**<간편결제시장의 후발주자 추격전략** by 도윤**>**

[가설설정]

대립가설: 간편결제 시장의 고객들은 **초기에 이미 썼던 결제시스템을 잘 바꾸지 않고, 계속 사용**한다는 특징을 가지고 있다.

(Hypothesis Setting Reasoning)

기술이 나중에 따라잡혀도 이미 **업계표준이 되어서 뒤따라오는 다른 회사들이 그 모듈을 계속 사용해야 되니까 계속 커질 수 밖에** 없다. 그리고 기술이 따라잡혀도 이미 정착되어 있으니 스위칭 코스트가 엄청 크다.

**1) 먼저 데이터로 해당 가설을 검증**

How? : 결제데이터 중 30번 이상(Normal 가정) 결제한 7000여명 개인의 결제일(Approval Date)변수 (2018 11월~2019 2월)와 Company\_Name(LT01)의 시차 연관성 분석

분석방법: Sequential Association Rule

동일한 고객이 시간이 지남에 따라서 어떤 소비형태를 보이는가에 대한 분석이다. 즉, 우리는 간편결제 내역에 대한 분석을 통해 “추후 발생시점에 따라 동일한 거래를 보이는지”에 대한 정보를 얻을 수 있다. 이러한 시차 연관성을 갖는 규칙들은 원인과 결과의 형태로 해석이 가능하므로 그 결과가 더욱 유용하다. 해당 시간 데이터의 경우 시간을 고려하여 데이터가 정리되므로, 시간의 간격은 같은 날짜(하루 단위)에 거래된 항목들은 동일한 시점의 거래로 정의되었다.

사용데이터셋: USER\_ID(고객번호), Approval Date(결제일), Company\_Name(간편결제 회사)

결과확인: A회사를 먼저 구입한 고객 중에 B사,C사,D사를 구매하는 비율(신뢰도:Confidence)이 모두 5%정도로 낮으면, 추후 발생시점에 따라 동일한 거래를 보인다고 볼 수 있다. 따라서 “사람들은 초기에 썼던 결제시스템을 계속 사용하지 않는다” 는 가설을 기각한다. 그리고 사용비율이 높은 브랜드들의 출시일이 빠른지 확인하여, 간편결제 시장에서는 업계 표준의 결제모듈을 만드는 First Mover가 시장 지배력을 유지한다고 해석한다. Ex) 국내최초 간편결제: 카카오페이

[EDA List]

1. Raw결제데이터 중복행 제거
2. Raw결제데이터 중 30번 이상 결제한 7천여 명 Panel들 Subset
3. 사용데이터셋 중 NA 제거

★Raw결제데이터 그대로 사용 (불필요한 컬럼 굳이 삭제 안시켜도 분석 시 필요한 컬럼 변수만 추출하면 될 것 같습니다.)

**2**) **MS 3%미만의 OO페이가 MS 30%이상의 마켓리더로 성장하기 위한 마케팅 전략(Seg, 4P) 제안**

**Target Variable**: 간편결제 승인 결제내역 Approval Price (LT01인 경우 Approval Original Price)

간편결제 시장의 고객들은 초기에 썼던 결제시스템을 계속 사용한다는 것이 Sequential Association Rule을 통해 검증되었다. 따라서 MS 3%미만의 OO페이가 MS 30% 이상의 마켓리더로 성장하기 위해선 First Mover가 되는 마케팅 전략을 제언해야 한다. Payment와 App 데이터 분석을 통해 마케팅 4P전략과 Segmentation 전략을 제언해보자.

[EDA List]

1. PA결제 240만개 데이터 1명당 다수 결제된 내역들 1행으로 EDA처리
2. 지리변수는 어차피 중복(같은 값)이므로, 빈도 처리 (자동 1개로 줄여짐)
3. 결제정보 수집 시 패널의 위치; 위도,경도 데이터 K평균 클러스터링 처리하여 최종군집의 무게중심에 위치 할당
4. Approval Origianl Price 평균 처리 (1인당 간편결제로 결제한 평균금액)
5. Approval Sotre 빈도 처리
6. 잔여한도 + 거래누적 금액 = 충전한도(Total Payment Limit) 파생변수 생성 (New Column 추가), 중복(같은 값)이므로 빈도 처리 (자동 1개로 줄여짐)
7. NA값 처리
8. PA앱 60만개 데이터 1명당 다수 기록된 내역들 1행으로 EDA처리

0-1 1명당 Max앱 사용시간인 카테고리만 남기기 (dplyr패키지에서 gather랑 spread 동사 활용)

1. **[Product]** “간편결제의 최대한도는 얼마가 적절할까?”  **[x변수]**: 파생변수 **충전한도**
2. **[Price] 간편결제의 수수료는 얼마를 책정해야 많은 이익이 남을까? (How get??)**

**->수수료는 카드사연계의 경우 카드사로부터 받습니다.**

1. **[Place]** “간편결제시스템을 **어떤 사용처(Approval Store)**에 유통하는게 좋을까?”

**[x변수]**: 사용처

->가맹점은 각 회사마다 다르기에 최적 사용처를 뽑을 수 없음 =>대안) 각 간편결제회사 별 고빈도 top5 사용처 막대그림을 그리고 서로 비교. 비슷한 결제환경의 최적 타회사(경쟁사)는 이러한 사용처에서 많이 거래가 이루어지는데 이런 가맹점을 추가로 영입하는 것이 어떠한지 묻기.

1. **[Promotion] 결제위치가 밀집된 최종군집**에 프로모션 전략 제언 ; **[x변수]**: 위도,경도

* *결제위치가 각자 개인의 집일 수도 있음(온라인마켓의 경우), 오프라인만?*

1. EDA된 결제데이터, 앱데이터를 Panel ID: Key값을 뽑아서 정렬 (두 데이터 1:1대응)
2. Payment데이터에 앱사용시간 Column 추가

->

->

1. **[간편결제 고객의 관심사에 따른 Segmentation 전략 제언] [x변수]**: Total Usage Time 카테고리별 앱 사용 시간 (엠브레인 기준 카테고리)

ex) 카드본인확인서비스를 이니시스가 최초로 모듈했다면, MS 3%미만의 OO페이가 MS 30%의 마켓리더로 성장하기 위해서는 “인공지능(앱 관심사) 결제 시스템”의 업계표준 모듈 개발에 착수해야 한다.

★EDA 추신: 패널데이터는 한 명의 행 데이터가 여러 번 기록되기 때문에 우리가 알고 있는 분석방법들을 사용할 수가 없다. Row 행들이 iid 가정을 위반하기 때문이다. Panel Regression이라고 다른 분석방법이 있는데, 이건 대학원과정에서 배우는거라 좀 힘들 것 같다. 그래서 데이터를 패널 1명당 1행으로 만들어 iid 가정을 성립할 수 있는 방향으로 EDA를 먼저 구성해보았다. 잘 토의하여, 논리적인 방법을 통해 EDA를 성공시킨다면 우리가 알고 있는 분석방법을 사용할 수 있을 것이다.

분석방법:

X가 연속형: 2번,5번,8번) 회귀분석, 신경망분석, 의사결정나무, 랜덤포레스트, 그래디언트부스팅(target이 존재할때) 등 우리가 아는 모델들을 다 돌려보면서 결과값 가장 잘 나온 모델을 채택하면 된다. (시간이 없어서 모델 다 자세하게 못적었어유죄송..)엠브레인도 결국 “**어떤 문제를 정의했고 데이터를 통해서 어떤 insight 를 제시했나**”에 관심이 있을 것. 결국 문제 정의를 내일 회의에서는 확실히 정해야 할 것 같다. 문제만 잘 정의하면 사실 모델링은 그냥 돌리면 되기 때문이다.

X가 명목형: 4번) 범주형 회귀분석: [현재 Market Positioning] OO페이, MS 1-3위 간편결제사와 사용처간의 관계분석: Category group code-> Approval store 칼럼 분석 Eg. 네이버페이: 501(온라인결제)-> 301(배달의 민족), 1799(기타) – 분석 결과 ->사용처와의 표준화된 베타, 상대적 중요도 확인) -> 각 브랜드별 사용처의 차이 해석. Why? And OO페이의 현재 사용처 상황 파악 가능

X가 ?? 3번) 보류 ;

추가로 생각해볼 방법들

+)

USER ID를 키값으로 Raw결제데이터에 Area\_Code1,2, Area\_Name1,2 칼럼추가 지역시장 타겟팅, “**어떤 지역**으로 유통하는게 좋을까?” ; **[x변수]**: 지리변수 (도,특별 및 광역시 ; 시,군,구)