

32기 방학세미나

2팀

김동희
김민주
박채원
방건우
이동기
진재언

INDEX

1. 데이터 확인 및 EDA
2. 데이터 전처리
3. Feature Selection
4. 모델링
5. 최종 모델

1

데이터 확인 및 EDA

Train data 구조

```
1 train_data.info()
✓ 0.0s

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5455 entries, 0 to 5454
Data columns (total 96 columns):
#   Column  Non-Null Count  Dtype  
---  -
0   Target  5455 non-null    int64   
1   X1       3751 non-null    float64  
2   X2       5455 non-null    float64  
3   X3       5455 non-null    float64  
4   X4       5455 non-null    float64  
5   X5       5455 non-null    float64  
6   X6       5455 non-null    float64  
7   X7       1512 non-null    float64  
8   X8       5455 non-null    float64  
9   X9       5455 non-null    float64  
10  X10      5455 non-null    float64  
11  X11      5455 non-null    float64  
12  X12      5455 non-null    float64  
13  X13      5455 non-null    float64  
14  X14      5455 non-null    float64  
15  X15      5455 non-null    float64  
16  X16      5455 non-null    float64  
17  X17      5455 non-null    float64  
18  X18      5455 non-null    float64  
19  X19      3327 non-null    float64  
...
94  X94      5455 non-null    int64   
95  X95      5455 non-null    float64  
dtypes: float64(93), int64(3)
```



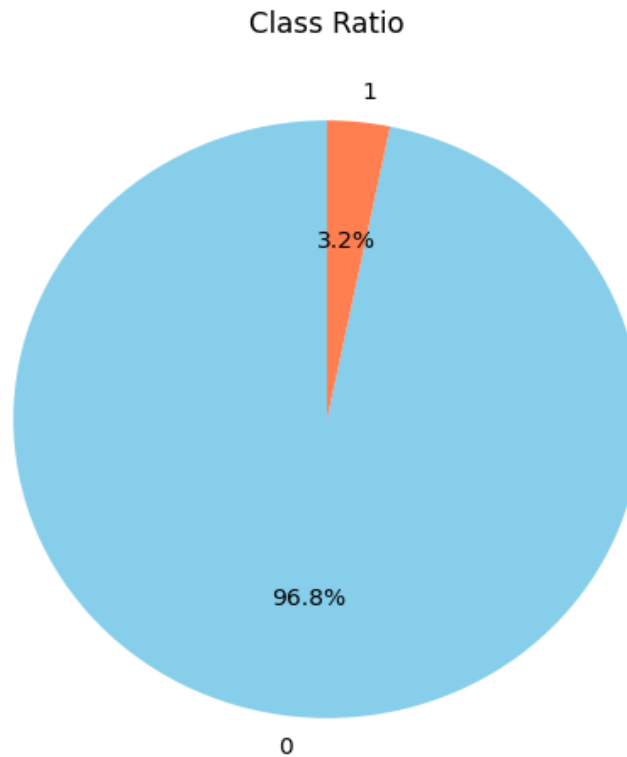
1. 'Target' 제외, **95개의 설명변수** 존재

2. 모든 변수가 **수치형**으로 구성

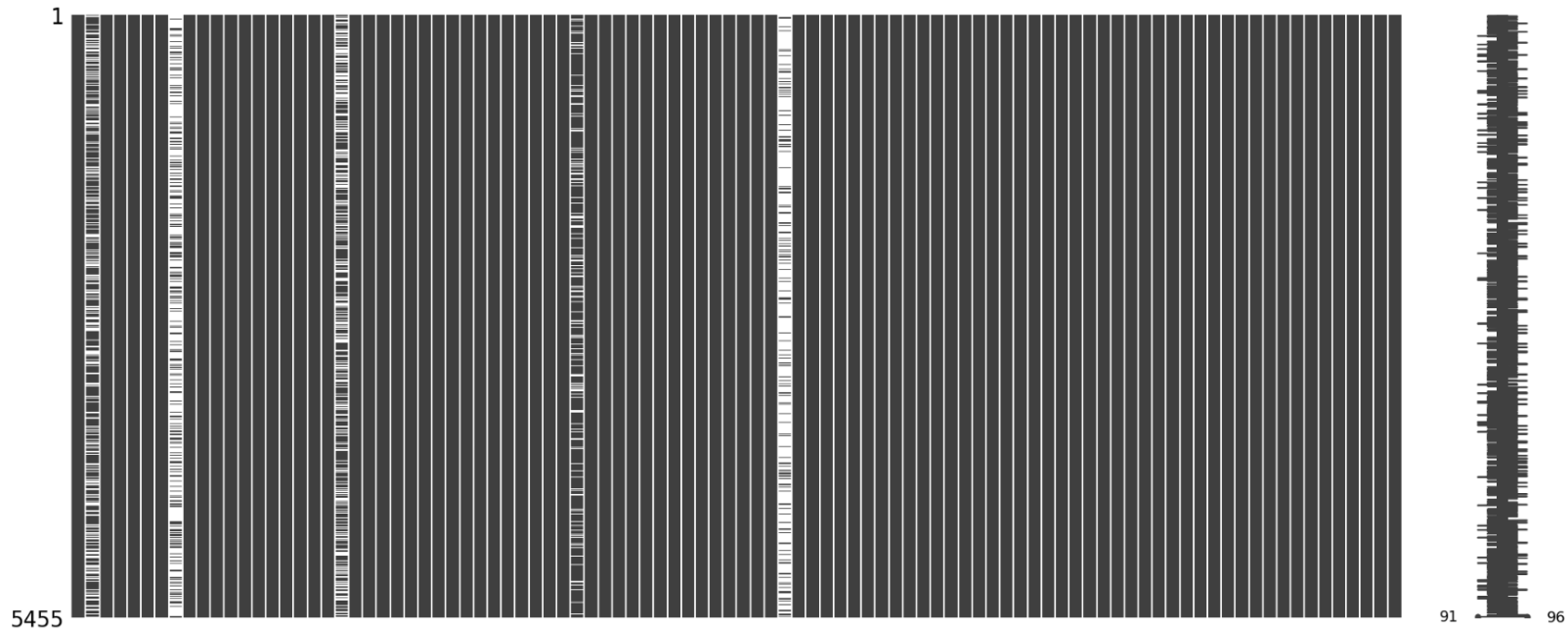
3. 모든 변수는 **masked data**

종속변수

Class0과 Class1의 비율이 97:3으로 **클래스 불균형**이 존재



결측치



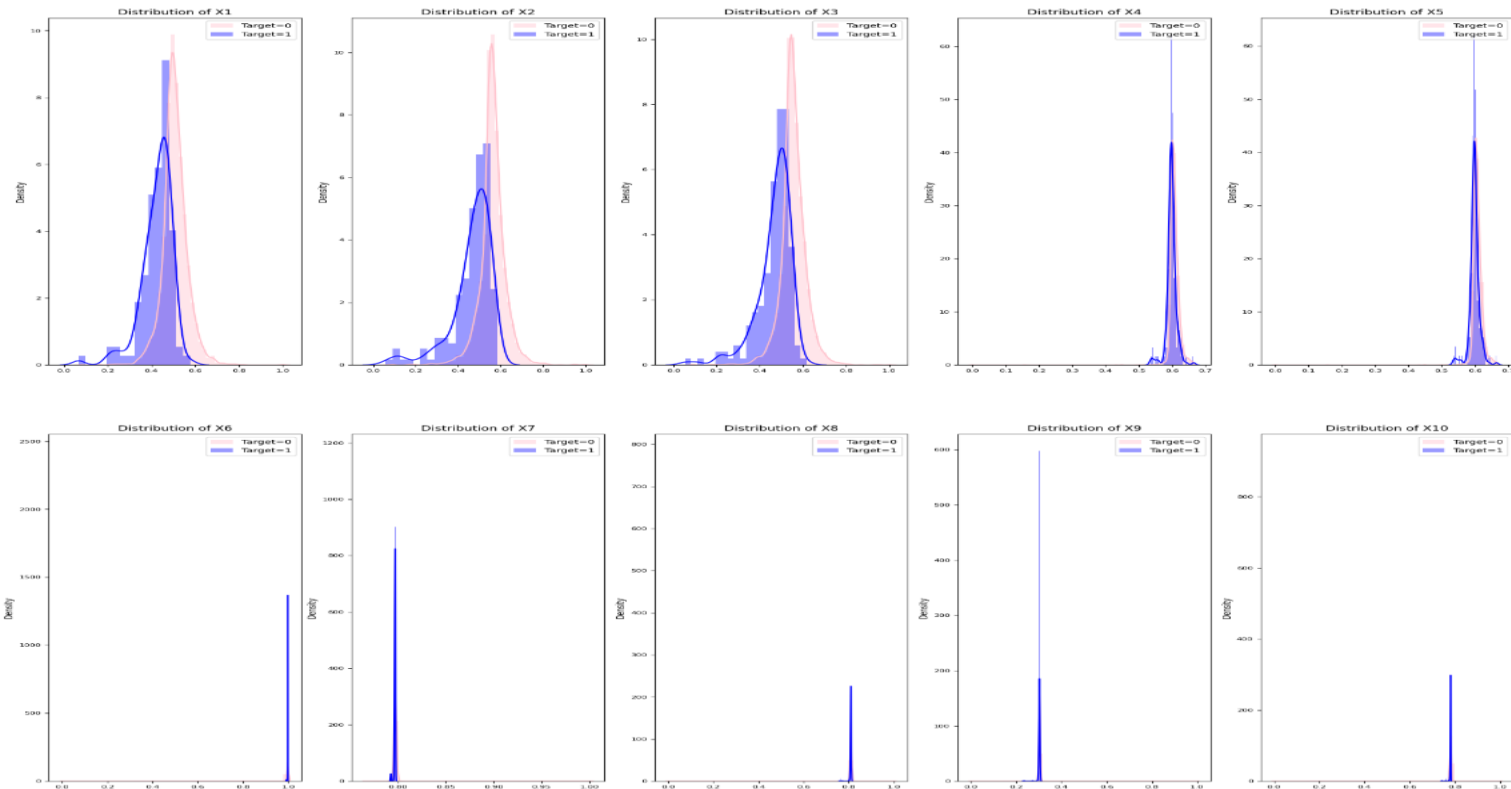
...

'X1', 'X7', 'X19', 'X36', 'X51' 5개의 변수에서 결측치가 존재



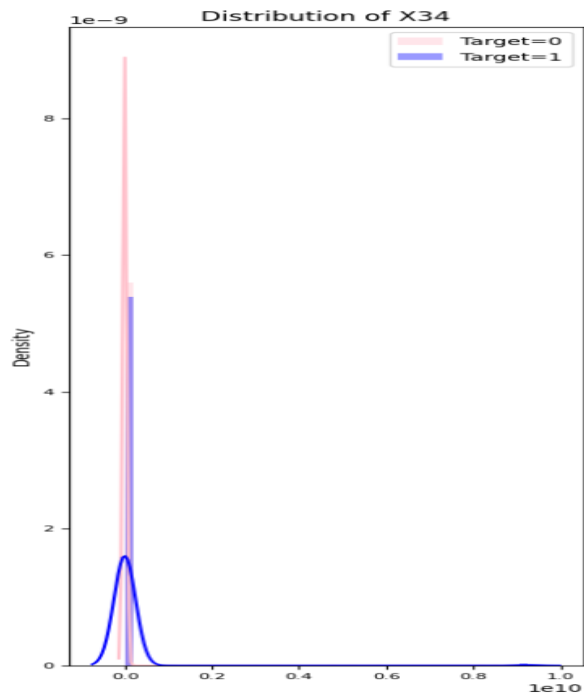
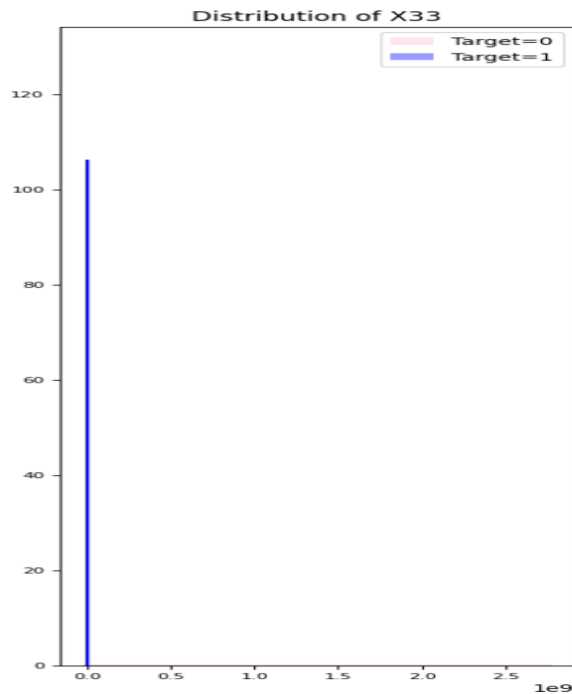
설명변수 분포

Train data의 설명변수 95개에 대해 분포 확인



이상치

이상치를 갖는 변수들이 여러 개 있음을 확인

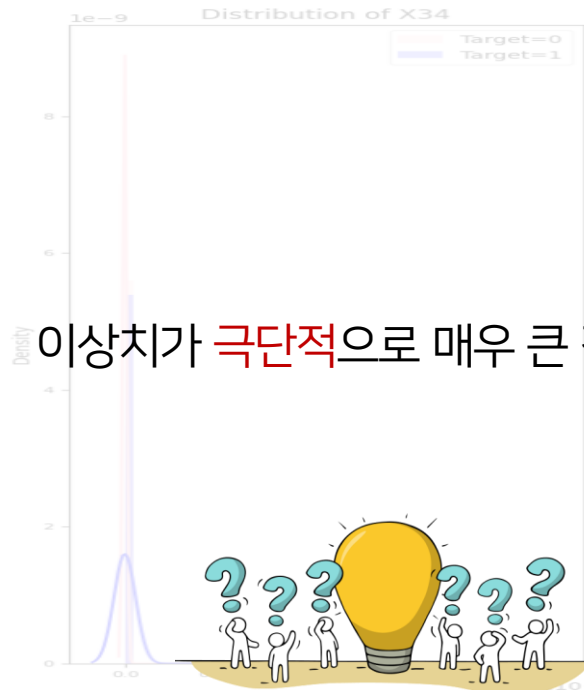
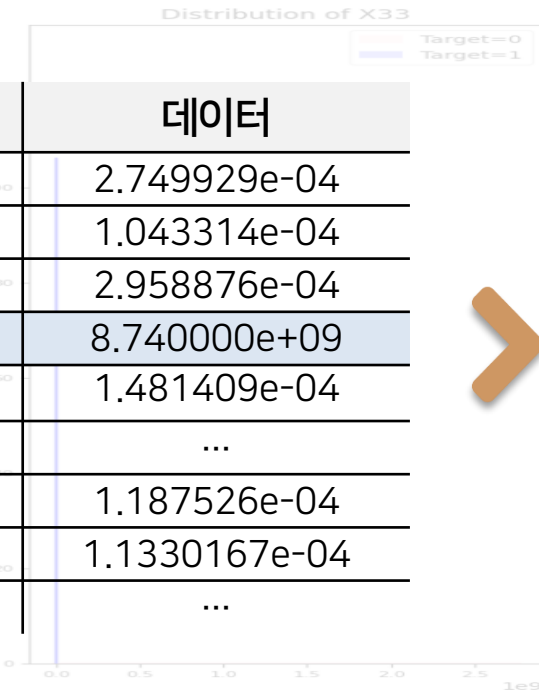


이상치

이상치를 갖는 변수들이 여러 개 있음을 확인

Ex) X71

index	데이터
0	2.749929e-04
1	1.043314e-04
2	2.958876e-04
3	8.740000e+09
4	1.481409e-04
...	...
24	1.187526e-04
25	1.1330167e-04
...	...

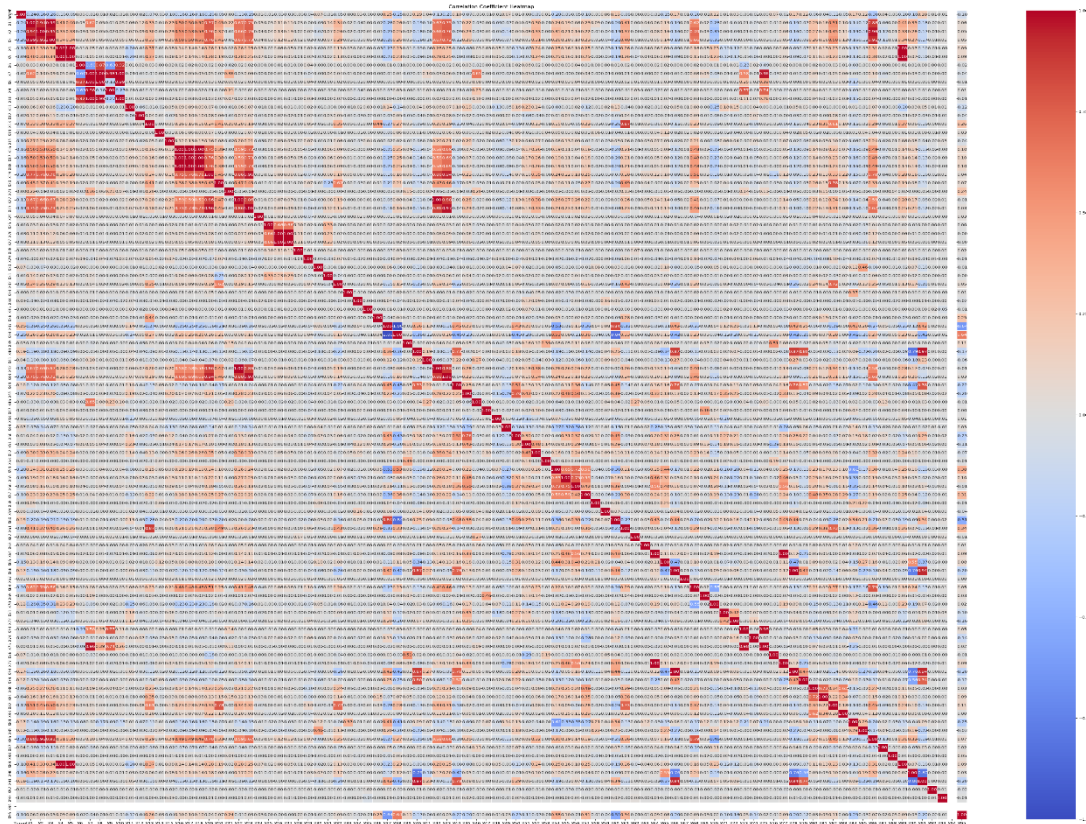


이상치가 극단적으로 매우 큰 값을 확인



설명변수 간 상관관계

상관계수가 높은 몇 개 변수들의 조합이 존재

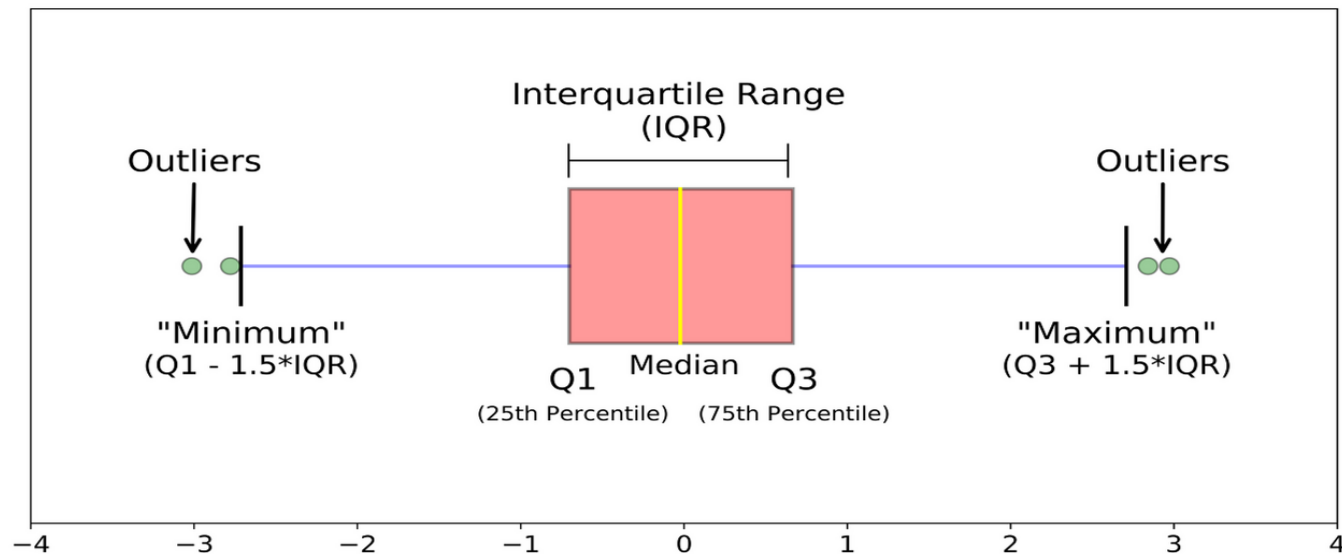


2

데이터 전처리

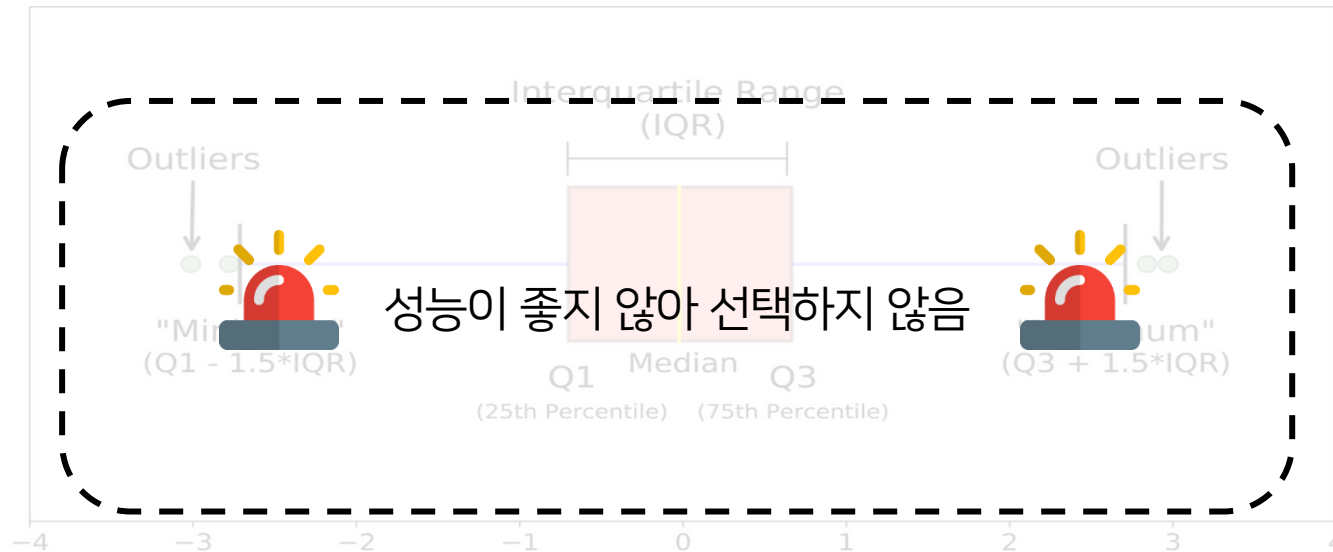
이상치 처리 | IQR(InterQuartile Range)

IQR을 이용하여 이상치를 제거하는 방법



이상치 처리 | IQR(InterQuartile Range)

IQR을 이용하여 이상치를 제거하는 방법



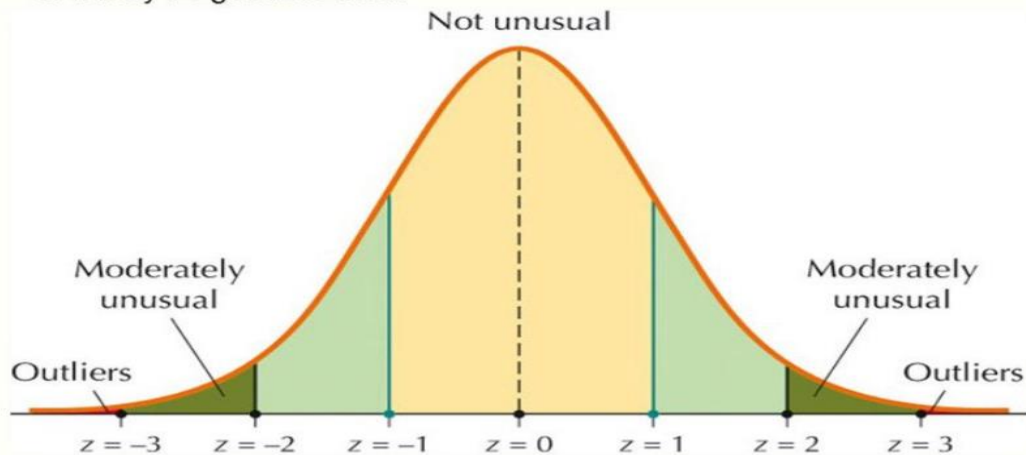
이상치 처리 | Z-score

데이터의 평균과 표준오차를 이용하여 이상치를 제거하는 방법

Detecting Outliers with z-Scores

28

An **outlier** is an extremely large or extremely small data value relative to the rest of the data set. It may represent a data entry error, or it may be genuine data.



이상치 처리 | Z-score

데이터의 평균과 표준오차를 이용하여 이상치를 제거하는 방법

Detecting Outliers with z-Scores

28

An **outlier** is an extremely large or extremely small data value relative to the rest of the data set. It may represent a data entry error, or it may be genuine data.



성능이 좋지 않아 선택하지 않음



이상치 처리 | 파생변수 생성

극단적으로 큰 값을 1, 나머지를 0으로 파생변수 생성

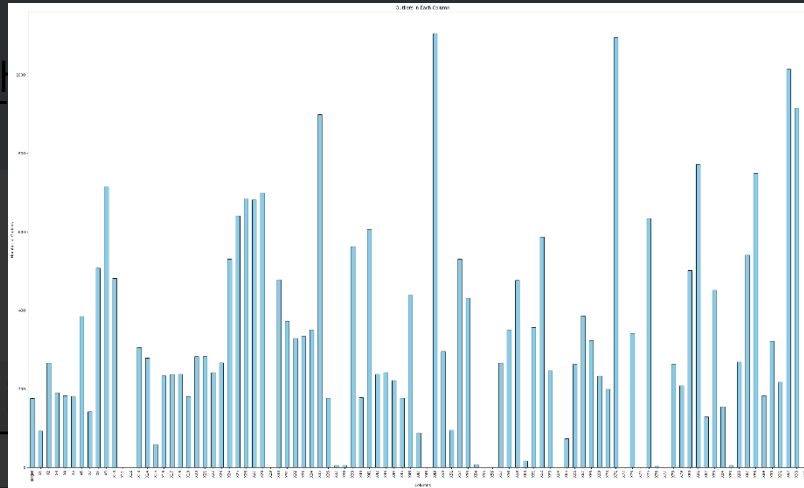
⋮

index	데이터	파생변수
0	2.749929e-04	0
1	1.043314e-04	0
2	2.958876e-04	0
3	8.740000e+09	1
...
24	1.187526e-04	0
25	1.1330167e-04	0
...



이상치 처리 | 파생변수 생성

0과 1로 파생변수를 생성한 이유



1	[IQR을 이용하여 이상치의 개수를 count 한 결과]	0
2	$1.043314e-04$	0
3	$2.958876e-04$	0
...	$8.74010e+09$	1
...
24	$1.187526e-04$	0
25	$1.1330167e-04$	0
...	이상치가 아니라고 의심!	...



이상치 처리 | 파생변수 생성

0과 1로 파생변수를 생성한 이유

극단적으로 큰 값을 1, 나머지를 0으로 파생변수 생성
0에 가까운 값들과 극단적으로 큰 값들로 이루어져 있기에,

scaling 하게 되면 0과 1 양쪽으로 수렴

↓

index	데이터	파생변수
0	2.749929e-04	0
1	1.045514e-04	0
2	2.958876e-04	0
3	8.740000e+09	1
...
24	1.187526e-04	0
25	1.1330167e-04	0
...

극단값을 1로, 나머지를 0으로 파생변수를 생성



이상치 처리 | 파생변수 생성

0과 1로 파생변수를 생성한 이유

극단적으로 큰 값을 1, 나머지를 0으로 파생변수 생성
0에 가까운 값들과 극단적으로 큰 값들로 이루어져 있기에,

scaling 하게 되면 0과 1 양쪽으로 수렴

index	데이터	파생변수
0	2.749929e-04	0
1	1.045514e-04	0
2	2.958876e-04	0
3	8.740000e+09	1
...
24	1.187526e-04	0
25	*Score 7898 -> 7475,	0

극단값을 1로, 나머지를 0으로 파생변수를 생성

드라마틱한 성능 향상을 보임을 확인함!



결측치 처리 | 'X1', 'X7', 'X19', 'X36', 'X51'

mean/mode/median

결측치가 있는 열의 나머지 값들의 평균/최빈값/중앙값을 구하여 결측치 대체

Hot Deck

데이터에서 해당 결측값을 다른 유사한 데이터의 값으로 채우는 방식

KNN

결측치를 가진 데이터 포인트와 근접한 k개의 이웃들을 찾아 결측치 추정 및 대체

Model based Imputation

데이터의 패턴과 관계를 고려하여 결측값을 추정
예측 모델을 사용하여 결측값을 대체하는 방법을 포함

결측치 처리 | 'X1', 'X7', 'X19', 'X36', 'X51'

mean/mode/median

결측치가 있는 열의 나머지 값들의 평균/최빈값/중앙값을 구하여 결측치 대체



(X1, X7, X19) 높은 상관관계를 보인 변수로 회귀식을 만들어 결측치 대체
(X36, X51) 해당 열의 나머지 값의 최빈값으로 결측치 대체

중앙값 < 평균 < **최빈값** 순으로 좋은 성능을 보였음!

결측치를 가진 데이터 포인트와 근접한 K개의 이웃들을 찾아 결측치 추정 및 대체

Model based Imputation

결측치가 없는 행을 모아 XGBoost 모델을 학습시키고,
해당 모델이 예측한 값으로 결측치 대체

결측치 처리 | 'X1', 'X7', 'X19', 'X36', 'X51'

mean/mode/median

결측치가 있는 열의 나머지 값들의 평균/최빈값/중앙값을 구하여 결측치 대체

Hot Deck

데이터에서 해당 결측값을 다른 유사한 데이터의 값으로 채우는 방식

KNN

'X1', 'X7', 'X19'

높은 상관관계를 보인 변수로
회귀식을 만들어 결측치 대체

⋮

'X36', 'X51'

해당 열의 나머지 값의
최빈값으로 결측치 대체

결측치 처리 | 'X1', 'X7', 'X19', 'X36', 'X51'

mean/mode/median

결측치가 있는 열의 나머지 값들의 평균/최빈값/중앙값을 구하여 결측치 대체

Hot Deck

데이터에서 해당 결측값을 다른 유사한 데이터의 값으로 채우는 방식

KNN

결측치를 가진 데이터 포인트와 근접한 k개의 이웃들을 찾아 결측치 추정 및 대체



Model based Imputation

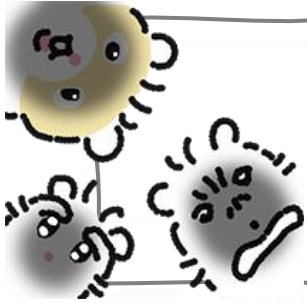
BUT 고차원데이터에는 적합한 방법이 아니기에, 고려하지 않음

해당 모델이 예측한 값으로 결측치 대체

결측치 처리 | 'X1', 'X7', 'X19', 'X36', 'X51'

mean/mode/median

결측치가 있는 열의 나머지 값들의 평균/최빈값/중앙값을 구하여 결측치 대체



결측치가 없는 행을 모아 XGBoost 모델을 학습시키고,
해당 모델이 예측한 값으로 결측치 대체

결측치를 가진 데이터 포인트와 근접한 k개의 이웃들을 찾아 결측치 추정 및 대체



Model based Imputation

데이터의 패턴과 관계를 고려하여 결측값을 추정
예측 모델을 사용하여 결측값을 대체하는 방법을 포함

결측치 처리 | 결론

mean/mode/median

결측치가 있는 열의 나머지 값들의 평균/최빈값/중앙값을 구하여 결측치 대체

LGBM을 기본 모델로 했을 때,

Hot Deck

데이터에서 해당 결측값을 다른 유사한 데이터의 값으로 채우는 방식



KNN

결측치를 가진 데이터 포인트와 근접한 이웃들을 찾아 결측치 추정 및 대체

Model based Imputation

결측치가 없는 행을 모아 XGBoost 모델을 학습시키고,
해당 모델이 예측한 값으로 결측치 대체

결측치 처리 | 결론

mean/mode/median

결측치가 있는 열의 나머지 값들의 평균/최빈값/중앙값을 구하여 결측치 대체

Hot Deck

(X1, X7, X19) 높은 상관관계를 보인 변수로 회귀식을 만들어 결측치 대체
(X36, X51) 해당 열의 나머지 값의 최빈값으로 결측치 대체

KNN

결측치를 가진 데이터 포인트와 근접한 K개의 이웃들을 찾아 결측치 추정 및 대체

비교적 좋은 성능을 보인

최빈값 (score 8177.5)과 Hot Deck (score 7648)을 이용한 방법을
결측치가 없는 해를 모아 XGBoost 모델을 학습시키고,
모델링 과정에서 시도해보기로 결정!



비용민감학습(Cost-sensitive learning)

모델이 학습할 때, 오분류의 유형에 따라 다른 가중치를 두는 기법으로,
데이터의 클래스 불균형이 심한 경우에 고려 가능

비용민감학습(Cost-sensitive learning)

모델이 학습할 때, 오분류의 유형에 따라 다른 가중치를 두는 기법으로,
데이터의 클래스 불균형이 심한 경우에 고려 가능

⋮

소수 클래스에 대해 더 높은 **가중치**를 부여



모델이 다수 클래스에 더 많이 적합되는 것을 방지!



비용민감학습(Cost-sensitive learning)

모델이 학습할 때, 오분류의 유형에 따라 다른 가중치를 두는 기법으로,
데이터의 클래스 불균형이 심한 경우에 고려 가능



소수 클래스

BUT

성능이 매우 좋지 않았음..



다수 클래스



샘플링 해보기로 결정!

모델이 다수 클래스에 더 많이 적합되는 것을 방지



샘플링 (Sampling)

클래스 불균형 데이터로 분류 모델을 학습하면 분류 성능이 저하되는 문제가 발생!

데이터의 클래스 **균형을 맞추기 위해 사용**되는 방법

⋮

Undersampling

불균형한 데이터 셋에서
높은 비율을 차지하던 클래스의
데이터 수를 줄임으로써
데이터 불균형을 해소

+

Oversampling

불균형한 데이터 셋에서
낮은 비율을 차지하던 클래스의
데이터 수를 늘림으로써
데이터 불균형을 해소

샘플링 (Sampling)

클래스 불균형 데이터로 분류 모델을 학습하면 분류 성능이 저하되는 문제가 발생!

데이터의 클래스 균형을 맞추기 위해 사용되는 방법

⋮

Undersampling

불균형한 데이터 셋에서
높은 비율을 차지하던 클래스의
데이터 수를 줄임으로써
데이터 불균형을 해소



+

Oversampling



불균형한 데이터 셋에서
데이터 제거로 인한
정보 손실로 인해
데이터 수를 늘림으로써
성능이 감소할 수 있음
데이터 불균형을 해소

샘플링 (Sampling)

클래스 불균형 데이터로 분류 모델을 학습하면 분류 성능이 저하되는 문제가 발생!

데이터의 클래스 **균형을 맞추기 위해 사용**되는 방법

⋮

Undersampling



불균형한 데이터 셋에서
과적합 가능성이 존재하며,
높은 비율을 차지하던 클래스의
이상치 및 노이즈에 민감함
데이터 수를 줄임으로써
데이터 불균형을 해소

+

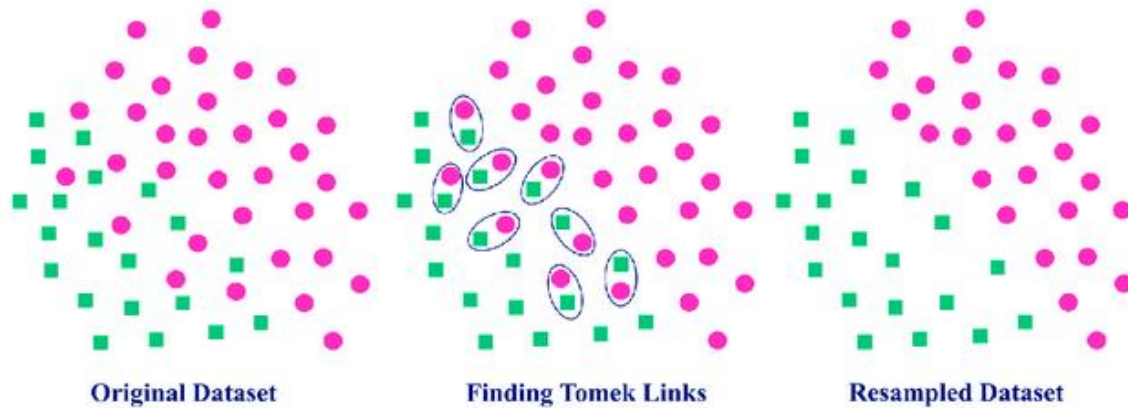
Oversampling

불균형한 데이터 셋에서
낮은 비율을 차지하던 클래스의
데이터 수를 늘림으로써
데이터 불균형을 해소

샘플링 | Undersampling

Tomek links

소수 클래스와 다수 클래스 사이에 있는 경계의 **노이즈 데이터를 제거**하여
다수 클래스의 모호한 관측치들을 제거하는 방법

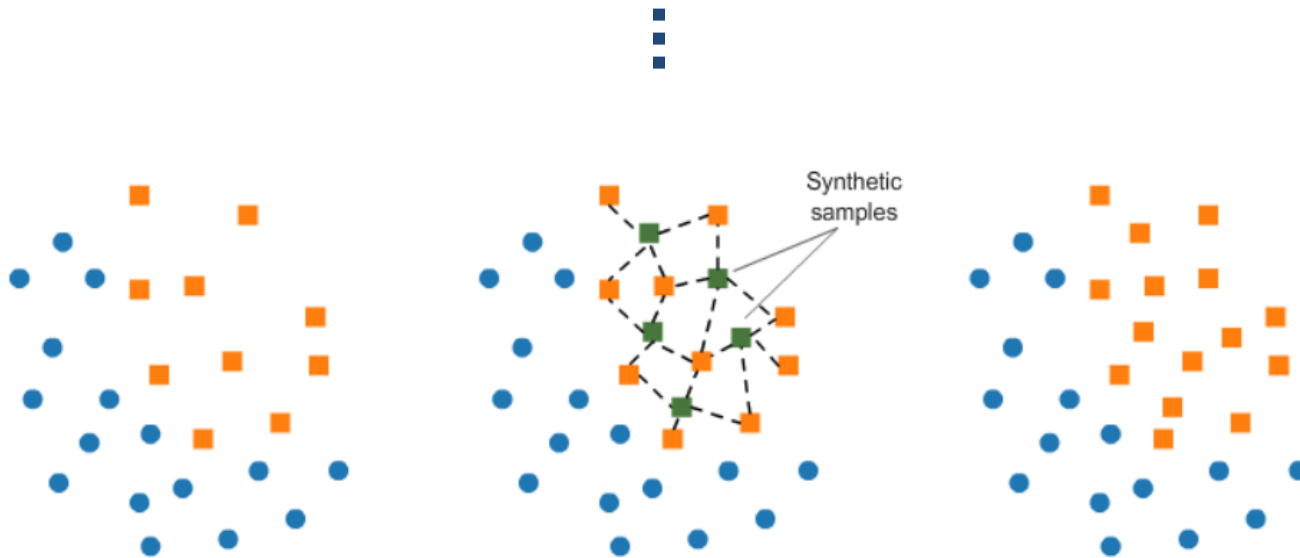


샘플링 | Oversampling

SMOTE (Synthetic Minority Over-sampling Technique)

낮은 비율 클래스 데이터들의 sample에

최근접 이웃(KNN)을 적용하여 새로운 데이터를 생성하는 방법

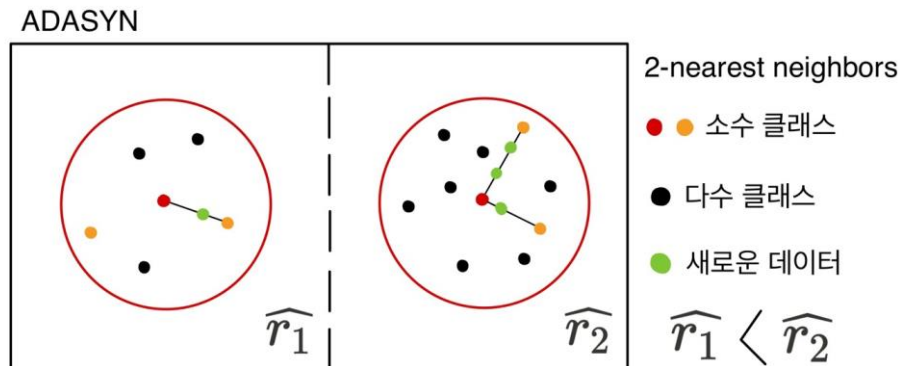


샘플링 | Oversampling

ADASYN (Adaptive Synthetic Sampling)

SMOTE와 같은 과정을 적용한 후,
 소수 클래스 데이터들 각각에 다수 클래스가 관측된 비율에 따라
 weight를 주어 데이터를 생성하는 비율에 차이를 주는 방법

⋮



Conditional GAN (CTGAN)

mode-specific 하게, 즉 입력 조건에 맞게 데이터를 생성하는 모델



열(column) 간의 관계 및 다양한 데이터 유형을 고려하여
데이터를 생성하기 때문에,
실제 데이터와 유사한 특성을 갖는 데이터를 생성할 수 있음

Conditional GAN (CTGAN)

tabular data GAN의 문제점 해결 가능

Continuous data

확률분포가 여러 개의 봉우리를 가지는
multi-modal distribution의 문제



**Model-specific
normalization** 으로 해결

Discrete data

category별로 빈도수가 모두 다른
성질(imbalance)의 문제



conditional generator 와
training-by-sampling 으로 해결

Conditional GAN (CTGAN)

tabular data GAN의 문제점 해결 가능

Continuous data

확률분포가 여러 개의 봉우리를 가지는

multi-modal distribution의 문제 BUT..



성능이 별로 좋지 않았음

Model-specific

normalization 으로 해결

Discrete data

category별로 빈도수가 모두 다른

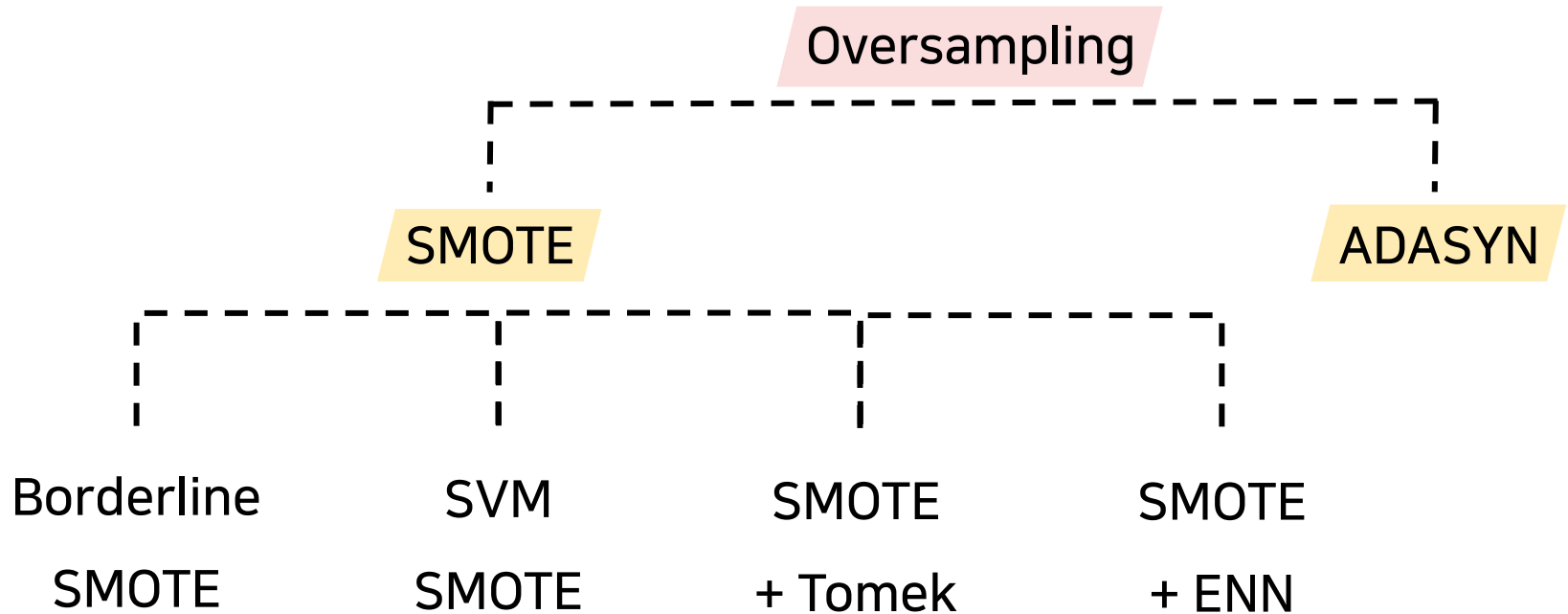
성질(imbalance)의 문제



conditional generator와

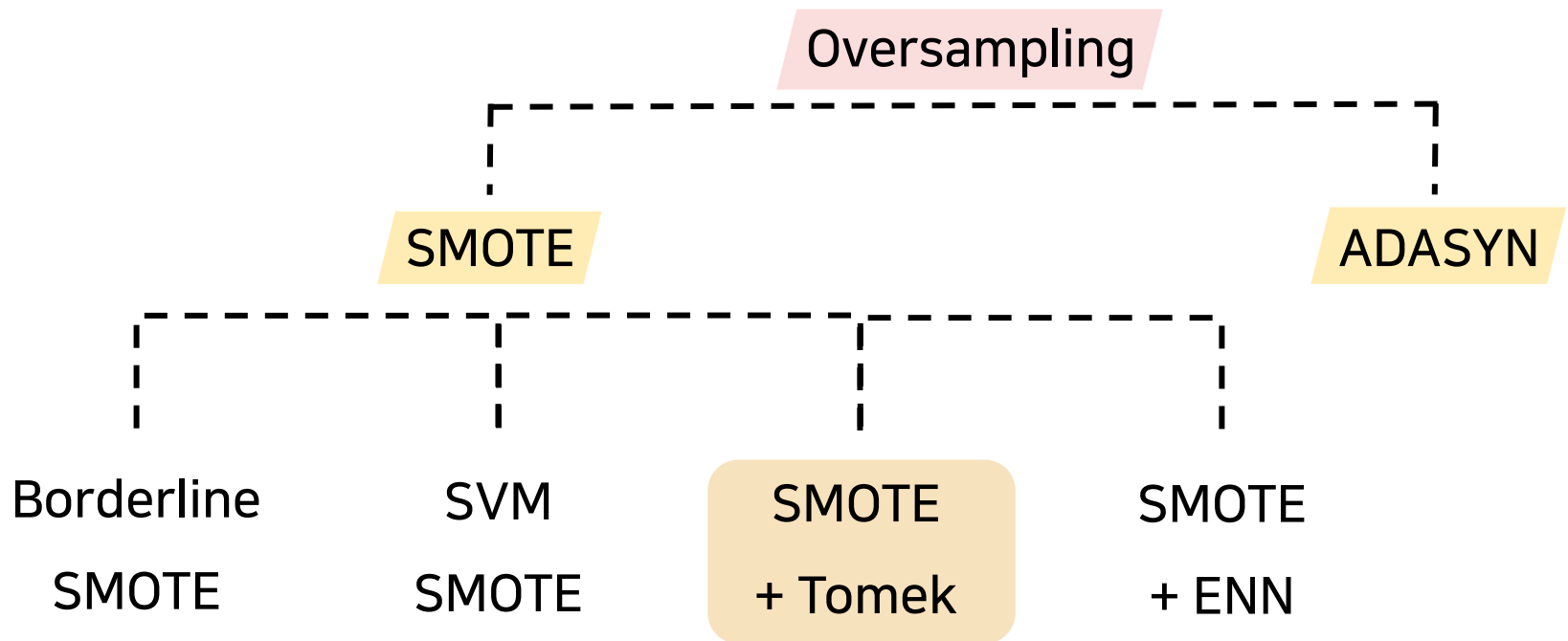
training-by-sampling 으로 해결

데이터 불균형 해결



Imblearn 패키지는 SMOTE / ADASYN과
관련하여 다양한 방법들을 제공

데이터 불균형 해결



이 중 성능이 가장 좋았던 (score 6485)

SMOTE-Tomek 방법을 활용!

3

Feature Selection

3

Feature Selection

Feature Importance

Gain Importance

특성이 데이터를 분할할 때
불순도를 **"얼마나"** 감소시키는지



Split Importance

특성이 데이터를
"얼마나 자주" 분할하는지

3

Feature Selection

Feature Importance

Gain Importance

특성이 데이터를 분할할 때
불순도를 **"얼마나"** 감소시키는지



Split Importance

특성이 데이터를
"얼마나 자주" 분할하는지



강한 feature를 정확하게
평가할 수 있지만, 이해하기 어려움



이해하기 쉽고 직관적이지만,
강한 feature가 과소평가될 수 있음

3

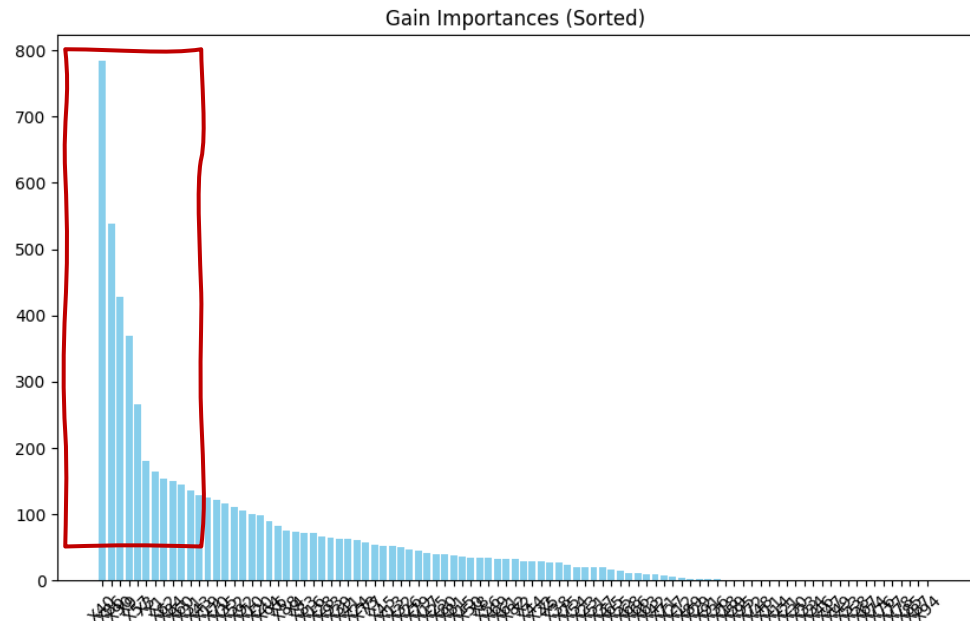
Feature Selection

Gain Importance

선택된 Feature

'X40', 'X86', 'X90', 'X9', 'X57', 'X3

변수명	Gain Importance
X40	784.1557428
X86	538.9504634
X90	427.4698201
X9	368.6355923
X57	265.220066
X3	180.3739585
...	...
X10	100.7839078
...	...



3

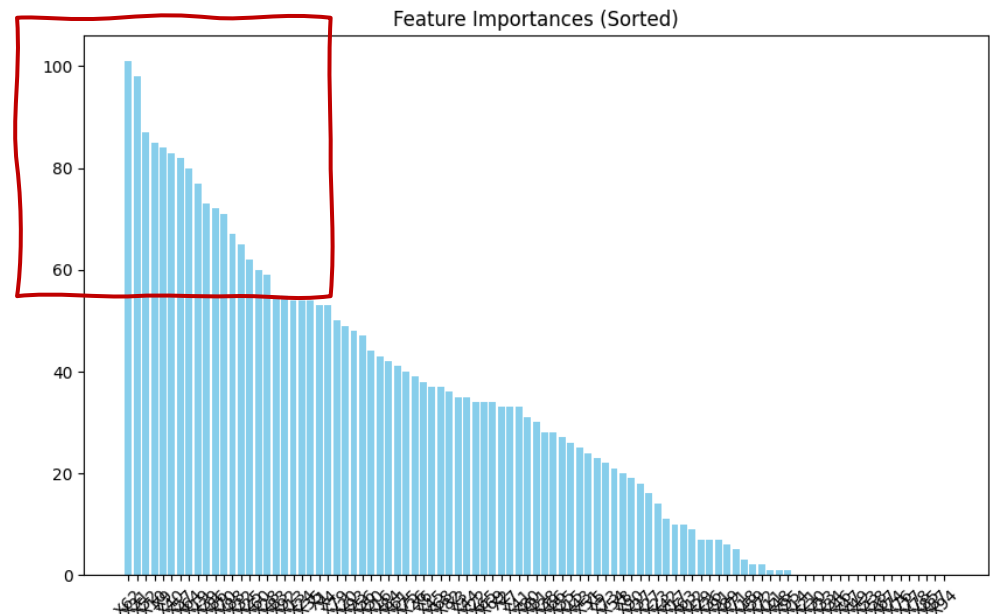
Feature Selection

Split Importance

선택된 Feature

'X62', 'X31', 'X52', 'X10', 'X9', 'X40', 'X57', 'X64', 'X19', 'X39', 'X86'

변수명	Split Importance
X62	101
X31	98
X52	87
X10	85
X9	84
X40	83
...	...
X83	55
...	...



3

Feature Selection

RFE CV (Recursive Feature Elimination with Cross-Validation)

특성 선택과 교차 검증을 결합하여 모델 성능을 향상시키는 기법



특성이 많고, 모델이 과적합되기 쉬운 상황에 유용



선택된 Feature

'X3', 'X4', 'X9', 'X13', 'X19', 'X20', 'X23', 'X25', 'X35', 'X36', 'X40',
'X52', 'X53', 'X57', 'X62', 'X64', 'X70', 'X76', 'X83', 'X88'

SHAP Value (SHapley Additive exPlanations)

게임이론을 기반으로, 각 특성이 모델의 예측에 어떻게 기여하는지를 설명



블랙박스 모델이나 복잡한 모델에서 모델 예측을 해석하기에 유용



선택된 Feature

'X40', 'X3', 'X10', 'X19', 'X41', 'X86', 'X9', 'X13', 'X61', 'X53', 'X36',
'X82', 'X23', 'X66', 'X91', 'X90', 'X68', 'X87', 'X92', 'X1'

3

Feature Selection

종속변수와의 상관관계

파생변수를 제외한 나머지 변수들에 대한 correlation matrix에서
종속변수와의 상관계수가 **0.15** 이상인 변수만 추출

⋮

변수명	상관계수	변수명	상관계수
X1	-0.195	X17	-0.159
X2	-0.265	X18	-0.158
X3	-0.256	X19	-0.154
X16	-0.159

종속변수와의 상관관계

파생변수를 제외한 나머지 변수들에 대한 correlation matrix에서
종속변수와의 상관계수가 0.15 이상인 변수만 추출

⋮

선택된 Feature

'X1', 'X2', 'X3', 'X16', 'X17', 'X18', 'X19', 'X23', 'X37', 'X38',
'X40', 'X43', 'X54', 'X60', 'X65', 'X66', 'X68', 'X78', 'X84',
'X86', 'X90', 'X91'

카이제곱 독립성 검정

두 가지 범주형 변수가 **관련될 가능성** 여부를 확인하는 데 사용되는 통계적 가설 검정

귀무가설(H_0): 두 변수는 독립이다.

대립가설(H_1): 두 변수는 서로 독립이 아니다.



이상치 처리에서 만든

파생변수와 **종속변수** 간의 카이제곱 독립성 검정 시행!

3

Feature Selection

카이제곱 독립성 검정

두 가지 범주형 변수가 **관련될 가능성** 여부를 확인하는 데 사용되는 통계적 가설 검정

```
1 count_table = pd.crosstab(train_data['X71'], train_data['Target'])
2
3 chi,p,df,expected = chi2_contingency(count_table,correction=False)
4 print(chi,p)

1.7537248508788525 0.18540915111442188

1 count_table = pd.crosstab(train_data['X72'], train_data['Target'])
2
3 chi,p,df,expected = chi2_contingency(count_table,correction=False)
4 print(chi,p)

2.031216191394643 0.1540972519723118

1 count_table = pd.crosstab(train_data['X74'], train_data['Target'])
2
3 chi,p,df,expected = chi2_contingency(count_table,correction=False)
4 print(chi,p)

11.95441747163369 0.0005451796975463964
```

P-value 0.05 이하인 변수들을 선택



선택된 Feature

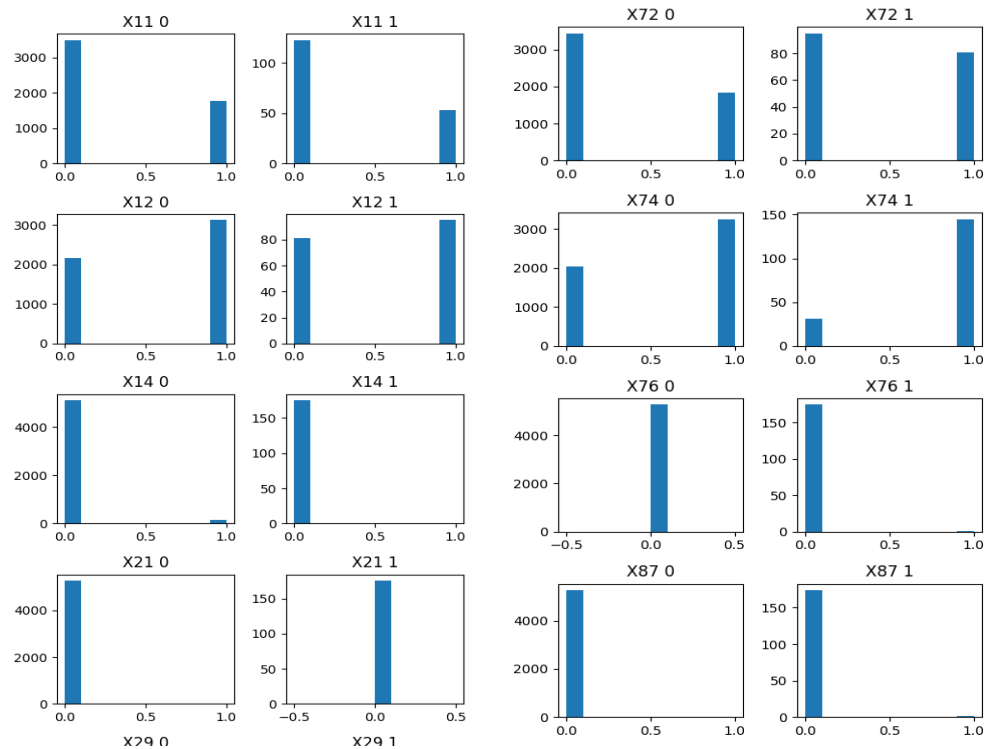
'X29', 'X48', 'X49', 'X59', 'X74'

3

Feature Selection

변수 별 분포 파악

파생변수 생성 후 변수 별 클래스 0, 1의 분포 확인



최종 Feature 선택

종속변수와의 상관관계, 카이제곱 독립성 검정,
추가적으로 분포를 보고 고려한 변수를 반영하여 최종 feature 선택!



선택된 26개의 최종 Feature

'X1', 'X2', 'X3', 'X16', 'X17', 'X18', 'X19', 'X23', 'X37', 'X38',
'X40', 'X43', 'X54', 'X60', 'X65', 'X66', 'X68', 'X78', 'X84',
'X86', 'X90', 'X91', 'X48', 'X49', 'X59', 'X74'



4

모델링

모델후보군

LGBM

XGBoost

Neural Networks

Random Forest

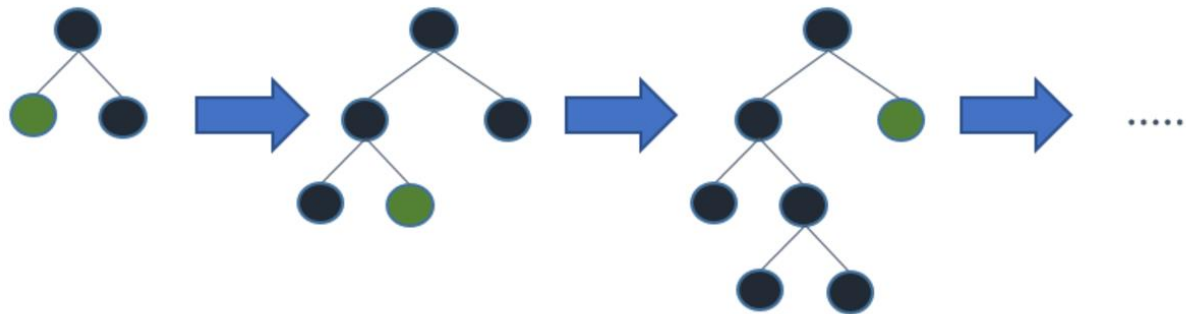
Extra Trees

KNN Classifier

LightGBM

트리 기반 Gradient boosting 모델로, 트리 기준 분할이 아닌
최대 손실 값을 갖는 리프를 기준으로 분할하여 예측 오류 손실을 최소화하는 방법

⋮

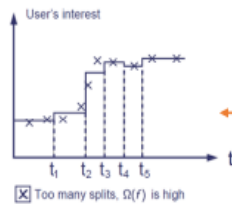
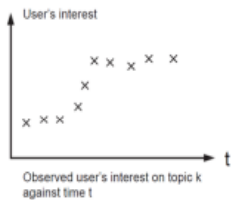


Leaf-wise tree growth

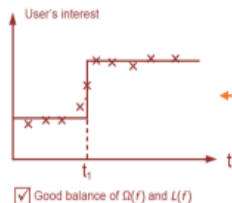
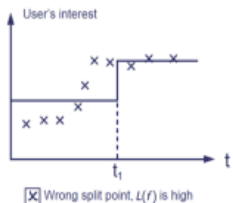
Xgboost (Extreme Gradient Boosting)

Gradient boosting을 병렬 학습이 지원되도록 구현한 라이브러리

■
■
■



Gradient Boosting



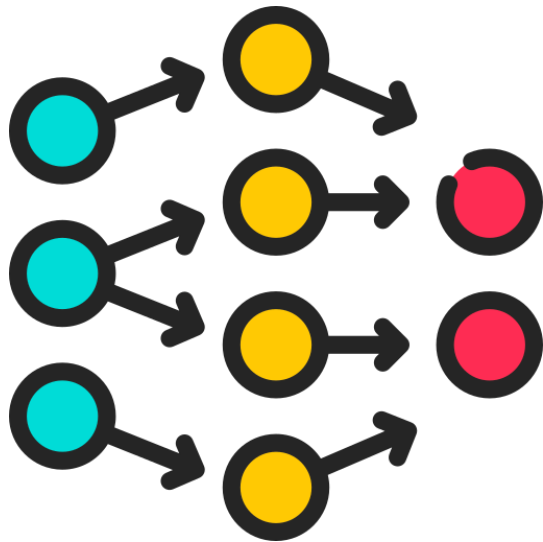
XGBoost



- ☑ 분류, 회귀문제에 모두 적용 가능
- ☑ 뛰어난 예측 성능 발휘
- ☑ 과적합 규제

Neural Networks (NNs)

각 퍼셉트론과 은닉층에서 이루어지는 **비선형 변환기법**을 통해,
높은 수준의 추상화를 시도하여 데이터를 처리하는 기법



Feature Selection의 중요성 감소
데이터의 **추상적 특징**을 포착할 수 있음

5

최종 모델

하이퍼파라미터 튜닝 결과 | Optuna

xgb_parameter	튜닝 결과
n_estimators	654
max_depth	12
min_child_weight	11
gamma	1
colsample_bytree	0.5
lambda	4.31742702882685
alpha	0.00989655020027596
subsample	0.8

lgbm_parameter	튜닝 결과
n_estimators	476
max_depth	15
min_child_samples	84
subsample	0.665433990013184
learning_rate	0.00012235676716719124



예측 하러 가보자..

+) NNs의 경우 은닉층 1레이어, 노드 16개를 사용!

하이퍼파라미터 튜닝 결과 | Grid Search

NNs의 경우

은닉층 1레이어에 대해 Grid Search 방법을 적용하여

최적의 노드 수를 탐색

NNs	개수
은닉층	1
노드	16



예측하러 가보자..

1. Model Ensemble

XGB & LGBM

트리모델을 통해 정형 데이터 분석



전통적으로 트리모델이
정형 데이터에 강함

Neural Networks

딥러닝 모델을 통해 정형 데이터 분석



딥러닝과 트리모델의 앙상블 조합이
경험적으로 효과가 좋다고 밝혀짐

1. Model Ensemble

XGB
Score : 3950

+

LGBM
Score : 3360

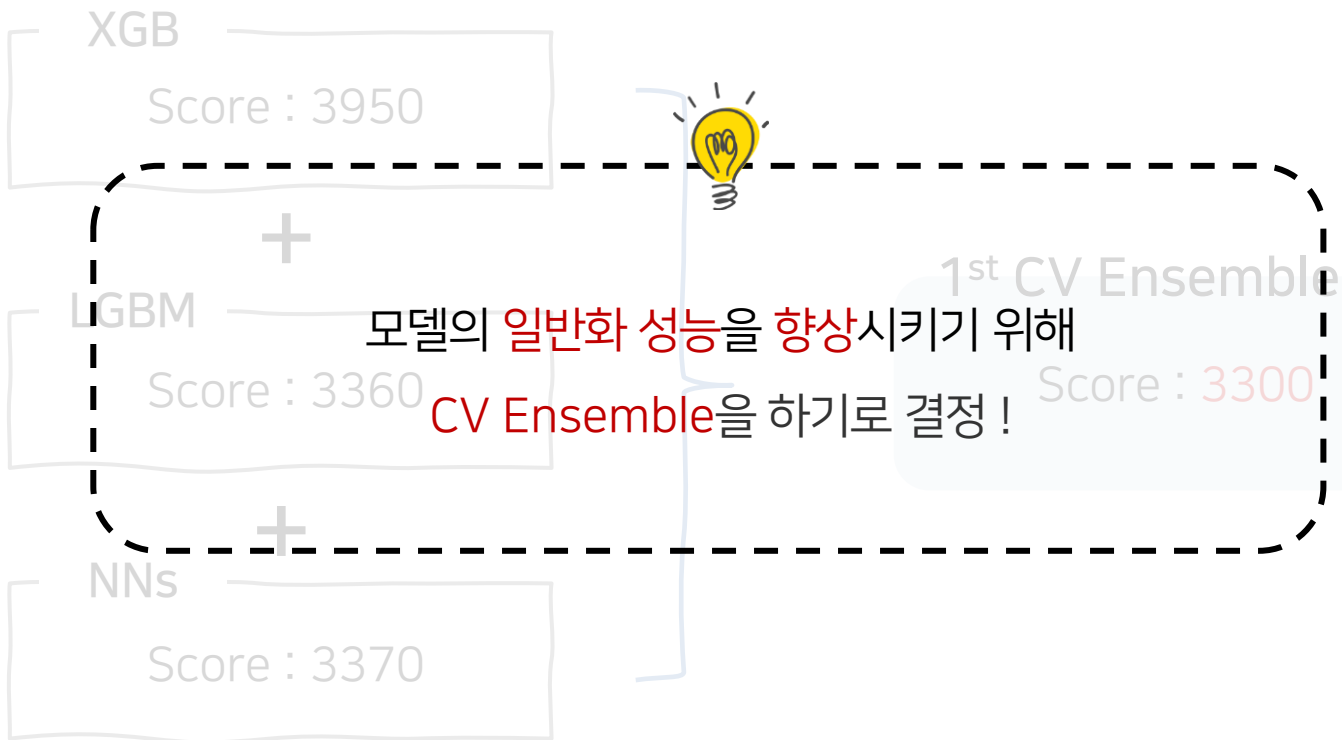
+

NNs
Score : 3370

1st CV Ensemble

Score : 3300

1. Model Ensemble



2. CV Ensemble

1st CV Ensemble

Score : 3300

2nd CV Ensemble

Score : 2350

3rd CV Ensemble

Score : 4720

4th CV Ensemble

Score : 3280

CV Ensemble

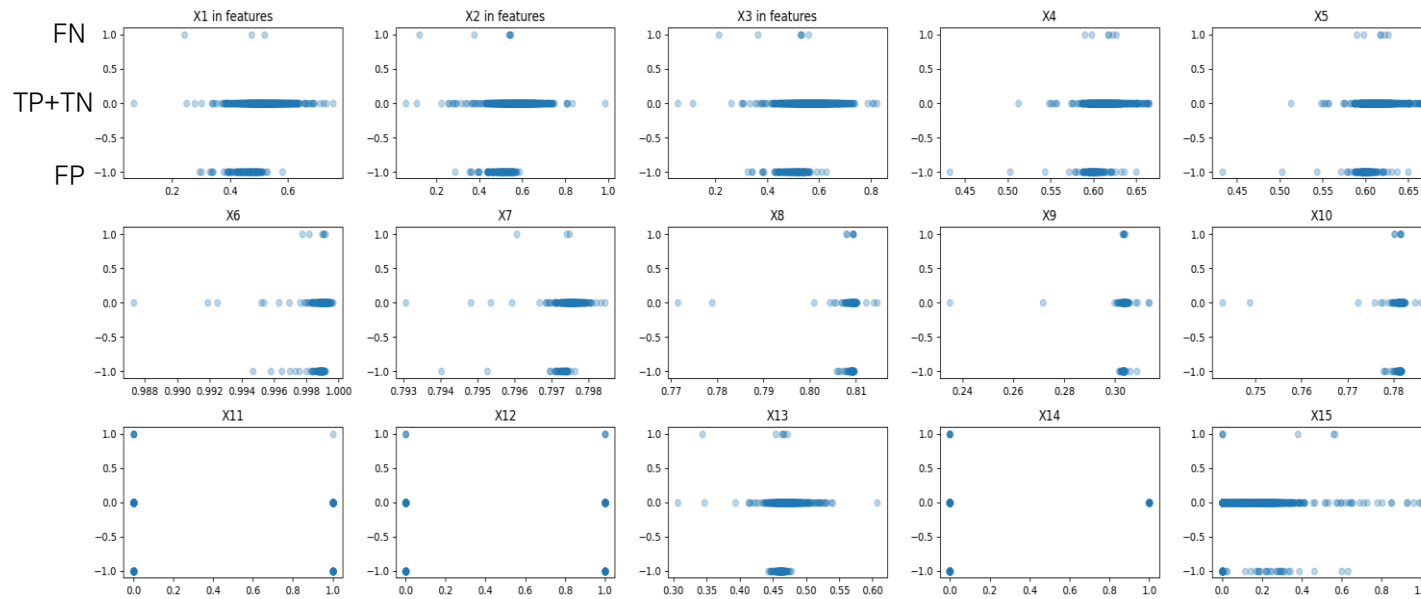
Score : 3412.5

Kaggle score : 1330



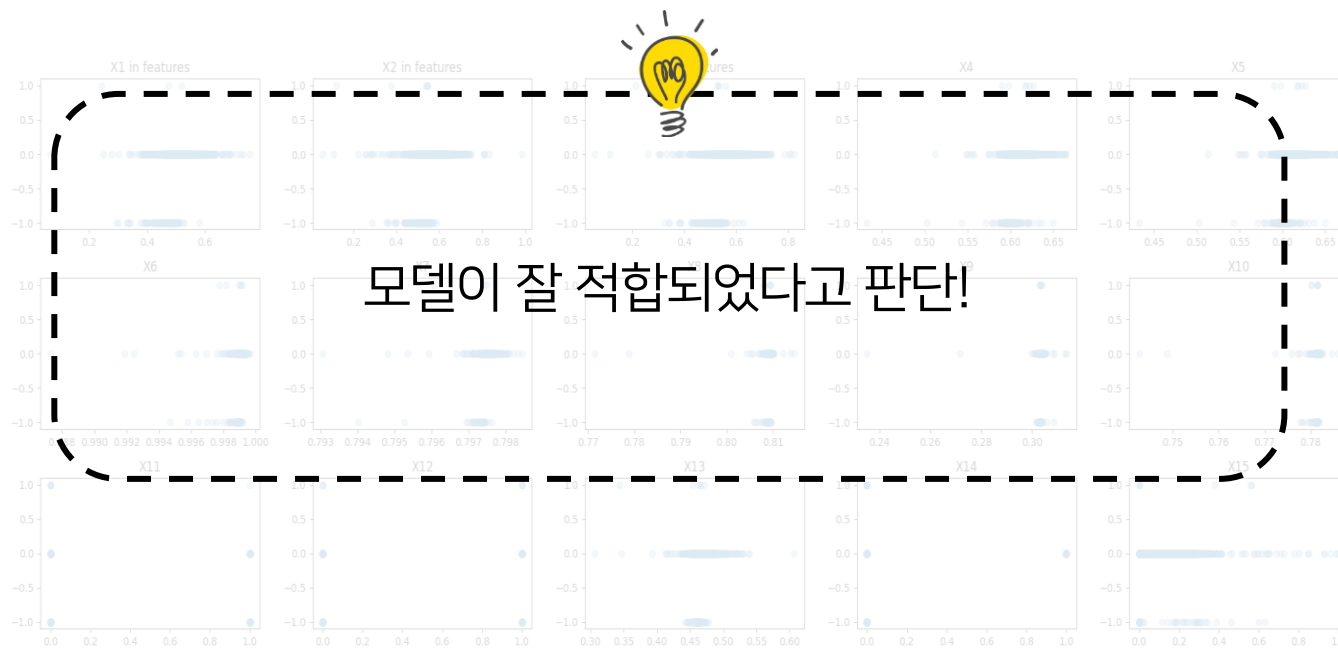
3. 잔차 확인

FN, TP, TN, FP 네 개의 분포가 많이 차이나지 않는 것을 확인

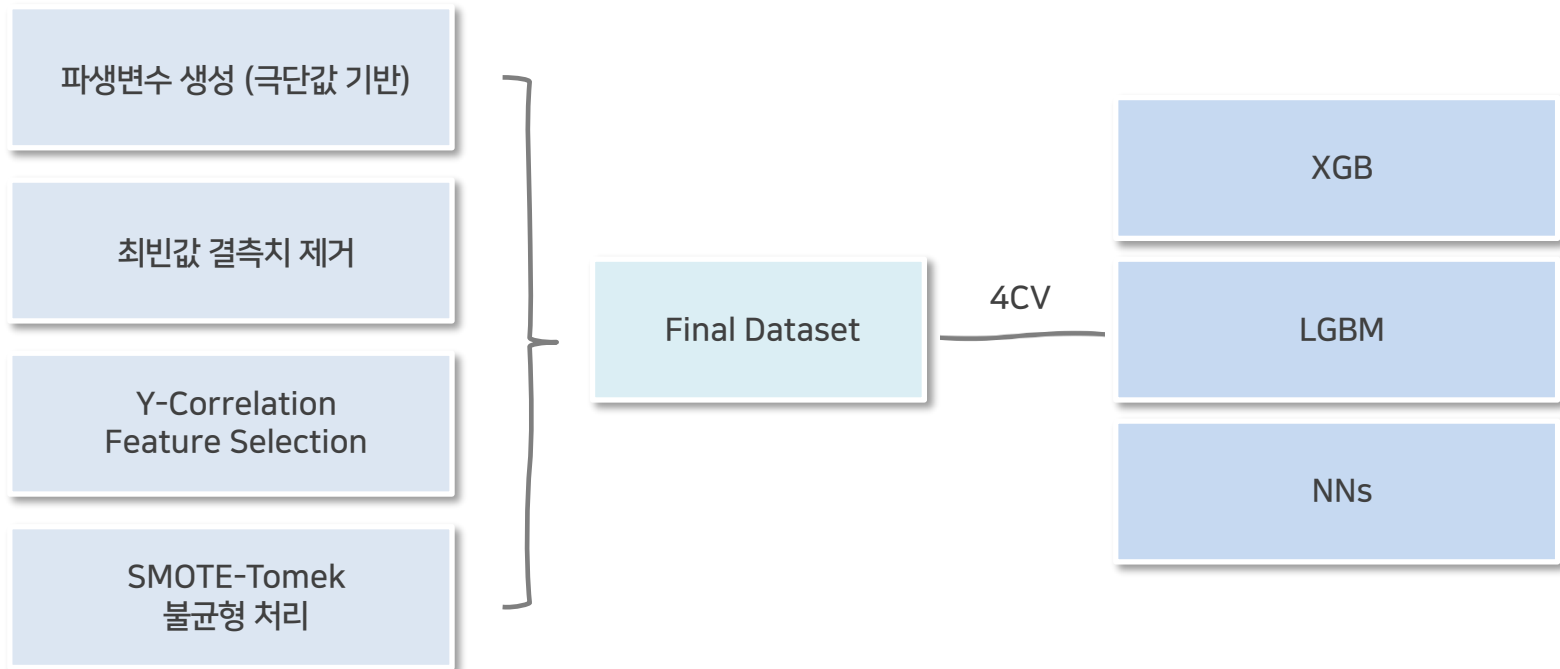


3. 잔차 확인

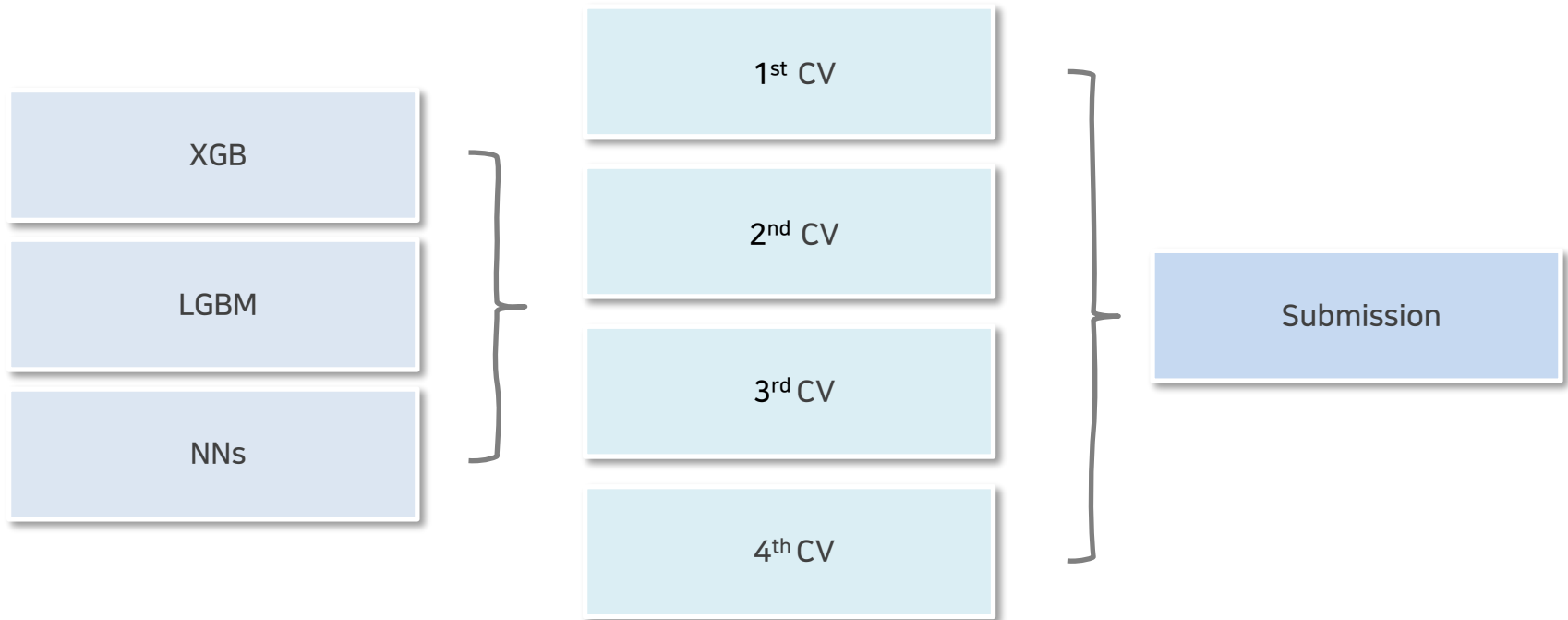
FN, TP, TN, FP 네 개의 분포가 많이 차이하지 않는 것을 확인



3. 최종 파이프라인



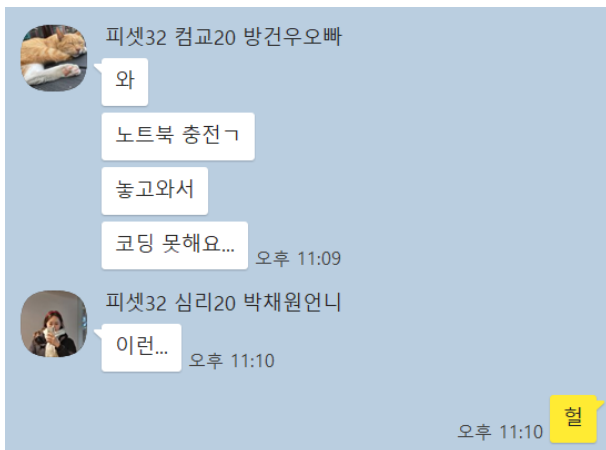
3. 최종 파이프라인



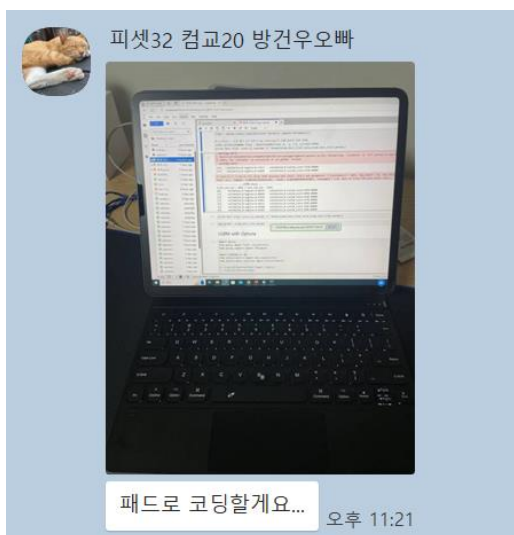
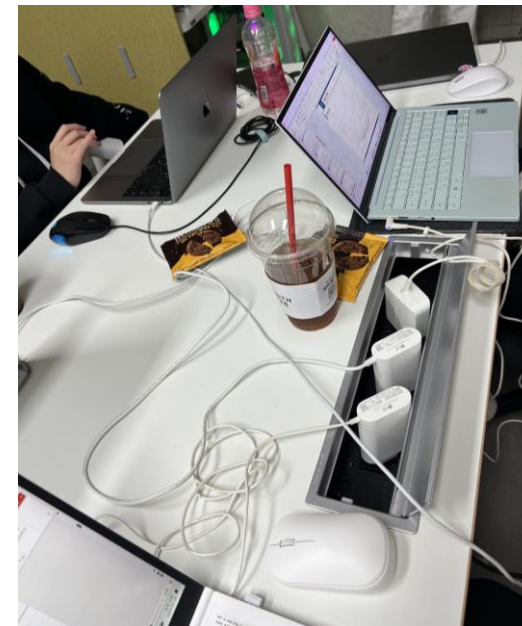
+

2팀의 이모저모

옹기종기 2팀 im 글웅스튜디오..



히든스시 돌아와주세요



충전기의 부재도 막지 못한 그의 코딩..
그렇게 시작된 성능싸움



헉헉한 나눔의 현장



웁저건우

+

화목2팀은요..



화목한 1, 2팀 ㅎㅎ



소문으로만 들던.. 봉천건우 영점
이제는.. 성천건우 (성능에 미친 건우)



개알 동기

캐글 대회가 처음이었는데 팀원들 덕분에 정말 많은 걸 배울 수 있었습니다. 팀원들이랑 친해지고 1팀이랑도 밥 같이 먹으면서 친해져서 정말 재밌었던 방학세미나였던 것 같습니다! 다음 학기에 더 친해져서 즐겁게 피셋 활동 마무리했으면 좋겠습니다!! 우리팀 다들 너무 너무 고생 많으셨습니다!!



김동희



김민주

일주일이 순식마다던 방학세미나.. 정말이네요.. 대회도 처음이고 아는게 많이 없어서 잘할 수 있을지 걱정이 많았는데 든든한 팀원들 덕분에 많이 배우고 무사히 잘 마칠 수 있었던 것 같아요. 팀원들이랑도 많이 친해지고, 32기 모든 분들과 가까워질 수 있어서 너무 !! 좋았습니다 ㅎㅎ 모두 고생 많으셨고, 다음 학기에도 잘 부탁드립니다 ❤️

진짜 어느새 정신차려보니 방세 마지막 날이 다가오고 있다니 믿기지 않네요. 할 줄 아는 게 많지 않기도 하고, 친해질 기회가 많이 없던 다른 팀원들이라 걱정이 많았는데 이제는 우리 팀원뿐 아니라 학회원 전부랑 친해질 수 있어서 정말 좋은 경험이었던 것 같아요! 우리 팀원들이 과연 어디까지 성능을 올릴 수 있을까... 싶기도 하고... 지켜보고 과정에 참여하면서 많이 배울 수 있었습니다! 다음 학기도 파이팅~~ ★



박채원

드디어 짧지만 알찼던 방학세미나가 끝났습니다.. 저 혼자만 했다면 절대 찍지 못했을 점수인데.. 여러분과 함께해서 이렇게 완벽한 모델을 만들 수 있었던 것 같네요. EDA 마스터 동희 형, 결측치의 대가 채원이, 균형잡기의 달인 재언이, 디자인과보다 디자인 잘하는 민주까지, 모두와 함께해서 정말 배울 점이 많았던 방세였던 것 같아요! 정말 몰랐던 사실들을 많이 알게 되어서 너무나 뿌듯했던 것 같습니다. 각자의 팀으로 가셔도 지금처럼 완벽한 모습을 보여줄 것이라 의심치 않습니다. 항상 화목했던 2팀 앞으로도 꼭 같이 성장하는 피셋이 됩시다 팻팅!



방건우



진재언

윗 기수 선배님(쫄X우)들이 이틀 밤을 샜다는 말을 듣고 참 막막했었는데 좋은 팀원들을 만나게 되어 좋은 결과(리더보드 1등)를 낼 수 있었던 것 같습니다! 캐글같은 대회를 한 번도 나가본 적이 없었는데 이번 기회를 통해 한층 더 성장할 수 있었던 것 같습니다. 일주일동안 또 글로벌융합학부 스튜디오에서 갇혀 있었는데, 덕분에 2팀과 전부 친해질 수 있었습니다 ㅎㅎ 남은 한학기도 다같이 화이팅해요!!!!!!



THANK YOU

