

하계 컨퍼런스 DA 2팀

2022.07.25

발표자 : 김소은

목차

1. 팀원 및 주제 소개

2. 활용 데이터 목록

3. 진행 계획

- 1) EDA 및 문제 정의, 주제 구체화
- 2) 파생변수 생성
- 3) 모델 선정 및 구현
- 4) 개인화 맞춤형 마케팅

1. 팀원 & 주제 소개



강영훈

경영학과 17

"컨퍼런스에 참여하면서
스스로 성장하는 것을
목표로 하면서
임하겠습니다!"



김소은

응용통계학과 19

"금융 데이터 분석은
처음인데 재밌게 공부하며
참여하겠습니다 ☺"



원민재

경영학과 17

" 훌륭한 팀원 분들과
배움과 성장을 도모하는
즐거운 시간이 되었으면
합니다. "

1. 팀원 & 주제 소개



✓ 대회 주제

: 고객 구매 데이터에 기반한 **예측 모델 개발** 및 **개인화 마케팅 전략** 제언

✓ 대회 목적

: 데이터 분야에서 경쟁력 있는 인재를 발탁하고 육성하기 위하여 마련된 대회로써, 빅데이터 분석 능력뿐 아니라 데이터 기반 비즈니스 로드맵을 제시할 수 있는 역량을 보유한 인재를 발굴하는 데 초점을 두고 있음

1. 팀원 & 주제 소개

소통 방식



CUAI 하계 DA 2팀

팀

- ★ 자기소개
- ▶ 공모전 및 분석 주제
- 🔥 매출 데이터 분석 스타디

RULE

- 📅 회의 일정 및 규칙

공모전

- 📅 제7회 롯데멤버스 빅데이터 경진대회 데이터셋

✓ 이슈 사항에 따라 비대면/대면 논의 후 미팅 진행

✓ 각자의 활동 중 공유 사항은 노션으로 공유

2. 활용 데이터 목록

데이터 테이블

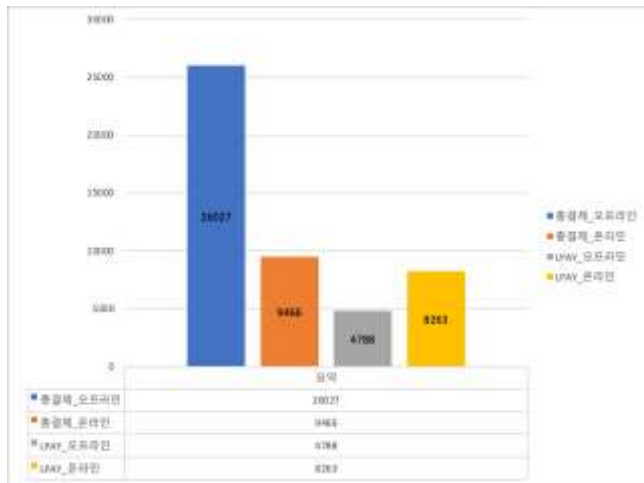
| 번호 | 테이블 명 | 설명 | 파일 명 |
|----|-----------|-----------------|--------------------------------|
| 1 | Demo | 고객 데모 정보 | LPOINT_BIG_COMP_01_DEMO.csv |
| 2 | 상품 구매 정보 | 유통사 상품 구매 내역 | LPOINT_BIG_COMP_02_PDDE.csv |
| 3 | 제휴사 이용 정보 | 제휴사 서비스 이용 내역 | LPOINT_BIG_COMP_03_COP_U.csv |
| 4 | 상품 분류 정보 | 유통사 상품 카테고리 마스터 | LPOINT_BIG_COMP_04_PD_CLAC.csv |
| 5 | 점포 정보 | 유통사/제휴사 점포 마스터 | LPOINT_BIG_COMP_05_BR.csv |
| 6 | 결제 이력 | 결제 이력 내역 | LPOINT_BIG_COMP_06_LPAY.csv |

롯데 멤버스 제공 데이터 6 종

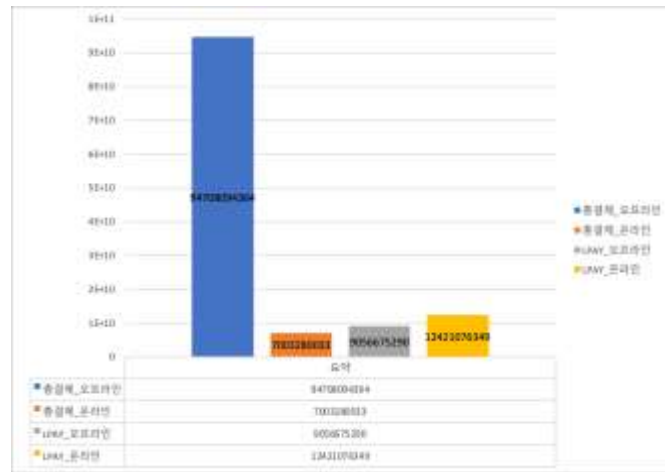
3. 진행 계획

1) EDA 및 문제 정의, 주제 구체화

결제 유형별 고객 수



결제 유형별 소비 금액

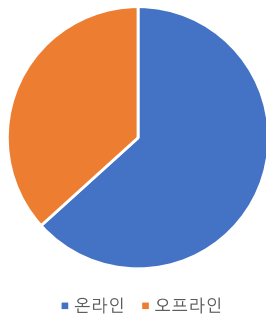


- ✓ LPAY 이용자 수 및 이용 금액은 전체 이용자 수와 전체 금액에 비해 현저히 낮은 수준.
- ✓ 또한 고객 수의 차이에 비해 소비 금액의 차이도 크게 나타나는 편.

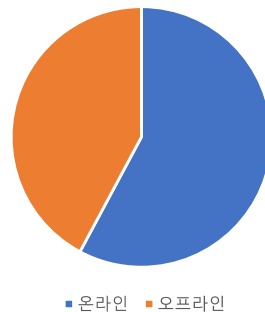
3. 진행 계획

1) EDA 및 문제 정의, 주제 구체화

온오프라인 별 LPAY 유저 수



온오프라인 별 LPAY 결제 금액



✓ 현재 LPAY는 오프라인 보다 온라인 결제 유저, 금액이 더 많은 상태.

3. 진행 계획

1) EDA 및 문제 정의, 주제 구체화

롯데멤버스, 엘페이 유료멤버십으로 그룹 생태계 구축

발행일 : 2022-01-23 09:00 | 지면 : 2022-01-24 | 4면



<롯데멤버스 앱포인트-엘페이>

롯데멤버스가 엘페이 유료멤버십을 내놔. 롯데그룹 통합 간편결제 엘페이를 활용해 온·오프라인 연계 생태계를 구축한다. 적립률을 최대 50% 높여 충성고객을 확보하고 쇼핑부터

'스마일클럽'vs롯데 '엘페이'유통공동신세계롯데멤버십경쟁 '활활'

이동민 기자 | 2022-01-24 15:34

신세계 통합 '스마일클럽' 할인규모 그대로 유지하며 추가
롯데 엘페이 5월엔 가입비 50% 할인해 추가가입 이벤트

온라인경매=온라인 특가, 거대 유통공동 신세계와 롯데가 멤버십 서비스를 확대하고 신규회원 유치에 적극적으로 나서면서 멤버십 경쟁에 가세
도 높이고 있다.



✓ LPAY 에 대한 롯데 멤버스의 관심과 집중도도 높아지고 있음.

→ LPAY의 이용률이 저조한 것을 문제로 인식하고 이를 LPAY가 주로 이용되는 온라인 플랫폼에서
활성화 할 수 있는 방법을 예측 모델을 통해서 찾고자 함.

3. 진행 계획

1) EDA 및 문제 정의, 주제 구체화

최종 선정 주제 :

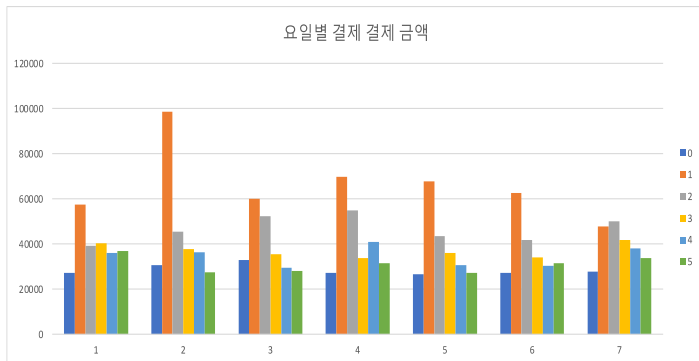
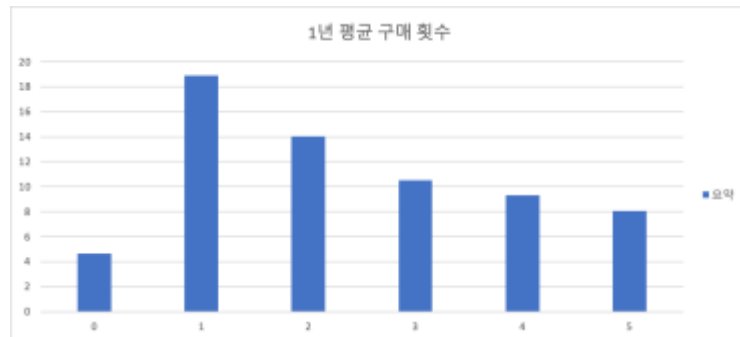
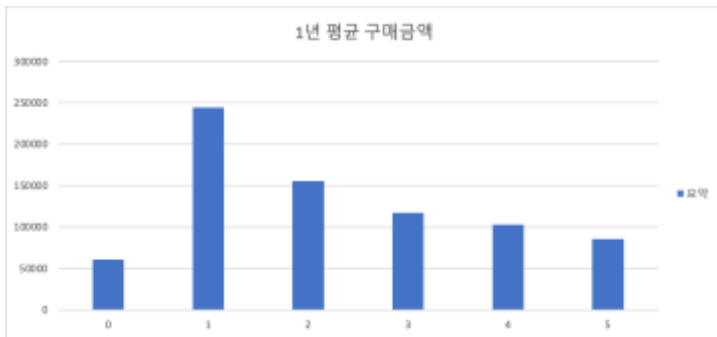
신규 가입 유저의 LPAY RFM 세그먼트 예측 모델링을 통한 LPAY 활성화 마케팅 방안 제시

1. LPAY 이용 내역으로 RFM 기준으로 세그먼트를 나눈다
2. 세그먼트 별 특성을 파악하고 주요 지표들을 이용해 파생 변수를 만든다.
3. 해당 변수들을 바탕으로 LPAY 미 이용 유저 중 LPAY 세그먼트를 예측하여 리스트화 한다
4. 해당 예측 모델링을 활용하여 타겟 마케팅 솔루션 적용

3. 진행 계획

2) 파생변수 생성

Lpay 이용 현황에 따른 RFM 별 파생변수 생성



추가 EDA 를 통해서 RFM 세그먼트 별로 차이가 드러나는 파생변수

탐색 과정 진행

3. 진행 계획

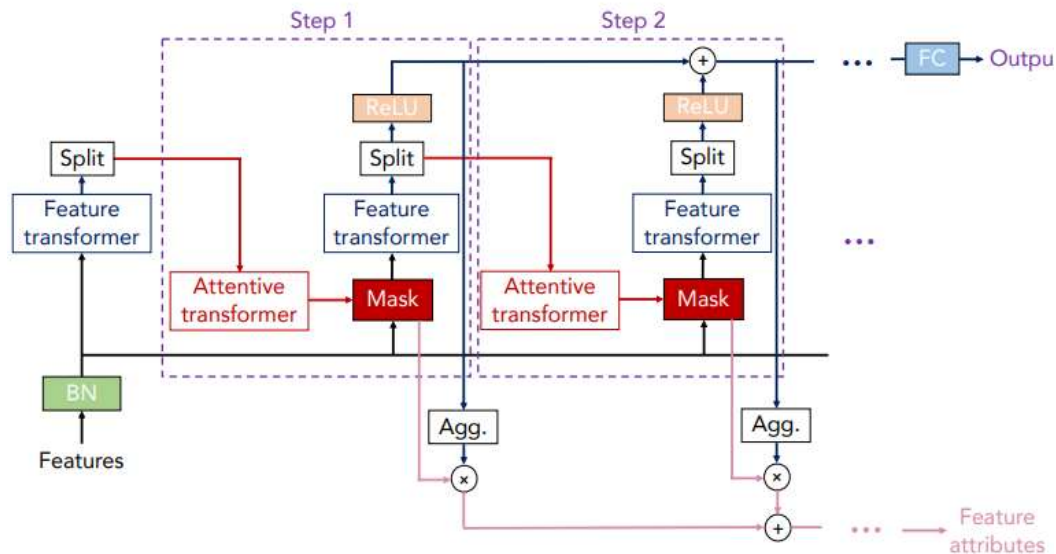
2) 파생변수 생성

| 항목 | 유효 | 비고 |
|---------------------------|------|---|
| 평균 구매 수량 | O | |
| 1회당 구매 금액 | O | 0 그룹과 유사성은 있긴 함. |
| 구매단가 | ▲ | 0 그룹이 타 그룹에 비해 상향, 1그룹과 타 그룹의 차이가 크진 않음. 하지만 차이가 분명 존재하긴 함. |
| 월별 추이 | X | 타 그룹과 비슷한 추이 발생 |
| 시간대별 인당 평균 거래 금액/ 횟수 | ▲ | 결제 금액에 있어서 추이는 존재하지만, 이를 어떻게 구분하고 활용할 수 있을지는 애매함. |
| 평일 / 주말별 인당 평균 결제 금액 / 횟수 | O | 그룹 1의 평일 결제 금액, 횟수가 가장 큼 |
| 요일별 인당 평균 결제 금액 / 횟수 | O | 평일/ 주말별 중에서 1개만 변수로 사용해야함. → 월요일 결제 금액이 두드러짐 |
| 성별 | O | 경제력이 있는 40대 특히 0 성별의 비중이 전체 기준보다 두드러짐 |
| 거주지별 | ▲, O | 217 에서 많은 비중을 차지하고는 있긴 하지만, 애매한 수준. |

✓ 파생변수 생성을 위한 추가 EDA 및 리스트업 작업 진행중

3. 진행 계획

3) 모델 선정

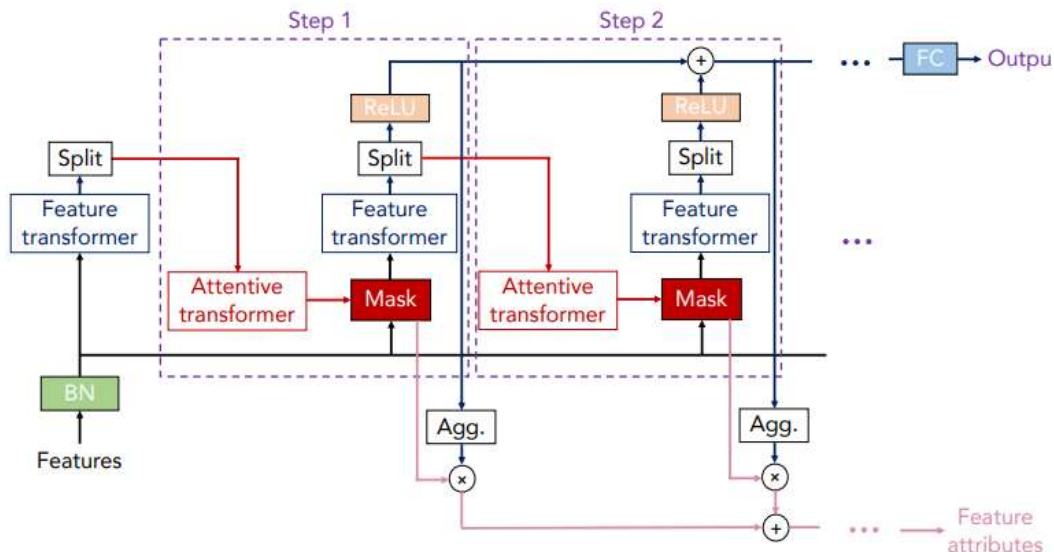


✓ TabNet 모델 선정

✓ 정형데이터에 적합한 딥러닝 모델

3. 진행 계획

3)모델 선정

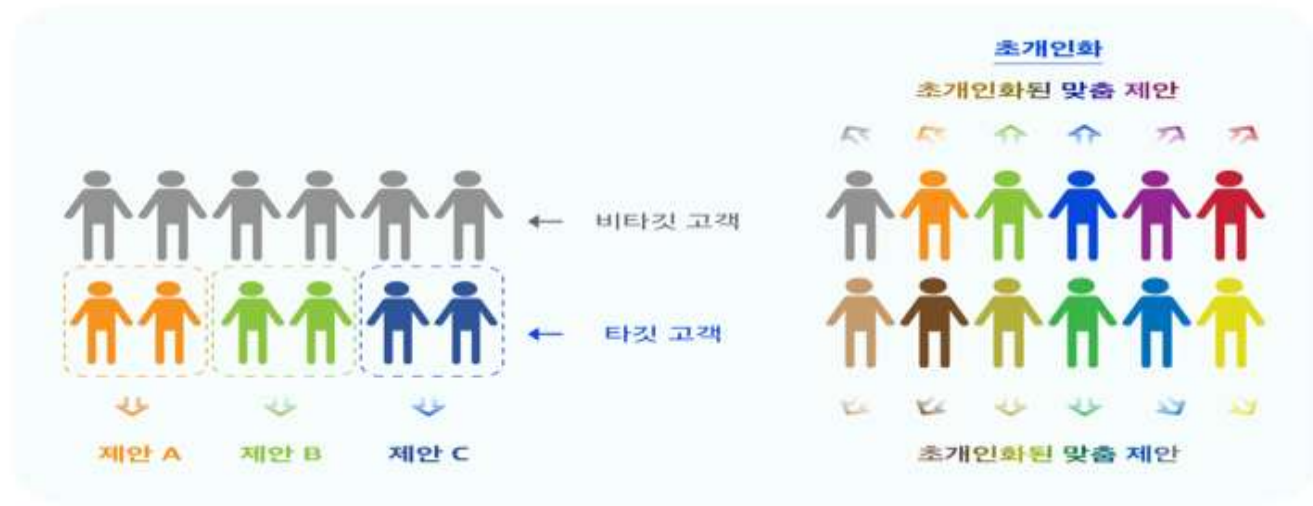


✓ TabNet 특징

- TabNet은 Feature의 전처리 없이 raw한 데이터를 입력으로 사용
- 성능과 해석력을 위하여 TabNet은 Sequential attention mechanism을 사용
- Gradient-descent 기반 최적화를 사용하여 End-to-End learning을 가능하게 함

3. 진행 계획

4)개인화 맞춤형 마케팅



- ✓ 현재 파생변수 생성 마무리 및 모델 선정과 구현 단계에 있음
- ✓ 개인화 맞춤형 마케팅 전략 도출 과정은 추후 보강 예정

감사합니다 !