

## [데이터 시각화&분석 취업캠프]

# 최종 프로젝트 기획안

기획안 작성일자: 2023.04.21(금)

조 명	3조: 클로바캐시(Clovacash)
조 원	조장: 남호제 조원: 박세형, 우정연, 이성민
프로젝트 주제 및 개요	<p><b>주제:</b></p> <p>Clovacash, 자신에게 맞는 카드를 찾으세요!</p> <p>- 소비 예측을 통한 소비자 맞춤 카드 추천 시스템</p> <p><b>개요:</b></p> <p>소비 패턴(사용처, 사용 금액, 횟수)에 따라 다음 달의 소비 현황을 예측해서 자신에게 맞는 카드를 발견할 수 있는 서비스</p> <p><b>1. 제안 배경</b></p> <p>1) 고객의 다양한 소비 성향의 증가</p> <p>카드 매출 데이터 지표만으로는 맞춤형 소비 패턴 분석이 어려움.<sup>1</sup></p> <p>2) 금융 회사마다 고객의 개인화를 점점 구체화하고 있음.<sup>2</sup></p> <p>3) 사용자 데이터 기반 맞춤형 추천이 큰 효과를 거두고 있음.<sup>3</sup></p> <p><b>2. 제안 목표</b></p> <p>고객 데이터를 기반으로 시중의 카드 상품 중 알맞은 상품을 추천하는 서비스를 개발하고자 함. 누구나 손쉽게 이용할 수 있는 추천 플랫폼</p>

<sup>1</sup> "정형 비정형 빅데이터의 융합분석을 위한 소비 트렌드 플랫폼 개발" <한국정보화진흥원 K-ICT 빅데이터센터, 비씨카드 빅데이터센터, 원광대학교 컴퓨터소프트웨어공학과>

<sup>2</sup> "2만 5000개 소비패턴 분석해서 혜택 제안 필요할 때 귀신같이 알려주는 '똑똑 카드'" <신한카드 2019년 5월>

<sup>3</sup> "현대카드, 데이터 기반 맞춤형 소비 컨설팅, 350만 고객의 선택을 받다." <현대자동차그룹 뉴스룸 2022년11월30일>

	<p><b>3. 대상</b></p> <p>맞춤형 카드를 찾고 있는 사용자 금융 추천 서비스를 도입하려는 기업</p> <p><b>4. 기대효과</b></p> <p>1) 고객 만족도 향상 고객의 소득 수준, 생활 습관 등 여러가지 조건을 고려하여 맞춤형 카드를 제공할 수 있음. 이를 통해 고객은 카드를 쉽게 찾아서 신청할 수 있음.</p> <p>2) 카드 발급 효율성 제고 무분별한 카드 발급으로 인한 카드 해지나 카드사 탈퇴에 따른 고객 이탈을 방지할 수 있음.</p> <p>3) 홍보 효과 카드 추천 서비스는 인터넷 상에서 홍보될 수 있고, 이를 통해서 브랜드 인지도 상승을 기대할 수 있음. 잘 알려지지 않았지만 좋은 혜택을 제공하는 카드 또한 잠재 고객에게 소개될 수 있음.</p> <p>4) 데이터 베이스 구축 본 프로젝트의 경우 카드 사용에 대한 고객 특성과 개별 성향을 반영한 데이터 베이스를 구축할 수 있음. 추후 카드 서비스 챗봇 개발 등 다양한 downstream task<sup>4</sup>의 기반이 될 수 있음.</p>
프로젝트 수행 방향	<p>● 데이터 설명</p> <p>1. 이용자 정보 (연령대, 성별) 및 카드 소비처 서울시민의 업종별 카드소비 패턴 데이터(2016-2022.8) &lt;서울특별시 빅데이터 캠퍼스&gt; + 온라인 스토어 데이터 + 페이크 데이터</p> <p>2. 카드사별 카드 및 제공 서비스 자료 &lt;카드 고릴라&gt; ※ 기타 개인정보는 일체 이용하지 않습니다.</p>

<sup>4</sup> Downstream Task: 최종적으로 해결하고자 하는 작업. 모델을 처음부터 새로 훈련하는 것이 아닌 미리 훈련된 모델로부터 학습을 시작하는 것

## ● 분석 및 개발

### 1. 데이터 수집

- 연령별, 성별 카드 이용건수, 이용금액, 사용처 수집
- 금융사별 카드 자료 (카드 혜택, 연회비 등), 제공 서비스 자료(카드 고릴라에서 웹 크롤링)

### 2. 데이터 전처리

- 제휴사 카테고리 통합. 연령, 지역, 시간대를 범주화 후 학습 데이터셋 분리
- 고객 속성 변수, 구매 패턴 변수화 및 지수화

### 3. 기본 통계 분석

- 소비 트렌드 분석 (데이터 예시)

카드 업종별 소비 분석을 통해 이용자 소비 성향 파악

	A	B	C	D	E	F	G	H	I
1	서울시인업	기준년월(Y)	고객 주소별	성별(GEDN)	연령대별(A)	카드이용금	카드이용건수계(USECT_CORR)		
2	SS013	201906	14037	F	50대	48288	65		
3	SS048	201608	156830	M	30대	132792	10		
4	SS016	202009	32925	M	50대	100600	25		
5	SS001	201805	214245	F	30대	213876	40		
6	SS044	201705	279472	M	30대	155930	5		
7	SS048	201911	220797	M	40대	20120	15		
8	SS058	202006	152728	F	40대	285201	91		
9	SS020	201711	83127	F	60대	15090	15		

[그림 1] 블록별 성별연령대별 카드소비패턴.csv

- 카드 상품 분석 (카드 고릴라)

각 카드사별 카드 상품, 정보(연회비, 혜택 등) 현황 파악

1	Title	Image	card_crop	View_URL	sale1	sale2	sale3	sale4	sale5	in_for	in_for6	Lnth		
2	신한카드 M	https://api.c	신한카드	https://www	모두드림	10% 할인	마트/편의점	10% 할인	식음료	10% 할인	해외결	15,000원	전월실적 30만원 이상	
3	신한카드 D	https://api.c	신한카드	https://www	모두드림	0.7% 적립	다해드림	2.1% 적립	생거드림	3.5% 적립	해외결	8,000원	전월실적 없음	
4	신한카드 D	https://api.c	신한카드	https://www	4개경유사	10% 할인	자랑서비스	10% 할인	편의점/커피	5% 할인	해외결	10,000원	전월실적 30만원 이상	
5	신한카드 D	https://api.c	신한카드	https://www	모두드림	0.7% 적립	다해드림	2.1% 적립	생거드림	3.5% 적립	해외결	10,000원	전월실적 없음	
6	신한카드 D	https://api.c	신한카드	https://www	모두드림	최대 1.1%	다해드림	최대 3.3%	생거드림	최대 5.5%	해외결	30,000원	전월실적 없음	
7	신한카드 T	https://api.c	신한카드	https://www	캐시백	1% 캐시백	대한항공	1,500원/1회	기프트콘	1회/년	제공	202,000원	전월실적 없음	
8	신한카드 B	https://api.c	신한카드	https://www	대중교통	200~900원	택시	10% 할인	KTX	10% 할인	해외결	10,000원	전월실적 30만원 이상	
9	아시아나 신	https://api.c	신한카드	https://www	1,000원	당	1.5마일	가	1,000원	당	3마일	추가	국내 전월실적 50만원 이상	
10	신한카드 H	https://api.c	신한카드	https://www	전국가맹점	0.2~2.0%	특정가맹점	1~5% 적립	현대오일뱅크	60원/L	적점	해외결 8,000원	전월실적 없음	
11	알뜰교통카드	https://api.c	신한카드	https://www	대중교통/택	10% 할인	마트/편의점	10% 할인	전기차 충전	50% 할인	해외결	12,000원	전월실적 30만원 이상	
12	신한카드 R	https://api.c	신한카드	https://www	모든주유소	40~150P/L	모든송전소	10~50P/L	모든가맹점	0.2~2% 적립	해외결	32,000원	전월실적 50만원 이상	
13	신한카드 T	https://api.c	신한카드	https://www	1,500원	당	1마일	적립	1,000원	당	1마일	적립	PP카드 무료 제공	해외결 120,000원
14	메리야드 폰	https://api.c	신한카드	https://www	메리야드	칩무로속박권	메리야드	칩6P/1천원	5메리야드	폰	카드	국내 전월실적 2만	해외결 없음	

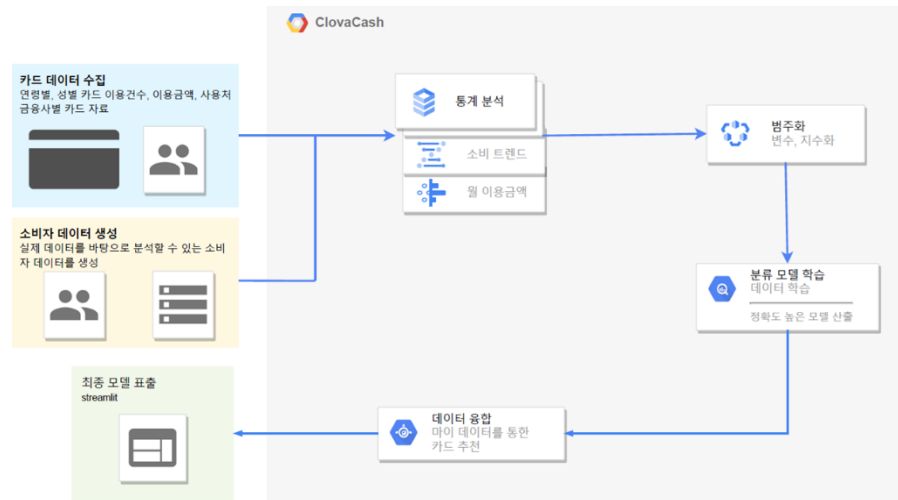
[그림 2] 카드 고릴라에서 크롤링한 카드 데이터

### 4. 이용자 소비 패턴 예측

- 분류 모델 학습 (Decision Tree, XGBoost, LightGBM 등)
- 모델 정확도 및 교차 검증을 통해 가장 좋은 결과를 보인 모델을 선정
- 예측 모델을 통해 소비 패턴에 따른 카드 추천

## 5. Streamlit을 통한 플랫폼 게시

- 최종 모델 표출 (streamlit 활용)



### 프로젝트 조직 (구성원 및 역할)

#### ● 역할분담

남호제: 고객 데이터 수집, 보고 및 문서작업  
 우정연: 데이터 전처리 및 학습 모델링 구상  
 박세형: 카드 정보 데이터 수집, 크롤링 및 카드 데이터 관리  
 이성민: streamlit 제작, 크롤링 및 카드 데이터 관리

### 프로젝트 추진 일정

#### ● 일정

4/4 ~ 4/11: 주제 선정 및 일정 수립  
 4/12 ~ 4/14: 카드 데이터 크롤링, 소비자 데이터 분석  
 4/17 ~ 4/21: 데이터 학습 알고리즘 강화  
 4/24 ~ 4/28: 모델 학습 및 결과물 도출  
 5/1 ~ 5/9: 결과물 평가 및 발표자료 작성  
 5/10: 프로젝트 발표

## Reference

- 1) "정형 비정형 빅데이터의 융합분석을 위한 소비 트렌드 플랫폼 개발" <한국정보화진흥원 K-ICT 빅데이터센터, 비씨카드 빅데이터센터, 원광대학교 컴퓨터소프트웨어공학과>  
<https://koreascience.kr/article/JAKO201719558338199.pdf>
- 2) "2만 5000개 소비패턴 분석해서 혜택 제안 필요할 때 귀신같이 알려주는 '똑똑 카드'" <신한카드 2019년 5월>  
[https://dbr.donga.com/article/view/1202/article\\_no/9121](https://dbr.donga.com/article/view/1202/article_no/9121)
- 3) "현대카드, 데이터 기반 맞춤형 소비 컨설팅, 350만 고객의 선택을 받다." <현대자동차그룹 뉴스룸 2022년11월30일>  
<https://www.hyundai.co.kr/news/CONT0000000000066986>
- 4) "네이버 카페 카드 추천 게시물 2,262,877건"



- 5) 서울시민의 업종별 카드소비 패턴 데이터 <서울시 빅데이터 캠퍼스 2016.02-2022.08>  
[https://bigdata.seoul.go.kr/data/selectSampleData.do?r\\_id=P213&sample\\_data\\_seq=318&tab\\_type=&file\\_id=&sch\\_text=%EC%B9%B4%EB%93%9C&sch\\_order=U&currentPage=1](https://bigdata.seoul.go.kr/data/selectSampleData.do?r_id=P213&sample_data_seq=318&tab_type=&file_id=&sch_text=%EC%B9%B4%EB%93%9C&sch_order=U&currentPage=1)
- 6) 카드 고릴라  
[https://www.card-gorilla.com/home?gclid=CjwKCAjwov6hBhBsEiwAvrvN6JgVJovYOt3yMKH3\\_ET2oWsBscBnLFx3-KcVDL68KDwUBM7wtZ4xOB0Cy38QAvD\\_BwE](https://www.card-gorilla.com/home?gclid=CjwKCAjwov6hBhBsEiwAvrvN6JgVJovYOt3yMKH3_ET2oWsBscBnLFx3-KcVDL68KDwUBM7wtZ4xOB0Cy38QAvD_BwE)
- 7) "산학협력력 기술개발사업(일반 첫걸음) 최종보고서" <한림성심대학교 배종식 교수 2015년9월>
- 8) "텍스트 마이닝을 이용한 소비자 소비패턴 분석 기법 설계" <정은희, 이병관 2018년11월2일>

