

Credit Risk Model Stability

Home Credit 

Team 9

박언선 이강건

최용빈 한민규 홍재민

목차

1. 문제 정의
2. Feature Engineering
3. Team9 Baseline
4. 하이퍼 파라미터 튜닝
5. Q&A

문제 정의 풀어야 하는 문제

이 대회 목표는 대출 상환 능력이 좋은 고객들을 예측하는 것!
신용 기록이 없는 사람*은 대출을 거절 받을 가능성이 높으므로, 대출 상환 능력이 있는 고객과 없는 고객을 잘 판단해야 함

* 젊은 나이어거나 현금을 선호할 수 있음

대출 상환 능력을 예측하는 모델 개발

고객의 신용 위험을 예측하기 위한 효과적인
모델을 개발해야 함

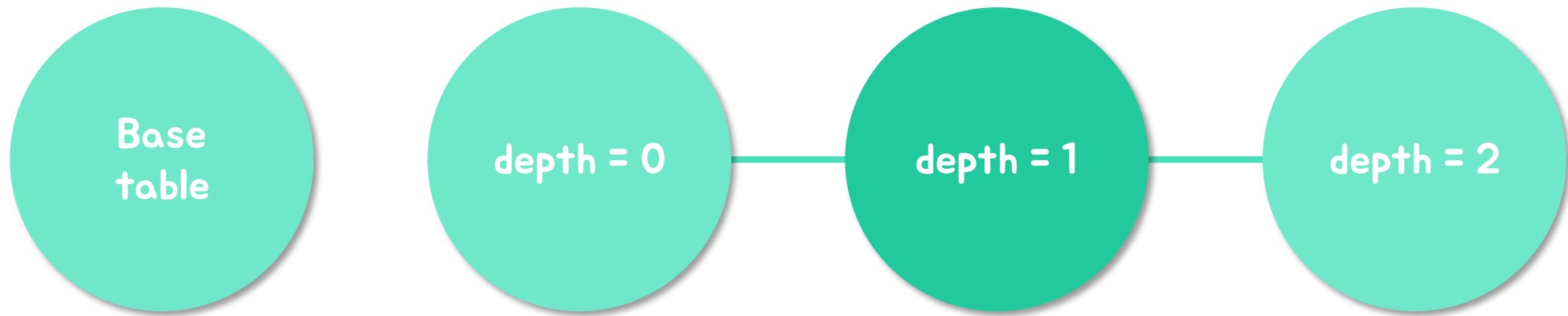
How? 주어진 고객 정보와 대출 상환 기록을
활용하여 모델을 훈련!

모델 안정성 확보

모델은 시간이 지나도 안정적으로 작동해야 함

Why? 실제로 고객의 행동은 지속적으로 변하
므로, 모델의 안정성을 유지하기 위해 정기적인
업데이트가 필요!

문제 정의 데이터



문제 정의 데이터

Base
table

case_id	date_decision	MONTH	WEEK_NUM	target
0	"2019-01-03"	201901	0	0
1	"2019-01-03"	201901	0	0
2	"2019-01-04"	201901	0	0
3	"2019-01-03"	201901	0	0
4	"2019-01-04"	201901	0	1

문제 정의 데이터

depth=0일 경우 Base table과 바로 join 가능!

depth=0

Internal Source

static_0

External Source

static_cb_0

미래
테스트 데이터에
존재하지 않을 수
있습니다.

문제 정의 데이터

depth=1일 경우 Base table과 바로 join 불가능!

why? case_id의 여러 기록을 하나의 특성으로 요약하는 aggregation function을 사용해야 할 수 있음

depth=1

Internal Source

applprev_1
other_1
deposit_1
person_1
debitcard_1

External Source

tax_registry_a_1
tax_registry_b_1
tax_registry_c_1
credit_bureau_a_1
credit_bureau_b_1

문제 정의 데이터

depth=2일 경우 Base table과 바로 join 불가능!

why? case_id의 여러 기록을 하나의 특성으로 요약하는 aggregation function을 사용해야 할 수 있음

depth=2

Internal Source

applprev_2
person_2

External Source

credit_bureau_a_2
credit_bureau_b_2

문제 정의 데이터

위 테이블 그룹 내에는 많은 피처들이 있으며,
피처의 이름 끝에는 'P, M, A, D, T, L'*의 키워드가 있음

* P - Transform DPD (Days past due)
M - Masking categories
A - Transform amount
D - Transform date
T - Unspecified Transform
L - Unspecified Transform

테이블 그룹에서 피처들이 어떤 의미를 가지고 있는지,
어떤 피처들이 문제 해결에 도움이 될 지 찾아보려고 함!

문제 정의

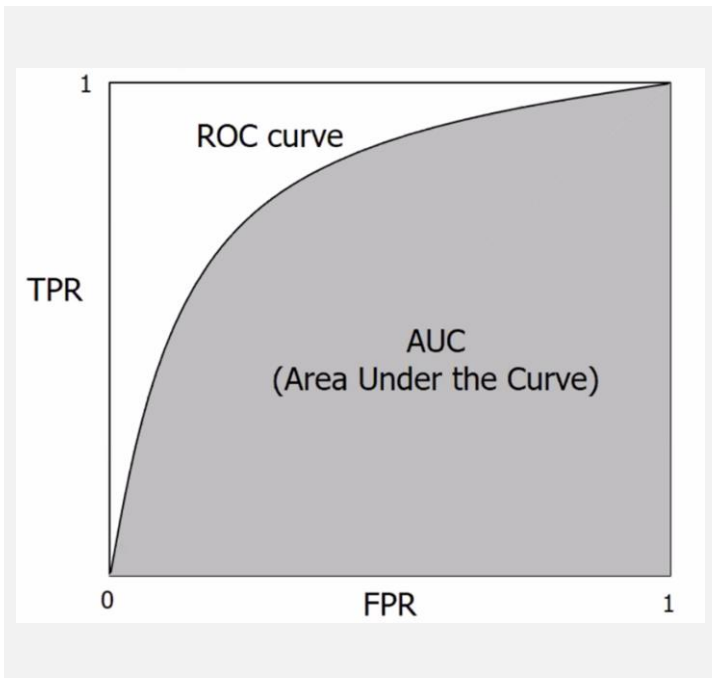
평가 지표

$$stability\ metric = mean(gini) + 88.0 \cdot \min(0, a) - 0.5 \cdot std(residuals)$$

문제 정의 평가 지표

$$\text{stability metric} = \text{mean}(\text{gini}) + 88.0 \cdot \min(0, a) - 0.5 \cdot \text{std}(\text{residuals})$$

$$\text{gini} = 2 \times \text{AUC} - 1$$



TPR (True Positive Rate = Sensitivity)

실제로 Positive인 것들 중에서 Positive를 맞춘 비율

$$\text{TPR} = \frac{TP}{TP + FN}$$

FPR (False Positive Rate = 1 - Specificity)

실제로 Negative인 것들 중에서 잘못해서 Positive로 판단한 비율

$$\text{FPR} = 1 - \frac{TN}{FP + TN}$$

AUC는 1에 가까울수록 모델의 성능이 우수함

지니계수가 0에 가까울수록 해당 Class의 불순도가 낮음

문제 정의 평가 지표

$$\text{stability metric} = \text{mean}(\text{gini}) + 88.0 \cdot \min(0, a) - 0.5 \cdot \text{std}(\text{residuals})$$

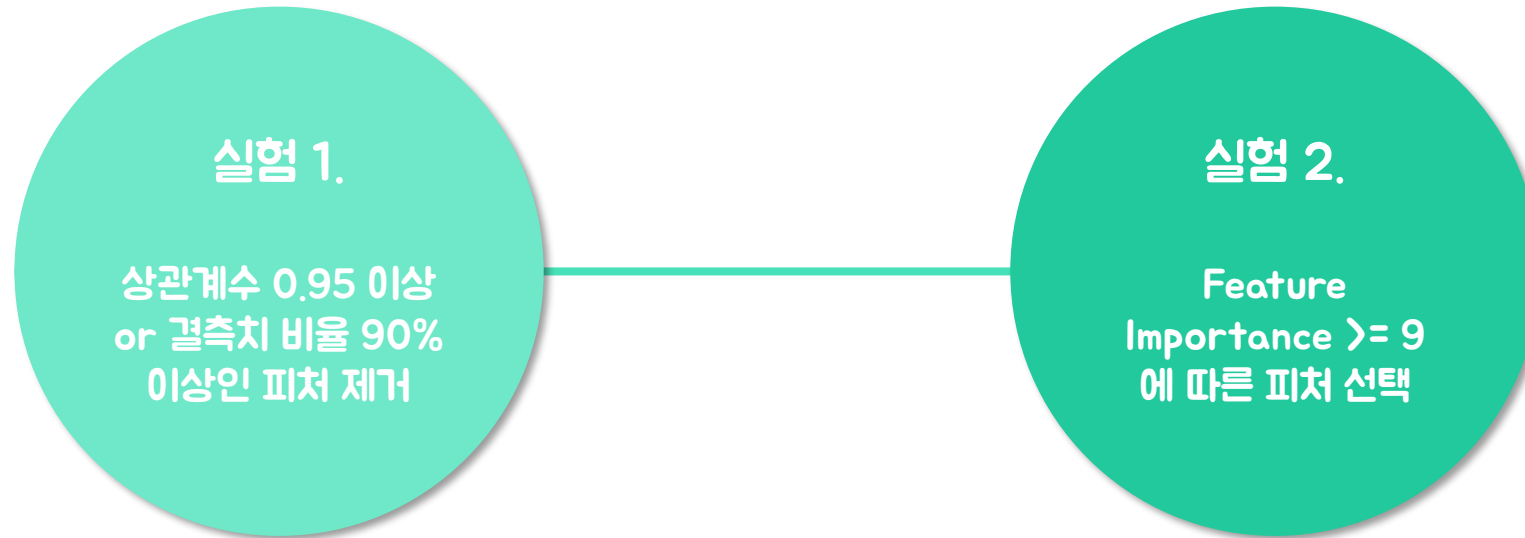
$$88.0 \cdot \min(0, a)$$

- $ax + b$: weekly 지니 점수를 적합한 선형 회귀 모형
- `falling_rate`은 $\min(0, a)$ 를 통해 계산됨
 - a 는 기울기이며, 예측 능력이 떨어지는 모델에 `penalty`를 주는 데 사용됨

$$0.5 \cdot \text{std}(\text{residual})$$

- `std(residuals)`란, 잔차의 표준편차
- 잔차는 실제값과 예측값의 차이이며, 그것의 표준편차임
- 즉, 각 값을 놓고 봤을 때, 실제값과 예측값의 차이가 퍼진 정도를 나타냄
- 따라서, 이 값이 높다면, 모델이 적합하지 않다는 의미임

Feature Engineering `applprev_1`



Feature Engineering applprev_1

실험 1, 2의 기준 $\text{Score} = \text{basetable} + \text{static_0} + \text{static_cb_0} + \text{applprev_1}$

The AUC score on the train set is: 0.7884195763038669

The AUC score on the valid set is: 0.7741817535554472

The AUC score on the test set is: 0.7695572002271935

The stability score on the train set is: 0.5533121690468635

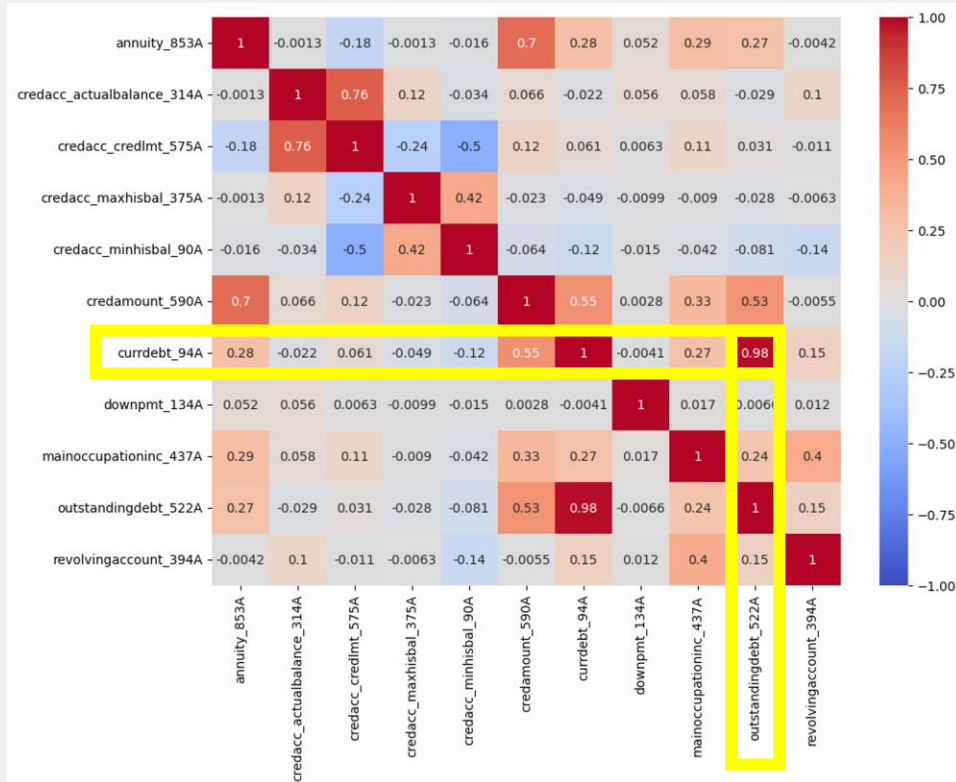
The stability score on the valid set is: 0.5176205072833473

The stability score on the test set is: 0.5030679701952991

위 Score를 기준으로 실험 진행!

Feature Engineering applprev_1

실험 1. 상관계수 0.95 이상 or 결측치 비율 90% 이상인 피처 제거



currdebt_94A : 이전 신청의 현재 부채
outstandingdebt_522A : 고객의 이전 신청에 대한 미결제 부채 금액



의미도 비슷하고, 상관계수도 높으므로
1개로 추려도 될 것 같다고 판단

currdebt_94A 결측치 : 2246512개
outstandingdebt_522A 결측치 : 2258268개



결측치가 더 많은 **outstandingdebt_522A** 제거

Feature Engineering `applprev_1`

실험 1. 상관계수 0.95 이상 or 결측치 비율 90% 이상인 피쳐 제거

Variable Name	Missing Percentage(%)
<code>revolvingaccount_394A</code>	95.45
<code>credacc_maxhisbal_375A</code>	95.10
<code>credacc_actualbalance_314A</code>	95.10
<code>credacc_transactions_402L</code>	95.10
<code>credacc_status_367L</code>	95.10
<code>credacc_minhisbal_90A</code>	95.10
<code>isdebitcard_527L</code>	92.91

결측치가 90% 이상인 피쳐들 제거

Feature Engineering applprev_1

Training LightGBM

The AUC score on the train set is: 0.7884
The AUC score on the valid set is: 0.7741
The AUC score on the test set is: 0.7695

The stability score on the train set is: 0.5533
The stability score on the valid set is: 0.5176
The stability score on the test set is: 0.5030

기준 Score

The AUC score on the train set is: 0.7804
The AUC score on the valid set is: 0.7708
The AUC score on the test set is: 0.7661

The stability score on the train set is: 0.5377
The stability score on the valid set is: 0.5102
The stability score on the test set is: 0.4938



기준 Score보다 낮은 것으로 확인

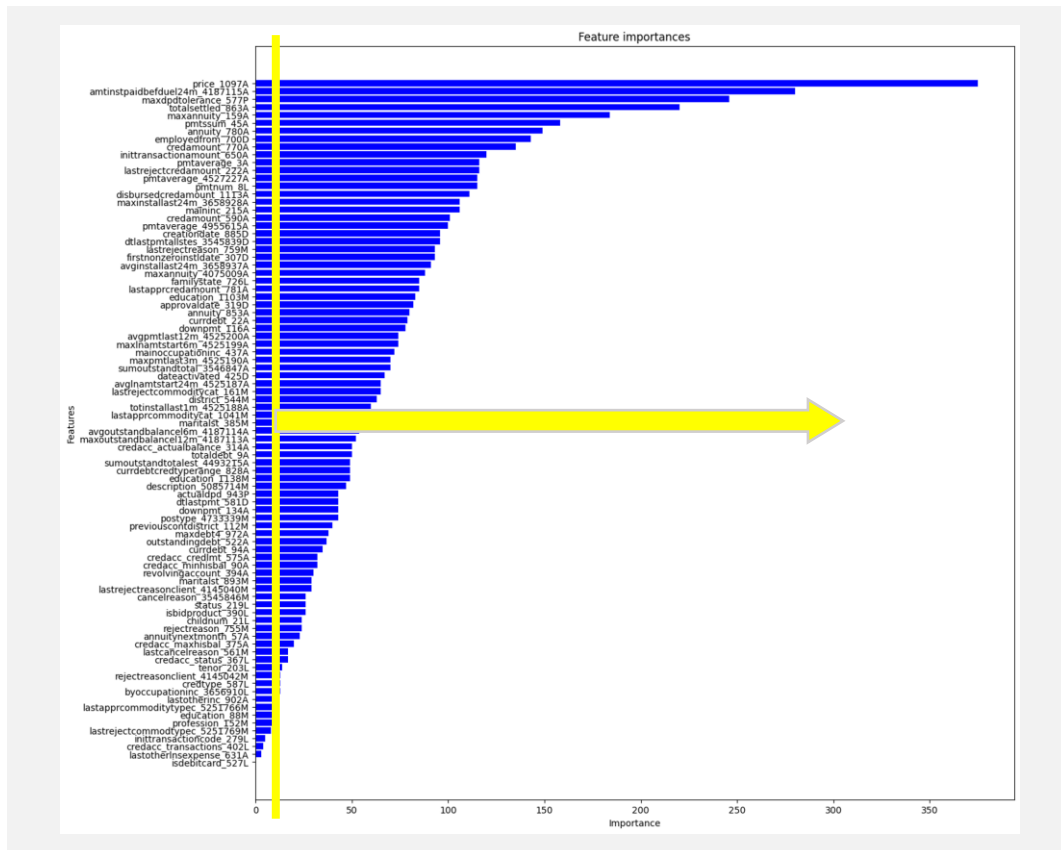
Feature Engineering applprev_1

성능 하락 예상 원인

1. 다중공선성을 염려하여 변수를 제거했으나, 모델의 성능에는 큰 영향을 미치지 않음
(참고: <https://pabii.com/news/256775/>)
2. LGBM에서 이미 결측치를 처리해주기 때문에 결측치가 많은 피처를 그냥 제거하는 것보다 다른 방법으로 처리해주는 게 더 좋을 것 같음
3. 상관계수가 높았던 **currdebt_94A**와 **outstandingdebt_522A** 중에서 결측치가 더 많은 **outstandingdebt_522A**를 제거했지만, Feature Importance에서는 **currdebt_94A**가 모델에 더 중요하게 보였음

Feature Engineering applprev_1

실험 2. Feature Importance ≥ 9 에 따른 피처 선택



Feature importance가 200, 100, 50, 25, 15, 10, 9, 8, 7, 5, 3 이상인 변수들로 실험 진행



11가지 성능 비교 결과,
가장 좋은 성능을 보인 Feature importance ≥ 9

Feature Engineering applprev_1

Training LightGBM

The AUC score on the train set is: 0.7884
The AUC score on the valid set is: 0.7741
The AUC score on the test set is: 0.7695

The stability score on the train set is: 0.5533
The stability score on the valid set is: 0.5176
The stability score on the test set is: 0.5030

기준 Score

The AUC score on the train set is: 0.7809
The AUC score on the valid set is: 0.7709
The AUC score on the test set is: 0.7666

The stability score on the train set is: 0.5382
The stability score on the valid set is: 0.5099
The stability score on the test set is: 0.4950



기준 Score보다 낮은 것으로 확인

Feature Engineering `applprev_1`

성능 하락 예상 원인

1. Feature Importance로 피처를 선택하면 계산량은 줄어든 수 있지만, 성능 향상을 보장하지는 않는 것 같음

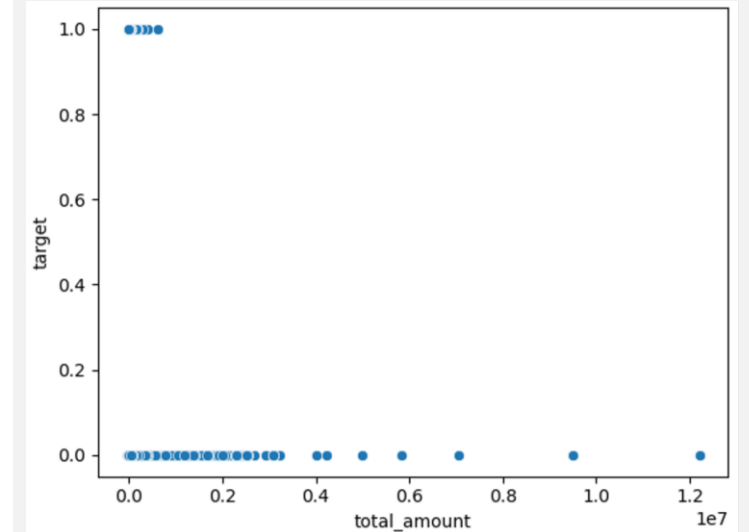
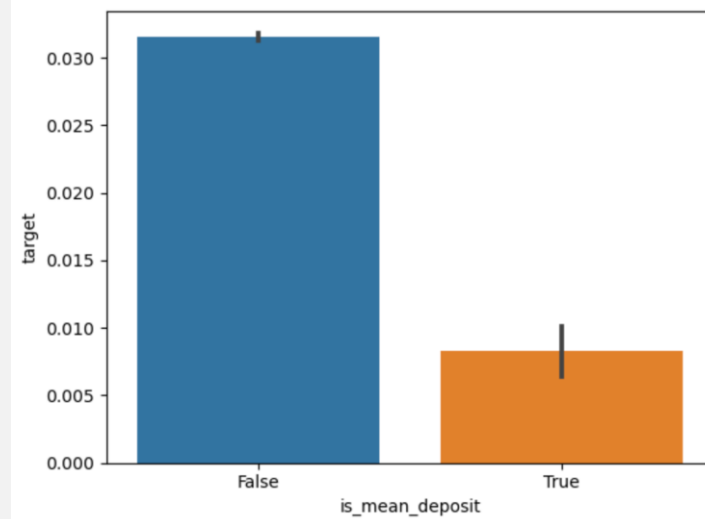
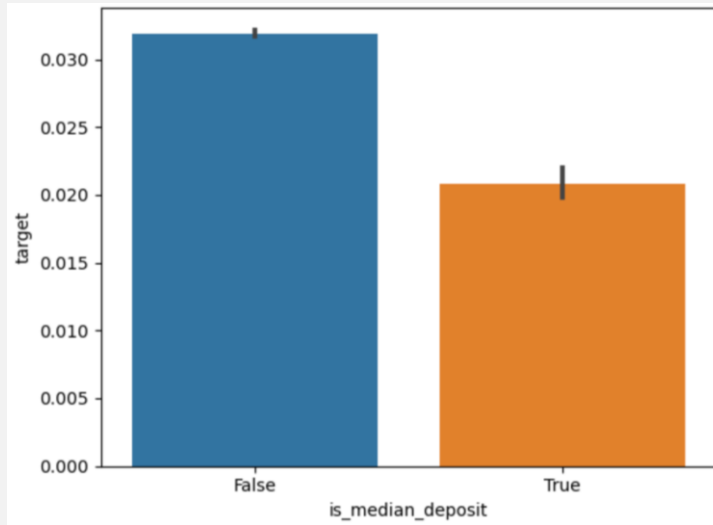
Feature Engineering deposit_1

실험 1.

amount_416A

Feature Engineering deposit_1

amount_416A 확인



예금 총액이 많을수록 대출 상환 능력이 좋을 것이다.

Feature Engineering deposit_1

Training LightGBM

The AUC score on the train set is: 0.7658777233931774

The AUC score on the valid set is: 0.7527407728341027

The AUC score on the test set is: 0.7488949126642412

The stability score on the train set is: 0.5011252123586621

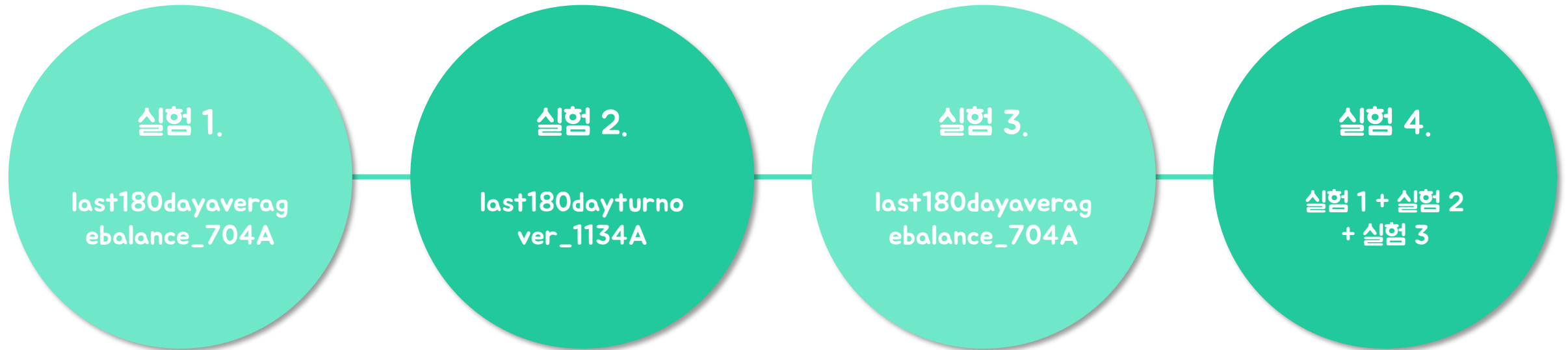
The stability score on the valid set is: 0.46959297053655136

The stability score on the test set is: 0.45844926348477866



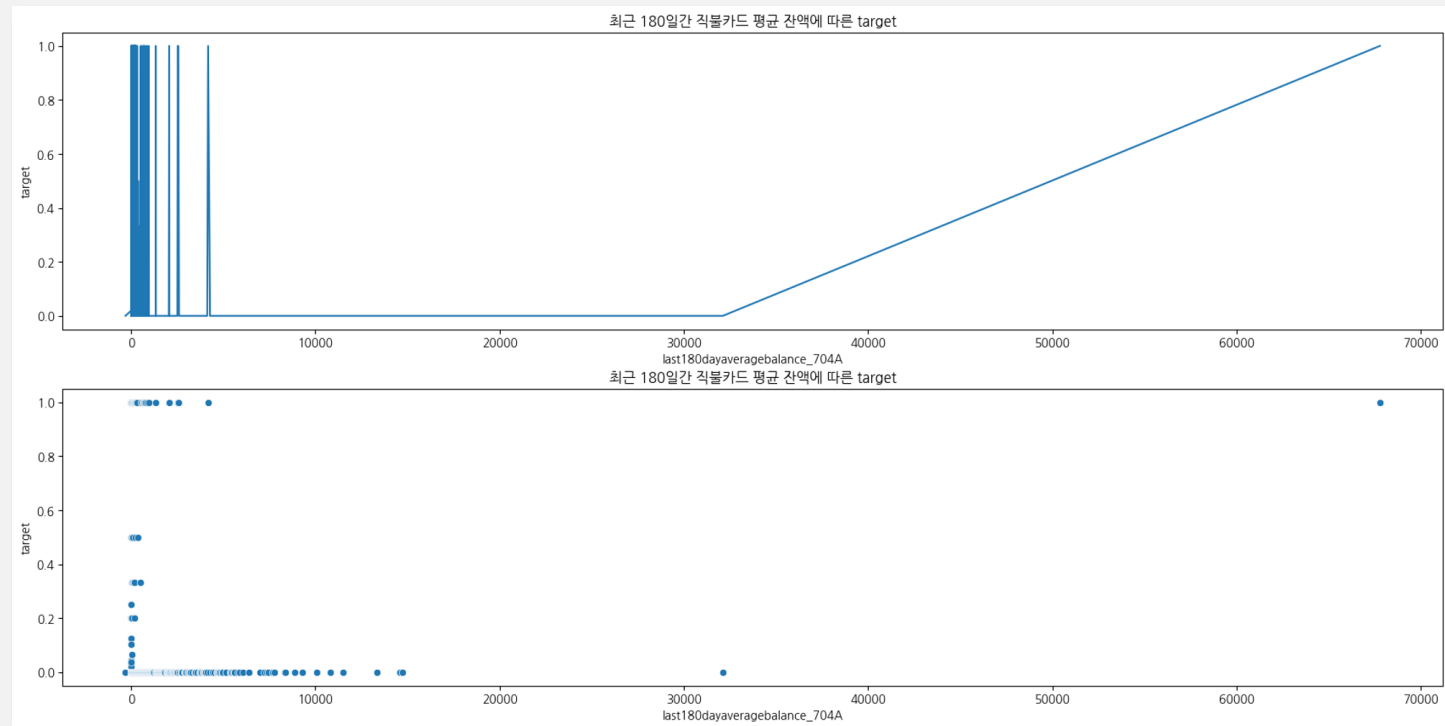
Competition Host의 성능보다 높은 것으로 확인

Feature Engineering `debitcard_1`



Feature Engineering debitcard_1

실험 1. last180dayaveragebalance_704A



최근 180일간 직불카드 평균 잔액이 높을수록 대출 상환 능력이 좋을 것이다.

Feature Engineering `debitcard_1`

최근 180일간 직불카드 평균 잔액이 높을수록 대출 상환 능력이 좋을 것이다.

case_id 별로
`last180dayaveragebalance_704A` 피처의 최댓값을 계산



`max_last180dayaveragebalance_704A` 라는
새로운 피처 추가

`num_group10`이 0인 경우 대출을 신청한 사람



`num_group10`이 0인 행만을 필터링하여 선택

Feature Engineering `debitcard_1`

Training LightGBM

The AUC score on the train set is: 0.7938575492487889

The AUC score on the valid set is: 0.7799847800484085

The AUC score on the test set is: 0.7793604701300043

The stability score on the train set is: 0.5630342059402479

The stability score on the valid set is: 0.5291618023471317

The stability score on the test set is: 0.5185566113039861

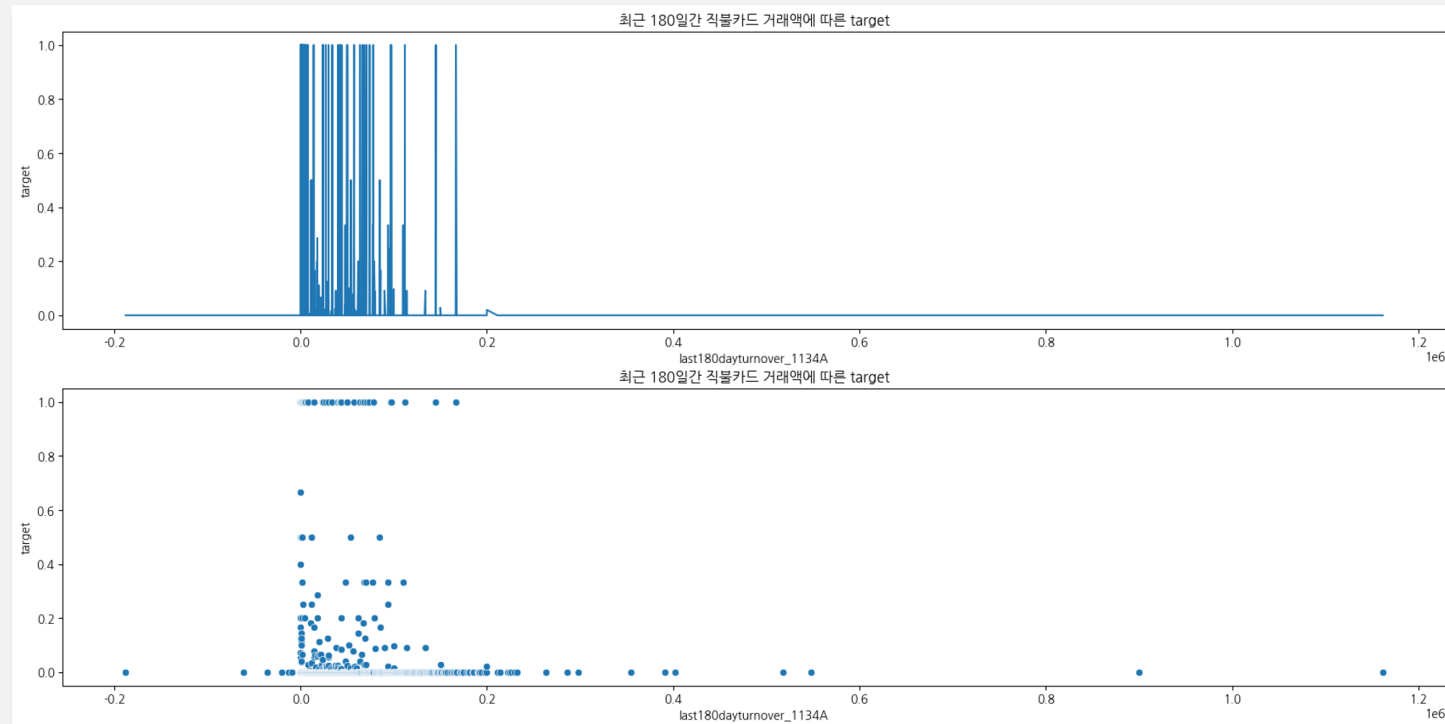
Public LB Score 0.408



Competition Host의 Public LB Score 0.36184보다 높은 0.408로 확인

Feature Engineering debitcard_1

실험 2. last180dayturnover_1134A



최근 180일간 직불카드 거래액이 높을수록 대출 상환 능력이 좋을 것이다.

Feature Engineering `debitcard_1`

최근 180일간 직불카드 거래액이 높을수록 대출 상환 능력이 좋을 것이다.

case_id 별로
`last180dayturnover_1134A` 피처의 최댓값을 계산



`max_last180dayturnover_1134A` 라는
새로운 피처 추가

`num_group10`이 0인 경우 대출을 신청한 사람



`num_group10`이 0인 행만을 필터링하여 선택

Feature Engineering `debitcard_1`

Training LightGBM

The AUC score on the train set is: 0.7903744958632996

The AUC score on the valid set is: 0.778437129806643

The AUC score on the test set is: 0.777866692209503

The stability score on the train set is: 0.5560140740734937

The stability score on the valid set is: 0.5254653892452124

The stability score on the test set is: 0.5141697796095965

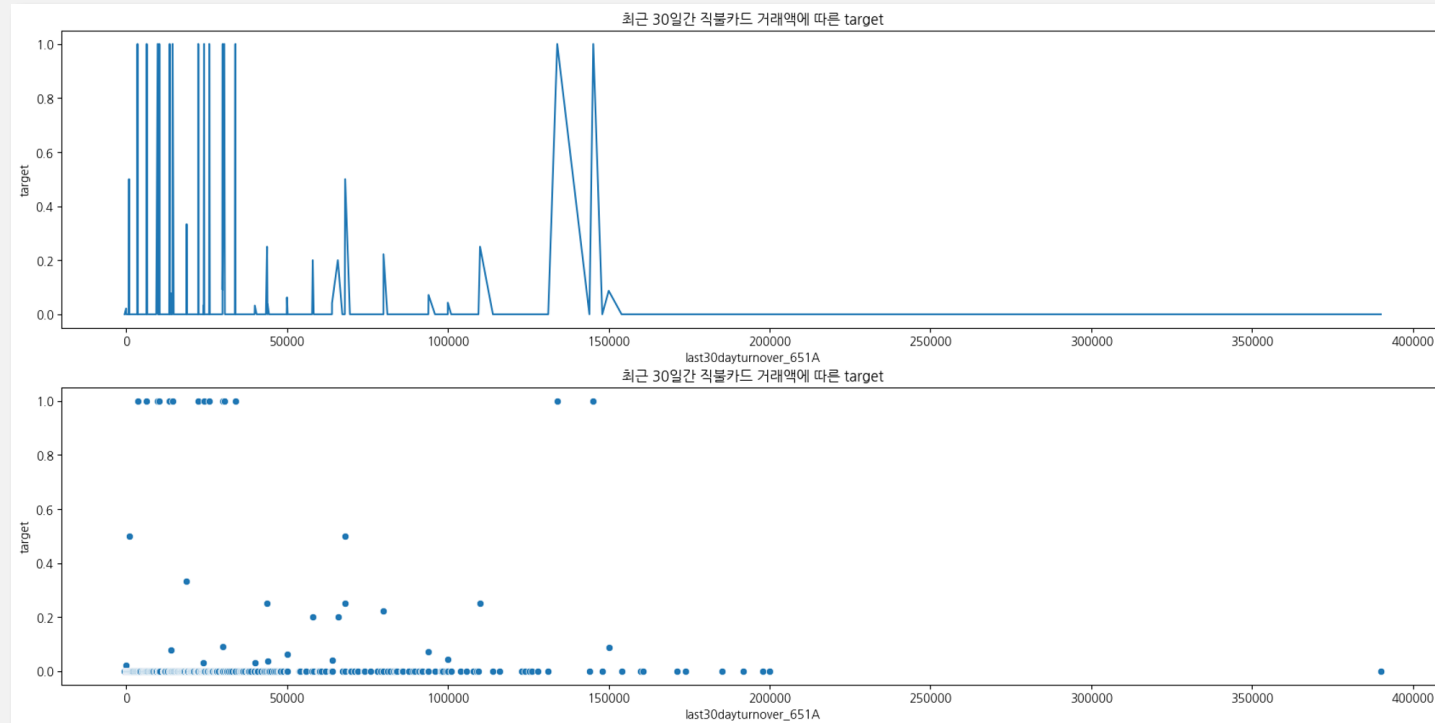
Public LB Score 0.401



실험 1의 Public LB Score 0.408보다 낮은 0.401로 확인

Feature Engineering debitcard_1

실험 3. last30dayturnover_651A



최근 30일간 직불카드 거래액이 높을수록 대출 상환 능력이 좋을 것이다.

Feature Engineering `debitcard_1`

최근 30일간 직불카드 거래액이 높을수록 대출 상환 능력이 좋을 것이다.

case_id 별로
`last30dayturnover_651A` 피처의 최댓값을 계산



`max_last30dayturnover_651A` 라는
새로운 피처 추가

`num_group10`이 0인 경우 대출을 신청한 사람



`num_group10`이 0인 행만을 필터링하여 선택

Feature Engineering `debitcard_1`

Training LightGBM

The AUC score on the train set is: 0.7951274908849022

The AUC score on the valid set is: 0.780320460563193

The AUC score on the test set is: 0.7798273371401692

The stability score on the train set is: 0.5662192390285257

The stability score on the valid set is: 0.5293071109206136

The stability score on the test set is: 0.5195888524915884

Public LB Score 0.401



실험 2의 Public LB Score 0.401과 같은 0.401로 확인

Feature Engineering debitcard_1

실험 4. 실험 1 + 실험 2 + 실험 3

The AUC score on the train set is: 0.7952969544547559

The AUC score on the valid set is: 0.7803852912343862

The AUC score on the test set is: 0.7795258905223396

The stability score on the train set is: 0.5664720960096692

The stability score on the valid set is: 0.5297959610430781

The stability score on the test set is: 0.5199922909775262

Public LB Score 0.416



실험 중 가장 높은 0.416으로 확인

Feature Engineering `credit_bureau_b_1`

실험 1.

A로 끝나는 피쳐들

Feature Engineering `credit_bureau_b_1`

A로 끝나는 피처 확인

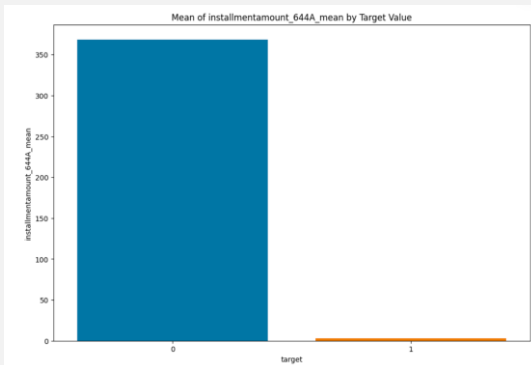
`installmentamount_644A`

`maxdebtptimevalodued_3940955A`

`overdueamountmax_950A`

`residualamount_1093A`

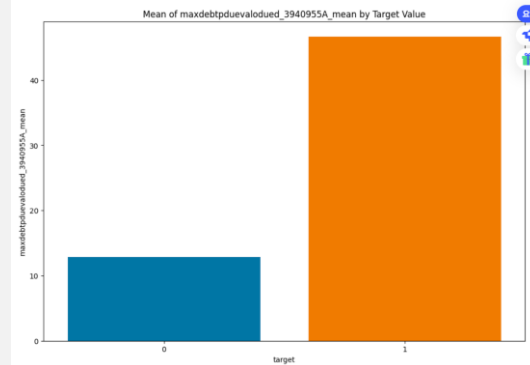
담보가 있는 신용 계약이 종료되었을 때,
고객이 지불해야 했던 할부 금액



0

1

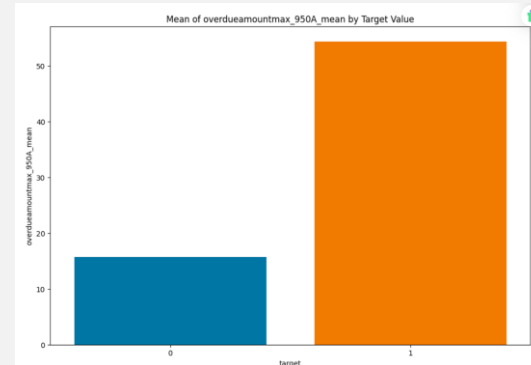
최대 부채 시점에서의 연체 일수



0

1

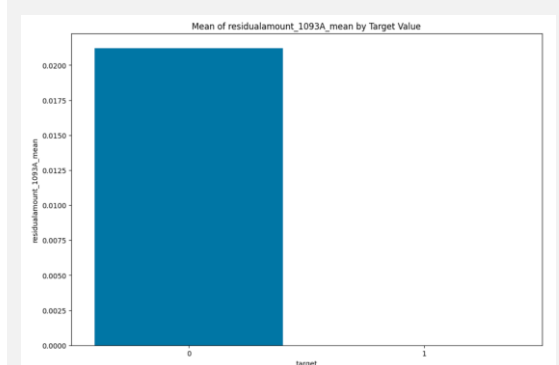
활성 상태인 계약에 대해 발생한
최대 연체 금액



0

1

보증 계약이 종료된 후에도 지불되어야 하는
잔여 금액



0

1

위 피처들 모두 1에서 더 높은 수치를 보여야 하는게 맞다고 생각했지만 그렇지 않았음
단 피처들간의 차이가 뚜렷했기 때문에 일단 모델 학습을 위해 사용함

Feature Engineering `credit_bureau_b_1`

실험 1. A로 끝나는 피쳐들

case_id 별로
`installmentamount_644A` 피쳐의 평균값을 계산



`mean_installmentamount_644A` 라는
새로운 피쳐 추가

case_id 별로
`overdueamountmax_950A` 피쳐의 평균값을 계산



`mean_overdueamountmax_950A` 라는
새로운 피쳐 추가

case_id 별로
`maxdebtptimevalodued_3940955A` 피쳐의 평균값을 계산



`mean_maxdebtptimevalodued_3940955A` 라는
새로운 피쳐 추가

case_id 별로
`residualamount_1093A` 피쳐의 평균값을 계산



`mean_residualamount_1093A` 라는
새로운 피쳐 추가

Feature Engineering `credit_bureau_b_1`

Training LightGBM

The AUC score on the train set is: 0.7959127430623194

The AUC score on the valid set is: 0.7817237975200586

The AUC score on the test set is: 0.7809793796338961

The stability score on the train set is: 0.5677708522471693

The stability score on the valid set is: 0.5324082871096333

The stability score on the test set is: 0.5234179010962219



Competition Host의 성능보다 높은 것으로 확인

Feature Engineering

tax_registry_a_1

amount_4527230A
정부 등록부에 의해 추적된
세금 공제 금액

recorddate_4527225D
세금 공제 날짜

name_4527232M
고용주 이름

tax_registry_b_1

amount_4917619A
정부 등록부에 의해 추적된
세금 공제 금액

deductiondate_4917603D
세금 공제 날짜

name_4917606M
고용주 이름

tax_registry_c_1

pmtamount_36A
신용 조사 기관 지불에 대한
세금 공제 금액

processingdate_168D
세금 공제 날짜

employername_160M
고용주 이름

- 피처의 의미가 너무나 비슷함
 - tax_registry_c_1은 Test dataset에는 존재하지 않음
- 따라서, tax_registry_a_1, tax_registry_b_1의 **세금 공제 금액**에 대해 실험을 진행함!

Feature Engineering `tax_registry_a_1, tax_registry_b_1`

실험 1.

세금 공제 금액에 대해
6가지 aggregation을
진행한 후,
성능이 가장 좋은 피처를
선택하여 성능 확인

Feature Engineering `tax_registry_a_1, tax_registry_b_1`

실험 1의 기준 $\text{Score} = \text{basetable} + \text{static_0} + \text{static_cb_0} + \text{tax_registry_a_1} + \text{tax_registry_b_1}$

The AUC score on the train set is: 0.7761621373057691

The AUC score on the valid set is: 0.764713454606537

The AUC score on the test set is: 0.7646944974248484

The stability score on the train set is: 0.4808720906573738

The stability score on the valid set is: 0.4558193297115025

The stability score on the test set is: 0.4554321441139577

위 Score를 기준으로 실험 진행!

Feature Engineering `tax_registry_a_1`

실험 1. 세금 공제 금액에 대해 6가지 aggregation 진행한 후, 성능이 가장 좋은 피처 선택

`amount_4527230A`

"median, max, min, mean, std, max - min"

6가지 aggregation 성능 비교 결과, 가장 좋은 성능을 보인 max!

The AUC score on the train set is: 0.7711752249033994

The AUC score on the valid set is: 0.7563006708463171

The AUC score on the test set is: 0.7531399784513242

The stability score on the train set is: 0.5100012488285078

The stability score on the valid set is: 0.47933573584298134

The stability score on the test set is: 0.46752867185969915

Feature Engineering `tax_registry_b_1`

실험 1. 세금 공제 금액에 대해 6가지 aggregation 진행한 후, 성능이 가장 좋은 피처 선택

`amount_4917619A`

"median, max, min, mean, std, max - min"

6가지 aggregation 성능 비교 결과, 가장 좋은 성능을 보인 std!

The AUC score on the train set is: 0.7707515209821105

The AUC score on the valid set is: 0.7541753109172304

The AUC score on the test set is: 0.7510086198127395

The stability score on the train set is: 0.5151804153351858

The stability score on the valid set is: 0.47812670958016956

The stability score on the test set is: 0.4667787972501702

Feature Engineering `tax_registry_a_1, tax_registry_b_1`

`amount_4527230A, amount_4917619A` 2가지 피처를 선택하여 성능 확인

The AUC score on the train set is: 0.7761
The AUC score on the valid set is: 0.7647
The AUC score on the test set is: 0.7646

The stability score on the train set is: 0.4808
The stability score on the valid set is: 0.4558
The stability score on the test set is: 0.4554

기준 Score

The AUC score on the train set is: 0.7716
The AUC score on the valid set is: 0.7563
The AUC score on the test set is: 0.7527

The stability score on the train set is: 0.5133
The stability score on the valid set is: 0.4797
The stability score on the test set is: 0.4689



기준 Score보다 낮은 것으로 확인

Feature Engineering `tax_registry_a_1, tax_registry_b_1`

성능 하락 예상 원인

1. 선택한 피처가 적어서 학습이 덜 된 것 같음
2. 여전히 기준 Score 보다는 낮지만, Feature Engineering을 진행하니 성능이 점점 좋아지고 있음
3. 카테고리과 날짜 피처들도 포함한다면 성능이 더 좋아질 것으로 기대됨

Feature Engineering static_0

A로 끝나는 피쳐 확인

실험 1.

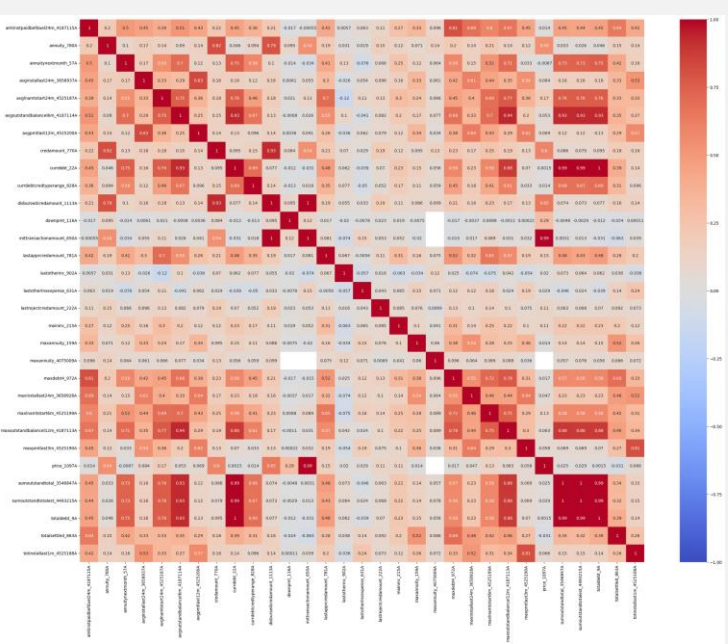
다중공선성 해결 후
성능 확인

Feature Engineering static_0

실험 1. 다중공선성 해결 후 성능 확인

Features	null_values(%)	VIF Factor_median	VIF Factor_mean
currdebt_22A	0.0003	133748.683191	131455.787029
totaldebt_9A	0.0002	133689.683697	131418.887390
sumoutstandtotal_3546847A	29.1636	90.690663	27.007894
maxoutstandbalancel12m_4187113A	50.2374	24.695424	23.277949
credamount_770A	0.0000	19.607894	20.234593

- currdebt_22A : 클라이언트의 현재 부채 금액
- sumoutstandtotal_3546847A, sumoutstandtotalest_4493215A : 총 미수금액 합계
- avgoutstandbalancel6m_4187114A : 지원자의 지난 6개월간의 평균 미수 잔액
- Corr, VIF, Description에서 수치적으로나 의미적으로 겹치는 변수임!



Feature Engineering static_0

currdebt_22A 제거 후

Features	null_values(%)	VIF Factor_median	VIF Factor_mean
sumoutstandtotal_3546847A	29.1636	90.660449	27.004763
totaldebt_9A	0.0002	87.850205	23.905561
maxoutstandbalancel12m_4187113A	50.2374	24.695307	23.277879
credamount_770A	0.0000	19.588772	20.213374
avgoutstandbalancel6m_4187114A	55.0691	17.697916	17.179364

- sumoutstandtotal_3546847A : 총 미수금액 합계
- sumoutstandtotal_4493215A : 총 미수금액 합계
- avgoutstandbalancel6m_4187114A : 지원자의 지난 6개월간의 평균 미수 잔액
- totaldebt_9A : 총 부채 금액
- Corr, VIF, Description에서 수치적으로나 의미적으로 겹치는 변수임!

Feature Engineering static_0

sumoutstandtotal_3546847A 제거 후

Features	null_values(%)	VIF Factor_median	VIF Factor_mean
maxoutstandbalance12m_4187113A	50.2374	24.640662	22.954094
credamount_770A	0.0000	19.567358	20.191911
avgoutstandbalance6m_4187114A	55.0691	17.588852	16.943869
disbursedcredamount_1113A	0.0000	17.413316	18.356587
totaldebt_9A	0.0002	9.361572	8.322158

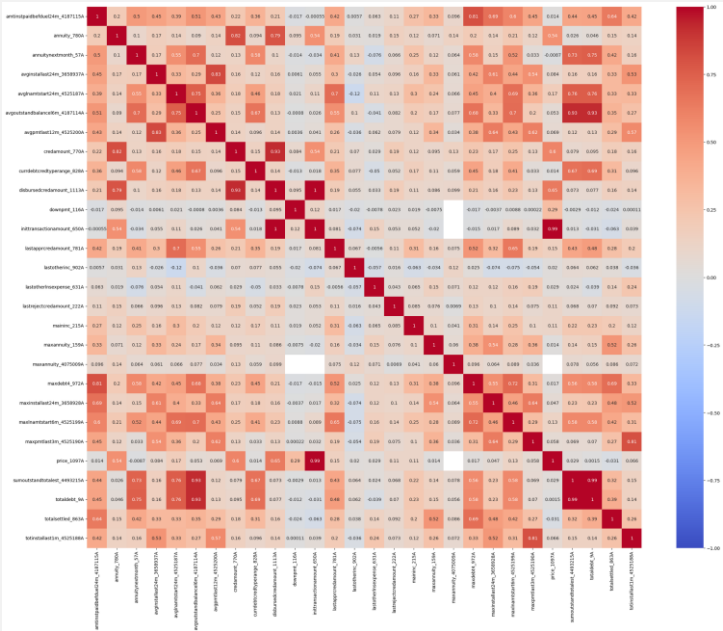
- maxoutstandbalance12m_4187113A : 지난 12개월 동안의 최대 미결제 잔액
- avgoutstandbalance6m_4187114A : 지원자의 지난 6개월간의 평균 미수 잔액
- sumoutstandtotal_4493215A : 총 미수금액 합계
- totaldebt_9A : 총 부채 금액
- Corr, VIF, Description에서 수치적으로나 의미적으로 겹치는 변수임!

Feature Engineering static_0

maxoutstandbalance12m_4187113A 제거 후

Features	null_values(%)	VIF Factor_median	VIF Factor_mean
credamount_770A	0.0000	19.566812	20.190878
disbursedcredamount_1113A	0.0000	17.413313	18.355320
avgoutstandbalance6m_4187114A	55.0691	9.135368	8.941729
totaldebt_9A	0.0002	9.097207	8.172612
annuity_780A	0.0000	8.820446	9.001235

- credamount_770A : 대출 금액 또는 신용 카드 한도
- disbursedcredamount_1113A : 통합 후 지급된 신용 금액
- Corr, VIF, Description에서 수치적으로나 의미적으로 겹치는 변수임!

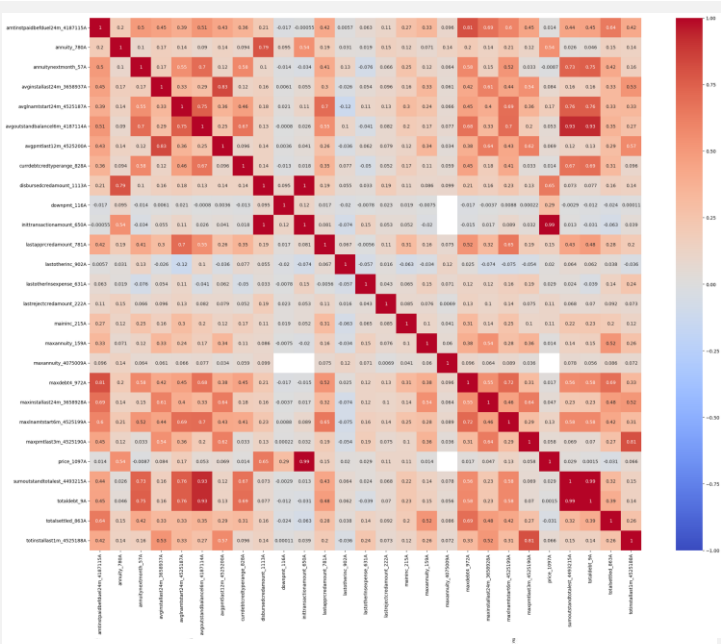


Feature Engineering static_0

credamount_770A 제거 후

Features	null_values(%)	VIF Factor_median	VIF Factor_mean
avgoutstandbalancel6m_4187114A	55.0691	9.135317	8.941537
totaldebt_9A	0.0002	9.096568	8.171602
maxdebt4_972A	20.0244	7.782633	7.186453
maxlhamtstart6m_4525199A	67.6487	7.760913	11.003957
annuity_780A	0.0000	7.626031	7.879359

- VIF Factor_median과 VIF Factor_mean이 10 이하로 떨어짐



Feature Engineering static_0

Training LightGBM

The AUC score on the train set is: 0.7202236746873704

The AUC score on the valid set is: 0.7065152028171818

The AUC score on the test set is: 0.707965962649521

The stability score on the train set is: 0.4202933062480578

The stability score on the valid set is: 0.3848337644844201

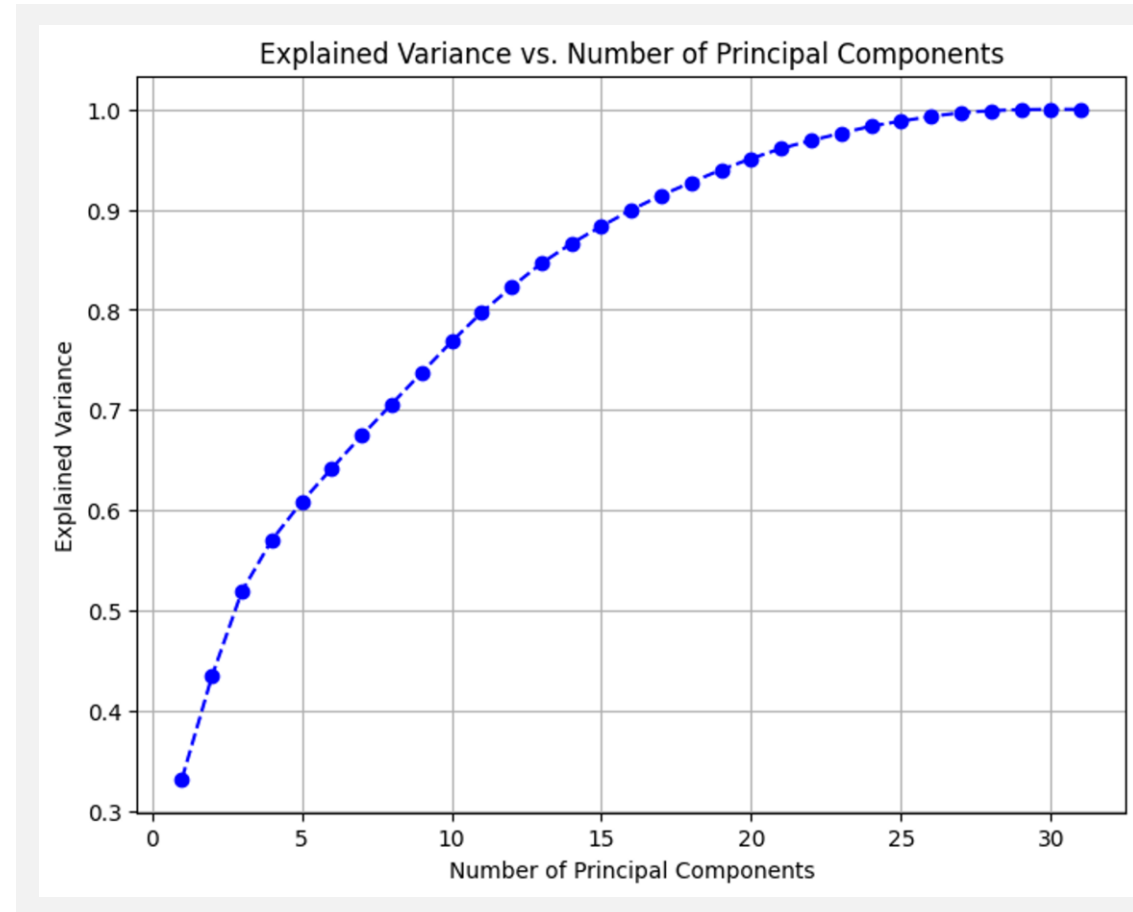
The stability score on the test set is: 0.39452605713465494



Competition Host 성능보다 낮은 것으로 확인

Feature Engineering static_0

PCA 진행



Feature Engineering static_0

Training LightGBM

누적 분산 비율 80% 이상인 경우 [component 7개]

The AUC score on the train set is: 0.6372467065811988

The AUC score on the valid set is: 0.6234002453323875

The AUC score on the test set is: 0.6217867547791558

The stability score on the train set is: 0.2515357027121558

The stability score on the valid set is: 0.21239044123991685

The stability score on the test set is: 0.20219981682074745



Competition Host 성능보다 낮은 것으로 확인

Feature Engineering static_0

Training LightGBM

eigenvalues 1 이상인 경우 [component 12개]

The AUC score on the train set is: 0.6821252533087547

The AUC score on the valid set is: 0.6638574653648868

The AUC score on the test set is: 0.6612163410959031

The stability score on the train set is: 0.3442657071221371

The stability score on the valid set is: 0.28918734332775253

The stability score on the test set is: 0.2932729648891077



Competition Host 성능보다 낮은 것으로 확인

Feature Engineering `static_0`

성능 하락 예상 원인

1. 다중공선성은 Linear Regression에서 Beta 계수를 신뢰할 수 없다는 문제점과 Tree Model에서 계산량을 늘린다는 단점이 있지만, 모델 성능이 향상되지는 않는 것 같음

Team9 Baseline

좋았던 성능의 피쳐들을 모아 하나로 합침!

The AUC score on the train set is: 0.7959298872757444

The AUC score on the valid set is: 0.7816695449396476

The AUC score on the test set is: 0.7806160802967705

The stability score on the train set is: 0.5684663870091166

The stability score on the valid set is: 0.53193349862031

The stability score on the test set is: 0.5223846969991214

 Public LB Score 0.417의 성능 달성! (+ 노트북 동메달 🏅)

하이퍼 파라미터 튜닝 `max_depth`

사용한 피처의 테이블 그룹

`applprev_1`

`other_1`

`deposit_1`

`person_1`

`debitcard_1`

`credit_bureau_b_1`

`tax_registry_a_1`

하이퍼 파라미터 튜닝 `max_depth`

The AUC score on the train set is: 0.8458365799366391

The AUC score on the valid set is: 0.8169373915904976

The AUC score on the test set is: 0.8154579894116637

The stability score on the train set is: 0.6731309416210751

The stability score on the valid set is: 0.5996755723249974

The stability score on the test set is: 0.6032876215320087

1. iteration이 다 돌 때까지 학습이 이루어졌기 때문에 iteration이 부족하다고 느낌
2. 하이퍼 파라미터 튜닝의 필요성을 확인하기 위해서 `max_depth = 4`로 바꾸어 다시 한번 학습을 진행함
3. early stopping에 의해 학습이 조기에 종료되었고 모델 성능이 향상됨
4. 극적인 성능 향상이 이루어진 건 아니지만 hyper-parameter tuning이 필요한 듯 보임

하이퍼 파라미터 튜닝 Optuna

사용한 피처의 테이블 그룹

applprev_1
other_1
deposit_1
person_1
debitcard_1
credit_bureau_b_1
tax_registry_a_1
tax_registry_b_1

하이퍼 파라미터 튜닝 Optuna

params

```
n_estimators : [100, 1000]  
  max_depth : [3, 15]  
  num_leaves : [20, 40]  
  min_child_samples : [5, 100]  
  learning_rate : [0.01, 0.3]  
  feature_fraction : [0.5, 1.0]  
  bagging_fraction : [0.5, 1.0]  
  bagging_freq : [1, 10]  
  scale_pos_weight : [15.4, 61.6]
```

하이퍼 파라미터 튜닝 Optuna

Score

	Only AUC	AUC + is_unbalanced	AUC + scale_pos_weight	Only gini	gini + is_unbalanced	gini + scale_pos_weight
AUC train	0.82737	0.84124	0.83747	0.85260	0.83927	0.82995
AUC valid	0.79911	0.80064	0.80173	0.80124	0.80093	0.80019
AUC test	0.79649	0.79885	0.79854	0.79900	0.79863	0.79660
gini train	0.63892	0.66818	0.66182	0.69507	0.66335	0.64522
gini valid	0.56946	0.57248	0.57492	0.57370	0.57209	0.57019
gini test	0.55754	0.56477	0.56066	0.56258	0.56109	0.55607

모델의 안정성을 고려하는 gini stability 에서 is_unbalanced or scale_pos_weight를 사용하는 경우
오히려 점수가 떨어짐

반면 AUC 를 최적화 metric 으로 사용하는 경우 점수가 높아짐
gini 만을 사용하는 metric 또는 AUC + scale_pos_weight 를 사용하는 metric 을 고려해봄

Q&A

Team 9

박언선 이강건

최용빈 한민규 홍재민

발표와 함께 하지 못한 실험들



Feature Engineering `other_1`



Feature Engineering other_1

other_1 피처 확인

amtdepositbalance_4809441A

amtdepositbalance_4809441A

amtdepositbalance_4809441A

target별로 피처를 평균을 냈을 때

0 : 10204.781335
1 : 2029.80945

0 : 2994.237252
1 : 1449.899868

0 : 3650.970136
1 : 1443.536012

고객별 예금 계좌 잔액과 입출금 금액에 대하여 target별로 나눠서 확인했을 때
target에 따라 차이가 어느 정도 차이가 존재함

Feature Engineering other_1

실험 1. amtdepositbalance_4809441A

case_id 별로
amtdepositbalance_4809441A 피처의 최댓값을 계산
↓
max_amtdepositbalance_4809441A 라는
새로운 피처 추가

Feature Engineering other_1

Training LightGBM

The AUC score on the train set is: 0.764122917660593

The AUC score on the valid set is: 0.7512157223309048

The AUC score on the test set is: 0.7483072129459662

The stability score on the train set is: 0.4976648127691175

The stability score on the valid set is: 0.4726726686264489

The stability score on the test set is: 0.4583643686935092



Competition Host의 성능보다 높은 것으로 확인

Feature Engineering other_1

실험 2. amtdepositincoming_4809444A, amtdepositoutgoing_4809442A

case_id 별로
amtdepositincoming_4809444A 피처의 최댓값을 계산



max_amtdepositincoming_4809444A 라는
새로운 피처 추가

case_id 별로
amtdepositincoming_4809442A 피처의 최댓값을 계산



max_amtdepositincoming_4809442A 라는
새로운 피처 추가

Feature Engineering other_1

Training LightGBM

The AUC score on the train set is: 0.7641
The AUC score on the valid set is: 0.7512
The AUC score on the test set is: 0.7483

The stability score on the train set is: 0.4976
The stability score on the valid set is: 0.4726
The stability score on the test set is: 0.4583

실험 1 Score

The AUC score on the train set is: 0.7677
The AUC score on the valid set is: 0.7529
The AUC score on the test set is: 0.7497

The stability score on the train set is: 0.5083
The stability score on the valid set is: 0.4750
The stability score on the test set is: 0.4614

↑ 실험 1의 성능보다 높은 것으로 확인

Feature Engineering other_1

실험 3. other_1에 있는 모든 feature의 최댓값

The AUC score on the train set is: 0.7677
The AUC score on the valid set is: 0.7529
The AUC score on the test set is: 0.7497

The stability score on the train set is: 0.5083
The stability score on the valid set is: 0.4750
The stability score on the test set is: 0.4614

실험 1 Score

The AUC score on the train set is: 0.7657
The AUC score on the valid set is: 0.7517
The AUC score on the test set is: 0.7489

The stability score on the train set is: 0.5027
The stability score on the valid set is: 0.4722
The stability score on the test set is: 0.4595

↓ 실험 2의 성능보다 낮은 것으로 확인

Feature Engineering `person_1`

실험 1.

일부 피쳐들만 선택

Feature Engineering `person_1`

실험 1. 일부 피쳐들만 선택

- `language1_981M` : 데이터가 균형적으로 분포
- `incometype_1044T` : 어떤 방식으로 돈을 버는지에 관한 데이터로 대출상환 능력을 나타낼 수 있는 지표일거라 판단
- `familystate_447L` : 보통 본인이 못갚으면 가족이 도와주거나 할 수도 있기 때문에 안정성에 영향을 미칠거라 판단
- `role_1084L` : 계약 `role`에 관한 데이터

Feature Engineering `person_1`

Training LightGBM

The AUC score on the train set is: 0.7736779050977972

The AUC score on the valid set is: 0.7590182535259641

The AUC score on the test set is: 0.7557274831181354

The stability score on the train set is: 0.5219091595835447

The stability score on the valid set is: 0.4860596382141017

The stability score on the test set is: 0.47811888561192684

Competition Host의 Score보다 낮은 것으로 확인