

Library

```
# Ignore the warnings •••
Cuda is Ready? True
Cuda Version: 12.5.1
Cudnn Version: 9

Tensorflow Version: 2.18.0
Keras Version: 3.7.0
Torch Version: 2.5.0+cu124
Torch Cuda Version: 12.4
Torch Cudnn Version: 90600

There are 1 GPU(s) available.
We will use the GPU: NVIDIA GeForce RTX 4070 Laptop GPU
```

Hyperparameters

```
# Data •••
```

Data Merge

```
# Data Loading •••
2022 (9968, 437)
2021 (9905, 454)
2020 (9776, 434)
2019 (9973, 511)
2018 (11747, 600)
```

```
[00:00, ?it/s]
[00:00, 4.29it/s]
[00:01, 2.50it/s]
4it [00:02, 1.80it/s]
```

100% | 5/5 [00:00<00:00, 59.24it/s]
Final Shape of df: (51369, 368)

Preprocessing

```
X_train, X_test, X_colname, \ •••
```

```
Preprocessing for Data Prepare...
Initial Shape: (51369, 368)
```

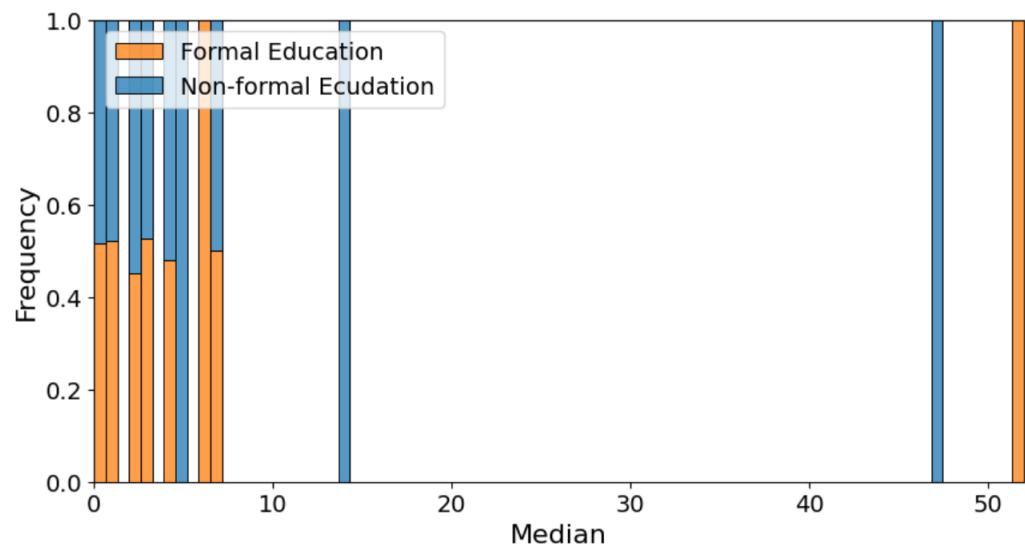
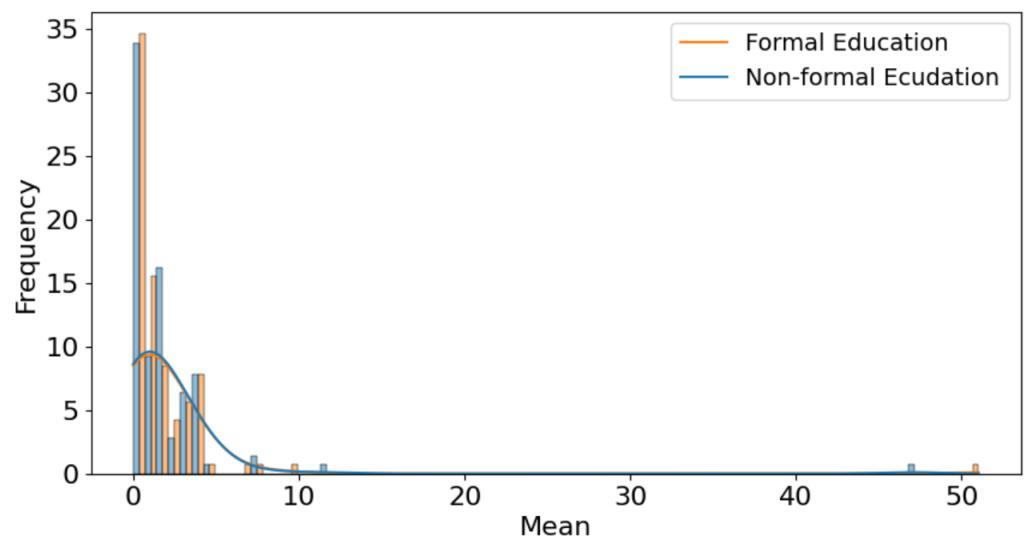
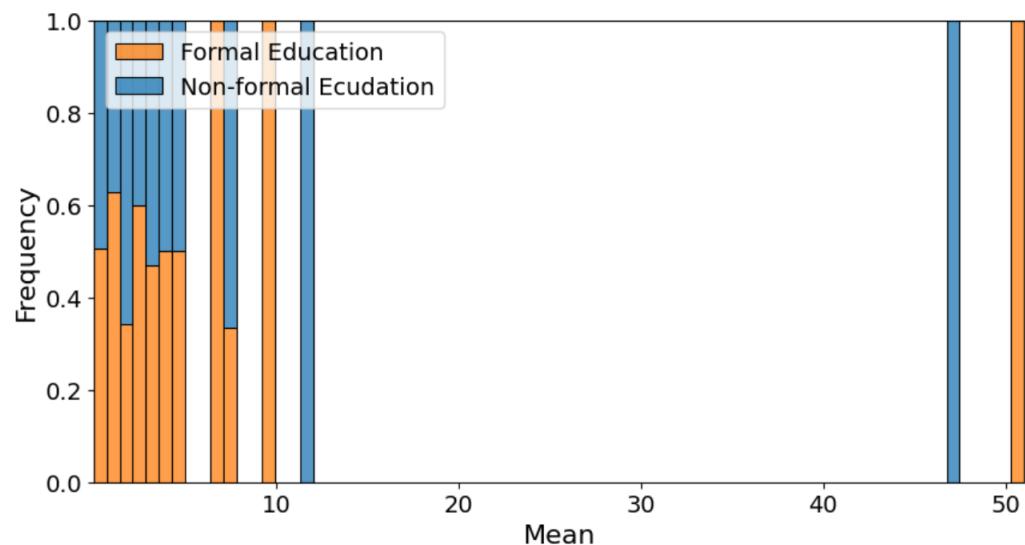
```
Deleting Non-meaning Variables...
Shape: (51369, 368)
```

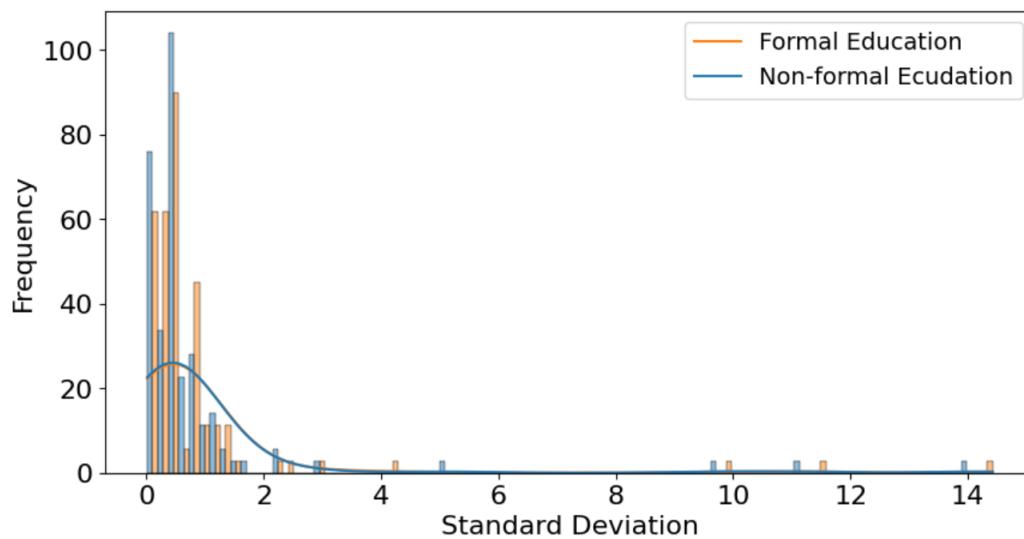
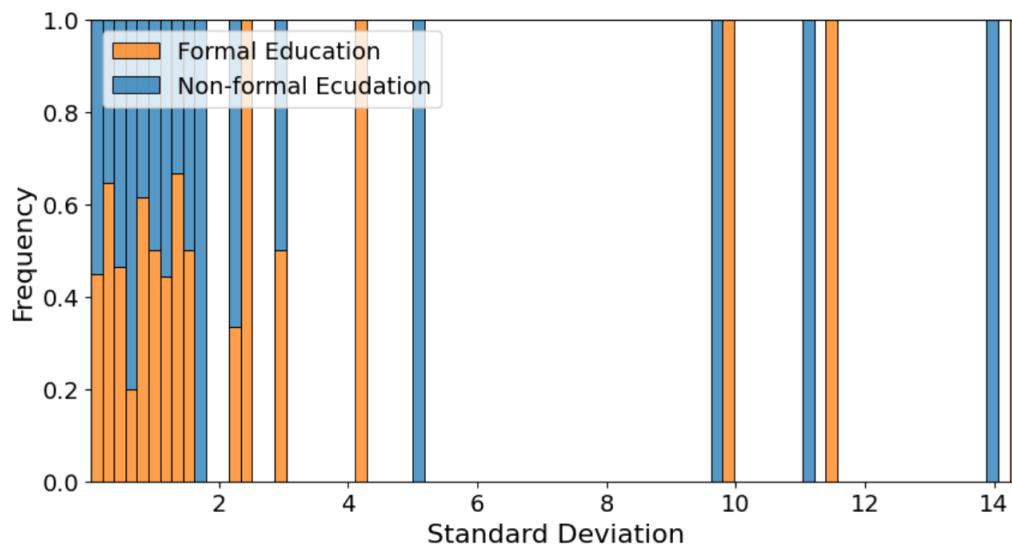
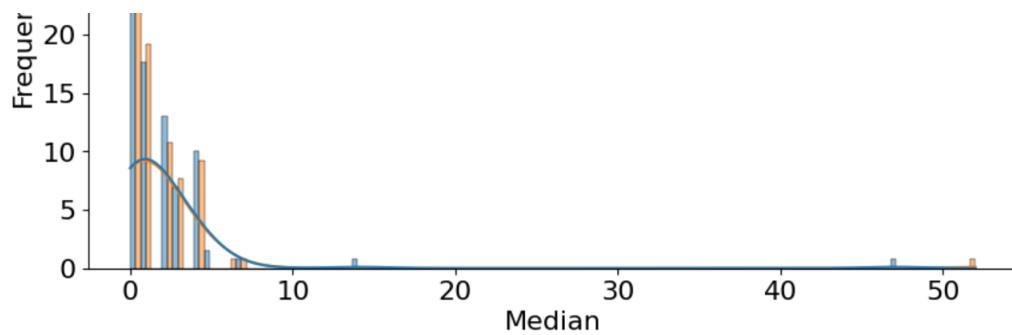
```
Preprocessing NAN...
Column Deleting...
```

```
These 79.07608695652173 %( 291 ) columns are containing more than 50.0 % null values!: ['문A2-1) 프로그램 유형_(1)학위취득을 위한 교육_1', '문A2-6) 연간참여시간_(1)학위취득을 위한 교육_1', '문A2-7) 자기부담학습비_(1)학위취득을 위한 교육_1', '문A1-2) 참여한 프로그램 수는 몇개입니까?_(2) 성인기초 및 문자해독교육_(문해교육)', '문A1-2) 참여한 프로그램 수는 몇개입니까?_(3) 직업능력향상교육', '문A1-2) 참여한 프로그램 수는 몇개입니까?_(4) 인문교양교육', '문A1-2) 참여한 프로그램 수는 몇개입니까?_(5) 문화예술스포츠교육', '문A1-2) 참여한 프로그램 수는 몇개입니까?_(6) 시민참여교육', '문A1-2) 참여한 프로그램 수는 몇개입니까?_(4) 인문교양교육', '문A2-5) 참여목적_(1)학위취득을 위한 교육_1', '문A2-1) 프로그램 유형_(2)성인기초 및 문자해독교육_1', '문A2-2) 프로그램 형태_(2)성인기초 및 문자해독교육_1', '문A2-3) 기관 유형_(2)성인기초 및 문자해독교육_1', '문A2-5) 참여목적_(2)성인기초 및 문자해독교육_1', '문A2-6) 연간참여시간_(2)성인기초 및 문자해독교육_1', '문A2-7) 자기부담학습비_(2)성인기초 및 문자해독교육_1', '문A2-8) 학습비 외부 지원여부_(2)성인기초 및 문자해독교육_1', '문A2-11) 프로그램만족도_(2)성인기초 및 문자해독교육_1', '문A2-11) 프로그램 불만족 요인(중복응답)__(2)성인기초 및 문자해독교육_1-1', '문A2-11) 프로그램 불만족 요인(중복응답)__(2)성인기초 및 문자해독교육_1-2', '문A2-11) 프로그램 불만족 요인(중복응답)__(2)성인기초 및 문자해독교육_1-3', '문A2-11) 프로그램 불만족 요인(중복응답)__(2)성인기초 및 문자해독교육_1-5', '문A2-11) 프로그램 불만족 요인(중복응답)__(2)성인기초 및 문자해독교육_1-7', '문A2-11) 프로그램 불만족 요인(중복응답)__(2)성인기초 및 문자해독교육_1-9', '문A2-1) 프로그램 유형_(3)직업능력향상교육_1', '문A2-2) 프로그램 형태_(3)직업능력향상교육_1', '문A2-3) 기관 유형_(3)직업능력향상교육_1', '문A2-6) 연간참여시간_(3)직업능력향상교육_1', '문A2-7) 자기부담학습비_(3)직업능력향상교육_1', '문A2-9) 학습비 외부 지원기관_(3)직업능력향상교육_1', '문A2-5) 참여목적_(3)직업능력향상교육_1', '문A2-10) 프로그램만족도_(3)직업능력향상교육_1', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_1-1', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_1-2', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_1-3', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_1-5', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_1-7', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_1-9', '문A2-1) 프로그램 유형_(3)직업능력향상교육_2', '문A2-2) 프로그램 형태_(3)직업능력향상교육_2', '문A2-5) 참여목적_(3)직업능력향상교육_2', '문A2-7) 자기부담학습비_(3)직업능력향상교육_2', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_2-1', '문A2-9) 학습비 외부 지원기관_(3)직업능력향상교육_2', '문A2-1) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_2-1', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_2-2', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_2-3', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_2-5', '문A2-8) 학습비 외부 지원여부_(3)직업능력향상교육_1', '문A2-1) 프로그램 유형_(3)직업능력향상교육_3', '문A2-2) 프로그램 형태_(3)직업능력향상교육_3', '문A2-5) 참여목적_(3)직업능력향상교육_3', '문A2-6) 연간참여시간_(3)직업능력향상교육_3', '문A2-7) 자기부담학습비_(3)직업능력향상교육_3', '문A2-9) 학습비 외부 지원기관_(3)직업능력향상교육_3', '문A2-10) 프로그램만족도_(3)직업능력향상교육_3-1', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_3-1', '문A2-10) 프로그램만족도_(3)직업능력향상교육_3-2', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_3-4', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_3-6', '문A2-11) 프로그램 불만족 요인(중복응답)__(3)직업능력향상교육_3-7', '문A2-6) 연간참여시간_(4)인문교양교육_1', '문A2-8) 학습비 외부 지원여부_(4)인문교양교육_1', '문A2-9) 학습비 외부 지원기관_(4)인문교양교육_1', '문A2-3) 기관 유형_(3)직업능력향상교육_3', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-1', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-2', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-3', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-5', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-7', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-9', '문A2-1) 프로그램 유형_(4)인문교양교육_1', '문A2-2) 프로그램 형태_(4)인문교양교육_1', '문A2-5) 참여목적_(4)인문교양교육_1', '문A2-7) 자기부담학습비_(4)인문교양교육_1', '문A2-9) 학습비 외부 지원기관_(4)인문교양교육_1', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-1', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-2', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-3', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-5', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-7', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_1-9', '문A2-1) 프로그램 유형_(4)인문교양교육_2', '문A2-2) 프로그램 형태_(4)인문교양교육_2', '문A2-5) 참여목적_(4)인문교양교육_2', '문A2-7) 자기부담학습비_(4)인문교양교육_2', '문A2-9) 학습비 외부 지원기관_(4)인문교양교육_2', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_2-1', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_2-2', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_2-3', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_2-5', '문A2-8) 학습비 외부 지원여부_(4)인문교양교육_2', '문A2-1) 프로그램 유형_(4)인문교양교육_3', '문A2-2) 프로그램 형태_(4)인문교양교육_3', '문A2-5) 참여목적_(4)인문교양교육_3', '문A2-6) 연간참여시간_(4)인문교양교육_3', '문A2-7) 자기부담학습비_(4)인문교양교육_3', '문A2-9) 학습비 외부 지원기관_(4)인문교양교육_3', '문A2-10) 프로그램만족도_(4)인문교양교육_3-1', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_3-1', '문A2-10) 프로그램만족도_(4)인문교양교육_3-2', '문A2-11) 프로그램 불만족 요인(중복응답)__(4)인문교양교육_3-3', '문A2-8) 학습비 외부 지원여부_(4)인문교양교육_2', '문A2-10) 프로그램만족도_(4)인문교양교육_2', '문A2-11)
```



```
Object X Features: []
Complete X Object Processing!
(41095, 114) (41095, 1)
(10274, 114) (10274, 1)
Data Split!
```





Number of significant features: 99 (0.868421052631579%)

		Formal Education	Non-formal Education	Statistics
문D1) 귀하께서는 작년에 평생교육 프로그램 관련 정보를 접한 적이 있습니까?	1.0000	8544 (26.04%)	24269 (73.96%)	$\chi^2=3388.53, p=0.00$
	2.0000	9573 (51.59%)	8983 (48.41%)	
문C1) 앞으로 참여하길 희망하는 프로그램은 무엇입니까? (중복응답)1	-1.0000	4599 (39.15%)	7148 (60.85%)	$\chi^2=2959.99, p=0.00$
	0.0000	2944 (20.49%)	11423 (79.51%)	
	1.0000	3 (37.50%)	5 (62.50%)	
...				
DQ10. 귀하는 현재 어떤 직업에 종사하고 계십니까?_11.0	0.0000	18108 (35.27%)	33235 (64.73%)	$\chi^2=0.00, p=1.00$
직업_11.0	0.0000	18108 (35.27%)	33235 (64.73%)	$\chi^2=0.00, p=1.00$
문C1) 앞으로 참여하길 희망하는 프로그램은 무엇입니까? (5)문화예술스포츠교육	1.0000	77 (35.16%)	142 (64.84%)	
DQ10. 귀하는 현재 어떤 직업에 종사하고 계십니까?_11.0	1.0000	9 (34.62%)	17 (65.38%)	
직업_11.0	1.0000	9 (34.62%)	17 (65.38%)	

522 rows × 3 columns

Comparison Statistics of X by Y class!

Ratio of Origin Y: 비형식교육 참여자

```
1           26629
0           14466
Name: count, dtype: int64
Ratio of Sapling Y: 비형식교육 참여자
0           27589
1           26629
Name: count, dtype: int64
```

Preprocessing of Scaling...

```
(54218, 114) (54218, 1) 0.0 1.0000000000000002
(10274, 114) (10274, 1) 0.0 1.0000000000000002
Complete!
```

Modeling

Machine Learing

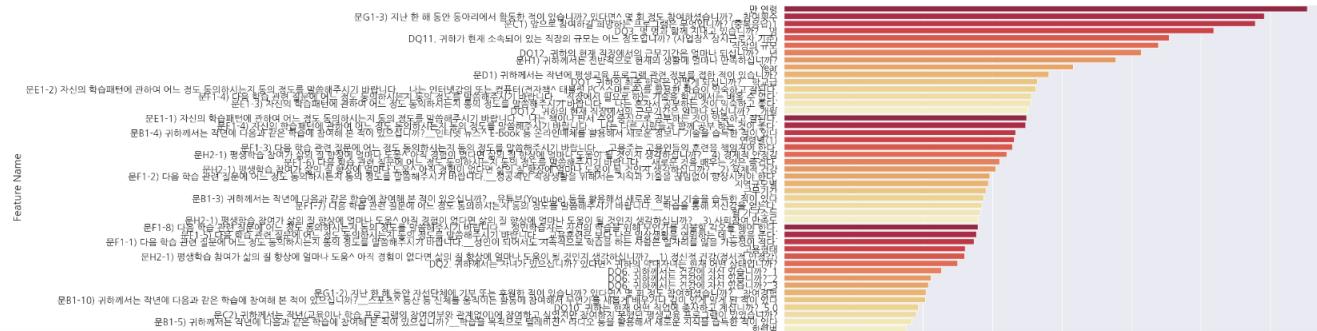
```
model_lr, Score_te_lr, Score_trte_lr = modeling_LogisticRegression(X_train, Y_train, X_test, Y_test, ***)
```

Explanations of Y:

	Features	Coef	Odds		Features	Coef	Odds
0	DQ6. 귀하께서는 건강에 자신 있습니까?_1	7.3674	1,583.4837		취업구분	-5.0130	0.0067
1	DQ6. 귀하께서는 건강에 자신 있습니까?_2	7.3625	1,575.7984	DQ8. 귀하께서는 현재 수입을 목적으로 일하고 계십니까?	-5.0130	0.0067	
2	DQ6. 귀하께서는 건강에 자신 있습니까?_4	5.5827	265.7927	문G1-3) 지난 한 해 동안 동아리에서 활동한 적이 있습니까? 있다면^ 몇 회 정도 참여하셨습니까?__참여횟수	-2.9082	0.0546	
3	DQ6. 귀하께서는 건강에 자신 있습니까?_3	5.3122	202.8015		건강에 대한 자신감	-2.5255	0.0800
4	DQ9. 귀하의 현재 고용형태는 어떻게 되십니까?_1.0	4.1259	61.9214		직장의 규모	-1.9258	0.1458
5	DQ9. 귀하의 현재 고용형태는 어떻게 되십니까?_2.0	3.6723	39.3406	DQ1. 귀하의 최종 학력은 어떻게 되십니까?__학교급	-1.3697	0.2542	
6	DQ7. 귀하의 주된 소득원천은 무엇입니까?_4	3.3260	27.8281	문H1) 귀하께서는 전반적으로 현재의 생활에 얼마나 만족하십니까?	-1.3038	0.2715	

```
model_rf, Score_te_rf, Score_trte_rf = modeling_RandomForestClassifier(X_train, Y_train, ***)
```

Explanations of Y:



```
model_xgb, Score_te_xgb, Score_trte_xgb = modeling_XGBClassifier(X_train, Y_train, ***)
```

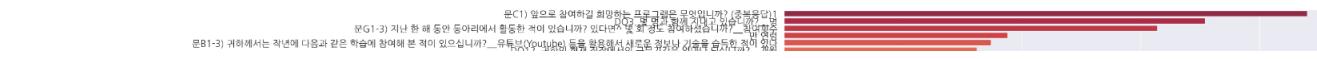
Explanations of Y:



```
model_lgbm, Score_te_lgbm, Score_trte_lgbm = modeling_LGBMClassifier(X_train, Y_train, ***)
```

```
[LightGBM] [Info] Number of positive: 26629, number of negative: 27589
[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.004130 seconds.
You can set `force_row_wise=true` to remove the overhead.
And if memory is not enough, you can set `force_col_wise=true`.
[LightGBM] [Info] Total Bins 3386
[LightGBM] [Info] Number of data points in the train set: 54218, number of used features: 111
[LightGBM] [Info] [binary:BoostFromScore]: pavg=0.500000 -> initscore=0.000000
[LightGBM] [Info] Start training from score 0.000000
```

Explanations of Y:



문B1-10) 귀하께서는 직년에 다음과 같은 학습에 참여해 보셨어 있습니까? DQ11) 직장에서 일하면서 동시에 학습을 하거나 학습과 별개로 학습시간을 갖고 다른 학습을 하는 경우는 어떤 경우인가요? DQ12) 직장에서 일하면서 동시에 학습을 하거나 학습과 별개로 학습시간을 갖고 다른 학습을 하는 경우는 어떤 경우인가요?

문H2-1) 평생학습 참여가 삶의 질 향상에 얼마나 도움이 아직 경험이 없다면 삶의 질 향상에 얼마나 도움이 있는지 혹은 참여한 사람에게 드는 도움은 어떤 것인가요? DQ10) 평생학습 참여가 삶의 질 향상에 드는 도움은 어떤 것인가요? DQ11) 평생학습 참여가 삶의 질 향상에 드는 도움은 어떤 것인가요? DQ12) 개인의 현재 환경에서의 문화기간은 얼마나 되십니까?

```
model_catb, Score_te_catb, Score_trte_catb = modeling_CatBoostClassifier(X_train, Y_train, ...)
```

```
Learning rate set to 0.46818
0: learn: 0.6291833      total: 12.7ms  remaining: 1.25s
1: learn: 0.5943113      total: 20.9ms  remaining: 1.02s
2: learn: 0.5598701      total: 28.2ms  remaining: 912ms
3: learn: 0.5119235      total: 33.1ms  remaining: 795ms
4: learn: 0.5041867      total: 38ms    remaining: 722ms
5: learn: 0.5000818      total: 43ms    remaining: 674ms
6: learn: 0.4912305      total: 47.5ms  remaining: 631ms
7: learn: 0.4844208      total: 52.3ms  remaining: 601ms
8: learn: 0.4816652      total: 56.3ms  remaining: 569ms
9: learn: 0.4692627      total: 60.3ms  remaining: 543ms
10: learn: 0.4667274     total: 64.2ms  remaining: 520ms
11: learn: 0.4627682     total: 67.9ms  remaining: 498ms
12: learn: 0.4612322     total: 72.2ms  remaining: 483ms
13: learn: 0.4504749     total: 76.2ms  remaining: 468ms
14: learn: 0.4486805     total: 80.7ms  remaining: 457ms
15: learn: 0.4466609     total: 85.1ms  remaining: 447ms
16: learn: 0.4453188     total: 89.6ms  remaining: 438ms
```

Deep Learning

MLP

```
# reshape ...
Y_train_reshape: (54218, 1) --> (54218, 2)
Y_test_reshape: (10274, 1) --> (10274, 2)
```

Model: "functional_7"

Layer (type)	Output Shape	Param #
input_layer_7 (InputLayer)	(None, 114)	0
dense_40 (Dense)	(None, 128)	14,720
dropout_42 (Dropout)	(None, 128)	0
dense_41 (Dense)	(None, 256)	33,024
total params	(None, 256)	0

CNN

```
# Reshape ...
Y_train_reshape: (54218, 1) --> (54218, 2)
Y_test_reshape: (10274, 1) --> (10274, 2)
X_train_reshape: (54218, 114) --> (54218, 114, 1)
X_test_reshape: (10274, 114) --> (10274, 114, 1)
```

Model: "functional_8"

Layer (type)	Output Shape	Param #
input_layer_8 (InputLayer)	(None, 114, 1)	0
conv1d_9 (Conv1D)	(None, 114, 128)	768
max_pooling1d_9 (MaxPooling1D)	(None, 57, 128)	0

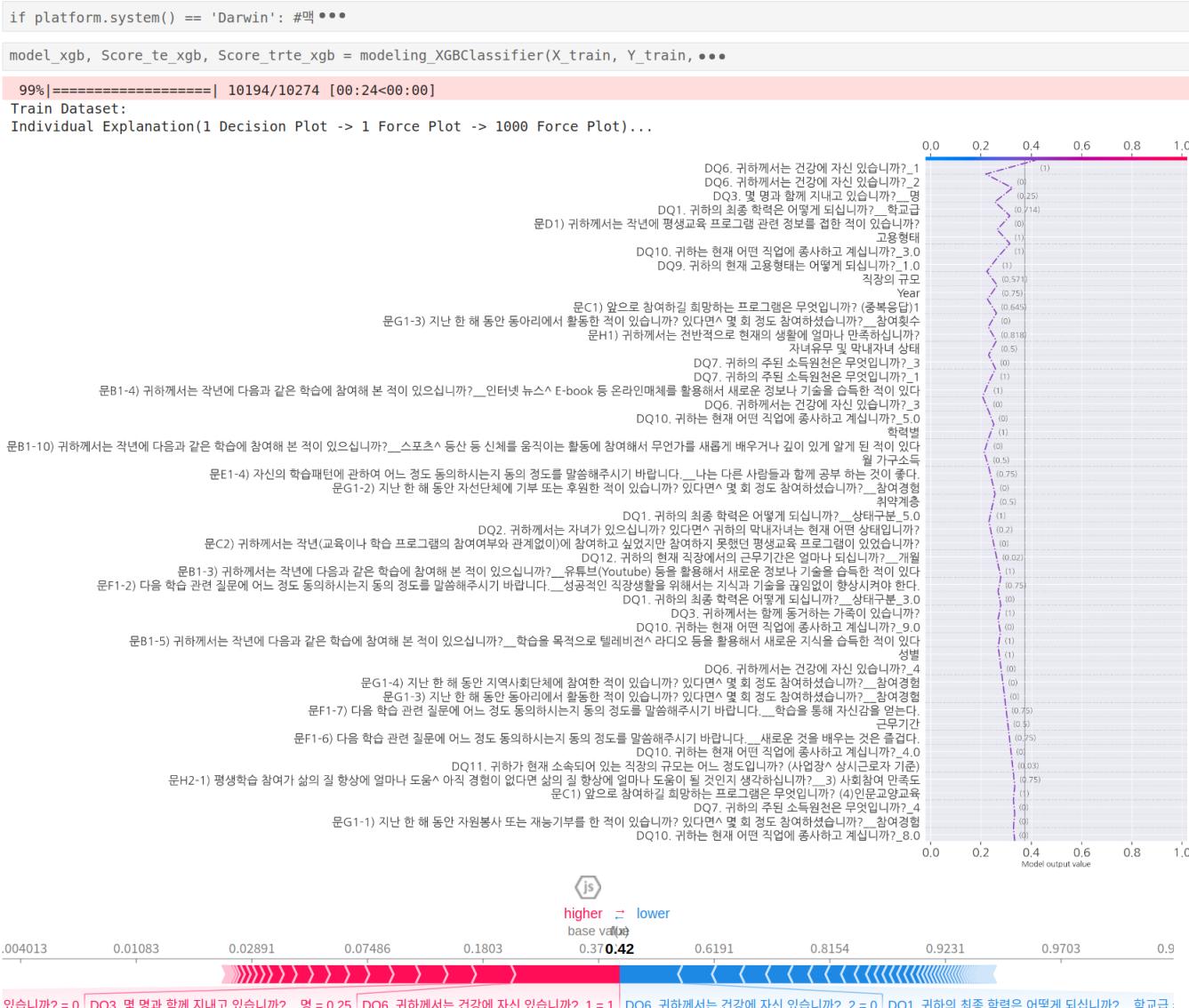
Summary

```
folder_location = os.path.join(os.getcwd(), 'Result') ...
```

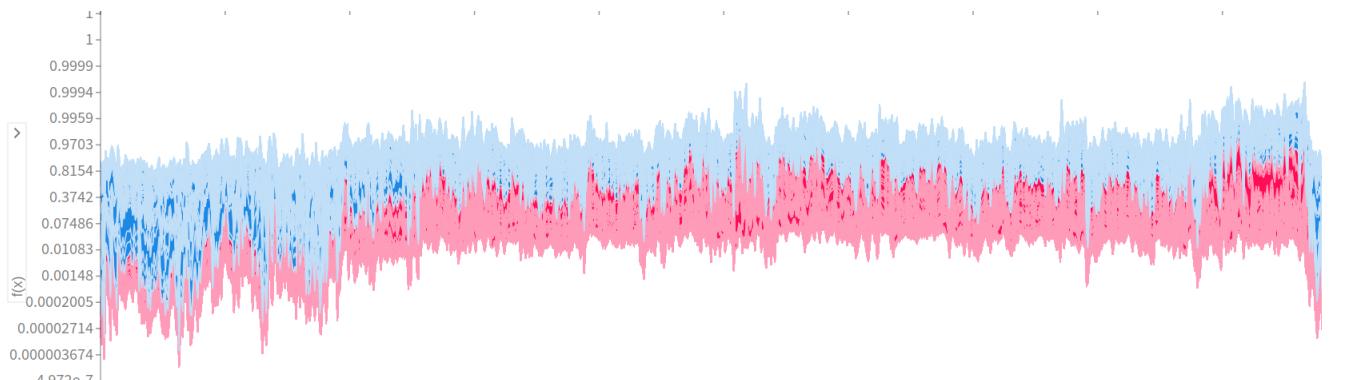
	Dataset	Algorithm	N	True Positive	True Negative	False Positive	False Negative	Precision	Recall	Specificity	F1-score	Accuracy	Balanced Accuracy	AUC
0	Test set	Logistic Regression	10274	5442	1878	1773	1181	0.754262	0.821682	0.514380	0.786530	0.712478	0.668031	0.750015
1	Test set	Random Forest	10274	5493	2027	1624	1130	0.771814	0.829382	0.555190	0.799563	0.731945	0.692286	0.784497
2	Test set	XGBoost	10274	5627	2001	1650	996	0.773258	0.849615	0.548069	0.809640	0.742457	0.698842	0.792795
3	Test set	LGBM	10274	5608	1960	1691	1015	0.768324	0.846746	0.536839	0.805631	0.736617	0.691793	0.789195

4	Test set	CatBoost	10274	5663	1946	1705	960	0.768594	0.855051	0.533005	0.809520	0.740607	0.694028	0.784960
5	Test set	MLP	10274	6623	0	3651	0	0.644637	1.000000	0.000000	0.783926	0.644637	0.500000	0.544952
6	Test set	CNN	10274	5093	2214	1437	1530	0.779939	0.768987	0.606409	0.774424	0.711213	0.687698	0.761822
Dataset	Algorithm	N	True Positive	True Negative	False Positive	False Negative	Precision	Recall	Specificity	F1-score	Accuracy	Balanced Accuracy	AUC	
0	Entire population	Logistic Regression	64492	27331	20789	10451	5921	0.723387	0.821936	0.665461	0.769519	0.746139	0.743698	0.829246
1	Entire population	Random Forest	64492	32122	29616	1624	1130	0.951876	0.966017	0.948015	0.958894	0.957297	0.957016	0.993683
2	Entire population	XGBoost	64492	29904	24678	6562	3348	0.820052	0.899314	0.789949	0.857856	0.846338	0.844632	0.927989
3	Entire population	LGBM	64492	28777	22968	8272	4475	0.776728	0.865422	0.735211	0.818680	0.802348	0.800316	0.887876
4	Entire population	CatBoost	64492	29526	23619	7621	3726	0.794842	0.887947	0.756050	0.838819	0.824056	0.821998	0.906228
5	Entire population	MLP	64492	33252	0	31240	0	0.515599	1.000000	0.000000	0.680390	0.515599	0.500000	0.570802
6	Entire population	CNN	64492	31547	22310	8930	1705	0.779381	0.948725	0.714149	0.855756	0.835096	0.831437	0.868469

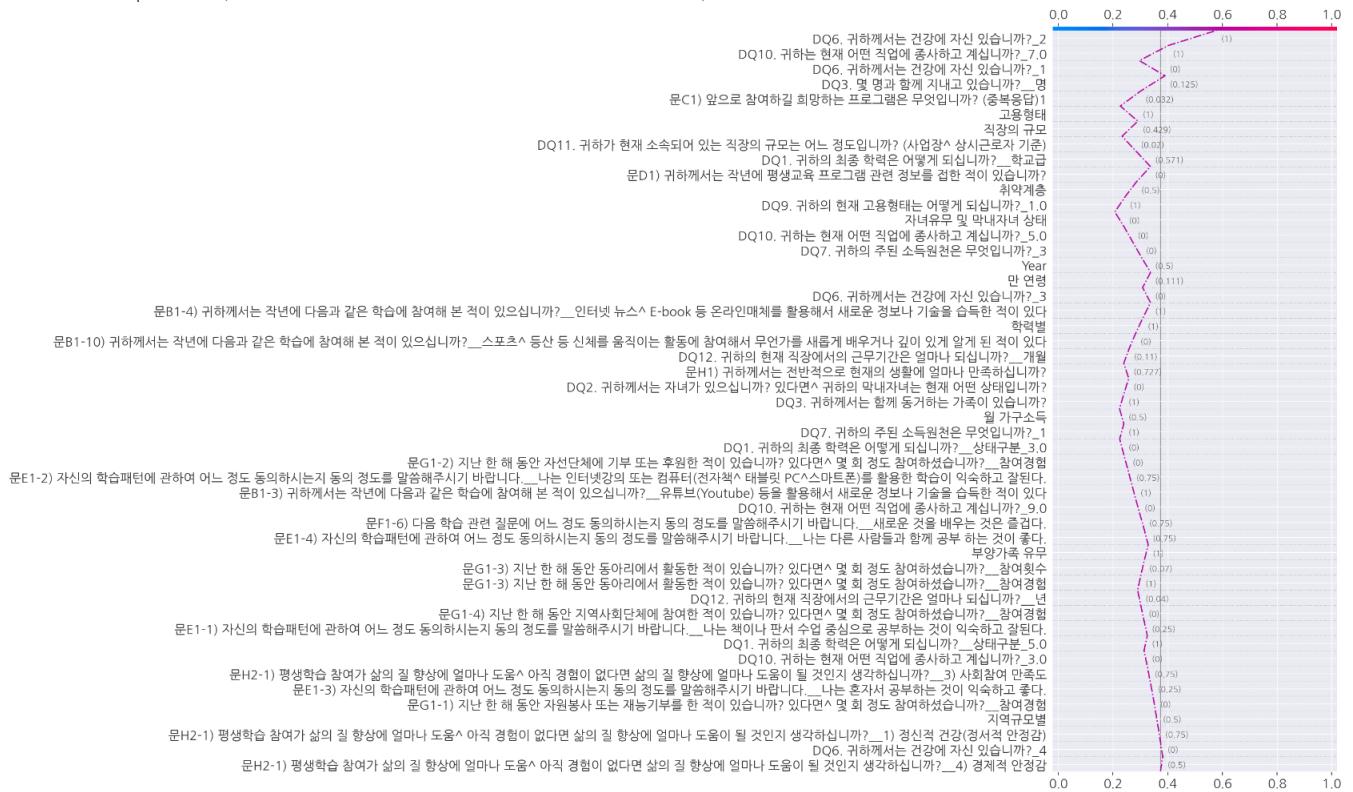
Explanation



있습니까? = 0 DQ3. 몇 명과 함께 지내고 있습니까? _명 = 0.25 DQ6. 귀하께서는 건강에 자신 있습니까? _1 = 1 DQ6. 귀하께서는 건강에 자신 있습니까? _2 = 0 DQ10. 귀하의 최종 학력은 어떻게 되십니까? _학교급



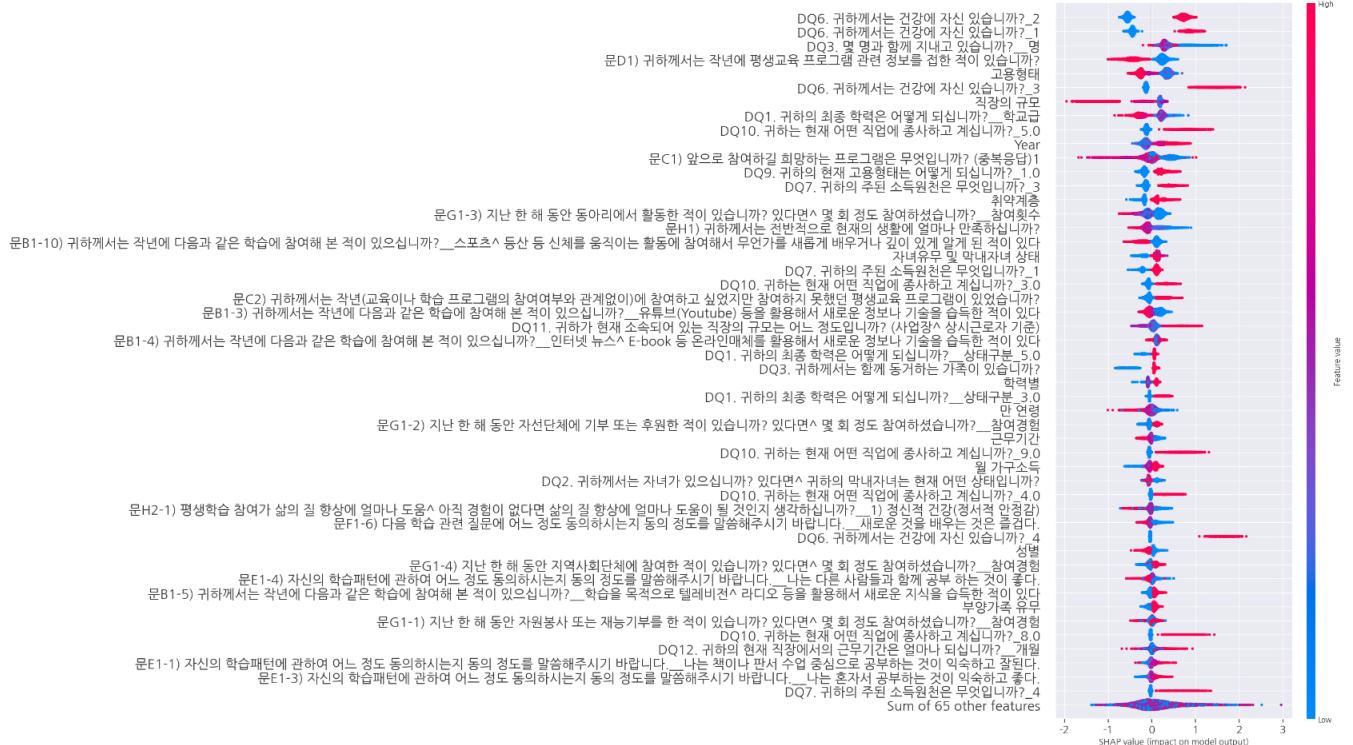
Test Dataset:
Individual Explanation(1 Decision Plot -> 1 Force Plot -> 1000 Force Plot)...



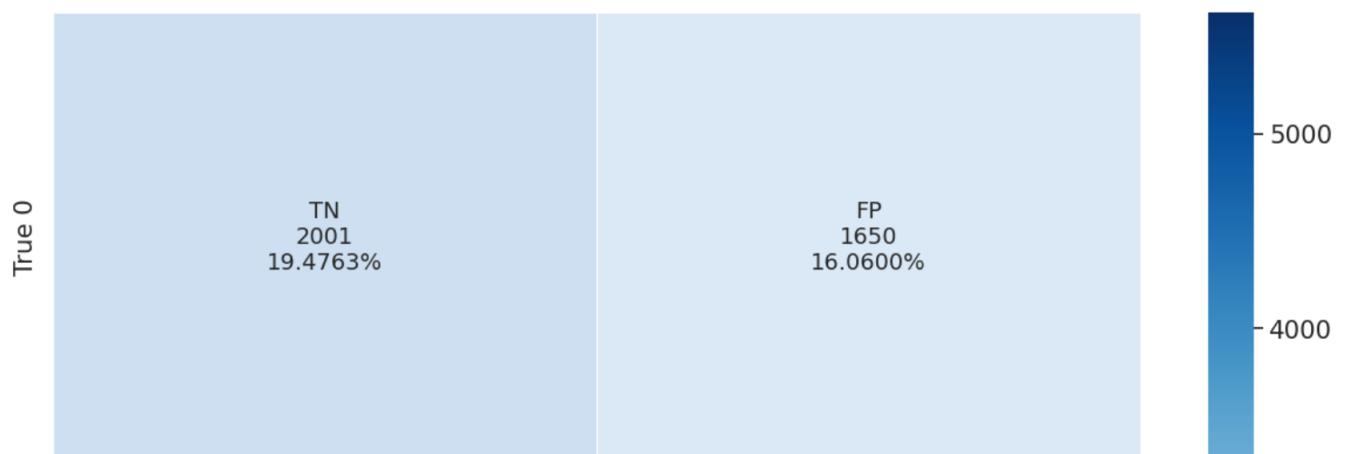


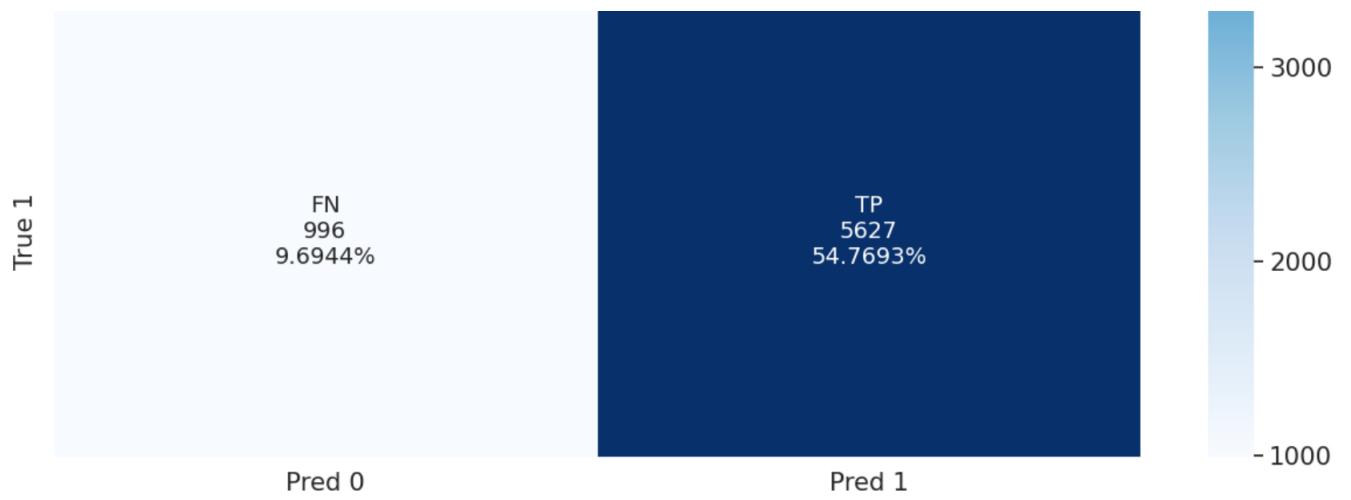
Test Dataset:

Total Explanation(Beeswarm Plot)... .

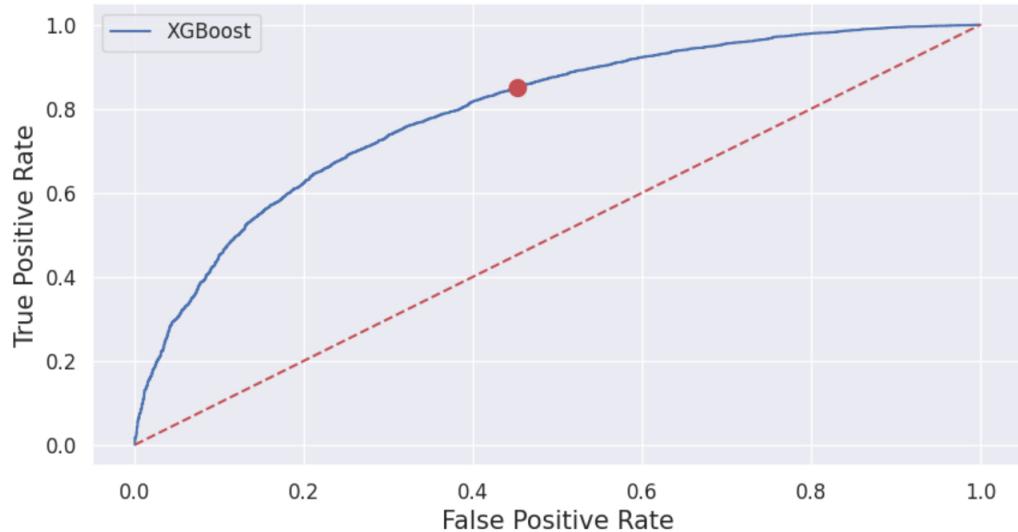


Performance:





AUC: 0.7927949432794665



```
XGBClassifier(base_score=None, booster=None, callbacks=None,
    colsample_bylevel=None, colsample_bynode=None,
    colsample_bytree=None, device=None, early_stopping_rounds=None,
    enable_categorical=False, eval_metric=None, feature_types=None,
    feature_weights=None, gamma=None, grow_policy=None,
    importance_type=None, interaction_constraints=None,
    learning_rate=None, max_bin=None, max_cat_threshold=None,
    max_cat_to_onehot=None, max_delta_step=None, max_depth=None,
    max_leaves=None, min_child_weight=None, missing=nan,
    monotone_constraints=None, multi_strategy=None, n_estimators=100,
    n_jobs=None, num_parallel_tree=None, ...)
```

	N	True Positive	True Negative	False Positive	False Negative	Precision	Recall	Specificity	F1-score	Accuracy	Balanced Accuracy	AUC
Test set	10274	5627	2001	1650	996	0.7733	0.8496	0.5481	0.8096	0.7425	0.6988	0.7928
	N	True Positive	True Negative	False Positive	False Negative	Precision	Recall	Specificity	F1-score	Accuracy	Balanced Accuracy	AUC
Entire population	64492	29904	24678	6562	3348	0.8201	0.8993	0.7899	0.8579	0.8463	0.8446	0.9280

```
# model_lgbm, Score_te_lgbm, Score_trte_lgbm = modeling_LGBMClassifier(X_train, Y_train, ...)
```

```
...
```

```
...
```

```
...
```