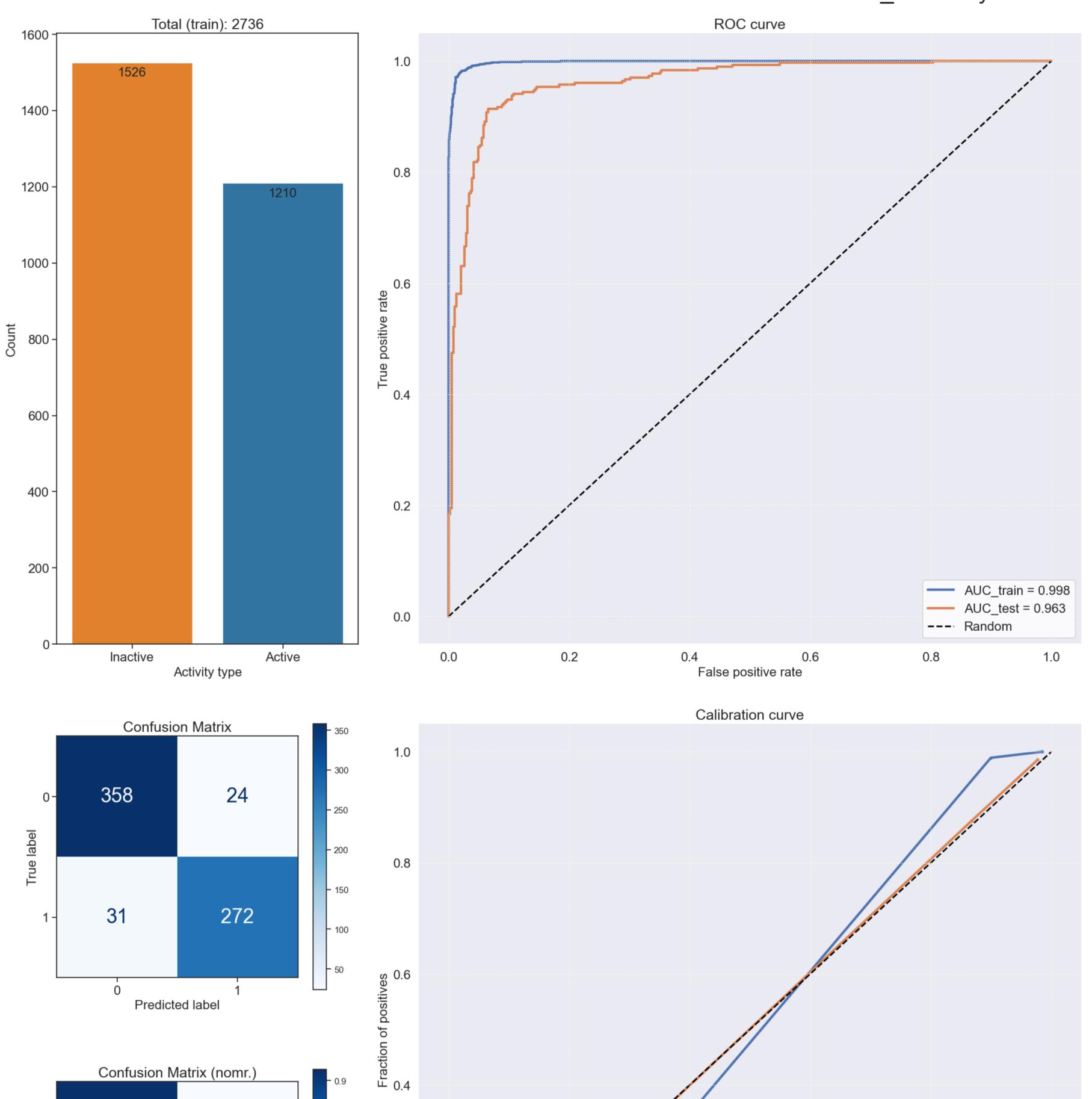
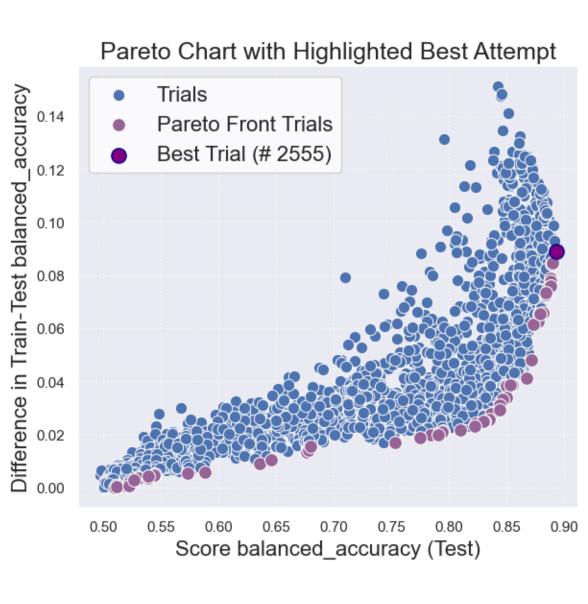
Protein: P08908 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	2.0621
colsample_bytree	0.7427
gamma	0.3588
lambda	3.7483
learning_rate	0.0171
max_depth	10.0
min_child_weight	1.0
n_estimators	954.0
reg_alpha	0.0039
reg_lambda	0.3138
scale_pos_weight	1.1786
subsample	0.564

Active/Inactive probability							
20	lna	ctive zone			Active :	zone	type active (303) inactive (382)
15							decoy (684)
Percent							
0	.0 0.).4 Prob	0. ability	6 0	.8	1.0

- 0.8

0.6

- 0.5

0.4

0.3

0.2

0.2

0.0

0.0

0.2

Predicted label

Confusion Matrix (nomr.)

Predicted label

0.063

0.9

0.94

0.1

0-

True label

Metric	Train	Test
AUC	0.9981	0.9628
accuracy	0.9803	0.9197
balanced accuracy	0.9804	0.9174
recall_(sens)	0.9818	0.8977
specificity	0.979	0.9372
precision	0.9738	0.9189
f1_score	0.9803	0.9196

--- train: 0.0226

8.0

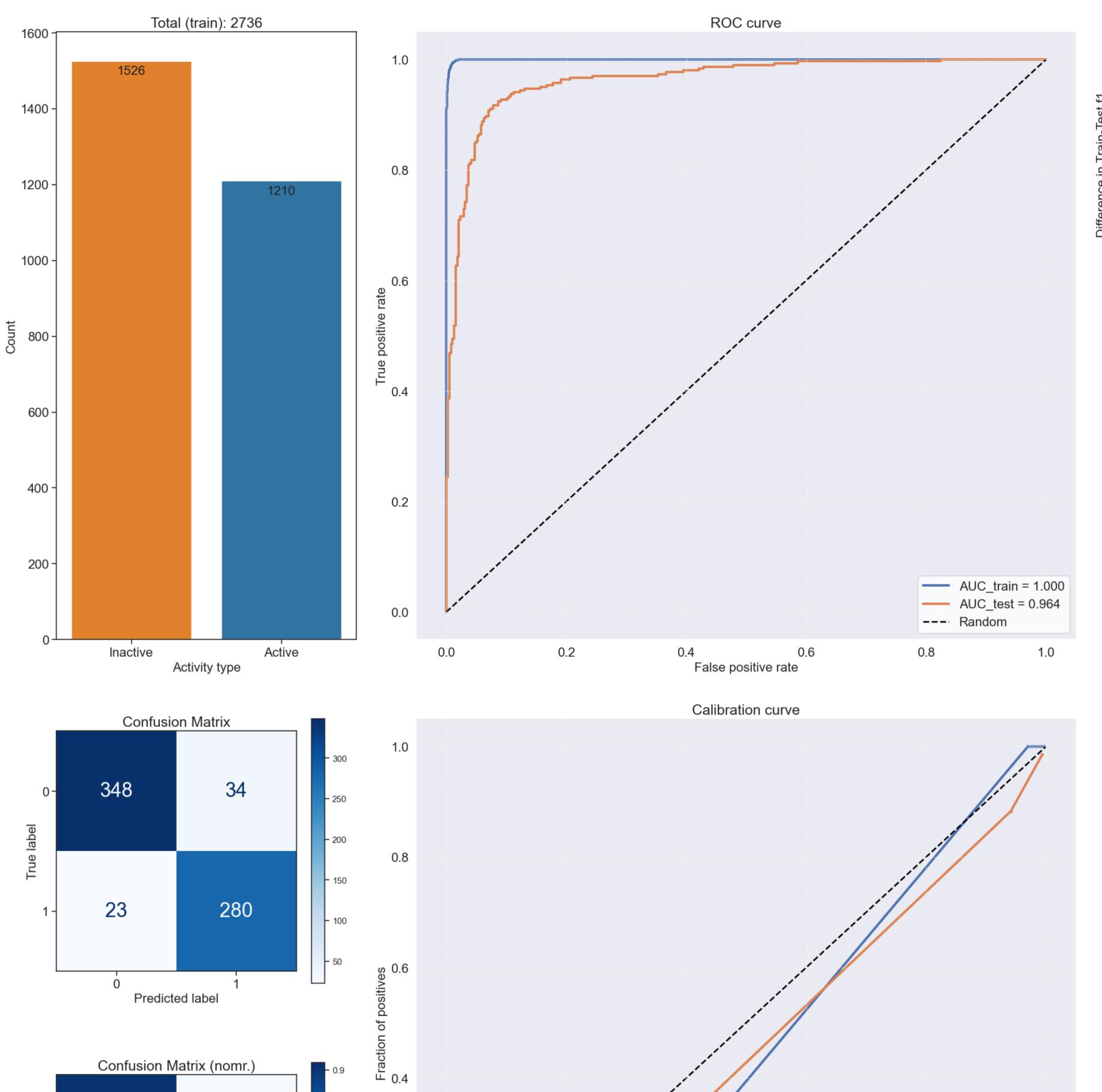
0.4 0.6 Mean predicted probability

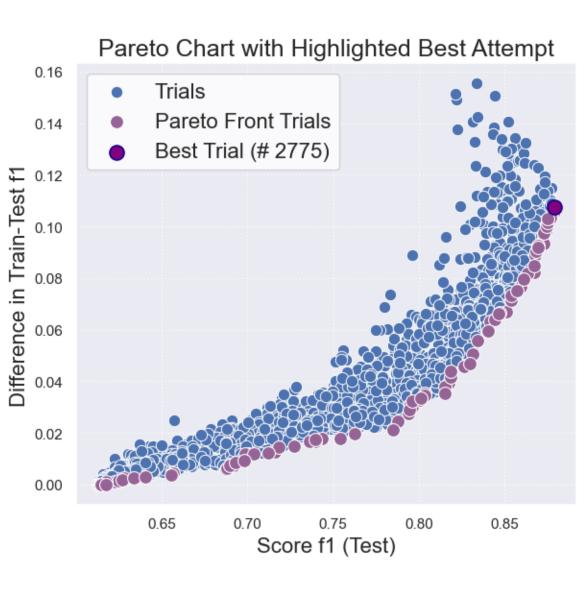
test: 0.0673

--- perfectly_calibrated

1.0

Protein: P08908 - metric: f1





XGBoost Parameters	Value
alpha	4.1316
colsample_bytree	0.7639
gamma	0.0422
lambda	4.8516
learning_rate	0.0423
max_depth	10.0
min_child_weight	1.0
n_estimators	628.0
reg_alpha	0.8244
reg_lambda	0.33
scale_pos_weight	3.445
subsample	0.5454

0.5			Active/Inactive	ve probability	,		
35		Inactive zone			Active z	zone = a	type ctive (303) nactive (382)
							ecoy (684)
25							
Percent 15							
10							
5							
0-	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

0.2

0.0

0.0

0.2

Confusion Matrix (nomr.)

Predicted label

0.089

0.92

0.91

0.076

0-

True label

Metric	Train	Test
AUC	0.9997	0.9638
accuracy	0.9857	0.9168
balanced accuracy	0.9872	0.9175
recall_(sens)	1.0	0.9241
specificity	0.9744	0.911
precision	0.9688	0.8917
f1_score	0.9858	0.9169

--- train: 0.0149

0.8

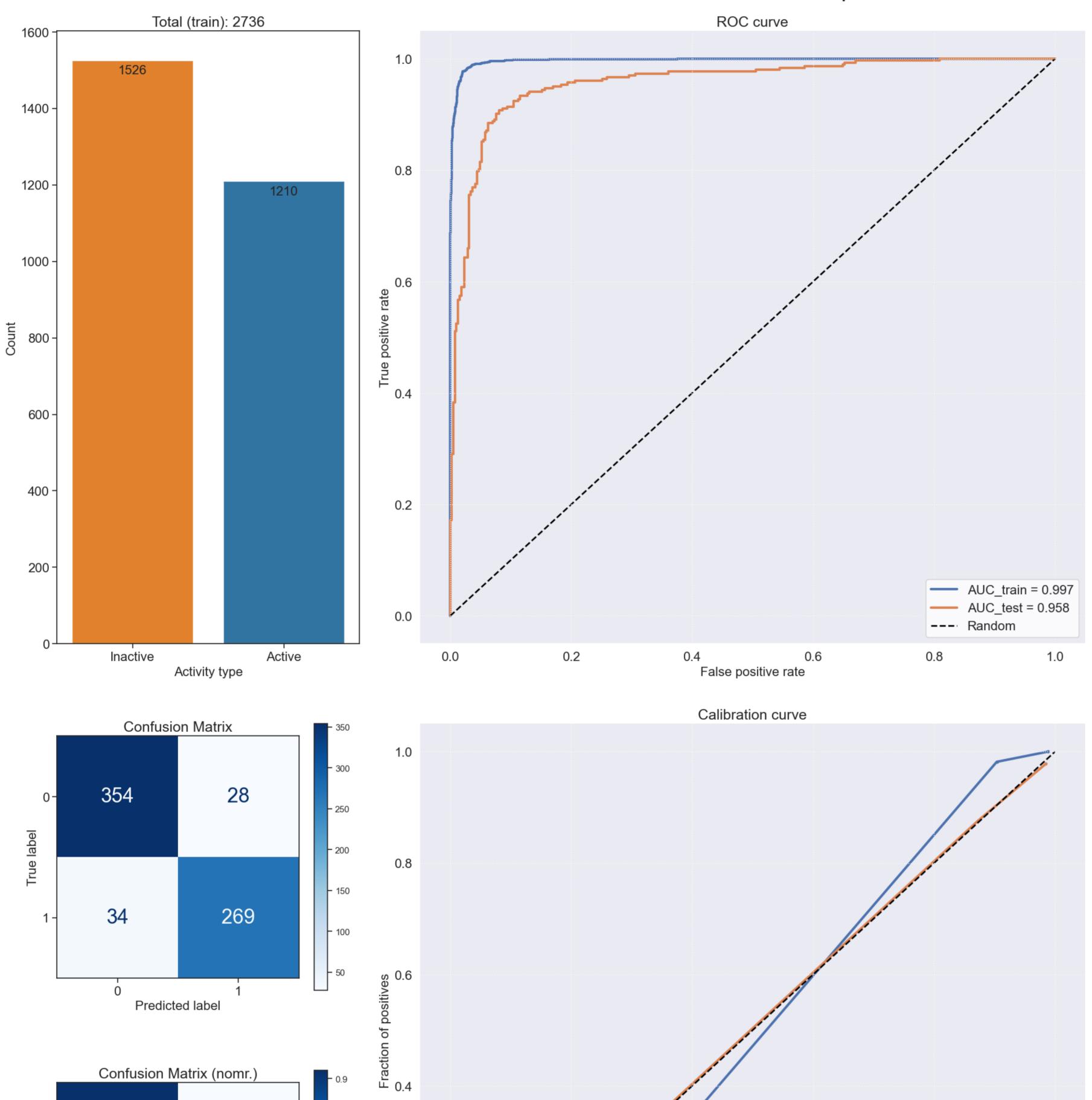
0.4 0.6 Mean predicted probability

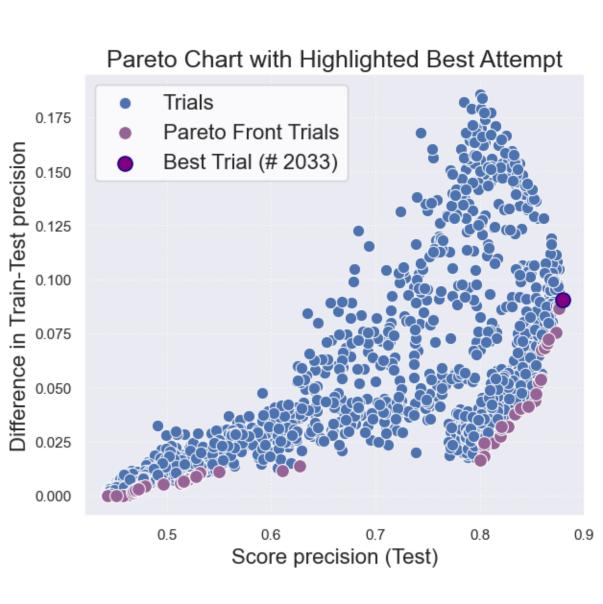
test: 0.0676

--- perfectly_calibrated

1.0

Protein: P08908 - metric: precision





XGBoost Parameters	Value
alpha	3.3505
colsample_bytree	0.9544
gamma	0.6268
lambda	3.5554
learning_rate	0.1184
max_depth	6.0
min_child_weight	2.0
n_estimators	511.0
reg_alpha	0.4403
reg_lambda	0.861
scale_pos_weight	1.1875
subsample	0.8002

		Active/Ina	active probability		
30 -	Inactiv	e zone		Active zone	type active (303)
					inactive (382) decoy (684)
25					
20					
Percent					
10 -					
10					
5					
0.0	0.2	0.4	0.6	0.8	1.0
		P	robability		

- 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

0.073

0.89

0.93

0.11

0-

True label

Metric	Train	Test
AUC	0.9971	0.9576
accuracy	0.977	0.9095
balanced accuracy	0.977	0.9072
recall_(sens)	0.9777	0.8878
specificity	0.9764	0.9267
precision	0.9705	0.9057
f1_score	0.977	0.9094

train: 0.0242

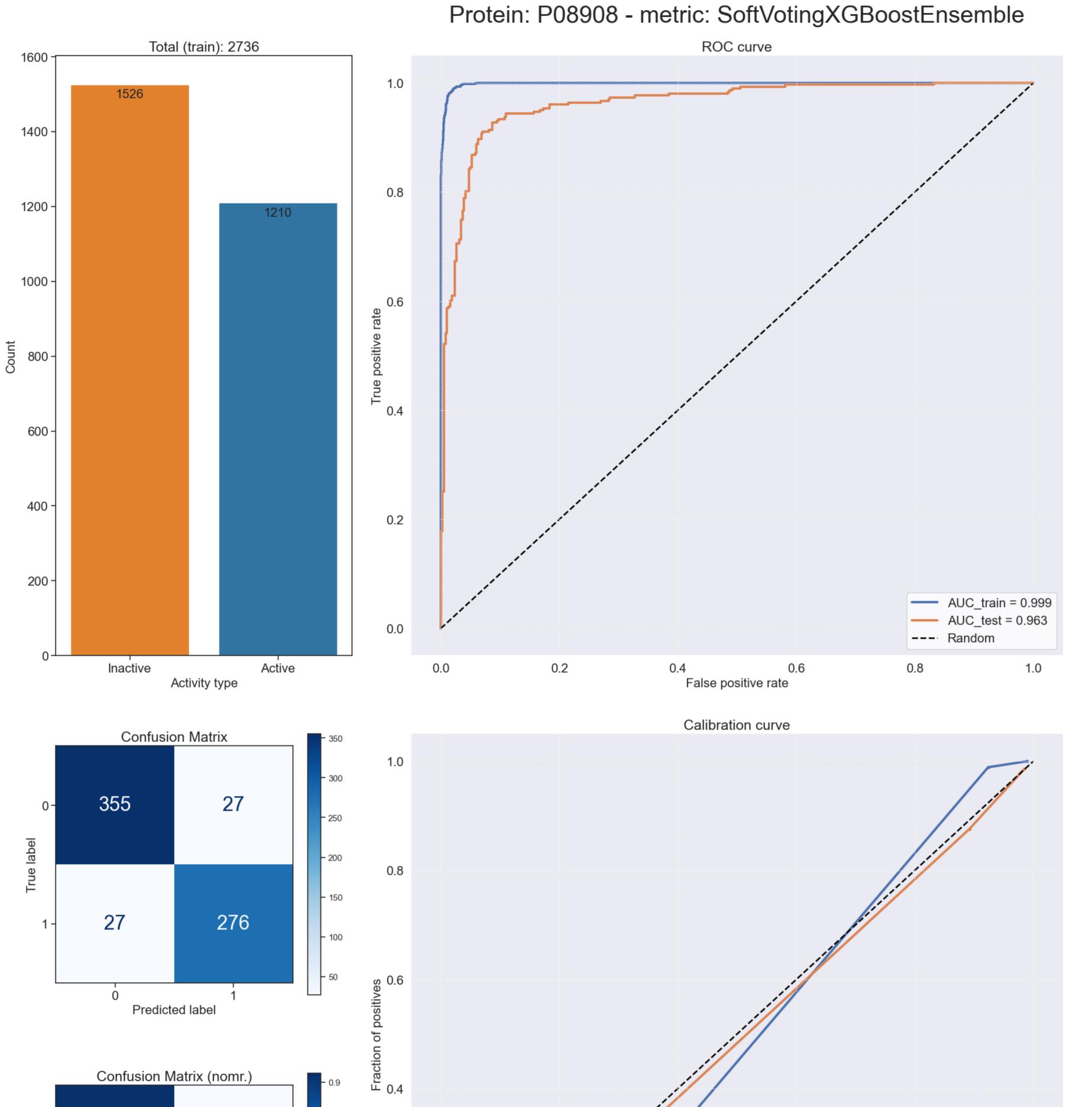
test: 0.0712

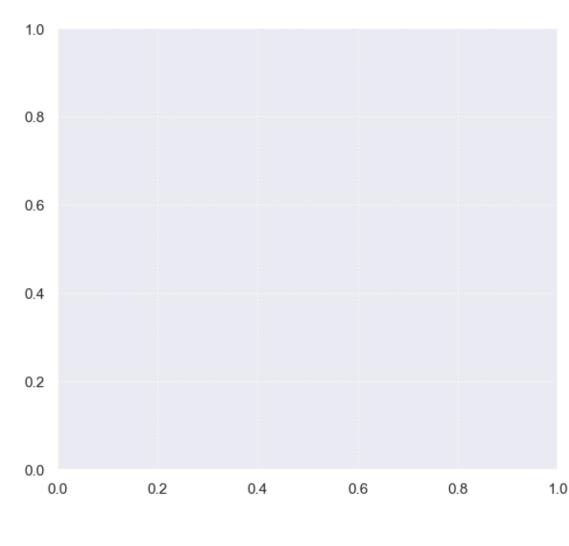
0.8

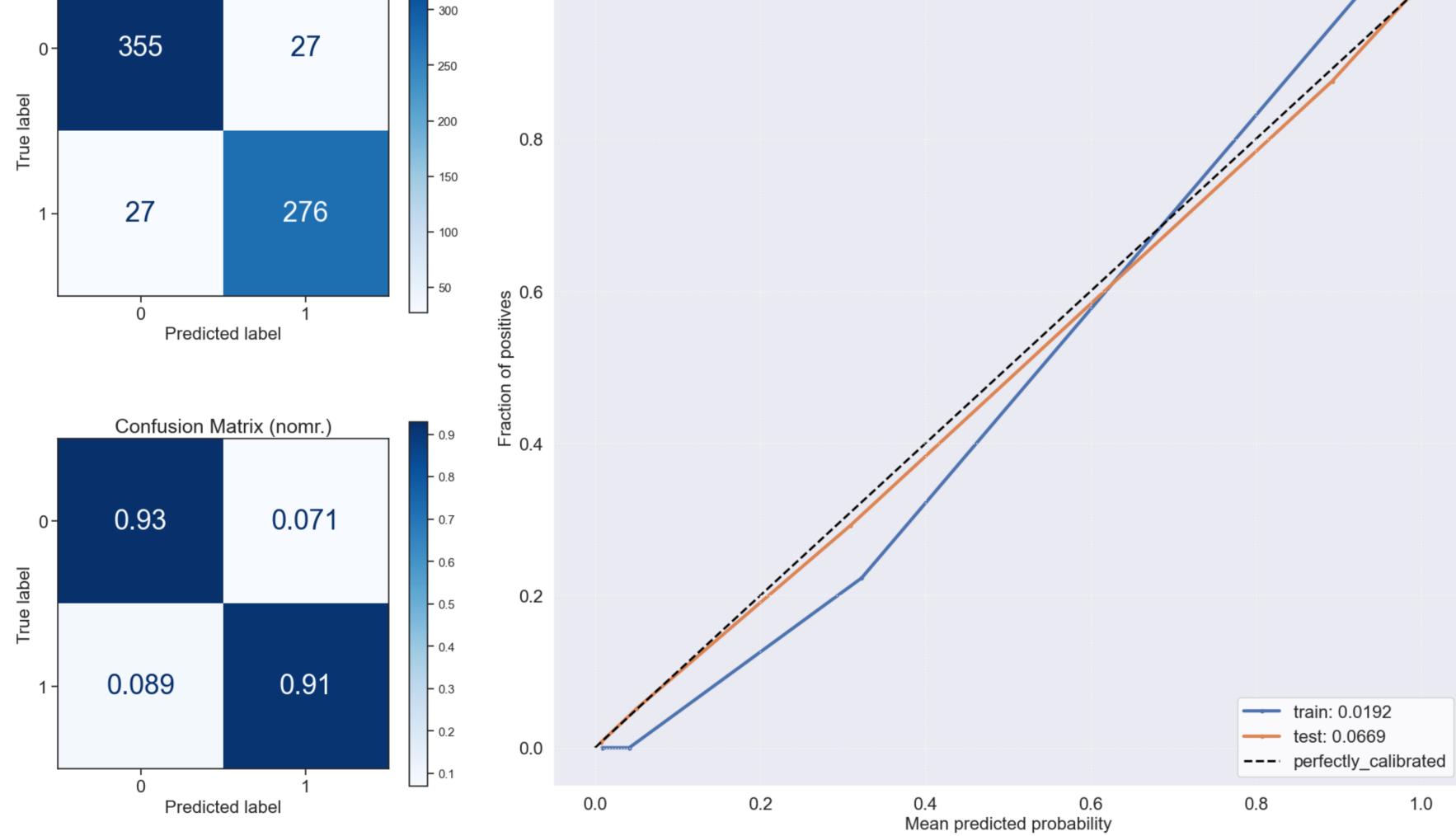
0.4 0.6 Mean predicted probability

--- perfectly_calibrated

1.0







		Active/I	nactive probability	у	
25	Inacti	ve zone		Active 2	type active (303) inactive (382) decoy (684)
20					
0	0 0.2	0.4	0.6 Probability		8 1.0

Metric	Train	Test
AUC	0.9988	0.9626
accuracy	0.9825	0.9212
balanced accuracy	0.9834	0.9201
recall_(sens)	0.9917	0.9109
specificity	0.9751	0.9293
precision	0.9693	0.9109
f1_score	0.9825	0.9212