

## Catch you!

이서연

강남대학교 데이터사이언스 전공

E-mail: tjdu5077@kangnam.ac.kr

## Introduction 서론

- 배경
  - 우리는 온라인 쇼핑몰을 이용할 때 '이런 상품은 어떠세요?' 라는 항목을 자주 보게 됨
  - 그러나 때로는 전혀 관심 없는 상품을 추천하기도 하는데, 어떤 상품의 추천이 고객과 기업의 입장에서 모두 좋을지 본 연구에서 답을 얻고자 함
- 목적 및 필요성
  - 고객과의 지속적인 관계 유지
  - 신규 고객 창출과 기존 고객 유지
- 활용 방안
  - 고객 관리 마케팅 활용

## Method 방법론

- 프로젝트 정의
  - 고객구매데이터에 기반한 고객 관리 및 상품 구매 예측과 추천
- 데이터 이해
  - 데이터 제공 : 제 7회 롯데멤버스 빅데이터 경진대회
  - 데이터 정의 : 고객 Demo, 상품 구매 정보, 제휴사 이용 정보, 상품 분류 정보, 점포 정보, 엘페이 이용, 총 6개의 csv파일
  - 데이터 기간 : 2021년 1월 ~ 12월 (총 12개월)
- 데이터 전처리
  - 데이터 정제 : 의미 없는 열 제거, 결측 값 제거
  - 데이터 통합 : 상품 구매 정보 데이터의 상품 코드 열에 따른 상품 분류 정보 열 추가
- 모델링
  - 11월 기준으로 상품 구매 데이터 분할
  - 신규구매/재구매 분석
    - 고객별로 구매 기록에 따른 old, new, both 디셔너리 생성 후 데이터 프레임 생성
  - 상품 구매 예측
    - 특정 시점 이전의 데이터에 대해 SVD 모델 학습 후 특정 시점 이후의 고객과 상품의 점수 예측
  - 상품 추천 시뮬레이션
    - 신규 구매 타겟, 재구매 타겟, 모든 상품 대상 추천

## Result 결과

- SVD 모델 학습 및 평가
  - Training time of model : 182.07 seconds
  - RMSE : 0.3640
- 상품 추천 시뮬레이션 평균 재현도
  - Reorder (재구매 타겟) : 33.5%
  - New (신규구매 타겟) : 18.6%
  - Total (모든 상품 대상) : 13%

	cust	실제주문	5개추천결과	TOP5추천_주문재현도
14253	M700807190	{쌀, 스낵류, 사탕/캔디, 과채음료, 생수, 수입과일, 요구르트, 차음료, 정리용...	[스낵류, 레토르트, 요구르트, 비스킷류, 용기라면]	1.0
10291	M505064484	{쌀, 스낵류, 호주산소고기, 조리기구, 소주, 나물류, 국산소고기, 봉지라면, 청...	[아이스크림, 요구르트, 소주, 열매채소, 국산과일]	0.2
141	M007349628	{소주, 맥주, 두부/콩나물류, 수입과일, 전통주}	[아이스크림, 맥주, 스낵케어, 베이커리, 수입과일]	0.4
4629	M228124869	{스낵류, 여아의류아우터, 유아동화, 맥주, 피트니스, 식당, 여성일반스포츠의류, ...}	[맥주, 홍삼/인삼가공식품, 베이커리, 용기라면, 초콜릿]	0.2
3034	M149105654	{스낵류, 조콜릿, 뿌리채소, 와인, 여성의류하의, 원두커피, 잎채소, 맥주, 모바...	[스낵류, 비스킷류, 수입과일, 열매채소, 요구르트]	0.6

## Discussion 토의

- 고객 구매 패턴
  - 구매 횟수 : 고객들은 평균 4회의 구매를 하며, 대부분의 경우 1~5회의 구매를 함
  - 상품의 다양성 : 평균 57개 종류의 상품을 구매함
- 신규구매/재구매
  - 전체 고객 수 26027명에서 신규 상품을 구매한 고객의 수는 17573명, 재구매한 상품이 있는 고객의 수는 18047명
  - 신규 상품 구매 고객의 경우 평균 385개 종류의 상품을 구매함
- 상품 추천 시뮬레이션
  - 기존에 구매한 기록이 없는 상품을 대상으로 추천해주었을 때 보다, 기존에 구매한 기록이 있는 상품을 대상으로 하여 추천해주었을 때 더 많이 구매할 것으로 나타남

## References 참고

- heehehe's study note, "빅데이터 분석 방법론", 2022-10-26, <https://heehehe-ds.tistory.com/entry/ADsP-2-1-%EB%8D%B0%EC%9D%B4%ED%84%B0-%EB%B6%84%EC%84%9D-%EA%B8%B0%ED%9A%8D%EC%9D%98-%EC%9D%B4%ED%95%B4-3-%EB%B9%85%EB%8D%B0%EC%9D%B4%ED%84%B0-%EB%B6%84%EC%84%9D-%EB%B0%A9%EB%B2%95%EB%A1%A0>
- 미래를 위한 취미, "구매 데이터 분석", 2022-08-18, <https://seethefuture.tistory.com/81>
- 남기백, 박상원, (2017), 머신러닝 기반 고객 재구매 상품 예측, 2017년 추계학술발표대회 논문집, 제 24권(제 2호), 421-423