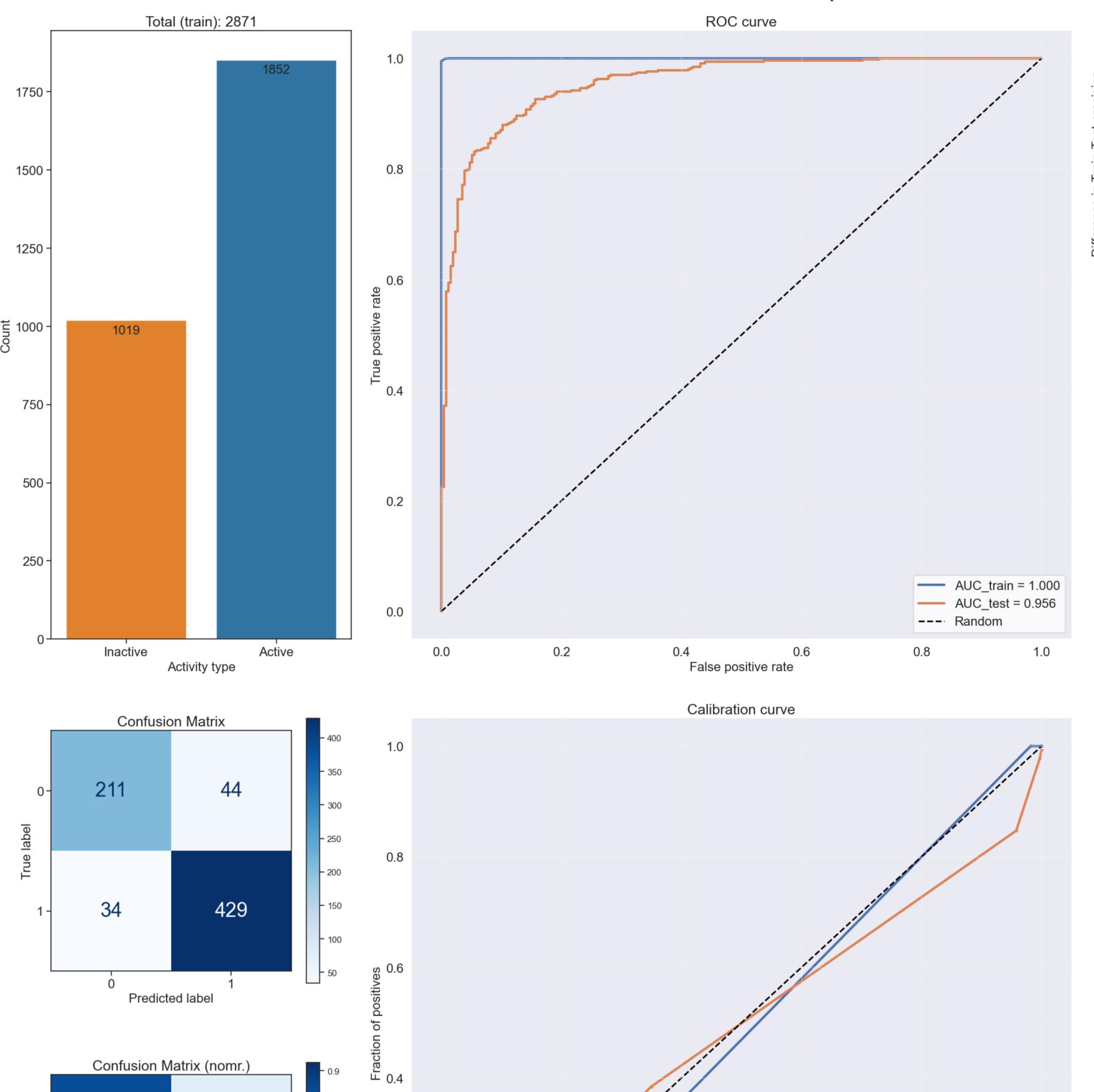
0.8

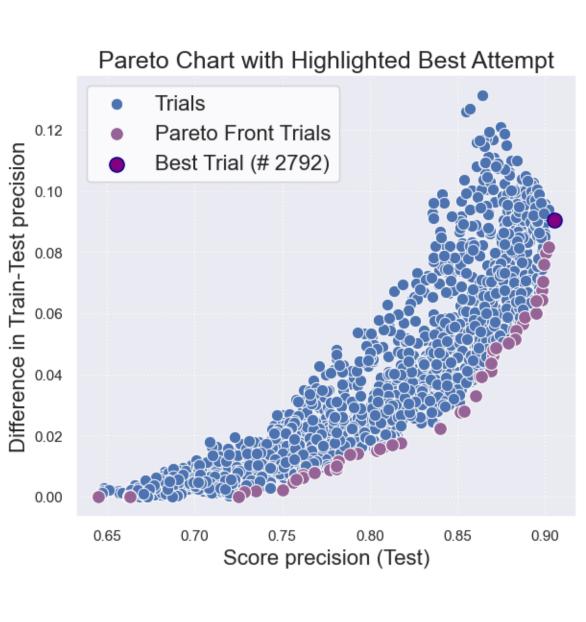
test: 0.0784

--- perfectly_calibrated

	Cuve/macuve pro	Dability						
Inactive zone		Active	zone ac	/pe tive (463)		Metric	Train	Test
			ina ina	coy (717)	*	AUC	0.9997	0.962
						accuracy	0.9861	0.880
						balanced accuracy	0.9806	0.859
						recall_(sens)	0.9995	0.930
						specificity	0.9617	0.788
						precision	0.9794	0.888
						f1_score	0.986	0.878
•	0.4 Probability		0.8	1.0				
	•							

Protein: P29274 - metric: precision





XGBoost Parameters	Value
alpha	3.3529
colsample_bytree	0.9582
gamma	0.1483
lambda	4.1225
learning_rate	0.2343
max_depth	4.0
min_child_weight	1.0
n_estimators	920.0
reg_alpha	0.7763
reg_lambda	0.0444
scale_pos_weight	1.7981
subsample	0.7581

Active/Inactive probability										
						decoy (717)				
40										
Percent										
20										
0	0.0	0.2				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

0.17

0.93

0.83

0.073

0-

True label

Metric	Train	Test
AUC	1.0	0.9562
accuracy	0.9965	0.8914
balanced accuracy	0.9953	0.877
recall_(sens)	0.9995	0.9266
specificity	0.9912	0.8275
precision	0.9952	0.907
f1_score	0.9965	0.8909

--- train: 0.005

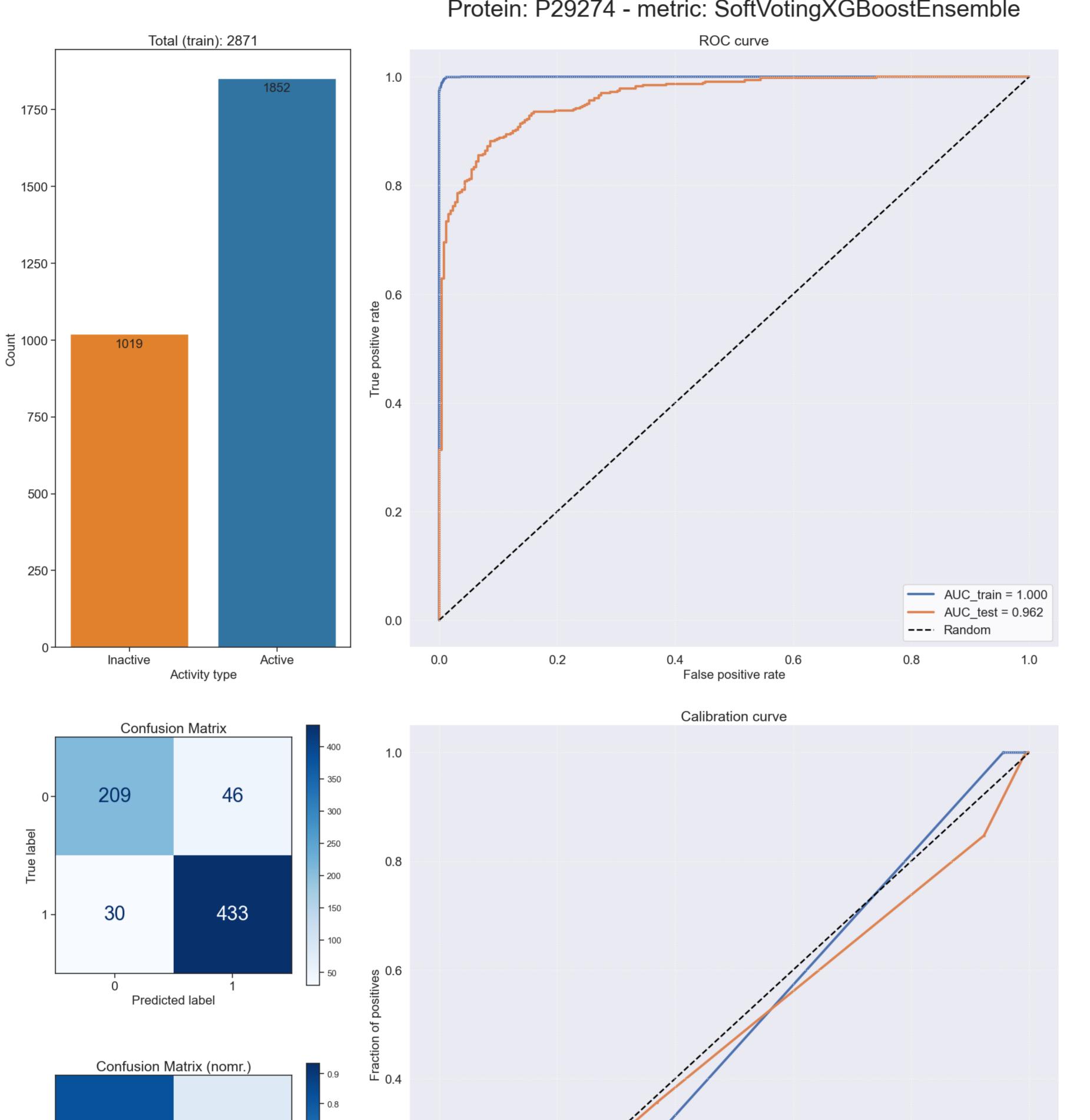
0.8

0.4 0.6 Mean predicted probability

test: 0.0845

--- perfectly_calibrated

Protein: P29274 - metric: SoftVotingXGBoostEnsemble



0.2

	perfectly_calibrated			
0.8	1.0			
	Metr	ic	Train	Test
		AUC	0.9999	0.9622
	ac	ccuracy	0.992	0.8942
	balanced ad	ccuracy	0.9889	0.8774
	recall	_(sens)	0.9995	0.9352
	sp	ecificity	0.9784	0.8196
	pr	ecision	0.9883	0.904
	f1	score	0.992	0.8933

train: 0.011

test: 0.0776

0.4 0.6 Mean predicted probability

0.8

0.6

0.4

0.2

0.0 0.0

0.2

0.6

	Active/Inactive probability								
50		Inactive zone			Active	:	type active (463) inactive (255) decoy (717)		
40									
Percen									
10									
0	0.0	0.2	0.4	0.	6 ().8	1.0		
			F	Probability					

0.82

0.065

Predicted label

0-

True label

0.18

0.94

- 0.6

- 0.5

- 0.4

- 0.3

0.2

- 0.1

0.2

0.0