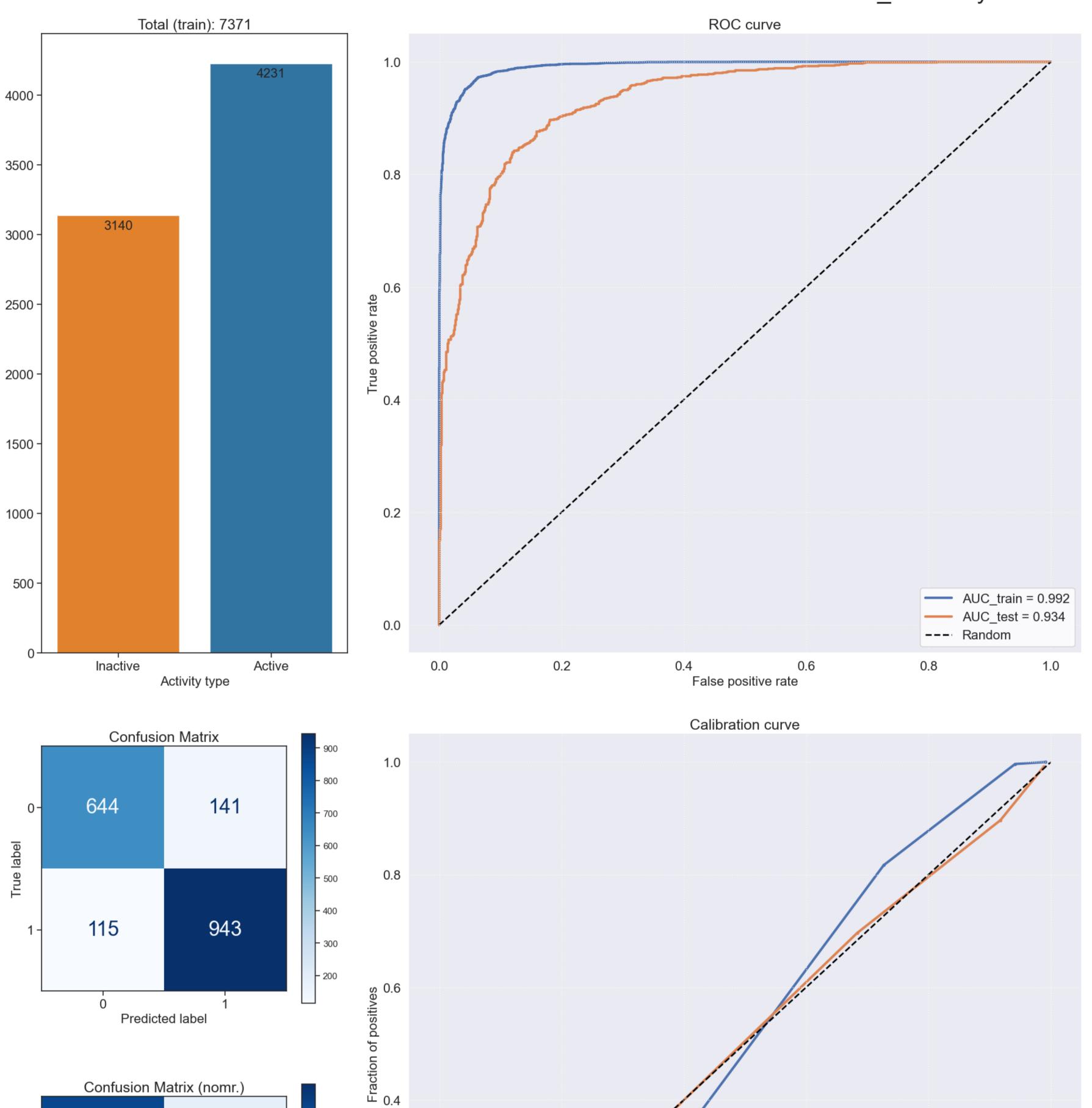
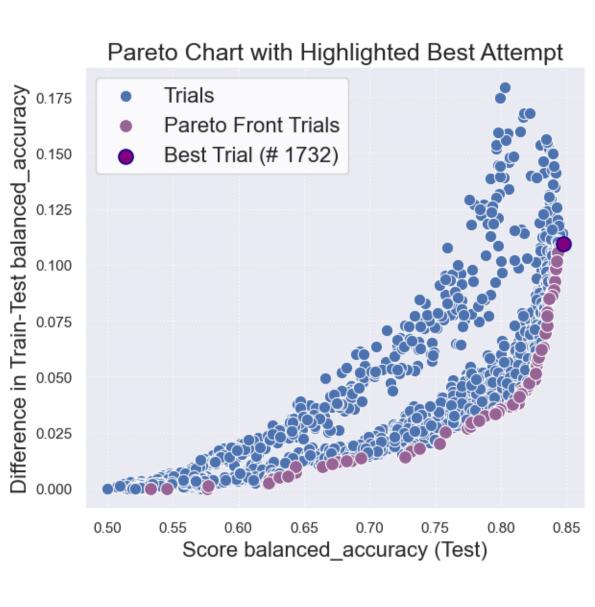
## Protein: P34972 - metric: balanced\_accuracy





XGBoost Parameters	Value
alpha	0.3607
colsample_bytree	0.5665
gamma	0.0488
lambda	4.3803
learning_rate	0.0339
max_depth	10.0
min_child_weight	1.0
n_estimators	920.0
reg_alpha	0.7525
reg_lambda	0.7749
scale_pos_weight	1.1031
subsample	0.5225

		Ac	ctive/Inactive	probability				
17.5	lna	ctive zone			Active z	zone	type active (	E
							inactive decoy (	
ent								
5.0								
2.5							<u>.                                    </u>	
0.0	0.0	.2 (	).4 Probak	0.6	0.	8	1	.0

Confusion Matrix (nomr.)

Predicted label

0.18

0.89

- 0.6

0.4

0.3

0.2

0.2

0.0

0.0

0.2

0.82

0.11

0-

True label

Metric	Train	Test		
AUC	0.9922	0.9337		
accuracy	0.9573	0.8611		
balanced accuracy	0.9546	0.8558		
recall_(sens)	0.9723	0.8913		
specificity	0.9369	0.8204		
precision	0.9541	0.8699		
f1_score	0.9572	0.8608		

--- train: 0.0416

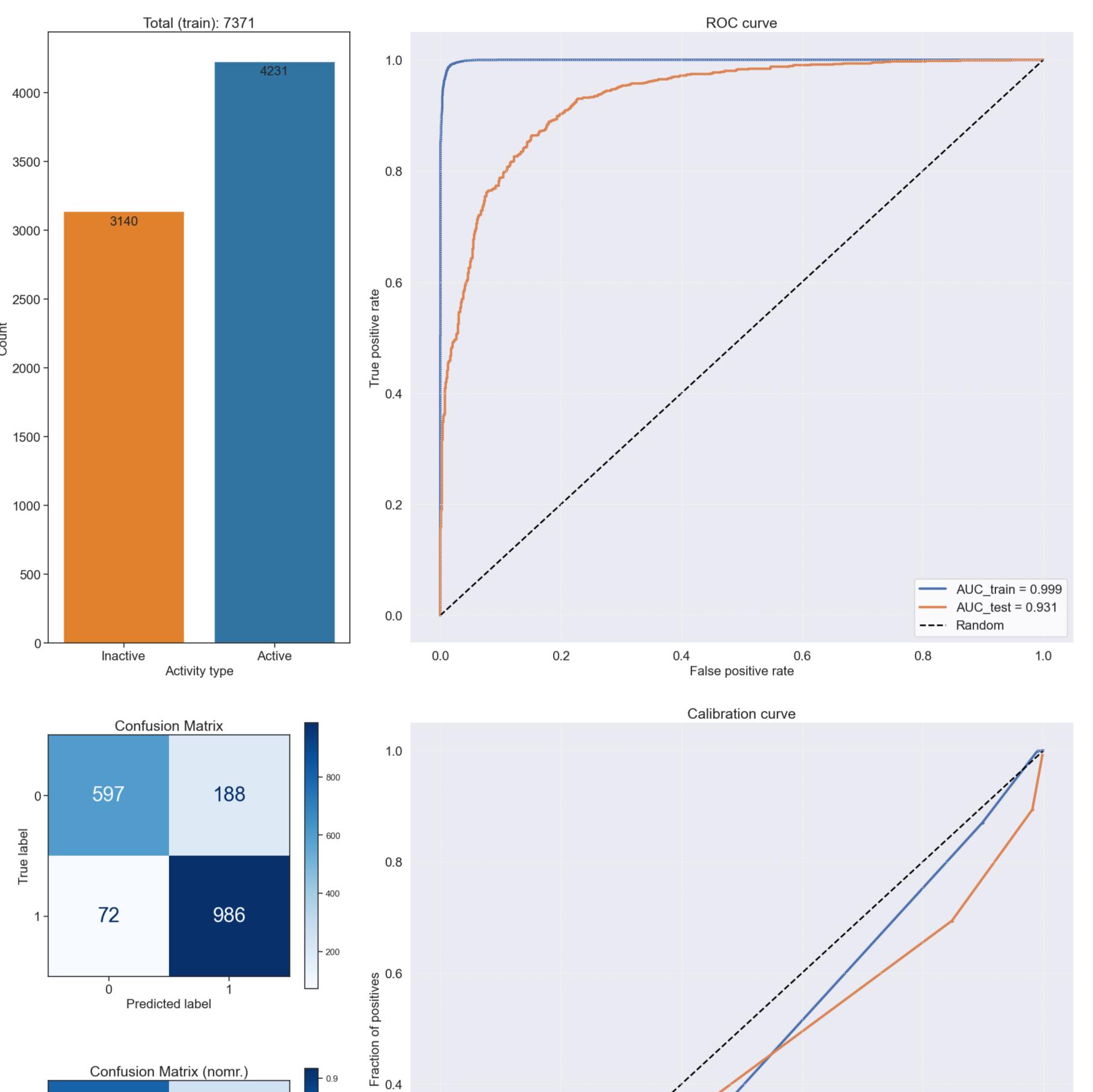
0.8

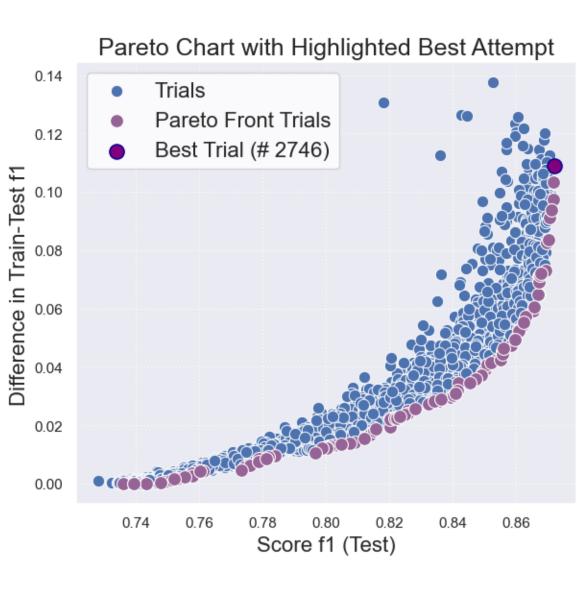
0.4 0.6 Mean predicted probability

test: 0.1011

--- perfectly\_calibrated

## Protein: P34972 - metric: f1





XGBoost Parameters	Value
alpha	4.8109
colsample_bytree	0.5117
gamma	0.1552
lambda	2.4148
learning_rate	0.0745
max_depth	10.0
min_child_weight	1.0
n_estimators	855.0
reg_alpha	0.0207
reg_lambda	0.7793
scale_pos_weight	3.7851
subsample	0.7871

	Active/Inactive probability								
40		Inactive zone			Active zone	type active (1058) inactive (785) decoy (1833)			
Percent 00 									
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

0.1

0.2

0.0

0.0

0.2

Confusion Matrix (nomr.)

Predicted label

0.24

0.93

0.76

0.068

0-

True label

Metric	Train	Test		
AUC	0.999	0.9306		
accuracy	0.9734	0.8589		
balanced accuracy	0.9688	0.8462		
recall_(sens)	0.9998	0.9319		
specificity	0.9379	0.7605		
precision	0.9559	0.8399		
f1_score	0.9733	0.857		

Metric	irain	lest
AUC	0.999	0.9306
accuracy	0.9734	0.8589
balanced accuracy	0.9688	0.8462
recall_(sens)	0.9998	0.9319
specificity	0.9379	0.7605
precision	0.9559	0.8399
f1_score	0.9733	0.857

--- train: 0.023

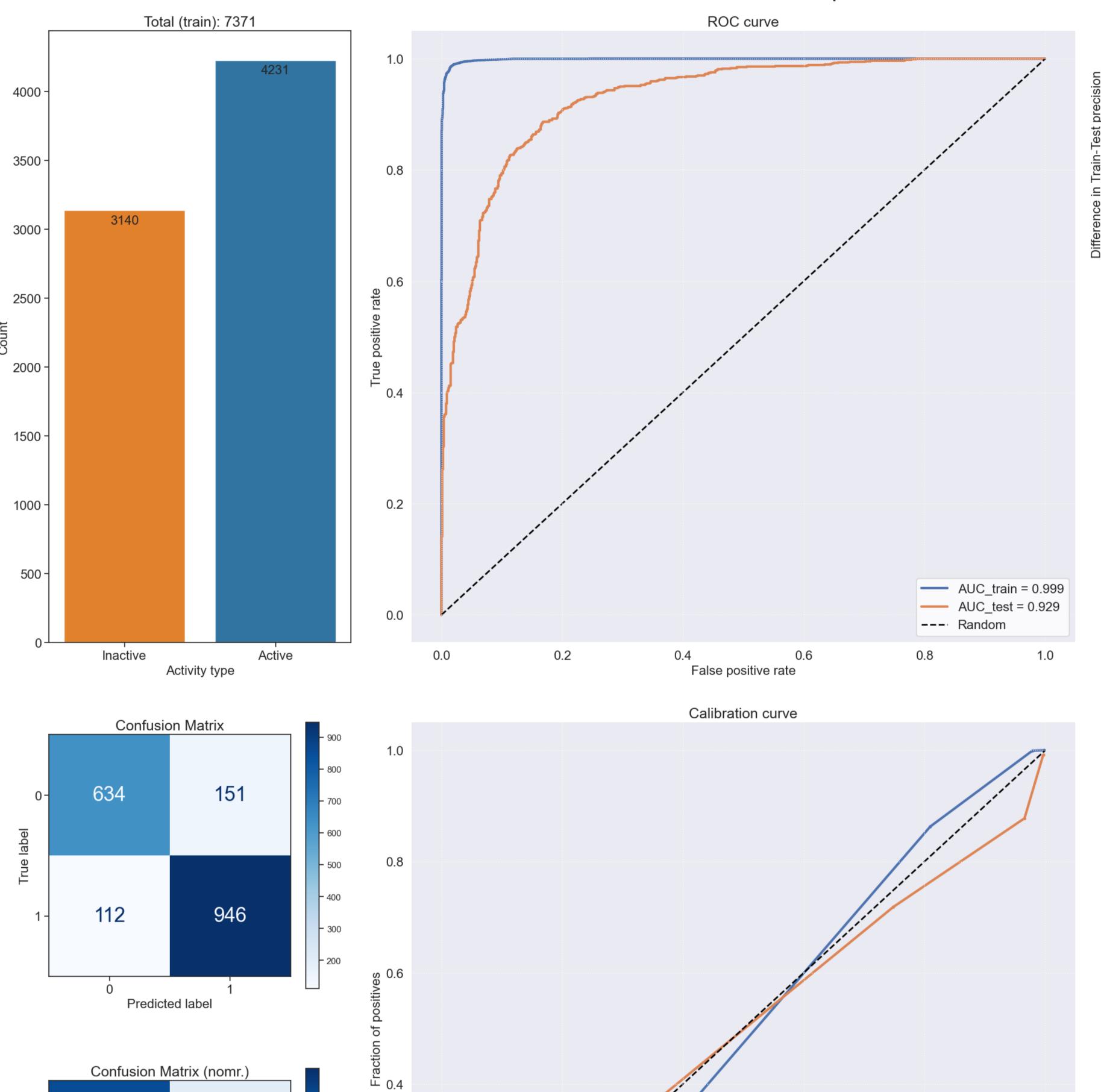
0.8

0.4 0.6 Mean predicted probability

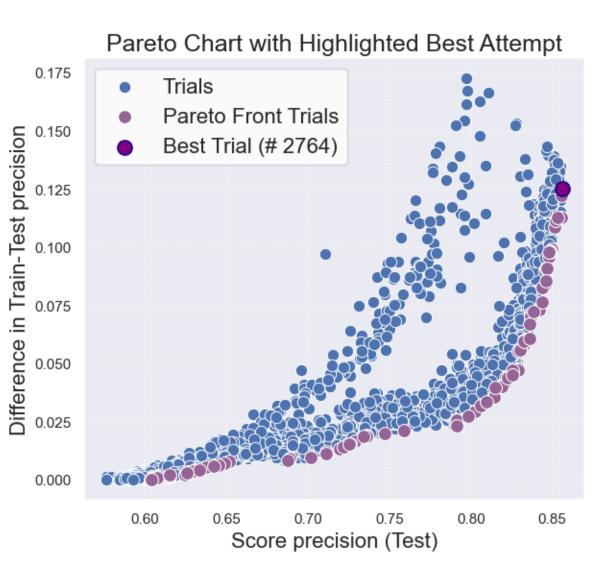
test: 0.1094

--- perfectly\_calibrated

## Protein: P34972 - metric: precision



0.2



XGBoost Parameters	Value
alpha	0.2912
colsample_bytree	0.7782
gamma	0.0068
lambda	2.8252
learning_rate	0.1463
max_depth	10.0
min_child_weight	1.0
n_estimators	358.0
reg_alpha	0.9308
reg_lambda	0.1068
scale_pos_weight	1.3928
subsample	0.6656

Active/Inactive probability								
35	lna	ctive zone			Active 2		type active (1058) inactive (785) decoy (1833)	
25								
- 20								
10								
5								
0	0.0 0.	2 0	0.4	0.6	6 0	.8	1.0	
	Probability							

0.2

0.0

0.0

0.4

0.3

0.2

0.81

0.11

Predicted label

0-

True label

0.19

0.89

Metric	Train	Test		
AUC	0.9988	0.9292		
accuracy	0.9824	0.8573		
balanced accuracy	0.9806	0.8509		
recall_(sens)	0.9927	0.8941		
specificity	0.9685	0.8076		
precision	0.977	0.8624		
f1_score	0.9823	0.8568		

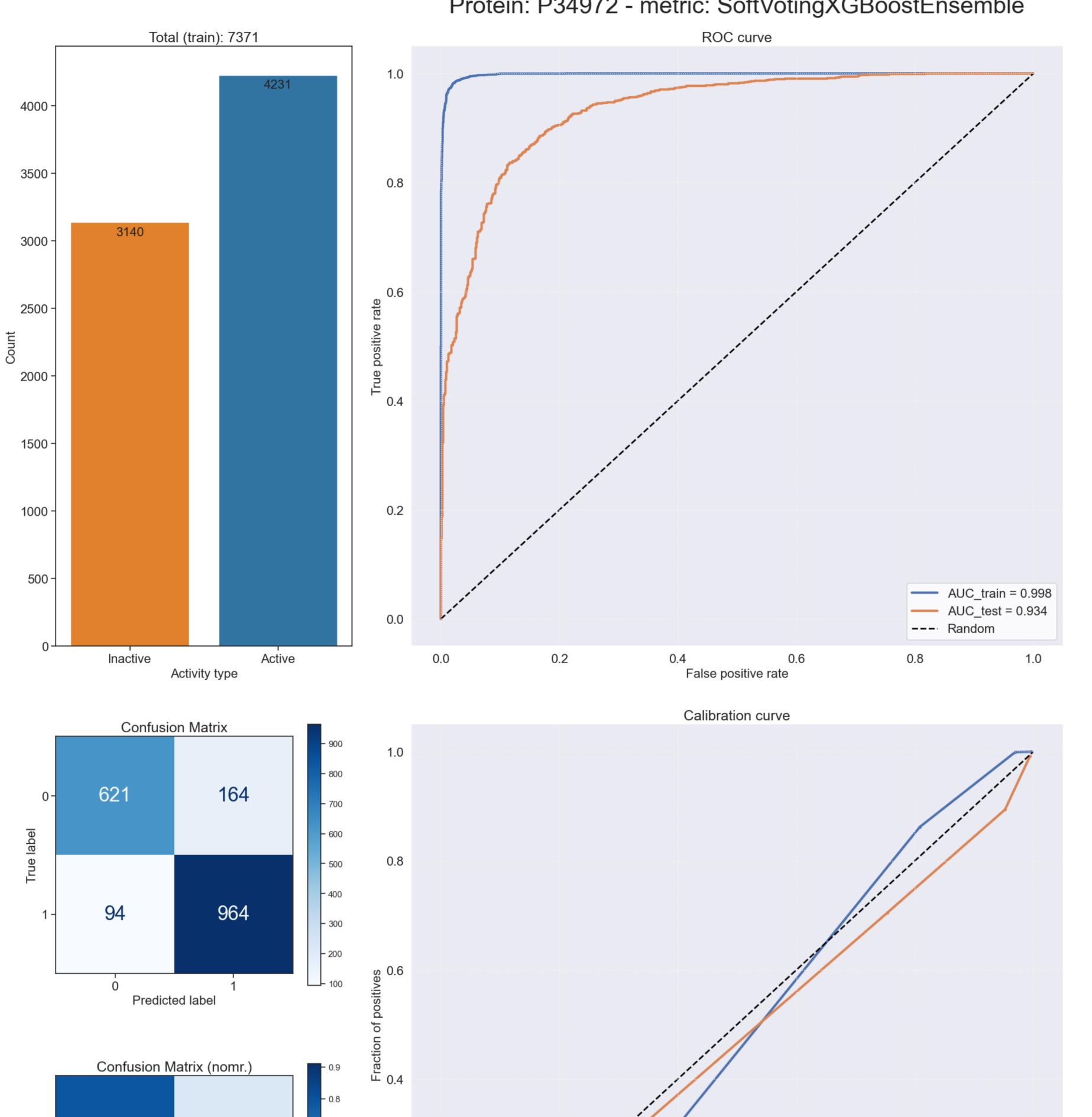
--- train: 0.0196

0.8

0.4 0.6 Mean predicted probability test: 0.1044

--- perfectly\_calibrated

## Protein: P34972 - metric: SoftVotingXGBoostEnsemble



0.2

0.4 0.6 Mean predicted probability

0.8	1.0		
	Metric	Train	Test
	AUC	0.9982	0.9339
	accuracy	0.9749	0.86
	balanced accuracy	0.9714	0.8511
	recall_(sens)	0.995	0.9112
	specificity	0.9478	0.7911
	precision	0.9625	0.8546
	f1 score	0.9748	0.859

train: 0.0254

--- perfectly\_calibrated

test: 0.101

0.8

0.6

0.4

0.2

0.0

0.0

0.2

0.6

Active/Inactive probability								
25	Ina	ctive zone			Active z		type active (1058) inactive (785)	
20							decoy (1833)	
Jercent 15								
10								
5								
0								
Ü	0.0	2	0.4 Probat	0.6	0.8	8	1.0	
			i iobai	Jility				

0.79

0.089

0-

True label

0.21

0.91

Predicted label

- 0.7

- 0.6

- 0.5

- 0.4

- 0.3

0.2

0.2

0.0