



신용평가 모델 개선을 통한 연체율 감소와 매출 증대

1. 추진배경
2. 현상 및 개선기회
3. 분석계획 및 결과
4. 개선안 및 적용방안
5. Learned Lesson

A1조

김도영(A) 김지은(A) 송민영 신어진 오진성 조유경 최수빈

추진배경

가계 대출 증가에 따른 금융당국의 대출규제로 인해 금융 취약계층이 제2금융권 유입
무차별적 카드발급과 다중채무자, 악성 연체자 비중이 늘어나며 자사 연체율 증가

금융당국 가계대출규제

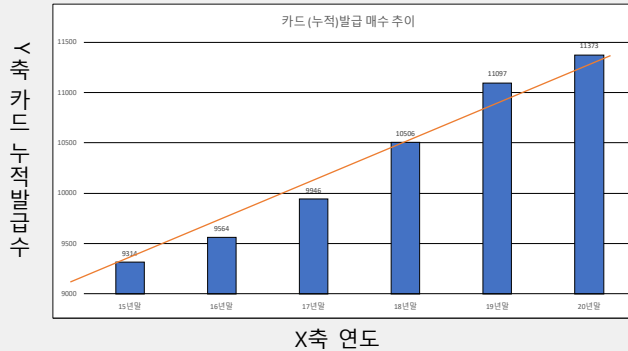
금융 당국의 가계대출 규제

- 총부채원리금상환비율(DSR) 관리 지표 시행
- 가계대출에 15% 가중치 부여하는 예대출 도입
- 업권별 가계부채 총량 규제

[출처] 금융감독원

상환능력이 낮은 대출수요자들이 제2금융권(카드사 등)으로 유입

카드사 경쟁과다



[출처] 한국은행

카드시장의 급속한 성장 및 과당경쟁으로 인한 무차별적 카드발급

카드론/리볼빙 사용자 증가

리볼빙 이월잔액

(단위: 억 원, 8개 전업카드사 기준)



[출처] 이투데이_금융감독원

연체 위험성 높은 다중채무자 증가, 잠재 부실 규모 증가

“신용점수 평가모델 개선을 통한 연체율 감소와 수익성 증가”

현상 및 개선기회

최종 목표 연체율 감소를 통한 약 500억 수익 증대

*현재 평가시점 기준으로 연체율 산출

KPI	연체율	현재	30%
		목표	25%

$$\text{연체율} = \frac{\text{총 연체액} \downarrow}{\text{총 대출액} \uparrow}$$

고위험군 유입 제재
저위험군 유입 유도

고객 분류 기준
신용평가체계

현상 및 개선기회

• 차주 : 돈을 빌리는 사람
* 싼파일러(Thin filer) : 금융정보가 적거나 없는 고객

	1. 등급제 → 점수제	2. 평가기준항목 재구성	3. 싼파일러 고객 확보
현상	등급 간 경계에 가까운 점수대 고객 분별력 ↓	차주의 대출상환능력 정확한 평가 X	비금융정보만 존재하는 싼파일러 고객 대응 X
개선 기회	등급 구간이 세분화되지 않음	현 모델의 평가 성능이 떨어짐	금융정보 위주로 구성되어 있음
개선 방향	신용평가 기준 범주형(등급제) → 연속형(점수제)	평가기준 항목과 항목별 가중치 변경	비금융정보를 이용해 싼파일러 중 고위험/저위험군 구분

잠재원인도출 및 우선순위 판별

잠재원인도출

연체율 증가

고객 특성

- 주거 정보
- 보유 재산정보
- 소득정보
- 인적사항

신용평가모델

- 상환이력정보
- 부채수준
- 신용거래 기간
- 신용형태정보

금융 리스크

- 리볼빙 이용금액
- 카드론 이용금액
- 신규 카드 발급수
- 악성 연체자수 및 비중
- 다중채무자수 및 비중

사회적 요인

- COVID-19 장기화에 따른 경기부진
- 저금리기조
- 대출 지원 정책
- 이자율에 따른 연체율 증가

잠재원인 우선순위

구분	상세 잠재원인	중요도	분석가능성	합계	선정
고객특성	주거정보	9	9	18	O
	보유 재산정부	9	9	18	O
	소득정보	9	9	18	O
	인적사항	9	9	18	O
신용평가모델	상환이력	9	9	18	O
	부채수준	9	9	18	O
	신용거래 기간	9	9	18	O
	신용형태정보	9	9	18	O
금융리스크	리볼빙/카드론이용여부	9	3	11	X
	신규카드 발급 수	3	3	6	X
	악성 연체자수	9	3	11	X
	다중채무자수	9	3	11	X
사회적요인	COVID-19장기화에 따른 경기부진	3	3	6	X
	저금리 기조	3	3	6	X
	대출지원정책	1	1	2	X
	이자율 따른 연체율 증가	1	1	2	X

* 9점 척도 기준

목적 및 분석계획

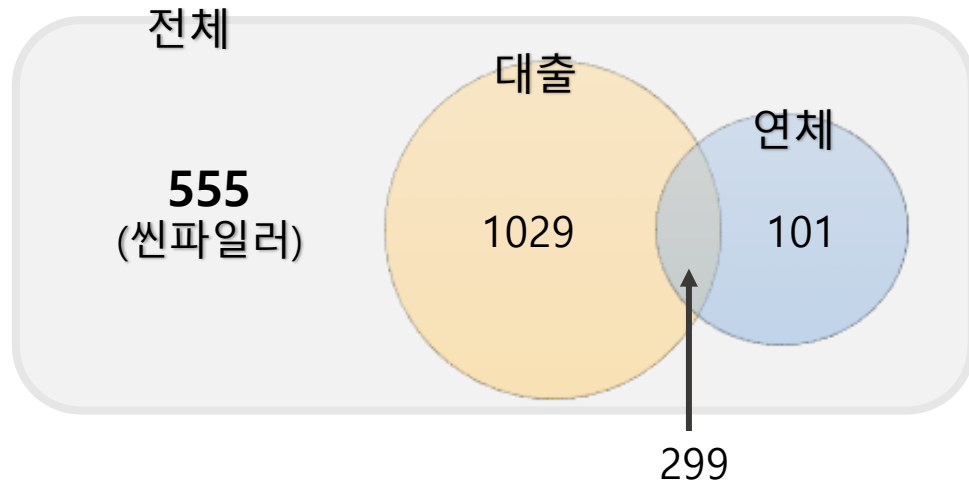
목적	분석방법	분석 내용
데이터 정제	<ul style="list-style-type: none"> 막대그래프 Boxplot XGBoost 	<ul style="list-style-type: none"> 명목형 변수 개수 확인 결측치 확인 및 연속형 변수 분포 확인 '무응답/모름' 결측치 처리
탐색적 분석	<ul style="list-style-type: none"> 히스토그램 막대그래프 원그래프 	<ul style="list-style-type: none"> 명목형 변수 개수 및 수익성 증가 확인 연체율과 다양한 x변수의 경향성 확인 중요 카테고리의 비율 확인
새로운 신용평가모델 제작	<ul style="list-style-type: none"> Decision Tree/ Random Forest / XGBoost Linear Regression 	<ul style="list-style-type: none"> 평가모델 제작 시 중요한 인자 추출 회귀계수 값을 이용하여 가중치 설정
잠재 고객 분류	<ul style="list-style-type: none"> XGBoost 	<ul style="list-style-type: none"> 실판일러 저위험군 분류
신용점수 조회서비스 개발	<ul style="list-style-type: none"> WEB 	<ul style="list-style-type: none"> 고객별 신용점수 조회서비스와 자사 대상 고객 데이터 관리 서비스 개발

데이터 개요 및 주요 정제

데이터 개요

Data	Description	Variables	Rows
차주	차주 정보 (연령, 성별, 직업군 등)	17	1,984
카드	차주별 카드발급 정보 (기관,종류,발급년월 등)	9	96,828
대출	차주별 대출정보 (기관,종류,대출년월 등)	11	58,775
연체	차주별 연체정보 (기관,종류,연체년월/사유 등)	11	10,784

차주 분포

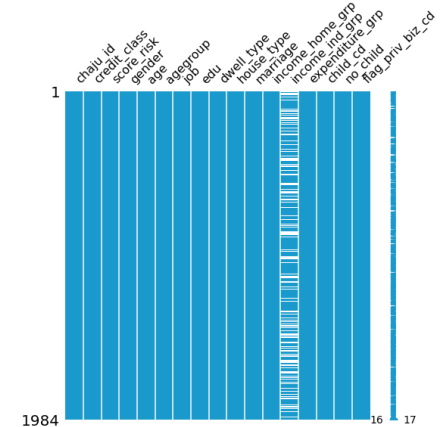


데이터 정제

개인 평균 수입(월) '모름/무응답'은 총 30.4% (응답자 608명/전체 1984명)

▶ 분류 모델링 기법 활용

사용모델	정확도
Xgboost	81%
Gredient Boosting	78%
Random Forest	72%
인공신경망	72%
KNN	69%
Decision Tree	57%



▶ 처리 결과

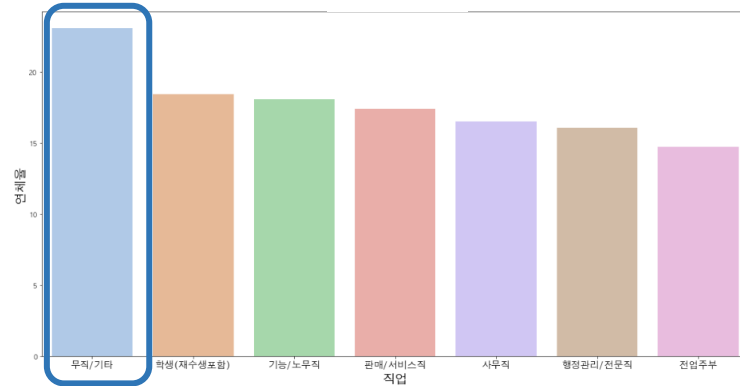
income_ho	income_in	expenditur
900-999만	300-399만	100-199만
200-299만	모름/무응답	99만원 이하
500-599만	500-599만	100-199만
500-599만	모름/무응답	99만원 이하
500-599만	500-599만	200-299만

→

income_ind_grp
300-399만원
500-599만원
500-599만원
100-199만원
300-399만원
100-199만원

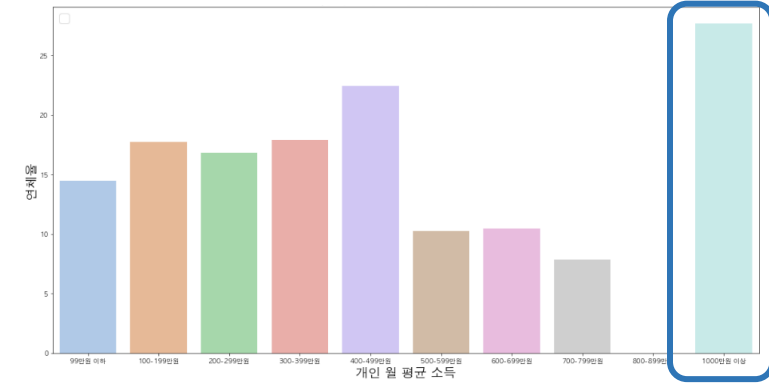
분석 결과_차주 특성 파악

• 직업군별 평균 연체율



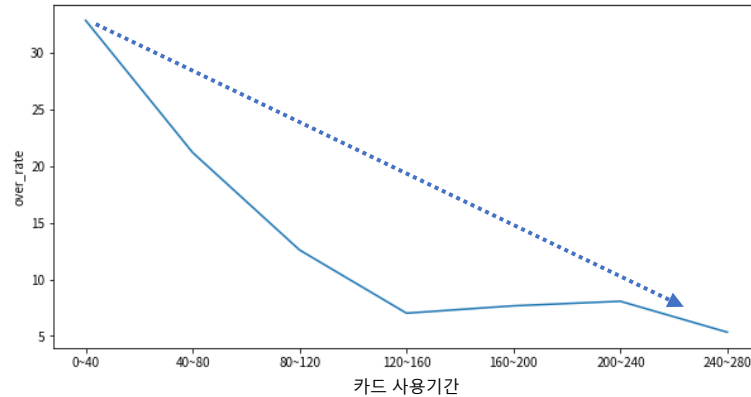
'무직/기타' 응답자의 평균 연체율 높음

• 개인 월평균 소득별 연체율



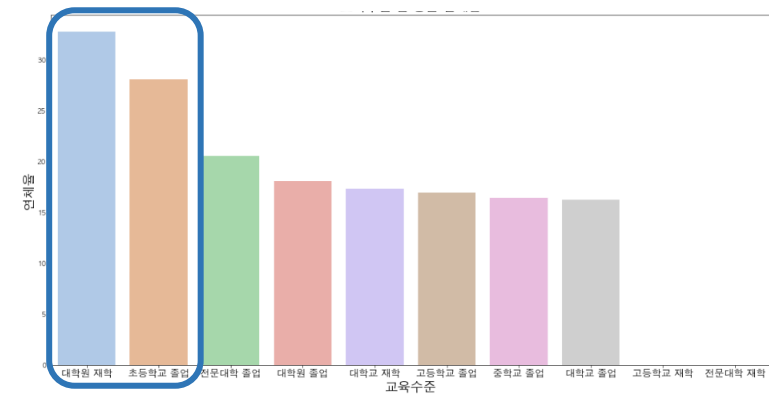
월 평균 1,000만원 이상 소득 연체율 높음

• 카드 사용기간 평균 연체율



카드 사용기간이 짧을수록 평균 연체율 높음

• 교육수준별 연체율



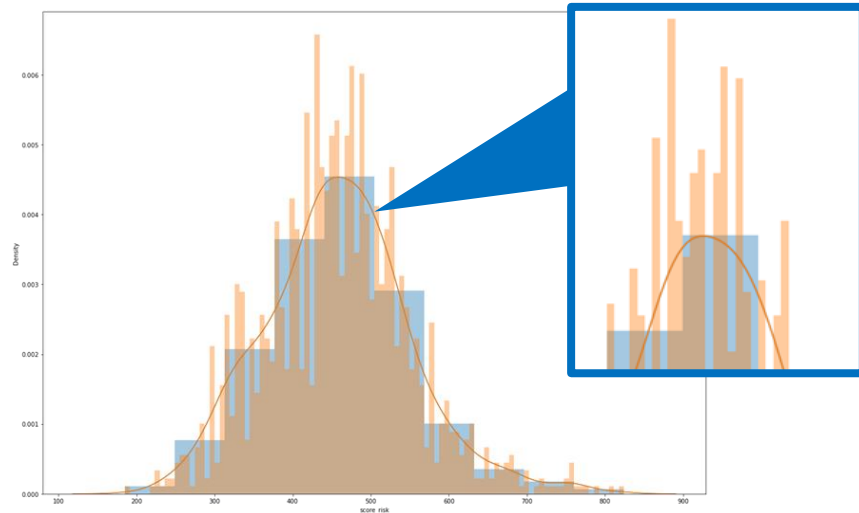
대학원 재학, 초등학교 졸업 평균 연체율 높음

분석 결과_신용점수 평가모델

분석결과 1

등급 내 세분화된 차주 분류 불가

등급으로 신용도 분류 시,
상이한 차주의 신용점수를 반영하지 못함



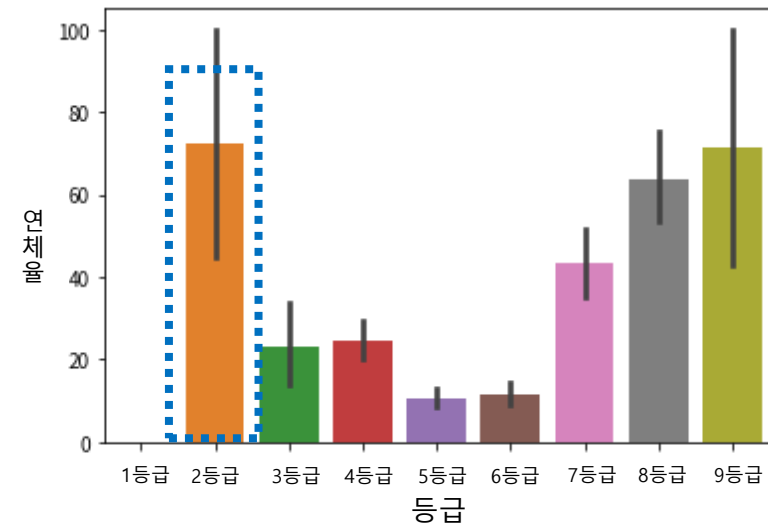
[등급과 점수별 차주 분포]

점수제 전환을 통해 세분화된 차주 구분이 필요함

분석결과 2

기존 평가모델에서 등급 내 이상분포 파악

현재 신용평가모형 기준,
2등급의 평균 연체율이 60%이상으로 비정상적 수치

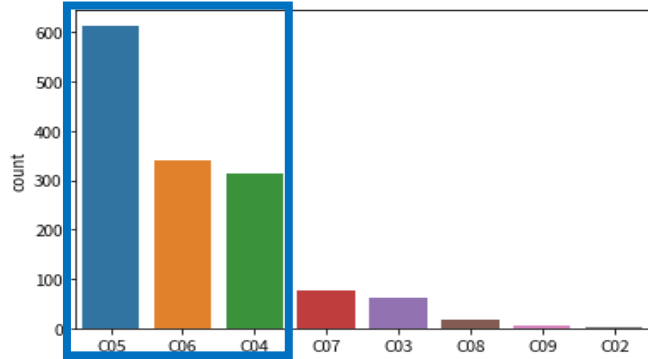


개인의 신뢰도와 관련된 금융 및 비금융정보를
평가 항목기준으로 채택할 예정

“차주 개별특성을 반영하는 새로운 신용점수 평가모델 개발”

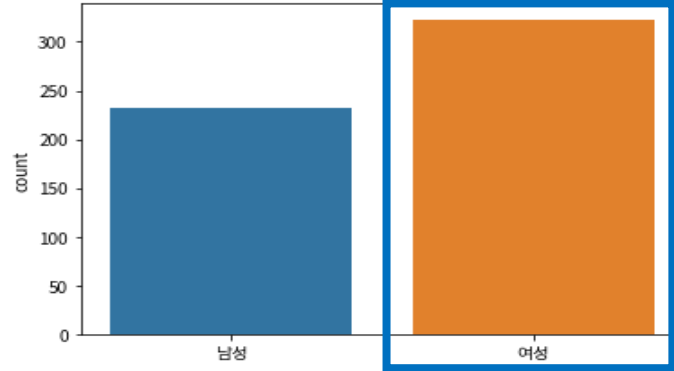
분석 결과_썬파일러 고객 분석

• 등급별 분포



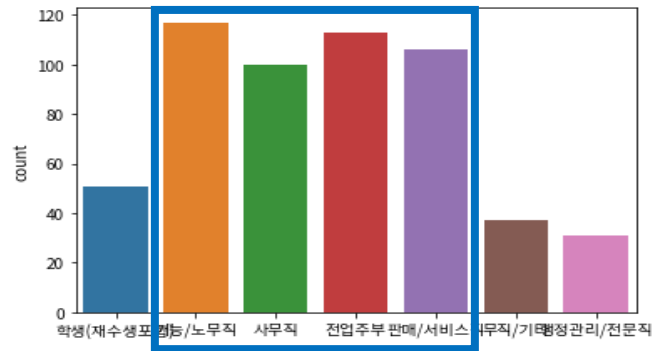
4~6등급에서 분포 높음

• 성별 분포



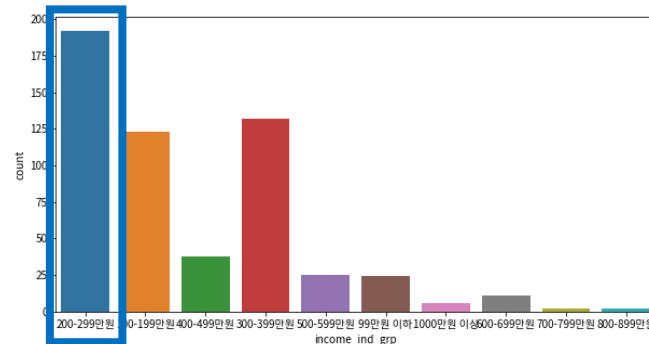
여성이 분포 높음

• 직업별 분포



기능/노무직, 전업주부, 판매/서비스, 사무직에서 분포 높음

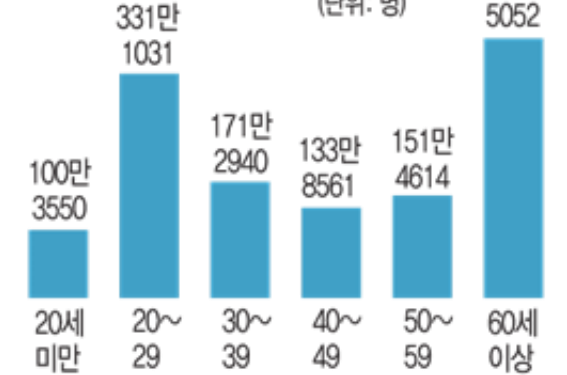
• 월평균 수입 분포



월평균 수입 200-299만원 분포 높음

연령대별 금융이력 부족자 현황

(단위: 명)



*2020년 6월 말 기준 (자료:나이스(NICE) 평가정보)

썬파일러는 매출을
증가시킬 수 있는 주요 잠재고객

→ 썬파일러 고객에 대해
더욱 심도 깊은 분석이 필요

개선안1. 신용평가 모델 등급제 → 점수제 전환

개선 사례



차주 A
신용점수 699점

개선 전

개선 후

4등급
(600~699점)

기존 4등급과
비슷한 수준

대출가능
금액

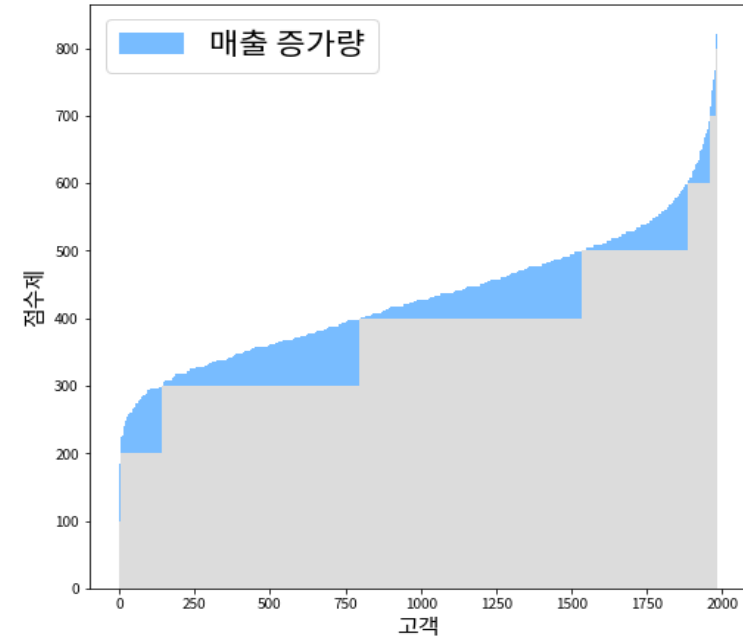
1,000만원 → 2,000만원

개선효과

[차주]
만족도 증가
금리 부담 감소

[자사]
정확한 차주 분석
리스크 방지

점수제 전환에 따른 매출 증가량



매출 기대 증가량 = 11.65 %

고위험-저수익군

낮은 점수부여 / 리스크(높은 이자율)

저위험-고수익군

높은 점수부여 / 혜택(낮은 이자율)

개선안2. 신용평가모델 평가항목 재구성

개선된 평가항목

구분	상세구분	항목
비금융 정보	차주특성	나이
		직업
		주택 형태
		주거 형태
		가계 월평균 소득
		개인 월평균 지출
		개인사업자 여부
금융 정보	카드 개설	카드 사용기간
	대출	대출 건수
		대출 금액(합계)
		대출 기간
	연체	연체 건수
		연체 금액(합계)
		연체 기간

유의한 금융 및 비금융정보를 평가항목에 포함

모델링

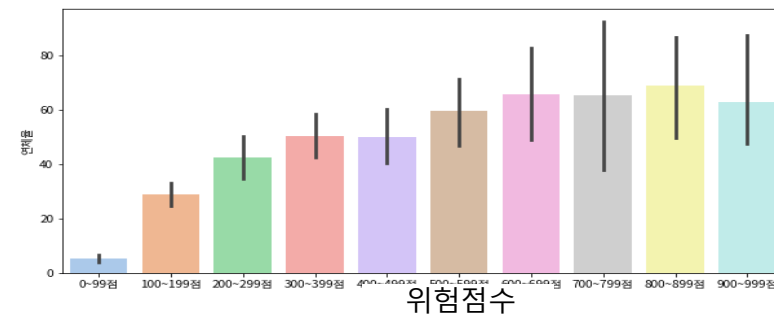
모델 기법

Decision Tree
Random Forest
Gradient Boosting
XG Boost

- 예측 모델링을 통해 주요 변수 도출 후 가중치 부여
- 변수 중요도 확인이 불가능한 변수의 경우, EDA에 근거하여 선정함

카드 개설 시 즉각 신용평가체제에 반영
정확도 상승 및 추후 신용평가에 사용

점수 구간별 평균 연체율

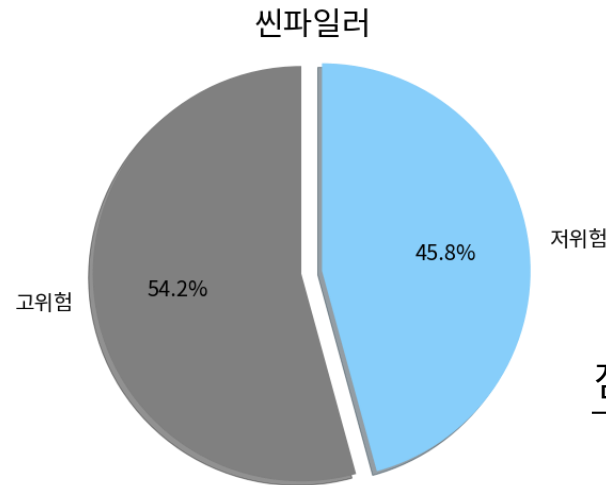


- ✓ 위험점수가 증가함에 따라 평균 연체율이 증가
→ 신용평가 모델의 성능이 개선됨

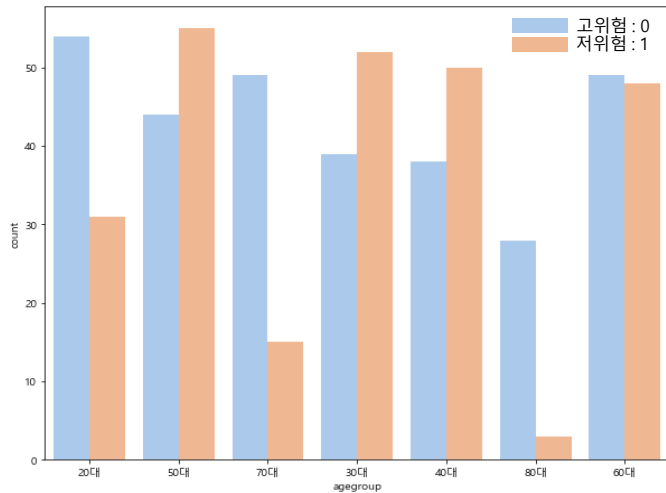
개선안3. 비금융정보를 활용한 실파일러 고객 분류

실파일러 고객 분류

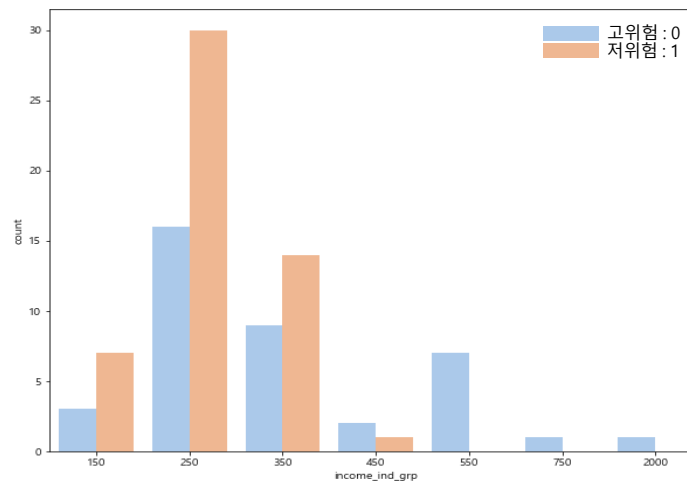
- XG Boost 모델링을 활용한 '저위험 - 고수익' 고객 분류
- 전체 실파일러 555명 중, '저위험 - 고수익'인 254명의 잠재고객 발굴



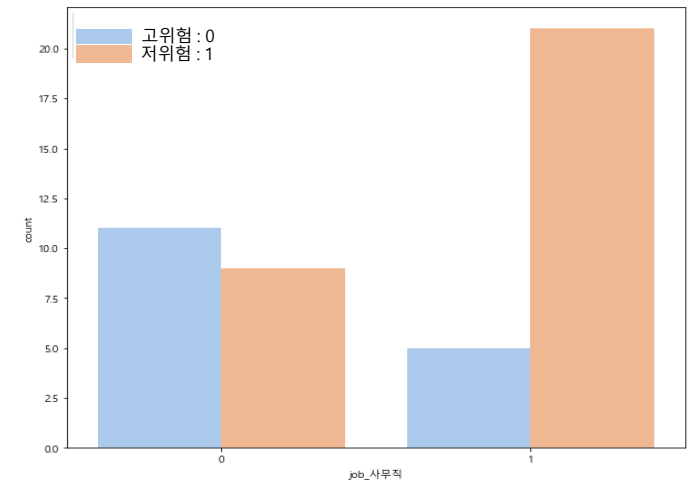
잠재고객 (저위험 고수익) 254명
실파일러 총 555 명



▶ 연령대별 저위험/고위험 고객 분류



▶ 개인 월수입 평균이 250만원 이상



▶ 30대 사무직이 저위험군 많음

KPI 달성 여부

S05-01016사 대출금 총액 : 381360000.0
 S05-01016사 연체금 총액 : 119460000.0

 S05-01016사 연체율 : 31.32%

신용점수 400점 이하를
 고위험군으로 책정 후
 본사의 연체율을 재도출

S05-01016사 대출금 총액 : 369790000.0
 S05-01016사 연체금 총액 : 88160000.0

 S05-01016사 연체율 : 23.84%

- ✓ 당사 연체율 약 7% 감소
- ✓ 700억 가량의 매출 증가 기대

발전 방향

신용점수 평가모델 고도화



통신사, '배달의 민족' 등
 고객의 삶에 밀접한 기업과의 제휴

차주의 소비 패턴 등 다양한
 비금융적 정보를 얻음

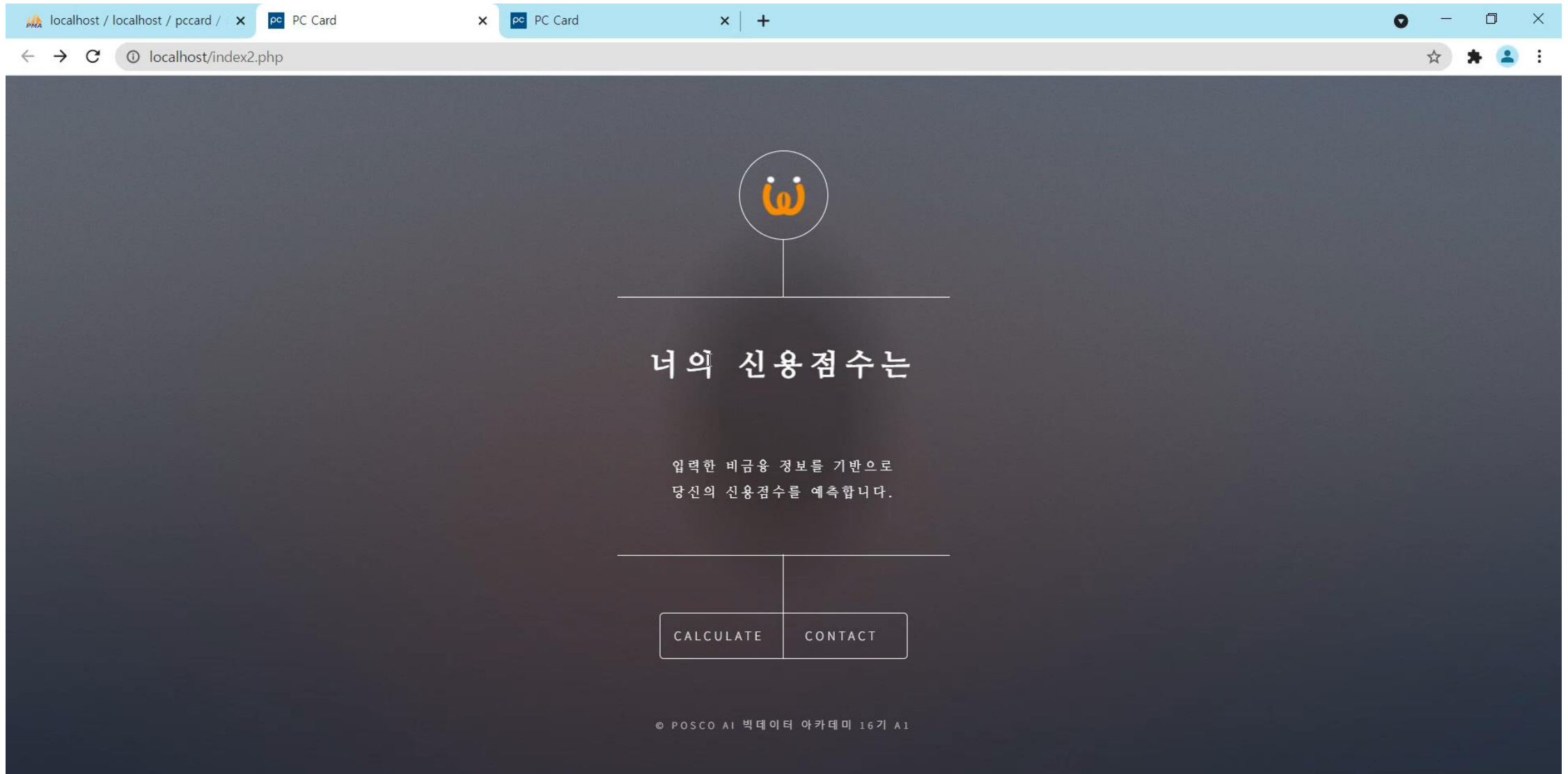
마이데이터 활용



마이데이터를 활용해 차주 정보 확보

- ✓ 고객에게 제휴 서비스를 제공함으로써 차주 만족도 증가 및 추가 유치 가능
- ✓ 특히, 씬파일러 차주들을 평가할 때 비금융지표로 활용될 수 있음

4.3 WEB



5. Learned Lesson



- 김도영

빅데이터 분석 프로젝트를 하며 결측치 대체 부분과 데이터 통합부분을 깊게 연구하며 결측치를 어떻게 채우느냐에 따라 모델에 중요한 피처가 될 수 있다는 것을 느꼈고 데이터에 필요한 도메인 지식이 매우 중요하며 상황에 맞는 모델을 사용하는 것이 중요하다는 것을 느꼈습니다.



- 송민영

학술적인 측면에서 데이터 전처리와 도메인 지식의 중요성을 깨달을 수 있었을 뿐만 아니라, 모든 팀원이 하나의 목표를 바라보고 함께 문제를 해결해 나가는 과정을 통해 협업을 배울 수 있었습니다.



- 최수빈

실제 현업의 데이터를 이용하여 분석을 진행하는 경험이 흔치 않아 새로운 경험이었습니다. 잘 모르는 금융분야를 처음 접하여 배경지식을 익히는 것에 어려움이 있었지만, 다양한 방법으로 인사이트를 얻으면서 해결할 수 있어 뜻깊었습니다. 또한 프로젝트를 진행하면서 데이터 전처리 부분이 가장 중요하다는 점을 깨닫게 되었습니다.



- 오진성

금융 데이터를 다루고 모델을 구현해보는 과정을 통해 많은 공부를 할 수 있었습니다. 데이터 정제를 통한 데이터 품질의 중요성에 대해 배울 수 있었고, 협업을 경험할 수 있는 소중한 기회였습니다.



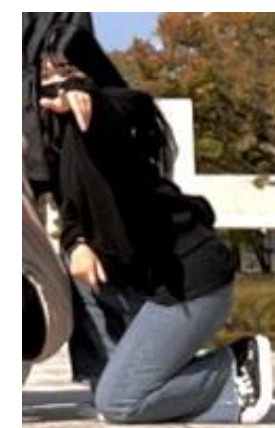
- 김지은

100을 준비했지만 1만 보여주는 기분이라 아쉬움이 있지만 새로운 것을 배우는 즐거움과 팀원들과 서로 격려하며 협력했던 모든 과정이 보람차고 즐거웠습니다.



- 조유경

실무 프로젝트를 통해서 배웠던 것들을 실제로 활용하는 방법을 배웠고, 동료들과 협업하여 프로젝트를 진행하는 경험을 할 수 있었습니다.



- 신어진

이번 프로젝트를 통해 빅데이터 과정에 배운 다양한 기법들을 실제 사례에 적용해 보는 기회를 갖게 되어 매우 뜻깊었던 경험이었던 것 같습니다.