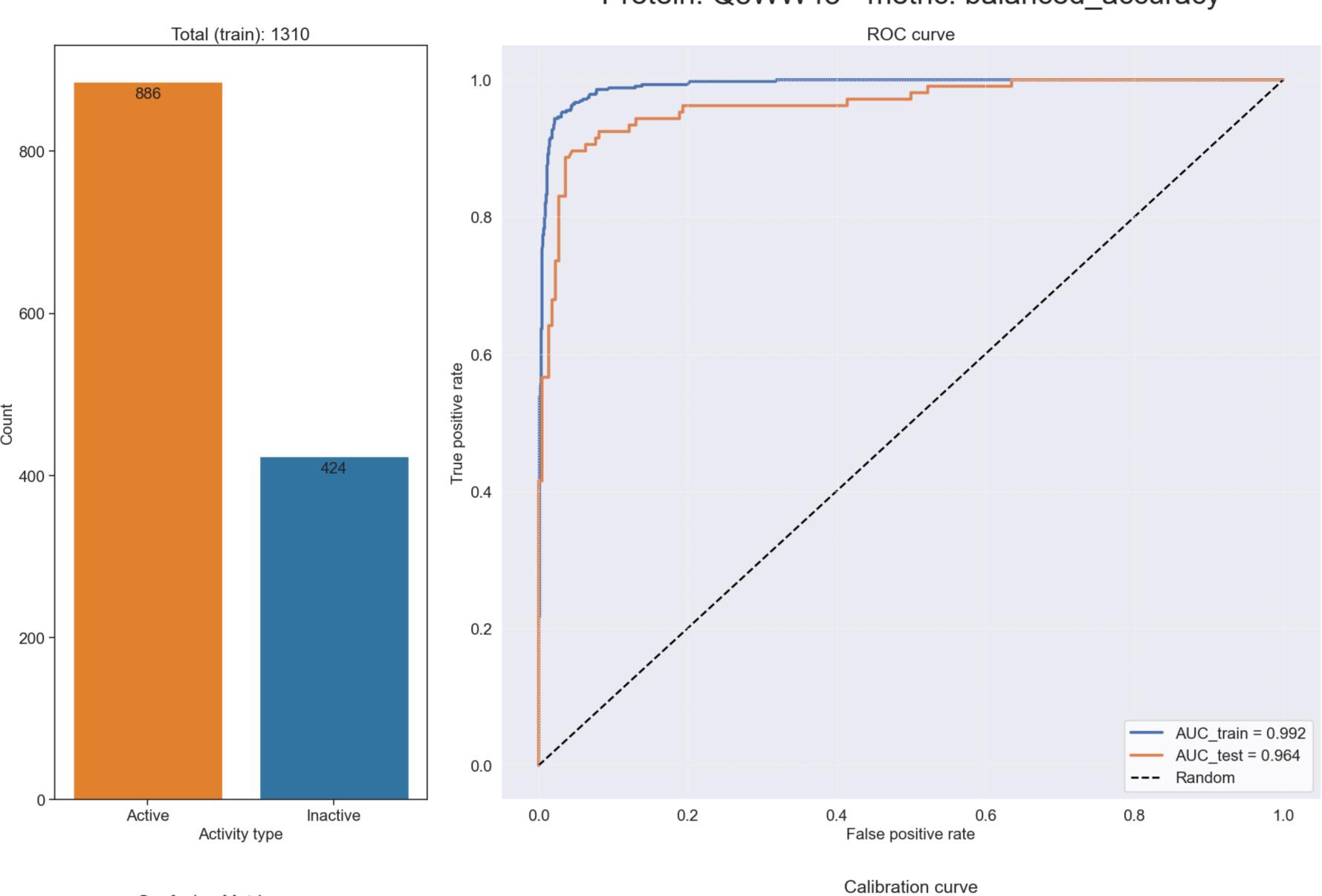
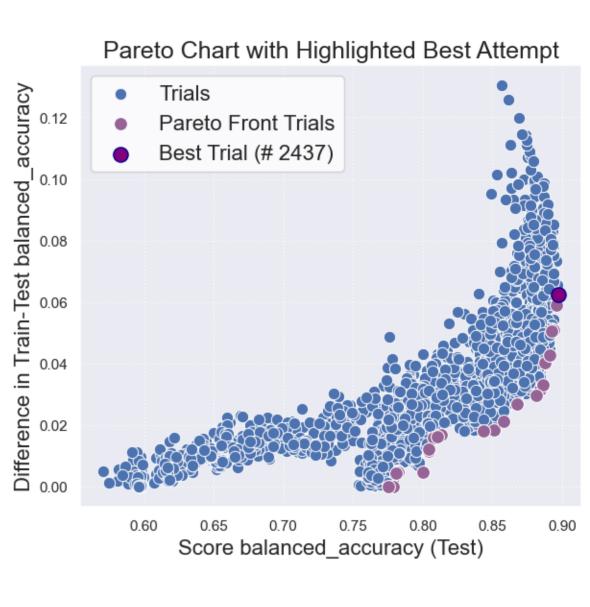
Protein: Q8WW43 - metric: balanced_accuracy





XGBoost Parameters	Value
alpha	4.2434
colsample_bytree	0.6024
gamma	0.4384
lambda	3.4211
learning_rate	0.0216
max_depth	8.0
min_child_weight	4.0
n_estimators	984.0
reg_alpha	0.6764
reg_lambda	0.6729
scale_pos_weight	1.3815
subsample	0.8151

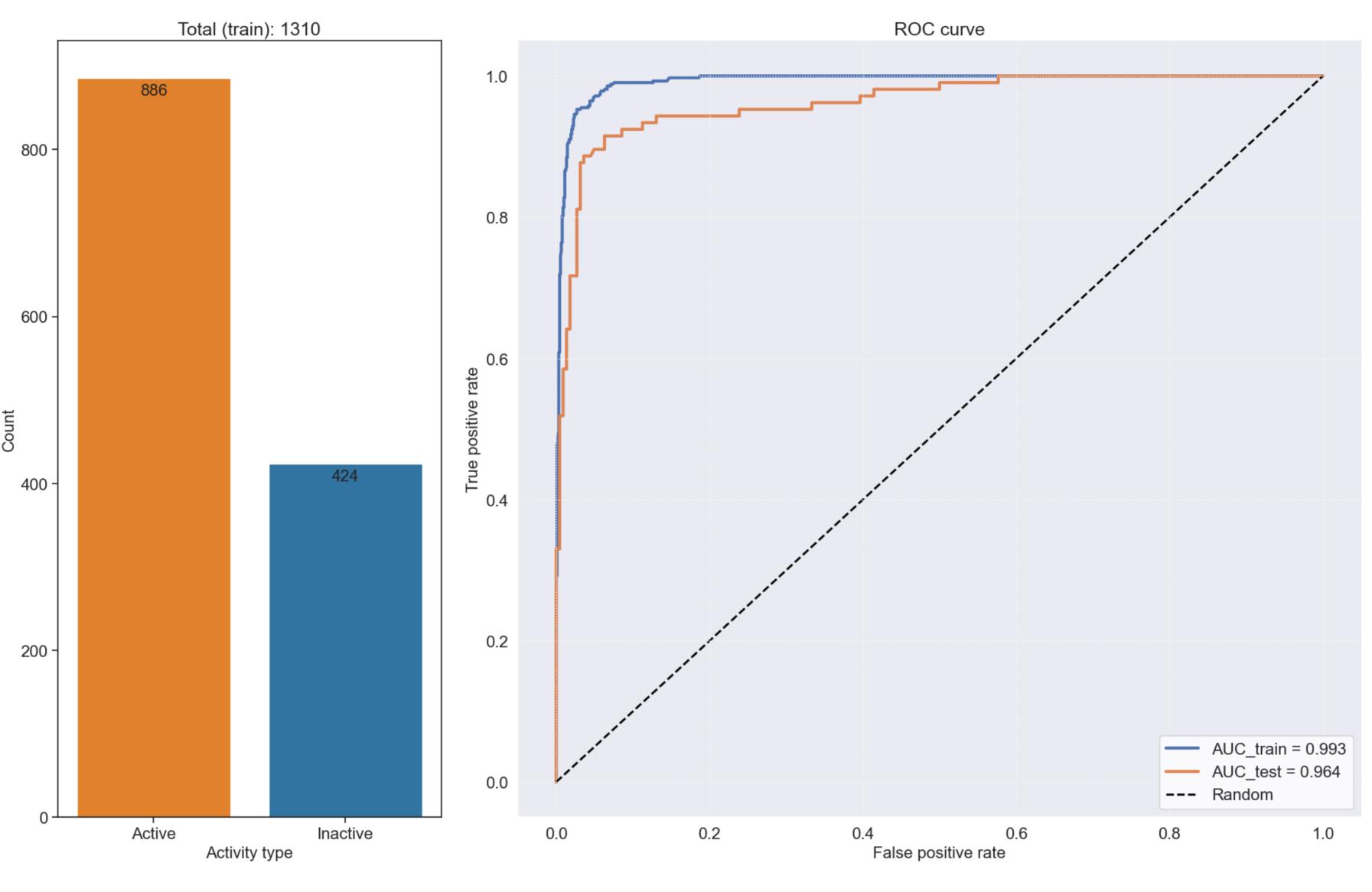
	Confucio	on Matrix	_				Calibration curve)		
label O	212	10	- 200 - 175 - 150 - 125	1.0						A Company of the Comp
True label	11	95	- 100 - 75 - 50 - 25	0.8 20.6						
	0 Predicte Confusion M		7	0.6 Dositives 0.4						
True label	0.95	0.045	- 0.8 - 0.6	0.2						
2 1-	0.1	0.9	- 0.4 - 0.2	0.0					train: 0 test: 0.	
	0 Predict	1 ed label		0.0	0.:	2	0.4 Mean predicted probab	0.6 pility	0.8	1.0

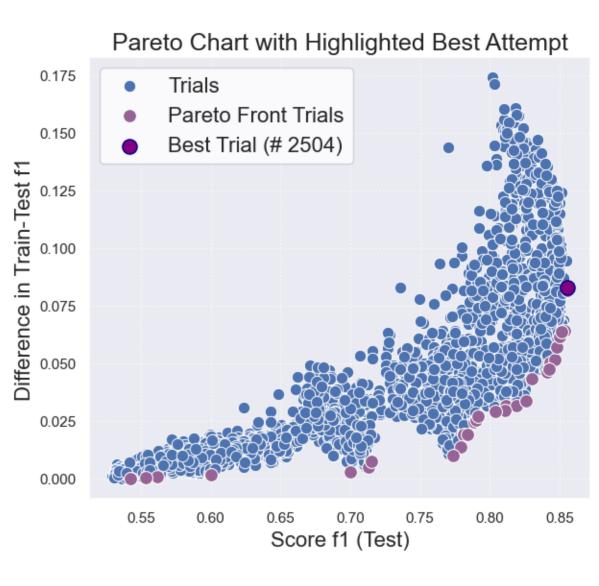
			Active/Inact	ive probability		
40 35 		Inactive zone			Active zone	type active (106) inactive (222) decoy (328)
30						
Percent 20						
⊕ 20 -····· 15 -····						
10	<u></u>					
5 						
	0.0	0.2	0.4 Prol	0.6 pability	0.8	1.0

Metric	Train	Test
AUC	0.9919	0.9637
accuracy	0.9611	0.936
balanced accuracy	0.9589	0.9256
recall_(sens)	0.9528	0.8962
specificity	0.965	0.955
precision	0.9287	0.9048
f1_score	0.9612	0.9359

Protein: Q8WW43 - metric: f1

Calibration curve





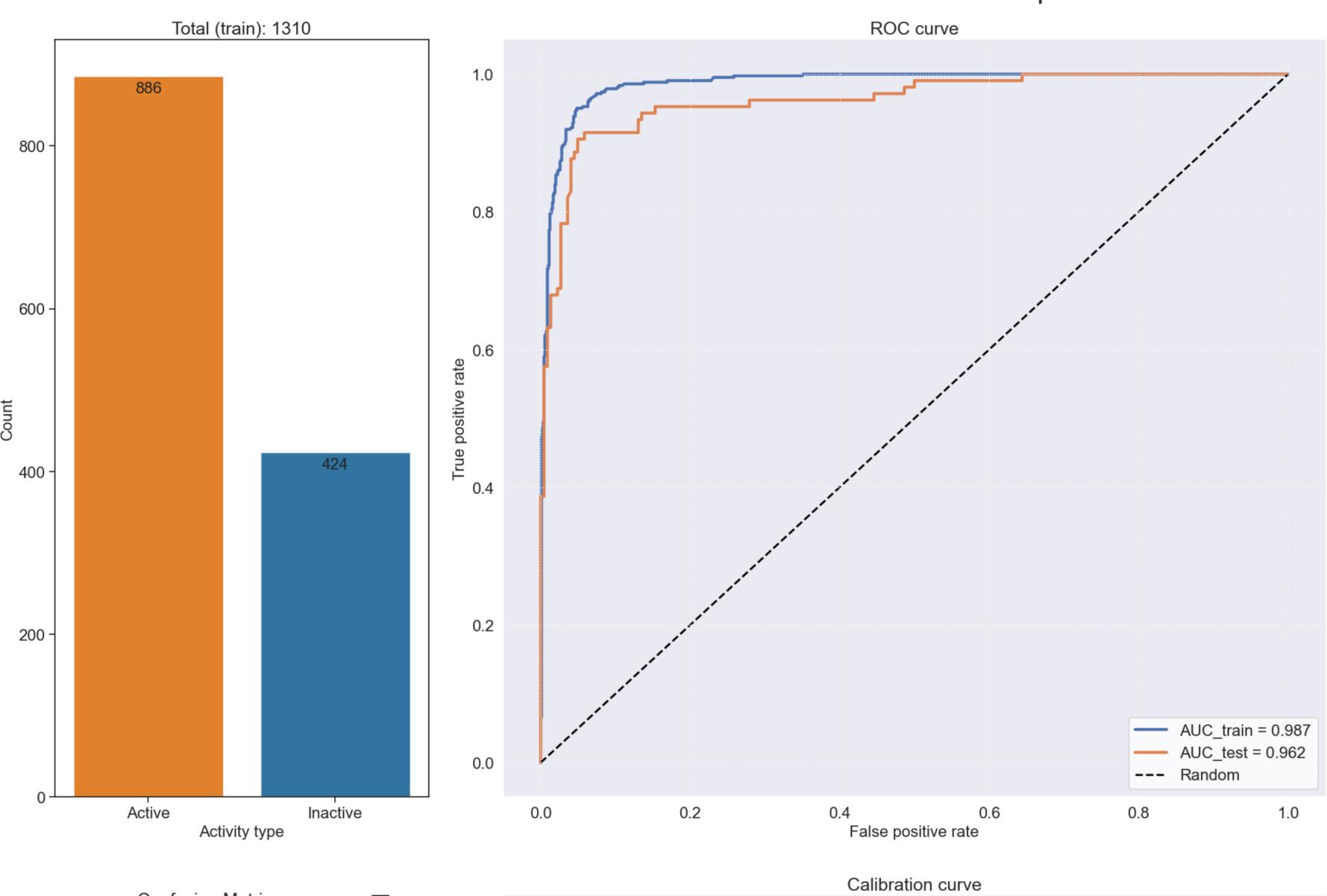
XGBoost Parameters	Value
alpha	3.073
colsample_bytree	0.7226
gamma	0.7225
lambda	0.9853
learning_rate	0.0168
max_depth	9.0
min_child_weight	3.0
n_estimators	709.0
reg_alpha	0.7171
reg_lambda	0.0202
scale_pos_weight	2.0469
subsample	0.6586

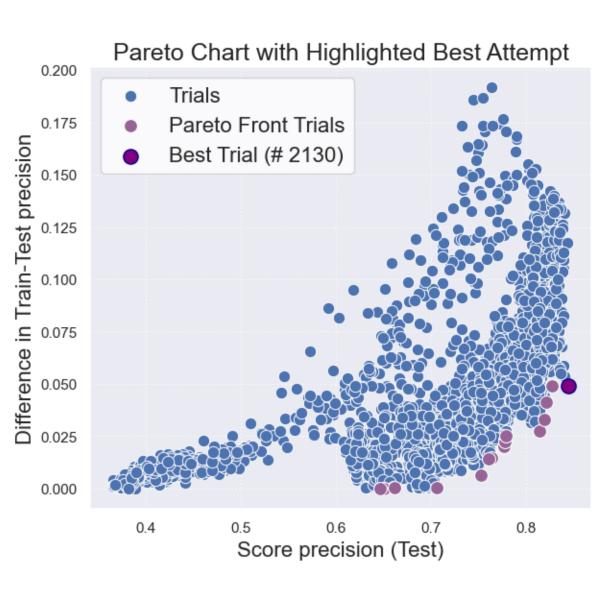
	Confusio	on Matrix	_				Calibration cu	rve		
	Contusio	on watrix	- 200							
<u> </u>	208	14	- 175 - 150 - 125	1.0						Jan .
label										
True	11	95	- 100 - 75 - 50	0.8						
L	0 Predicte	1 ed label	- 25	positives 9.0			,,,			
	Confusion M	latrix (nomr.)	- 0.9	Fraction of p						
- 0 -	0.94	0.063	- 0.8 - 0.7 - 0.6							
True label	0.1	0.9	- 0.5 - 0.4 - 0.3 - 0.2 - 0.1	0.2					train: 0.03 test: 0.05	
	0 Predicte	1 ed label			0.0	0.2	0.4 Mean predicted pro	0.6 bability	0.8	1.0

				Act	ive/Inactiv	e probab	ility		
35		Ina	ctive zone					Active zone	type active (106) inactive (222) decoy (328)
25									
- 20 20 21	-								
10									
5				T =					
	0.0	0	.2	0.4	4 Proba		.6	0.8	1.0

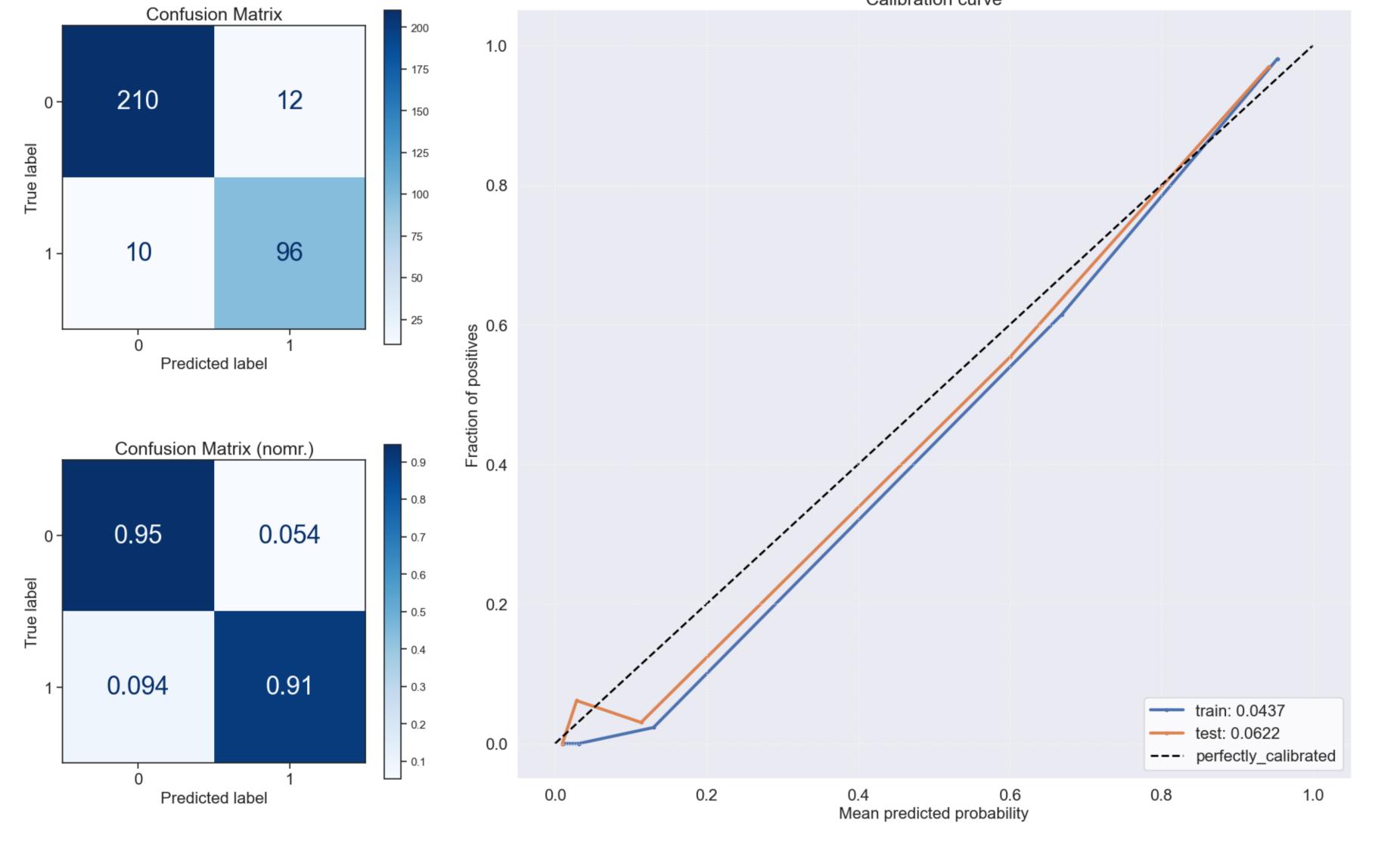
Metric	Train	Test
AUC	0.9926	0.9635
accuracy	0.9573	0.9238
balanced accuracy	0.9604	0.9166
recall_(sens)	0.9693	0.8962
specificity	0.9515	0.9369
precision	0.9053	0.8716
f1_score	0.9576	0.9241

Protein: Q8WW43 - metric: precision





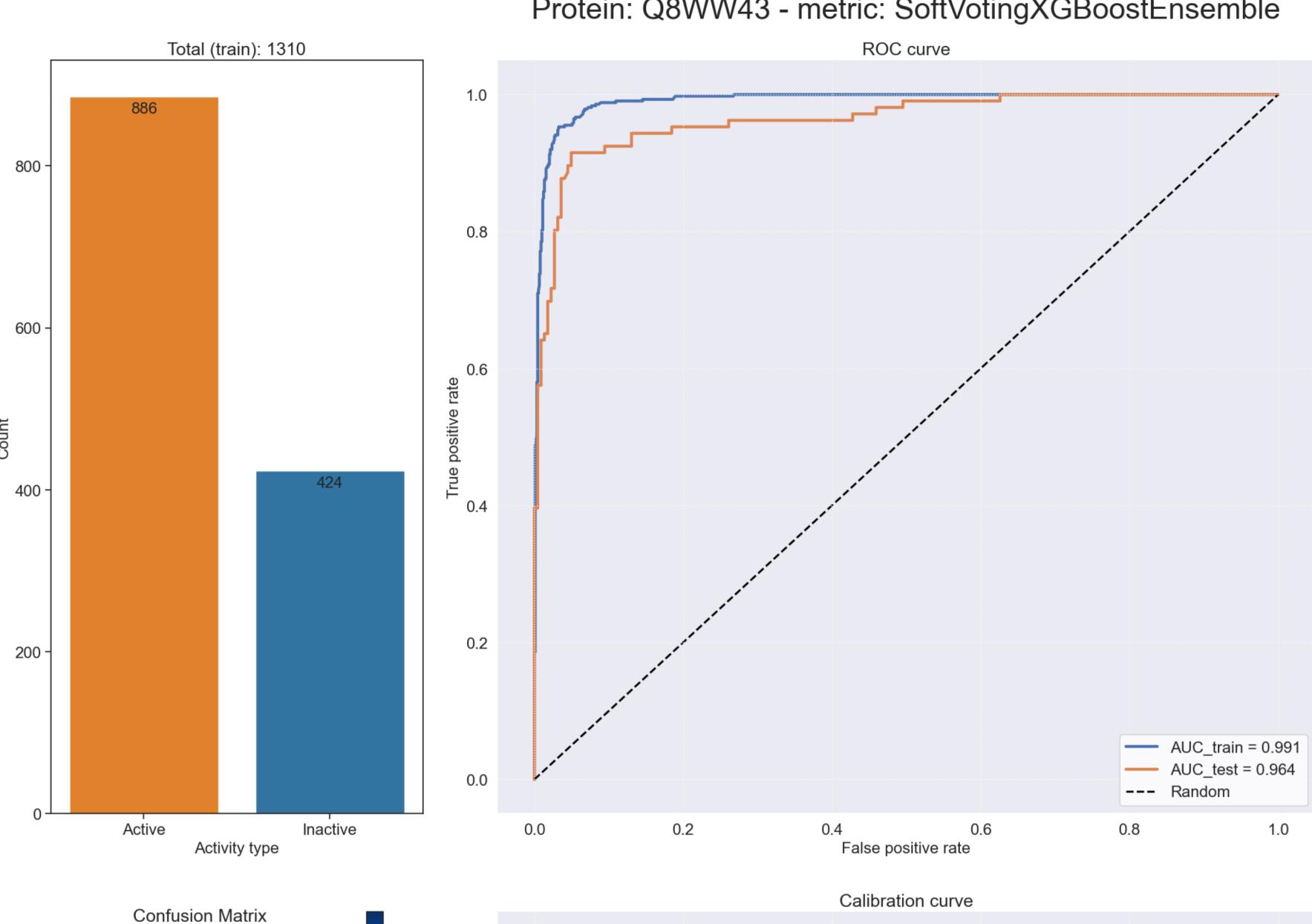
XGBoost Parameters	Value
alpha	2.4557
colsample_bytree	0.5156
gamma	0.0698
lambda	1.6488
learning_rate	0.023
max_depth	9.0
min_child_weight	3.0
n_estimators	367.0
reg_alpha	0.8701
reg_lambda	0.3158
scale_pos_weight	1.9598
subsample	0.5131

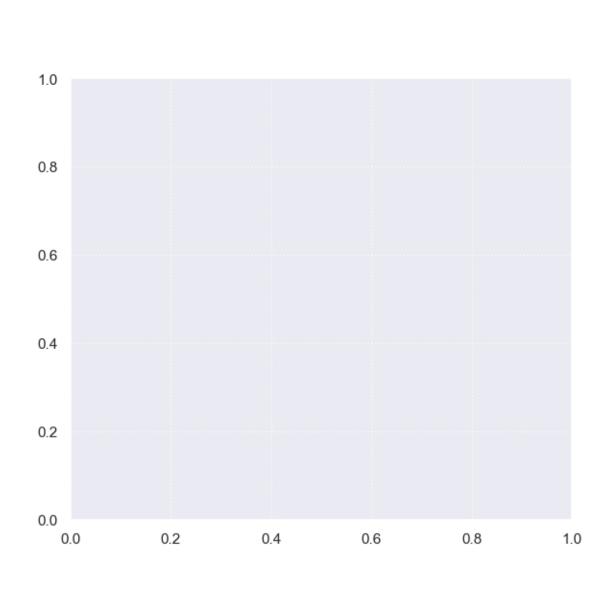


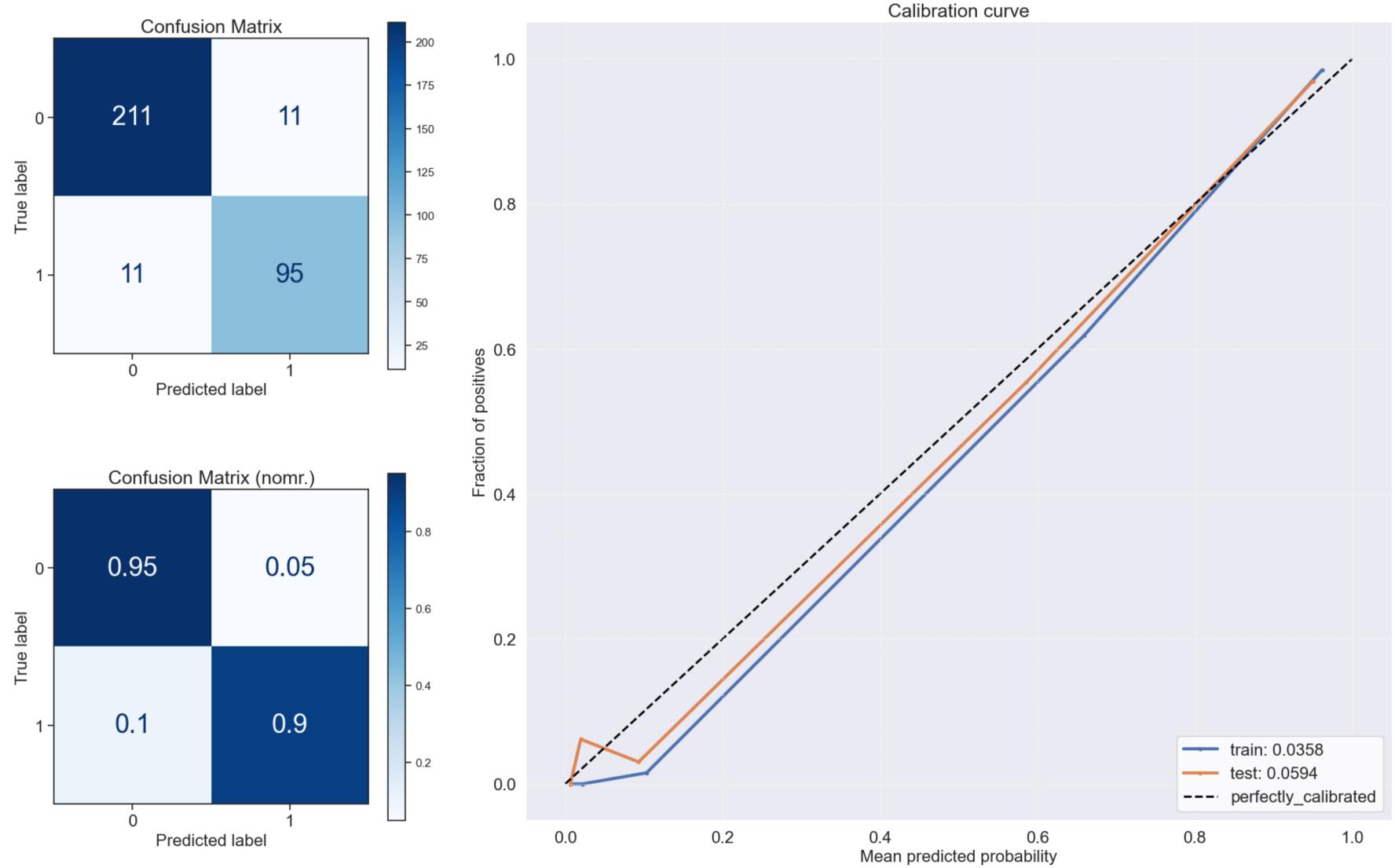
	:	:	Active/Ina	active probab	pility	:	:
25		Inactive zone			Acti	ve zone	type active (106) inactive (222) decoy (328)
Percent 0 -							
5							
0 —	0.0	0.2	0.4 P	robability	0.6	0.8	1.0

Metric	Train	Test
AUC	0.9868	0.9623
accuracy	0.9466	0.9329
balanced accuracy	0.9482	0.9258
recall_(sens)	0.9528	0.9057
specificity	0.9436	0.9459
precision	0.8899	0.8889
f1_score	0.947	0.9331

Protein: Q8WW43 - metric: SoftVotingXGBoostEnsemble







35				e probability		
30	П	ctive zone			Active zone	type active (106) inactive (222) decoy (328)
25						
Dercent						
10						
5						<u> </u>
0 —						<u>'</u>
	0.0	.2 ().4 Proba	0.6 ability	0.8	1.0

Metric	Train	Test
AUC	0.9911	0.964
accuracy	0.9542	0.9329
balanced accuracy	0.9545	0.9233
recall_(sens)	0.9552	0.8962
specificity	0.9537	0.9505
precision	0.9081	0.8962
f1_score	0.9545	0.9329