

AI_12_3팀 CP1 기획서

서론

• 프로젝트 목적

1. (상황 가정) 다음 분기에 신용 우수 고객 대상으로 기획 중인 새로운 신용카드 혜택에 참고할 인사이트 발견
2. 신용카드 이용자의 신용 등급을 예측하여 신용 등급에 맞는 신용카드 추천 시스템 구축

• 프로젝트 배경

◦ 데이터 분석 측면

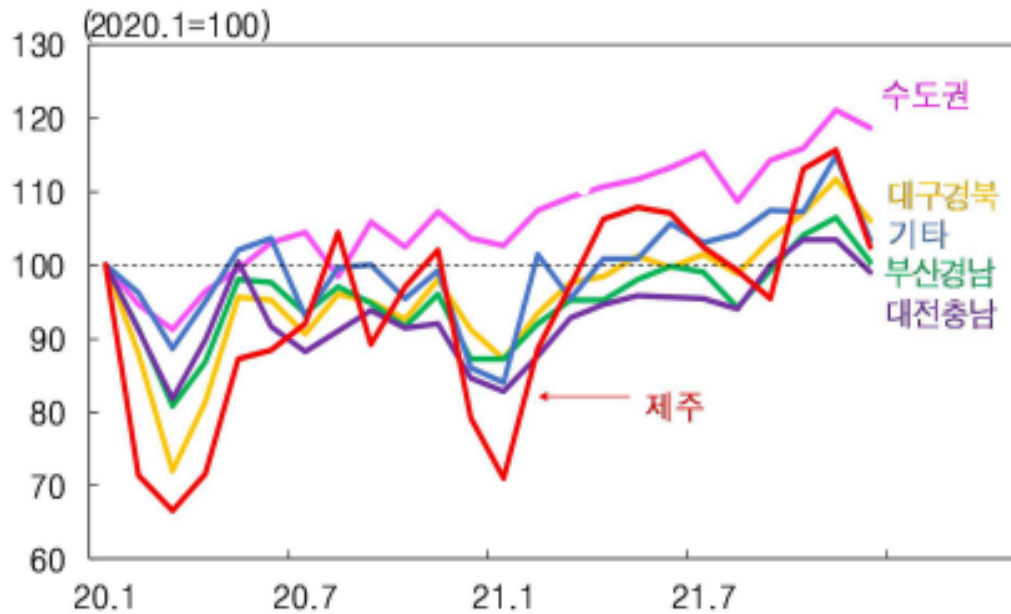
- 신용카드를 사용하면 사용자의 소비 성향이 데이터로 남는다. 이렇게 카드 회사에 전해지는 대표적인 데이터는 신용카드 사용 내역 중 비중이 가장 큰 업종은 무엇인지, 사용하는 시간은 주로 몇 시에 집중되는지, 소액 혹은 고액 결제가 많은지 등이 있다.

다양한 정보가 신용카드 회사에 전달되는 건 어느 기업과 협력을 진행하면 좋을 지 효율적으로 판단할 수 있게 도움을 준다. 이러한 측면에서 **신용 카드 결제 내역 뿐만 아니라 다른 정보들도 다각도로 분석하여 상품 기획에 활용한다면 또 다른 방면에서 신용카드 회사와 이와 관련된 업계에 긍정적인 영향을 미칠 수 있을 것이라 생각하여 카드 이용자 데이터 분석을 진행하였다.**

◦ 추천 서비스 측면

- 신용카드는 대표적인 비현금 지급수단의 하나이다. 신용카드 사용은 그 비중이 점점 확대되고 있고, 현재는 가게 소비에 있어서 현금 거래가 차지하는 비중보다 신용카드 거래가 차지하는 비중이 더 큰 것으로 추정된다. 이렇듯 결제하기 편한 장점 때문에 많은 사람들이 신용카드를 이용 측면에서만 바라본다면, 장점만 존재하는 것처럼 보인다.

하지만 올해 5월 한경 경제 뉴스 기사에 올라온 소식에 따르면, **“소비자가 카드 대금을 갚지 못해 대출로 돌려막는 ‘리볼빙’ 잔액이 올 들어 6조2000억원에 육박했다. 전년 대비 16% 늘면서 사상 최대 증가율을 나타냈다.”** 고 한다. 이 뿐만 아니라 한국은행에서 발표한 신용카드 지급 동향을 살펴보면 그 추세가 점점 증가하고 있어 **신용카드 이용자의 연체율을 낮출 수 있는 방안을 모색하는 것은 최근에도 중요한 이슈임을 다시 한 번 확인할 수 있었고, 최종적으로 이 연체 문제를 해결하는데 일조할 수 있는 신용 등급 예측 기반의 카드 추천 서비스를 기획하게 되었다.**



2021년 신용카드 지급결제 동향 그래프

출처: 한국은행

연구 및 개발 과정

• 개발에 필요한 데이터 소개

출처: 인공지능 플랫폼 DACON

<https://www.dacon.io/competitions/official/235713/data>

◦ features

index: 인덱스	DAYS_BIRTH: 출생일
gender: 성별	DAYS_EMPLOYED: 업무 시작일
car: 차량 소유 여부	FLAG_MOBIL: 핸드폰 소유 여부
reality: 부동산 소유 여부	work_phone: 업무용 전화 소유 여부
child_num: 자녀 수	phone: 전화 소유 여부
income_total: 연간 소득	email: 이메일 소유 여부
income_type: 소득 분류	occyp_type: 직업 유형
edu_type: 교육 수준	family_size: 가족 규모
family_type: 결혼 여부	begin_month: 신용카드 발급 월
house_type: 생활 방식	credit: 사용자의 신용카드 대금 연체를 기준으로 한 신용도

◦ shape

- train data: (26457, 20)

- test data: (10000, 19)

• 연구 및 개발에 활용한 기술 스택

1. 프로젝트에서 자신이 담당한 기술의 정의

- Python을 활용한 신용카드 데이터 EDA
- Tableau를 활용한 대시보드 제작 및 분석
- XGBoosting 모델을 통한 예측 모델 개발

2. 해당 기술(또는 연구 방법)의 최근 장점과 단점

- Seaborn & Matplotlib
 - 장점 : Python 코드로 쉽게 구현해 볼 수 있고, 코드를 작성할 줄 안다면 추가 비용이 들지 않는다.
 - 단점 : 다른 사람들에게 공유해서 보여 줄 수 있는 대시 보드 제작이 어렵다.
- Tableau
 - 장점 : GUI 를 활용하여 빠른 시각화가 가능하며, 화면 공유 및 화면 크기 조정 등 다방면으로 다른 사람과 소통하기에 적합한 구조를 갖추고 있다.
 - 단점 : 데이터베이스 서버를 활용하고 싶은 경우 상품 구매를 통해 진행이 가능하여 일정 비용이 지속적으로 발생한다.
- XGBoosting
 - 장점: XGB 모델은 NaN값 있어도 작동하고, 트리 기반 알고리즘이므로 모든 기능에 대해 라벨-인코딩을 진행해도 괜찮기 때문에 원-핫-인코딩을 사용해야 하는 모델들과 비교 했을 때 차원의 저주에 빠질 위험이 적다.
 - 단점: 기본 성능이 좋아서 test할 때 과적합 우려가 큰 모델이기 때문에 Tuning 시 파라미터 주의가 필요하다. → 너무 많은 횟수의 학습은 오히려 좋지 못하다.

3. 기간 내 구현 가능한 범위 설정

Aa 이름	📅 날짜	👤 담당자	📌 상태
<u>프로젝트 주제 선정</u>	@2022년 6월 24일 → 2022년 6월 27일		
<u>데이터 선정, 가공</u>	@2022년 6월 27일 → 2022년 6월 29일		
<u>EDA</u>	@2022년 6월 29일 → 2022년 7월 3일		
<u>Tableau 학습 및 대시보드 작성</u>	@2022년 6월 29일 → 2022년 7월 4일		
<u>모델 학습 및 성능 평가</u>	@2022년 6월 30일 → 2022년 7월 4일		
<u>Flask 학습 및 앱 개발</u>	@2022년 7월 4일 → 2022년 7월 7일		

- CP1 협업 방식
 - 09:00 시 게더 타운을 통한 간단한 이슈, 진행 과정 공유
 - 17:00 시 추가적인 회의 필요 시 게더 타운을 통한 공유
 - 지정 시간 외 의견은 디스코드 채널을 통해 공유

: 개인 기량에 따라 모든 과정을 구성원 모두 수행하되, 프로젝트에 반영되는 건 더 좋은 방향, 방안 첨가 하면서 프로젝트 보완함(모든 과정을 수행하는 이유는 아직 초급자인 구성원으로 이루어져 전체적인 내용을 구성원 모두 배워야 하기 때문)

: 다른 사람이 이해하기 쉽게 코드로 작성되는 부분은 주석을 표기하여 서로 공부할 수 있도록 진행

: 진행하면서 이해하기 쉽거나 내용이 상세히 잘 나타난 참고 사이트, 자료 등은 공유

⇒ 6/30까지 구글 시트에 공유

CP1 과정 및 결과 공유 문서 작성 (7월 6일 ~ 7월 7일)

- **결론(고찰):** CP1 프로젝트 전반에 대한 고찰 진행
 1. 연구 및 개발 통해 알게 된 점
 2. 프로젝트 결과를 사회 혹은 비즈니스에 어떠한 방식으로 활용해 볼 수 있는지 (※ 긍정적인 측면으로)
 3. 연구 및 개발을 진행하며 수행하지 못했던 부분 및 아쉬웠던 부분 (※ 이유를 포함하여 작성)
 4. 향후 자신의 프로젝트 주제를 이어서 진행할 연구자 혹은 개발자를 위한 방향성 및 기술적 측면 제언

⇒ 7/7 Urclass에 제출

참고자료

- 프로젝트 배경
 - 한경 경제 뉴스

카드사 리볼빙 16% 급증..."사실상 연체 늘어난 것"

소비자가 카드대금을 갚지 못해 대출로 돌려막는 '리볼빙' 잔액이 올 들어 6조2000억원에 육박했다. 전년 대비 16% 늘면서 사상 최대 증가율을 나타냈다. 22일 카드업계에 따르면 지난 3월 말 기준 신한·국민·삼성·현대·롯데·우리·하나 등 7개 전업

<https://www.hankyung.com/economy/article/2022052229191>



- 신용카드 데이터의 유용성 및 향후 과제(이세중, 손 원)
- 효성 FMS news room

[금융 트렌드] 금융권은 빅데이터를 어떻게 활용할까? - 효성FMS 뉴스룸


지금은 초연결 사회입니다. 그물망처럼 촘촘히 연결된 사회에서 모든 정보는 알게 모르게 연결선을 타고 퍼져 있습니다. 이러한 정보가 연결망에서 축적되면 사람들의 행동과 패턴을 파악할 수 있는 빅데이터가 되는데요. 이러한 빅데이터는 어떻게 https://www.hyosungfms.com/fms/promote/fms_news_view.do?id_boards=13395



○ 한국은행

2021년중 국내 지급결제동향 | 보도자료(상세) | 커뮤니케이션 | 한국은행 홈페이지

지급결제 동향

 <https://www.bok.or.kr/portal/bbs/P0000559/view.do?nttId=10069644&menuNo=200690&pageIndex=1>

