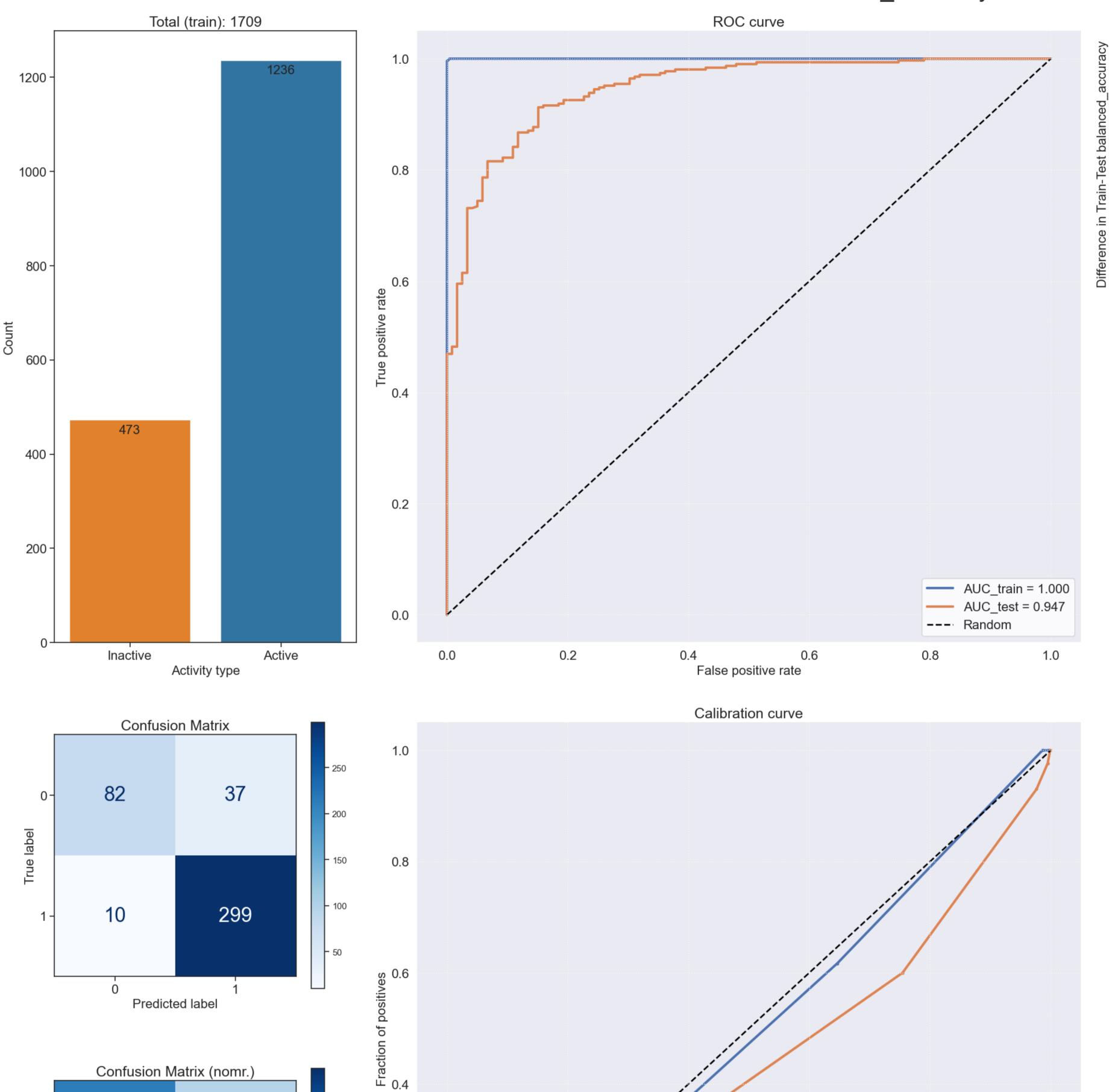
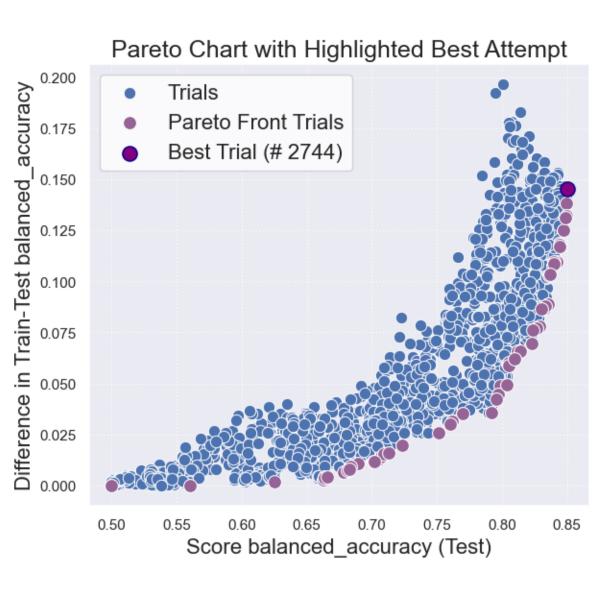
Protein: O43614 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	3.1625
colsample_bytree	0.5865
gamma	0.0291
lambda	2.012
learning_rate	0.0591
max_depth	10.0
min_child_weight	1.0
n_estimators	770.0
reg_alpha	0.3664
reg_lambda	0.4586
scale_pos_weight	1.51
subsample	0.6147

		Ac	tive/Inactive prob	pability		
50	lna	ctive zone			Active zone	type active (309) inactive (119)
40						decoy (427)
- 05						
20						
10						
0	0.0	.2 0	.4 Probability	0.6	0.8	1.0

Confusion Matrix (nomr.)

Predicted label

0.31

0.97

0.69

0.032

0-

True label

Metric	Train	Test
AUC	1.0	0.9469
accuracy	0.9965	0.8902
balanced accuracy	0.9937	0.8284
recall_(sens)	1.0	0.9676
specificity	0.9873	0.6891
precision	0.9952	0.8899
f1_score	0.9965	0.8855

--- train: 0.005

0.8

0.4 0.6 Mean predicted probability

test: 0.0835

--- perfectly_calibrated

1.0

			cuve/macuv	e probabil	ity			
50	Ina	active zone			Active :	:	type active inactive	
40							decoy	
ent								
20								
10								
0	0.0 0	0.2	0.4 Proba	0. ability	6 0	.8	1	.0

0.2

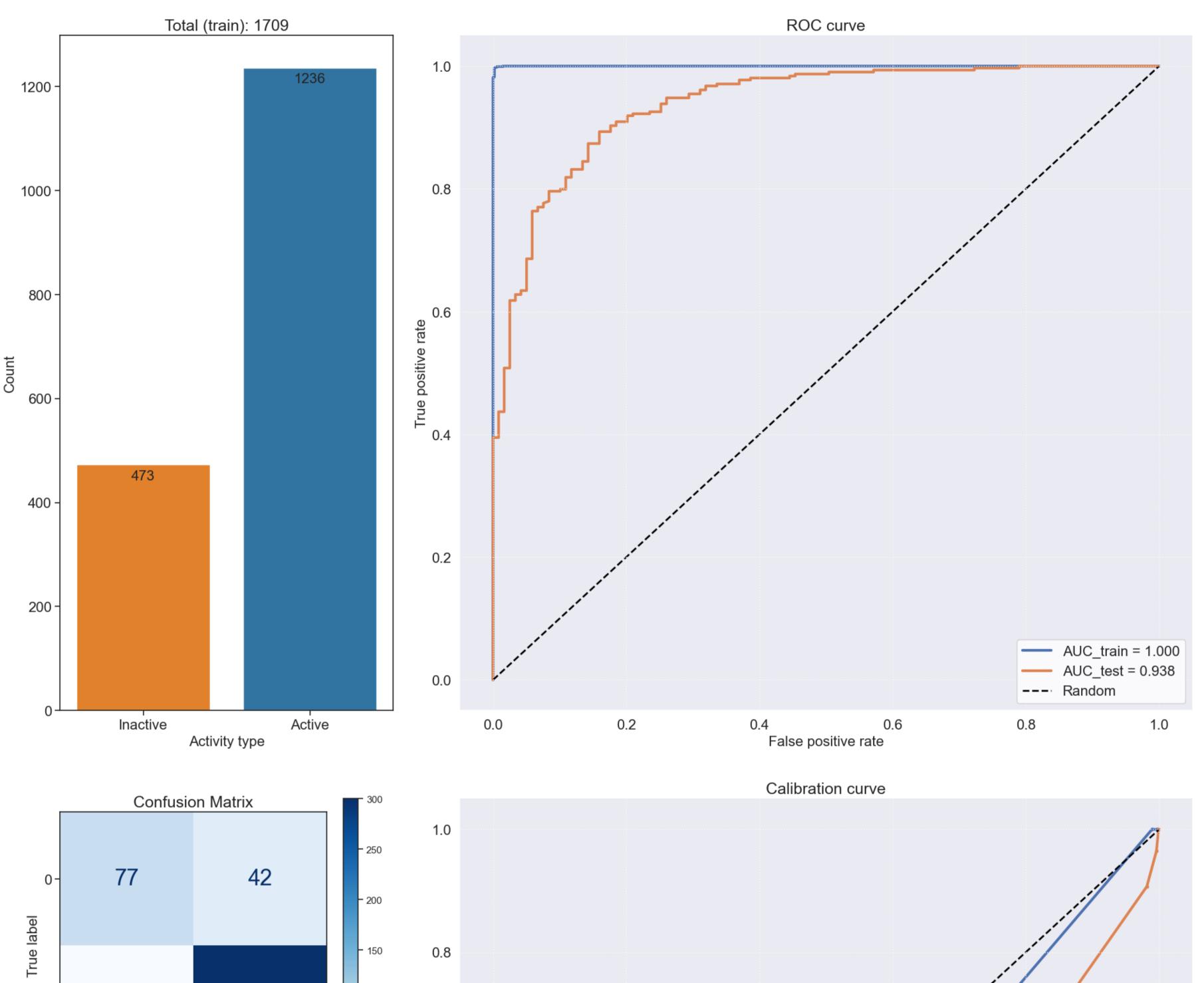
0.0

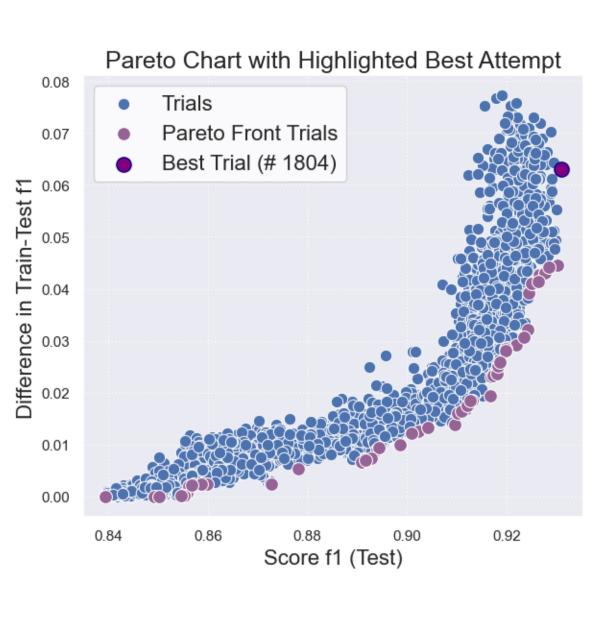
0.0

0.2

0.4

Protein: O43614 - metric: f1





XGBoost Parameters	Value
alpha	4.8206
colsample_bytree	0.5035
gamma	0.0464
lambda	1.1819
learning_rate	0.087
max_depth	10.0
min_child_weight	4.0
n_estimators	968.0
reg_alpha	0.2787
reg_lambda	0.1302
scale_pos_weight	2.8238
subsample	0.9695

	Predicted labe) <u> </u>	0	0.0	0.2	0.4 Mean predicted p	0.6 probability
		Act	ive/Inactive prob	pability	:		
60	Ina	ctive zone		Α	ctive zone	type active (309) inactive (119)	
50						decoy (427)	
40							
Percent 08							
20							
10							
0							
Ü	0.0	2 0.4	4 Probability	0.6	0.8	1.0	

100

- 50

0.4

0.2

Fraction of positives

0.2

0.0

300

0.35

0.97

9

0.65

0.029

0-

True label

Predicted label

Confusion Matrix (nomr.)

Metric	Train	Test
AUC	0.9999	0.9383
accuracy	0.9906	0.8808
balanced accuracy	0.9831	0.809
recall_(sens)	1.0	0.9709
specificity	0.9662	0.6471
precision	0.9872	0.8772
f1_score	0.9906	0.8743

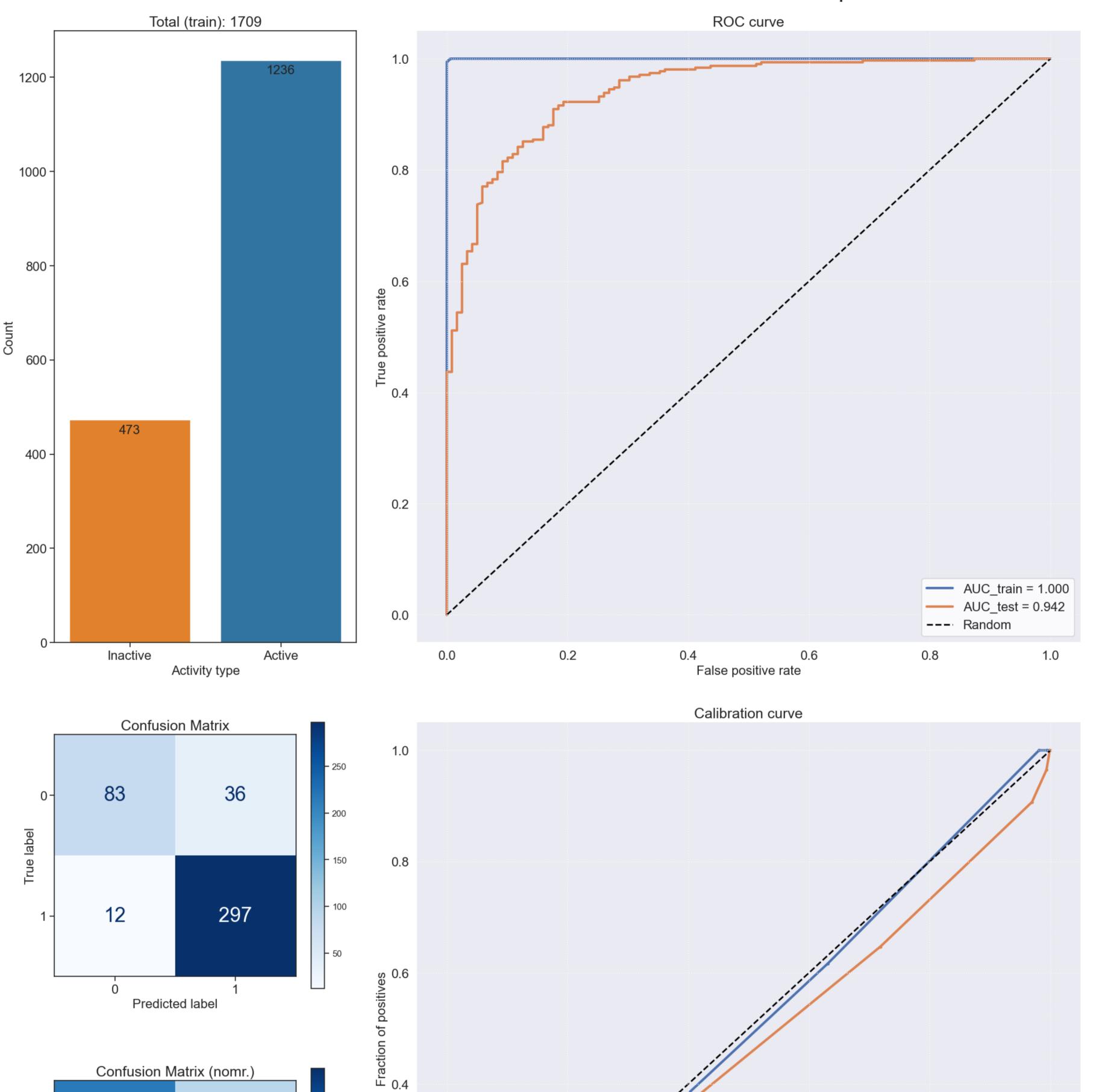
--- train: 0.0089

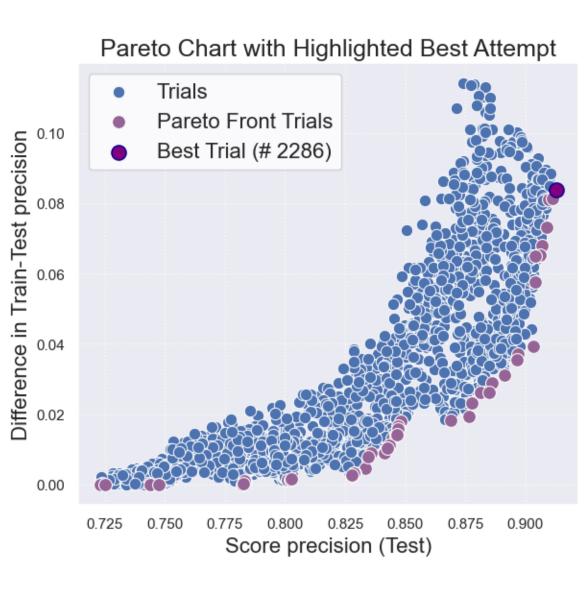
0.8

test: 0.0913

--- perfectly_calibrated

Protein: O43614 - metric: precision





XGBoost Parameters	Value
alpha	3.4813
colsample_bytree	0.9421
gamma	0.0962
lambda	2.2935
learning_rate	0.128
max_depth	6.0
min_child_weight	1.0
n_estimators	573.0
reg_alpha	0.6375
reg_lambda	0.1287
scale_pos_weight	1.1046
subsample	0.6379

		Ac	tive/Inactive	e probability			
50							type
00	Ina	ctive zone			Active z	cone	active (309) inactive (119) decoy (427)
40							
Percer							
20							
0-							
	0.0 0.	2 0).4 Probal	0.6 bility	0.	8	1.0

0.2

0.0

0.0

0.2

0.2

Confusion Matrix (nomr.)

Predicted label

0.3

0.96

0.7

0.039

0-

True label

Metric	Train	Test
AUC	1.0	0.9417
accuracy	0.9977	0.8879
balanced accuracy	0.9958	0.8293
recall_(sens)	1.0	0.9612
specificity	0.9915	0.6975
precision	0.9968	0.8919
f1_score	0.9977	0.8837

--- train: 0.006

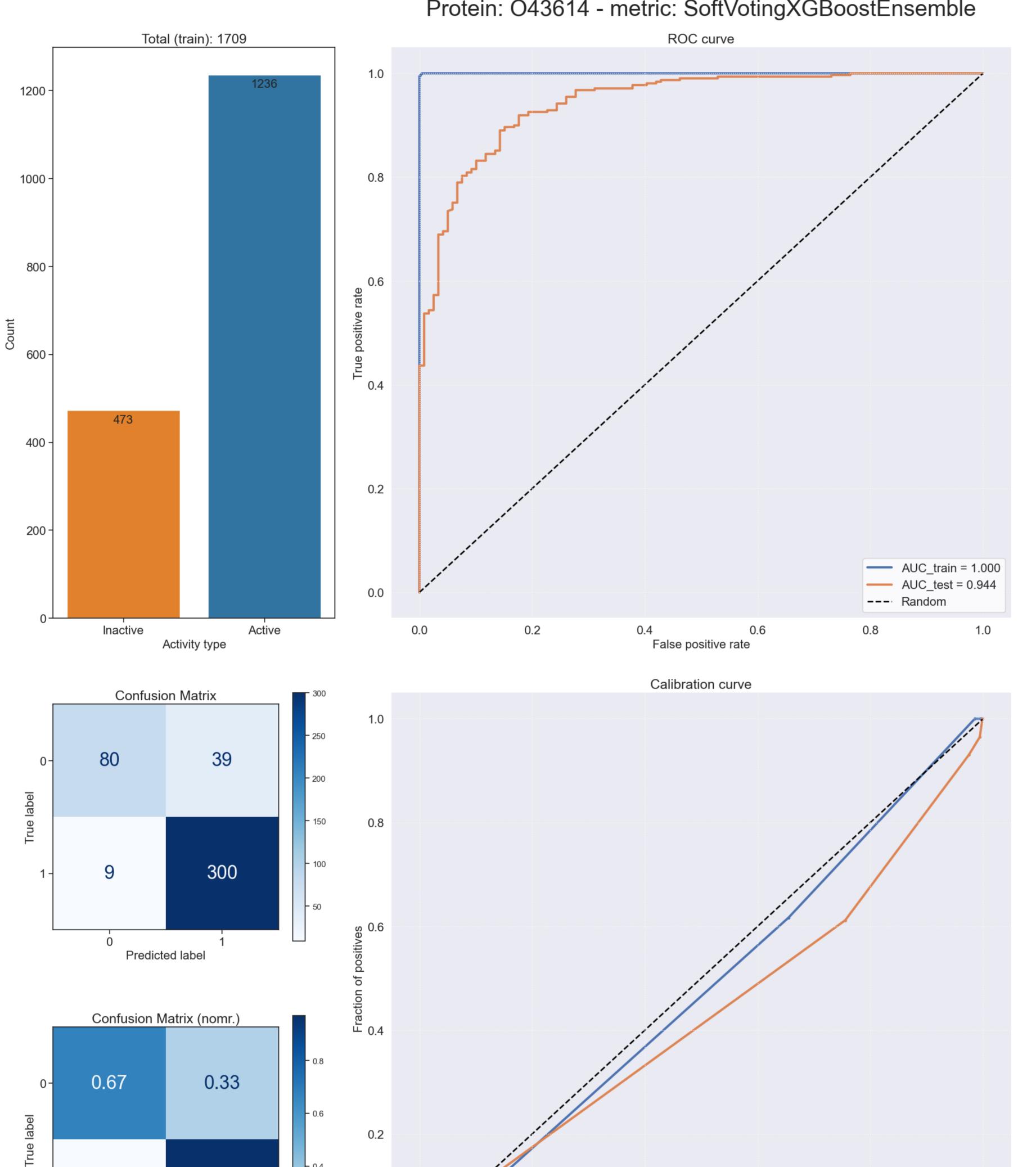
0.8

0.4 0.6 Mean predicted probability

test: 0.0847

--- perfectly_calibrated

Protein: O43614 - metric: SoftVotingXGBoostEnsemble



		Ac	tive/Inactiv	e probabil	lity		
50	Ina	ctive zone			Activ	e zone	type active (309) inactive (119) decoy (427)
40							
20							
10							
0	0.0	.2 0.	.4 Proba	0.0 bility	6	0.8	1.0

- 0.4

0.2

0.0

0.0

0.2

0.97

0.029

Predicted label

Metric	Train	Test
AUC	1.0	0.9443
accuracy	0.9941	0.8879
balanced accuracy	0.9894	0.8216
recall_(sens)	1.0	0.9709
specificity	0.9789	0.6723
precision	0.992	0.885
f1_score	0.9941	0.8824

train: 0.0063

test: 0.0847

0.8

0.4 0.6 Mean predicted probability

perfectly_calibrated

1.0

0.8

0.6

0.4

0.2

0.0

0.0

0.2