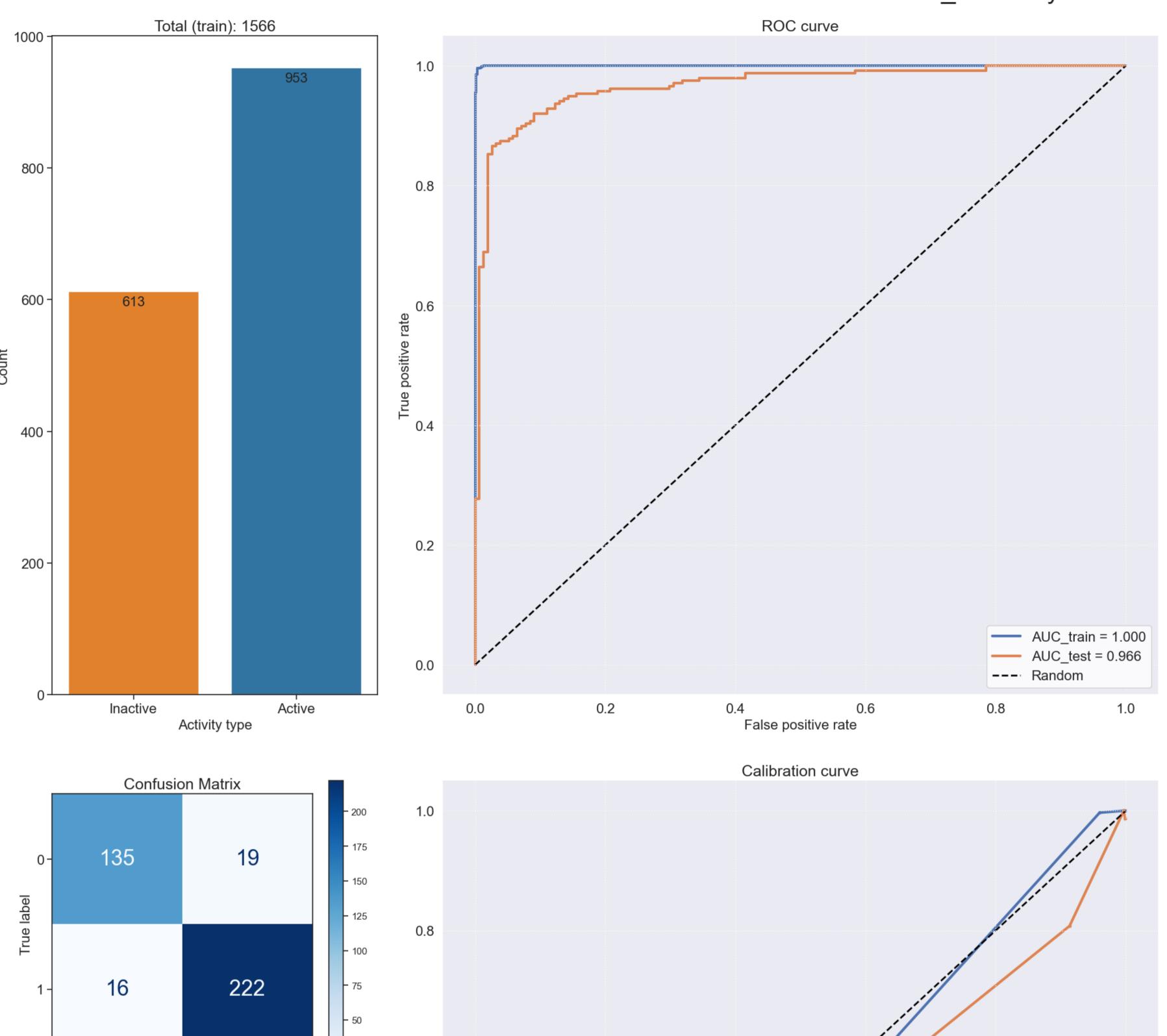
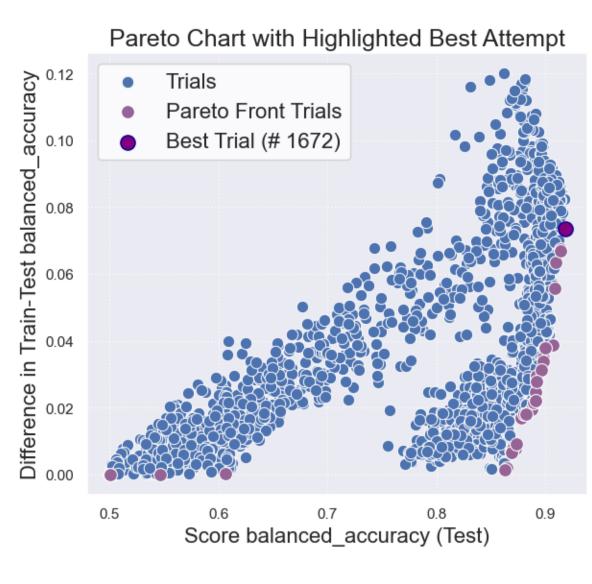
Protein: P00519 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	2.7297
colsample_bytree	0.6475
gamma	0.255
lambda	2.8088
learning_rate	0.2463
max_depth	8.0
min_child_weight	1.0
n_estimators	171.0
reg_alpha	0.7949
reg_lambda	0.3378
scale_pos_weight	2.7085
subsample	0.9924

Predicted label		0.0	0.2	0.4 Mean predic	0.6 eted probability	
		Active/Inactive	e probability			
70	Inactive a	zone		Active zone	type active (238)	
60 -					inactive (154)	
50						
Percent 6 04						
30						
20	L					
10						
0 —	0.0 0.2	0.4 Probal	0.6 bility	0.8	1.0	_

Fraction of positives

0.2

0.0

- 0.8

- 0.5

0.4

0.3

0.2

0.1

Predicted label

Confusion Matrix (nomr.)

0.12

0.93

0.88

0.067

0-

Metric	Train	Test
AUC	0.9999	0.9661
accuracy	0.9943	0.9107
balanced accuracy	0.9927	0.9047
recall_(sens)	1.0	0.9328
specificity	0.9853	0.8766
precision	0.9906	0.9212
f1_score	0.9942	0.9106

--- train: 0.0089

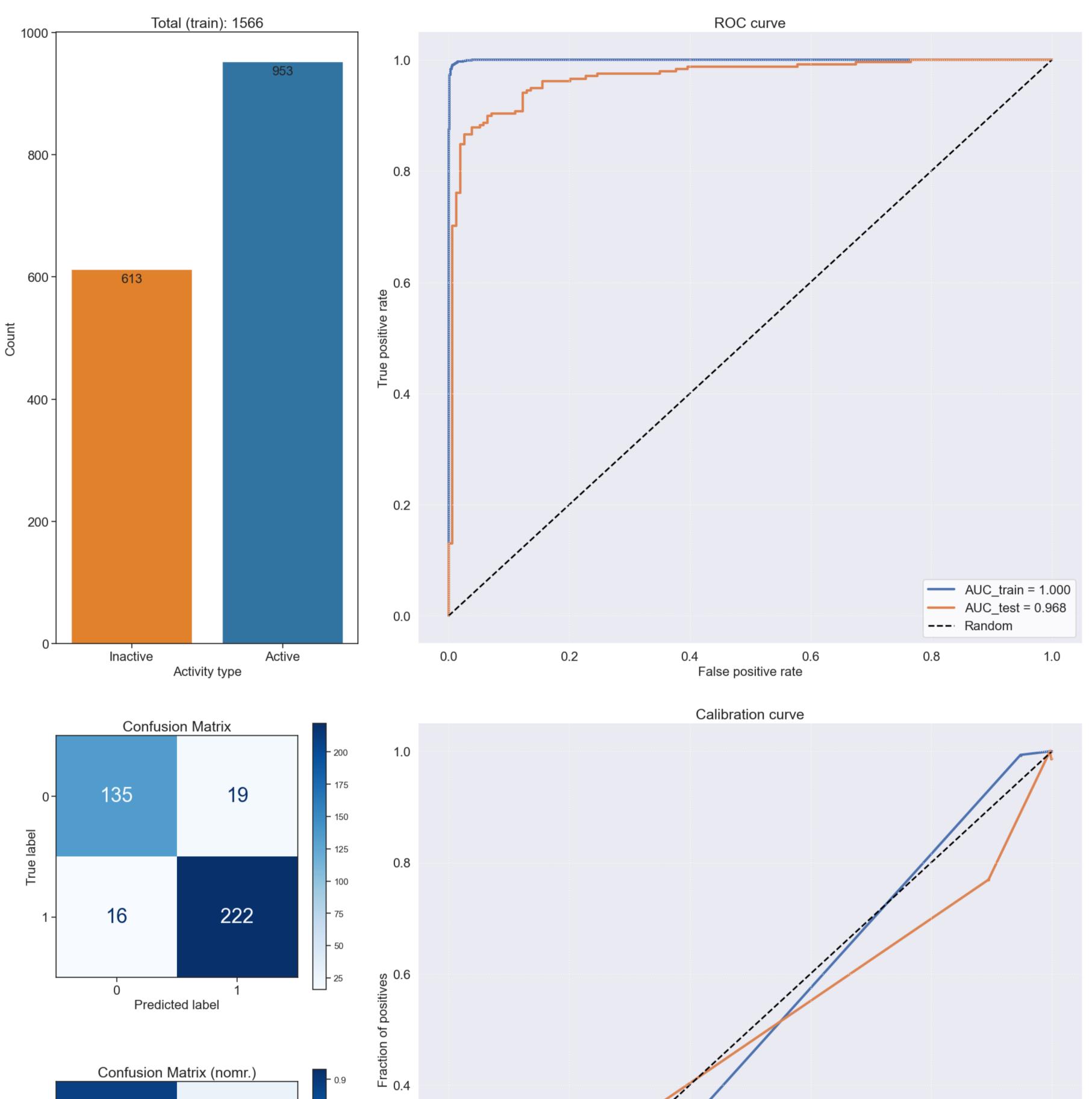
8.0

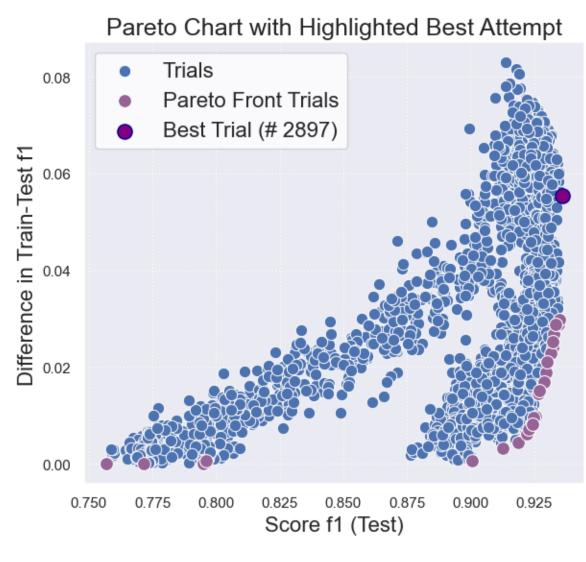
test: 0.0678

--- perfectly_calibrated

1.0

Protein: P00519 - metric: f1





XGBoost Parameters	Value
alpha	2.4855
colsample_bytree	0.6864
gamma	0.7694
lambda	1.4545
learning_rate	0.0198
max_depth	9.0
min_child_weight	2.0
n_estimators	982.0
reg_alpha	0.0708
reg_lambda	0.7613
scale_pos_weight	2.1136
subsample	0.8072

		Active/Inactive pr	obability		
70	Inactive zone		Acti	ve zone	active (238) inactive (154)
					decoy (391)
30					
Percent					
20					
10					
0—	0.0 0.2	0.4 Probability	0.6	0.8	1.0

- 0.8

- 0.6

- 0.5

0.4

0.3

0.2

0.1

0.2

0.0

0.0

0.2

Predicted label

Confusion Matrix (nomr.)

Predicted label

0.12

0.93

0.88

0.067

0-

True label

Metric	Train	Test
AUC	0.9996	0.9676
accuracy	0.9891	0.9107
balanced accuracy	0.987	0.9047
recall_(sens)	0.9969	0.9328
specificity	0.9772	0.8766
precision	0.9855	0.9212
f1_score	0.9891	0.9106

--- train: 0.0108

0.8

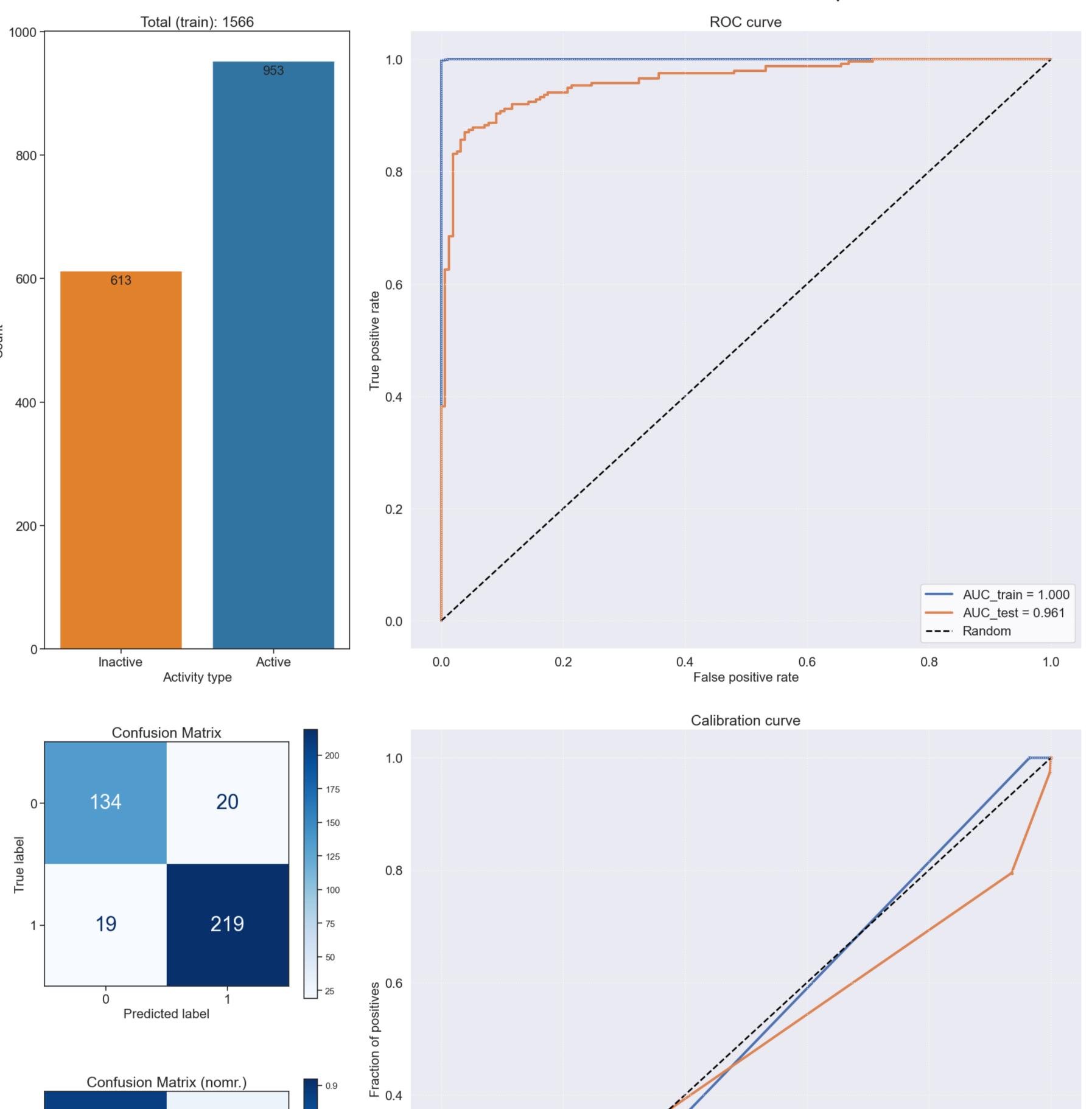
0.4 0.6 Mean predicted probability

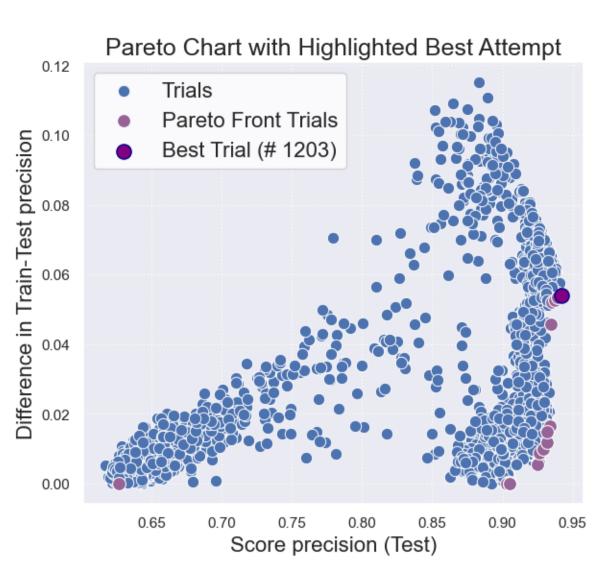
test: 0.0651

--- perfectly_calibrated

1.0

Protein: P00519 - metric: precision





XGBoost Parameters	Value
alpha	0.7322
colsample_bytree	0.9955
gamma	0.2702
lambda	2.0309
learning_rate	0.3139
max_depth	6.0
min_child_weight	2.0
n_estimators	551.0
reg_alpha	0.8069
reg_lambda	0.565
scale_pos_weight	1.7565
subsample	0.7405

	Ac	tive/Inactive probab	lity				
				type active (238) inactive (154) decoy (391)			
50							
Percent 08 04							
20							
0.	0 0.2 0	.4 0 Probability	.6 0.	.8 1.0			
	Trobability						

Confusion Matrix (nomr.)

Predicted label

0.13

0.92

- 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

0.87

0.08

0-

True label

Metric	Train	Test
AUC	1.0	0.9613
accuracy	0.9968	0.9005
balanced accuracy	0.9962	0.8951
recall_(sens)	0.999	0.9202
specificity	0.9935	0.8701
precision	0.9958	0.9163
f1_score	0.9968	0.9005

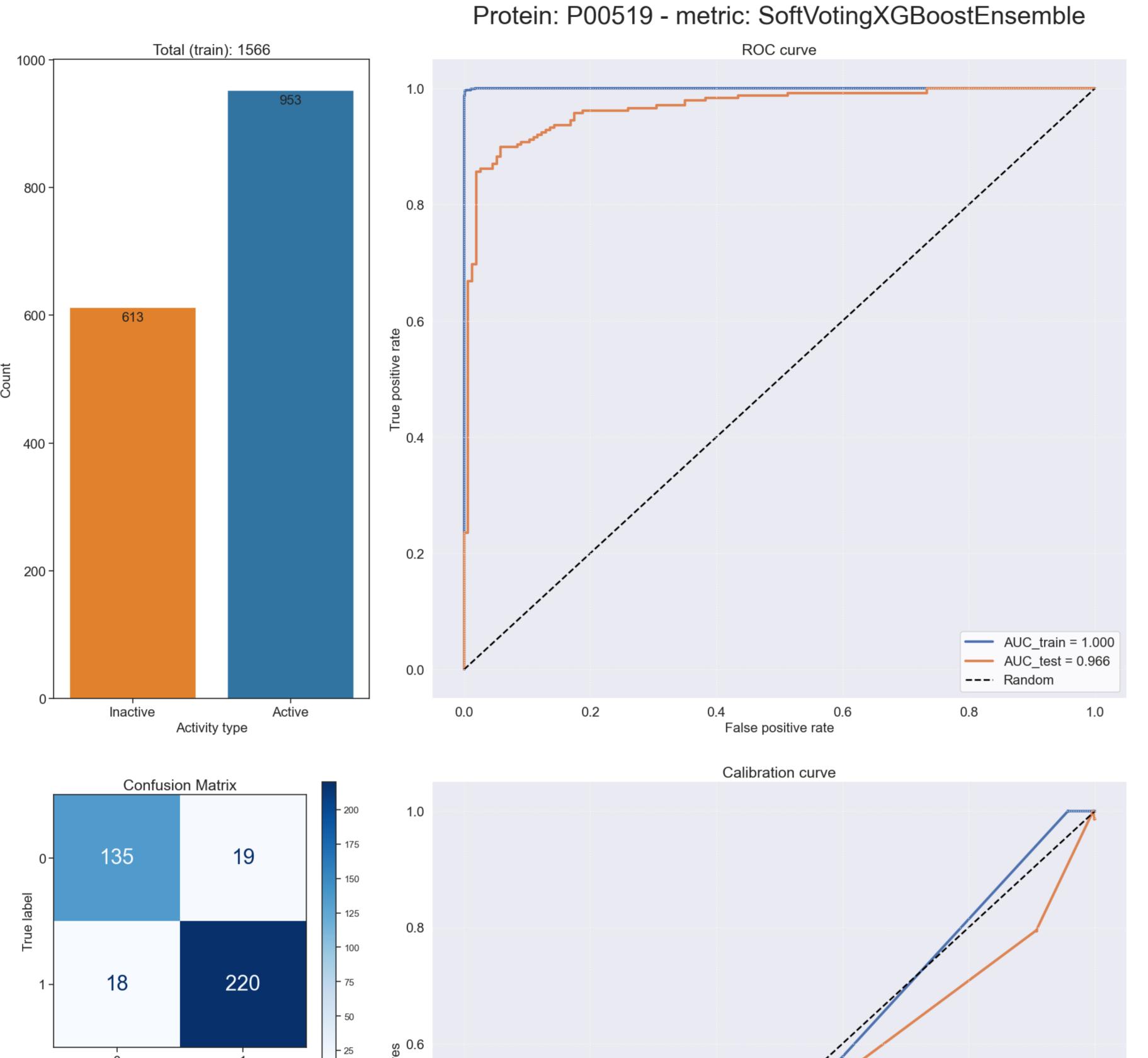
- train: 0.0047

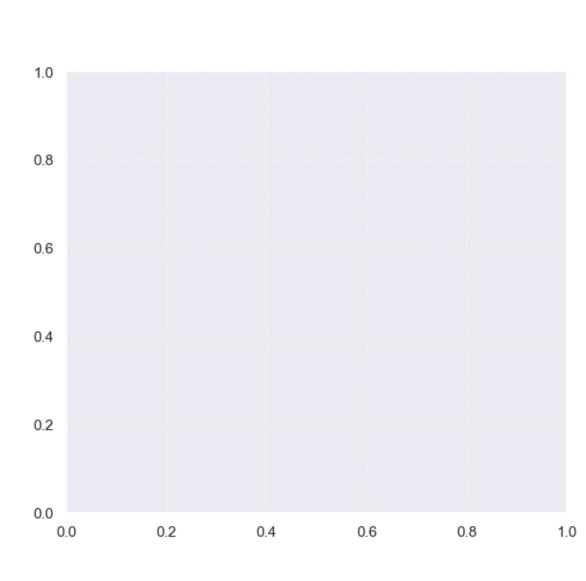
- test: 0.0802

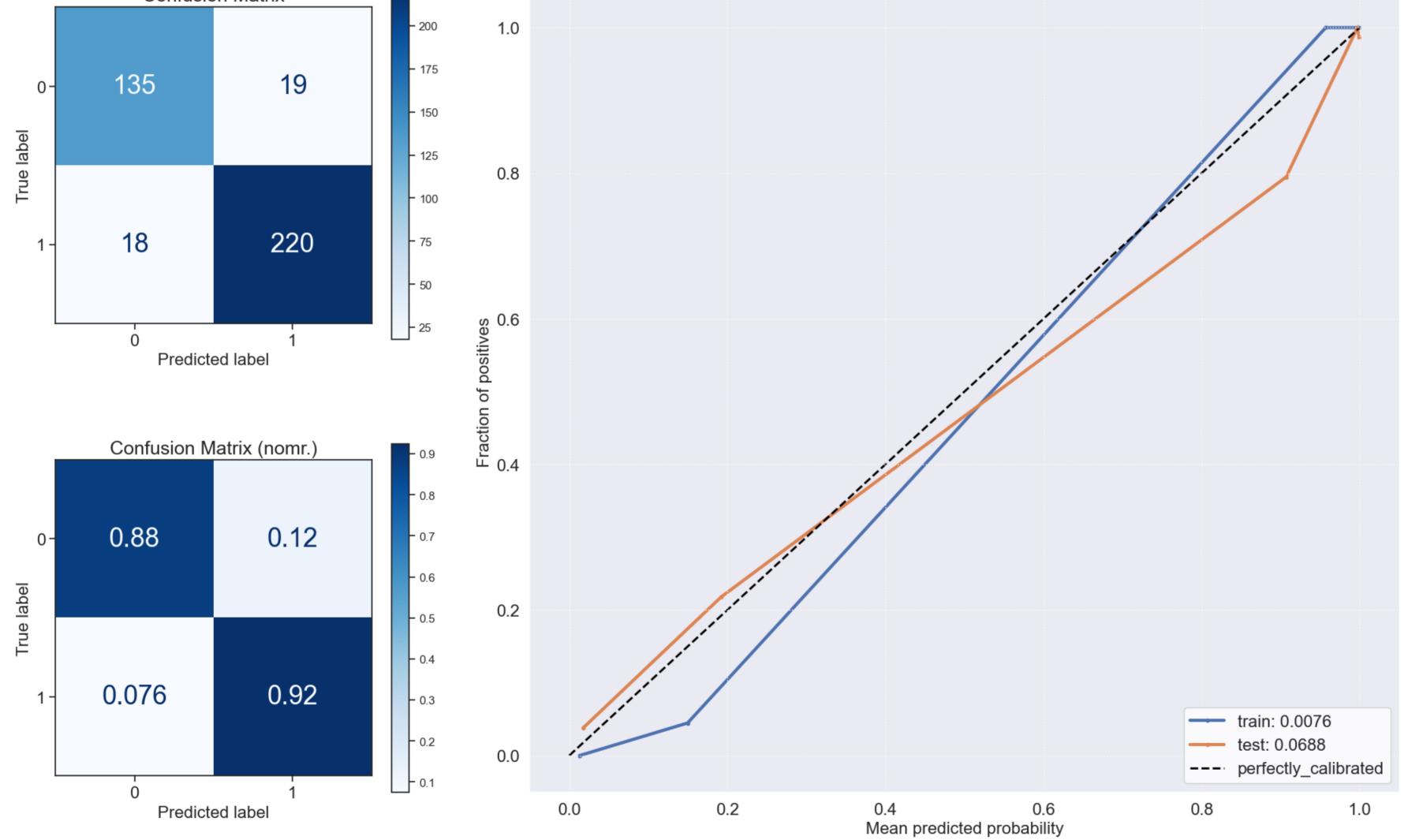
8.0

--- perfectly_calibrated

1.0







		Ac	tive/Inactive pro	obability		
70 -	In	active zone			Active zone	active (238) inactive (154)
50						
Percent						
00						
20						
0	0.0	0.2	0.4	0.6	0.8	10
	Probability					

Metric	Train	Test
AUC	0.9999	0.9657
accuracy	0.9936	0.9056
balanced accuracy	0.9921	0.9005
recall_(sens)	0.999	0.9244
specificity	0.9853	0.8766
precision	0.9906	0.9205
f1_score	0.9936	0.9056