

소비패턴을 통한 보험종목별 사고율 분석

고객 군집분석과 사고율 모델 해석을 중심으로

용앤리치 팀

심현구 | 강세현 | 최용원



목차

- 00 개요
- 01 탐색적 데이터 분석 및 전처리
- 02 군집분석
- 03 자동차보험 사고율 예측
- 04 보험종목별 사고율 분석
- 05 아이디어 및 기대효과

00 개요

분석주제 및 목적

분석 주제

소비패턴을 통한 보험종목별 사고율 분석

분석 목적

세분화·다양화된 소비 특성을 고려한 맞춤형 분석 서비스 제공

01

탐색적 데이터 분석
및 전처리

01 탐색적 데이터 분석 및 전처리

탐색적 데이터 분석(EDA)

수치 요약과 시각화를 사용하여 데이터를 탐색하는 프로세스



EDA의 목적

데이터에 대한 적절한 **전처리**
자료를 **요약** 제시하여 소비패턴을 파악

01 탐색적 데이터 분석 및 전처리

Raw 데이터셋 확인

85001 X 118

성별	연령	결혼예정스 코어	야외활동관 심성향등급	...	매출금액합 계_쇼핑_오 전	매출건수합 계_건강보조 식품_오후	...	보험종목	사고건수	손해액	보험료	...
1	40	1	7	...	6030600	352000	...	3	0	NA	NA	...
0	40	2	4	...	3022200	707200	...	1	0	0	69000	...
1	50	1	5	...	323300	723500	...	1	0	0	79000	...
0	20	6	6	...	2456700	32000	...	3	1	NA	NA	...
0	30	5	3	...	7708700	245200	...	3	0	NA	NA	...
1	60	1	2	...	442200	570000	...	1	0	7000	22000	...
1	50	2	5	...	6370200	398000	...	2	0	0	2000	...
...

인구통계학적 변수

카드매출 관련 변수

보험 관련 변수

01 탐색적 데이터 분석 및 전처리

계획과의 차이

문제 1

주어진 사고율 데이터를 모두 활용하려고 하였으나
자동차보험 사고율이 전부 결측치



장기보험 사고율 데이터로 학습시킨 머신러닝 모델로
자동차보험 사고율을 예측하여 대체

문제 2

대용량 데이터셋의 처리에 어려움이 예상되어
MZ세대의 데이터만을 활용하려고 하였으나
통제 가능한 사이즈의 데이터셋 확인



전체 데이터 활용 후,
파생변수 'MZ세대여부'를 생성하여
연령대에 따른 특성 파악

01 탐색적 데이터 분석 및 전처리

계획과의 차이

문제 1

주어진 사고율 데이터를 모두 활용하려고 하였으나
자동차보험 사고율이 전부 결측치



장기보험 사고율 데이터로 학습시킨 머신러닝 모델로
자동차보험 사고율을 예측하여 대체

문제 2

대용량 데이터셋의 처리에 어려움이 예상되어
MZ세대의 데이터만을 활용하려고 하였으나
통제 가능한 사이즈의 데이터셋 확인



전체 데이터 활용 후,
파생변수 'MZ세대여부'를 생성하여
연령대에 따른 특성 파악

01 탐색적 데이터 분석 및 전처리

계획과의 차이

문제 3

생명보험의 '사고건수' 값의 99.9%가 0으로 구성
사고율이 전부 0으로 계산되는 문제 발생



사고율 계산식을 구성하는 '경과계약건수'와 '사고건수'를
각각 산출 후 사고율을 사후적으로 계산

문제 4

데이터셋이 매우 편향되어 있어
AI 모델 학습에 부적합



통계적 데이터 증폭을 고려해보았으나
현실 설명력이 미흡하다고 판단
→ AI 모델링 대신 통계적 모델링 선택

01 탐색적 데이터 분석 및 전처리

계획과의 차이

문제 3

생명보험의 '사고건수' 값의 99.9%가 0으로 구성
사고율이 전부 0으로 계산되는 문제 발생



사고율 계산식을 구성하는 '경과계약건수'와 '사고건수'를
각각 산출 후 사고율을 사후적으로 계산

문제 4

데이터셋이 매우 편향되어 있어
AI 모델 학습에 부적합



통계적 데이터 증폭을 고려해보았으나
현실 설명력이 미흡하다고 판단
→ AI 모델링 대신 통계적 모델링 선택

01 탐색적 데이터 분석 및 전처리

결측치 처리

변수명	결측치 수	처리 방법
손해액	52970	* 간접 대체
보험료	54210	* 간접 대체
자동차국외산구분코드	30791	자동차보험에서만 활용
자동차보험가입경력	30791	자동차보험에서만 활용
거주지_시군구	697	삭제
직장지_시군구	442	삭제

* 머신러닝 모델을 이용하여 예측한 값으로 간접적으로 대체

범주형 자료 처리

거주지_광역	직장지_광역	수도권거주	수도권직장
경기도 고양시	서울특별시 종로구	1	1
경기도 시흥시	인천광역시 연수구	0	1
경상남도 양산시	부산광역시 해운대구	0	0
제주특별자치도 제주시	제주특별자치도 제주시	0	0

파생변수 생성

파생변수명	변수 설명
매출건수_항목명_Sum	각 품목의 오전과 오후의 매출건수의 합
매출금액_항목명_Sum	각 품목의 오전과 오후의 매출금액의 합
매출건수_오전여부	오전 매출건수가 오후 매출건수보다 많으면 1
매출금액_오전여부	오전 매출금액이 오후 매출금액보다 많으면 1
총매출건수_Sum	전체 매출건수의 총합
총매출금액_Sum	전체 매출금액의 총합

01 탐색적 데이터 분석 및 전처리

전처리 완료 데이터셋 확인

변수명
연령
추정라이프스테이지
미혼스코어
결혼예정스코어
출산예정스코어
미취학자녀부모스코어
초등학생자녀부모스코어
중고생자녀부모스코어
대학생자녀부모스코어
자녀결혼예정스코어
전업주부스코어

인구통계학적변수

변수명
1인가구페르소나등급
온라인소비선호등급
얼리어답터성향등급
프리미어성향등급
배달식품관심성향등급
반려동물보유추정등급
야외활동관심성향등급
해외여행관심성향등급
인테리어스타일링관심성향등급
인테리어모델링관심성향등급
건강관리관심성향등급

생활성향관련변수

변수명
매출건수합계_항목명_Sum *
매출금액합계_항목명_Sum *
매출건수합계_항목명_오전여부 *
매출금액합계_항목명_오전여부 *
총매출건수_Sum
총매출금액_Sum

카드매출관련변수

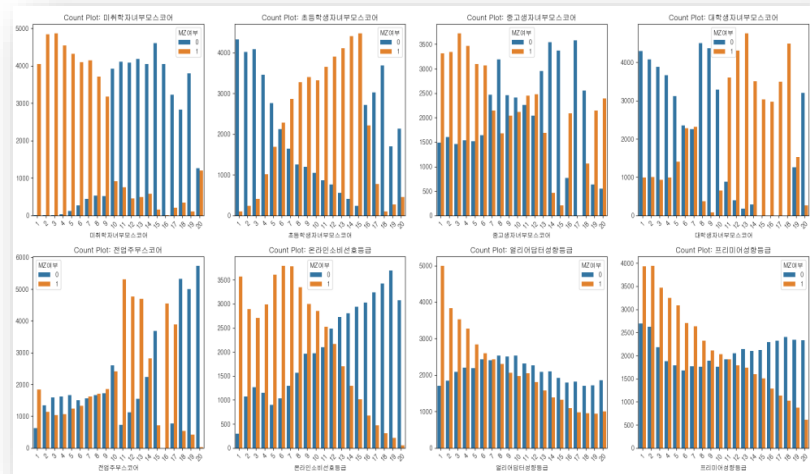
변수명
수도권거주여부
수도권직장여부

공간관련변수

* 항목명은 (요식 ~ 요양원)의 매출건수 / 매출금액 각각의 합계를 나타냄

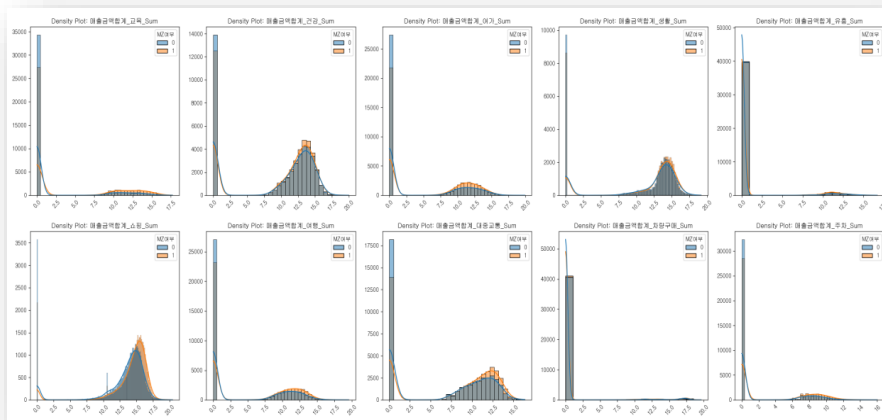
01 탐색적 데이터 분석 및 전처리

연령대에 따른 EDA



인구통계학적 변수

MZ세대와 기성세대가
대부분 상반되는 모습

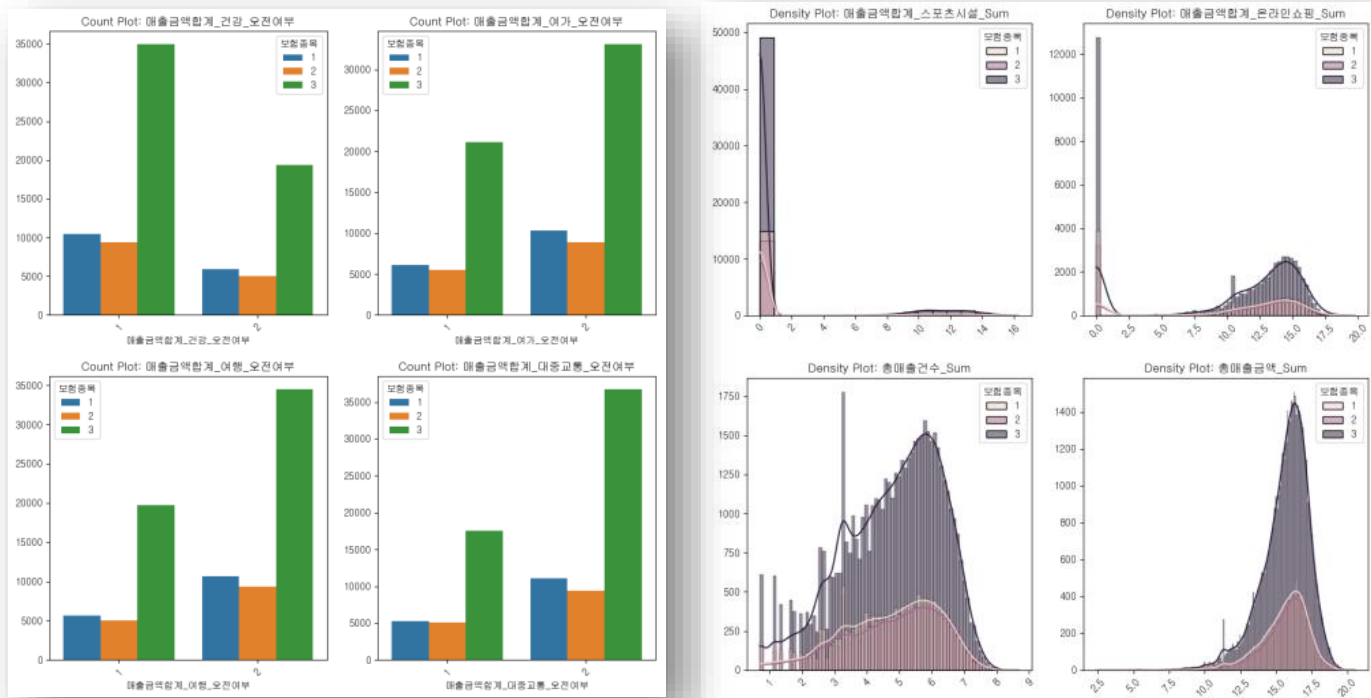


매출 관련 변수

MZ세대와 기성세대가
대부분 유사한 분포

01 탐색적 데이터 분석 및 전처리

보험종목에 따른 EDA



매출 관련 변수

생명보험, 장기보험, 자동차보험이
매출금액, 매출건수 등에서 유사한 분포

01 탐색적 데이터 분석 및 전처리

보험종목에 따른 EDA



연령 및 보험 종목을 통한 일차원적인 구분으로는 소비자의 다양한 지출 특성을 파악하는데 **한계**

매출패턴을 보다 상세하게 파악할 필요성 확인

매출 관련 변수

생명보험, 장기보험, 자동차보험이

매출금액, 매출건수 등에서 유사한 분포

카드 매출 관련 변수를 바탕으로 한 **군집분석** 시도

02
군집분석

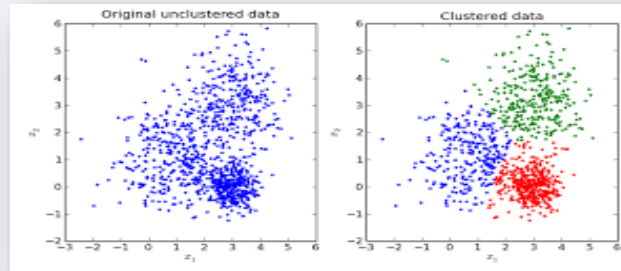
02 군집분석



군집분석(Clustering Analysis)

각 데이터의 유사성을 측정하여 다수의 그룹으로 나누고

그룹 간의 특성을 확인하는 비지도학습 기법



군집분석을 통해 숨겨진 매출패턴을 보다 상세하게 파악하고자 함

02 군집분석

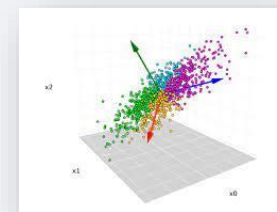
군집분석 과정 소개

1. Data Scaling

PCA에서 설명가능한 분산의 왜곡 발생을 방지하고
군집분석 시 거리 계산에 단위에 따른 차이를 없애기 위해
변인들의 범위를 맞춤

2. PCA

정보량이 95% 이상이 되도록
데이터의 **차원을 축소**해
손실을 최소화(5%)



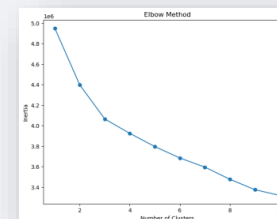
3. K-Means Clustering

특정한 임의의 지점을 선택해
그룹 중심점을 설정
→ **중심에 가장 가까운 포인트 선택**



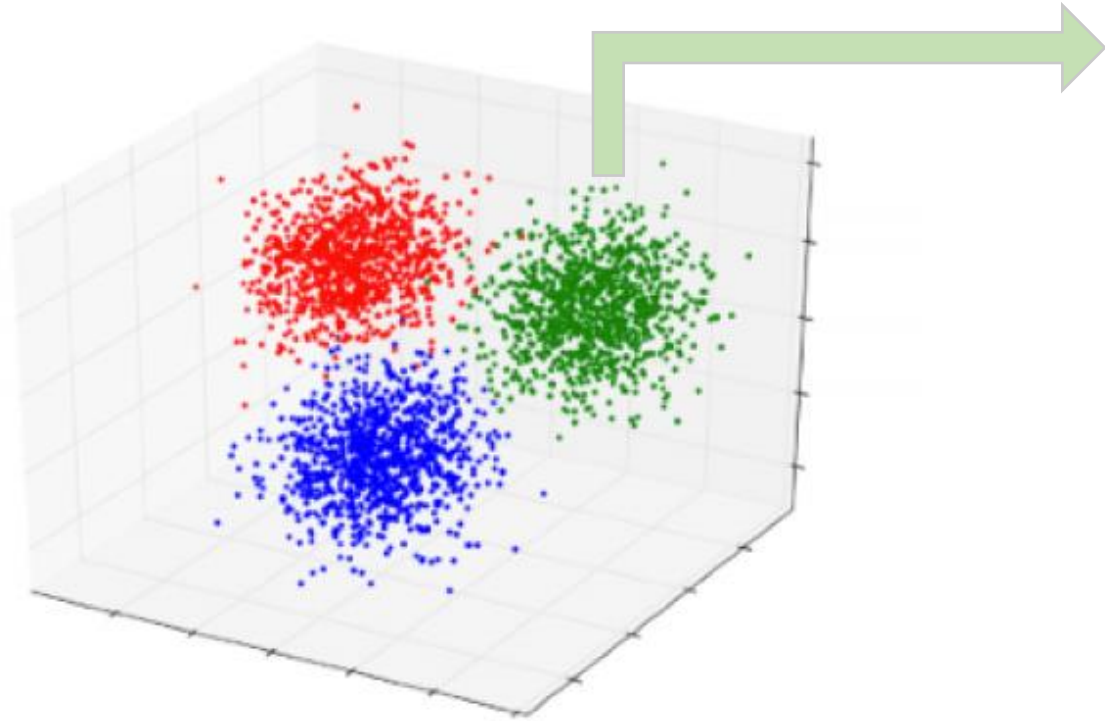
4. Elbow Method

클러스터링 시 적절한
그룹 개수를 선택하는 방법
(3개로 결정)

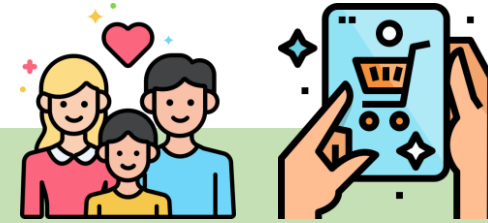


02 군집분석

군집분석 결과



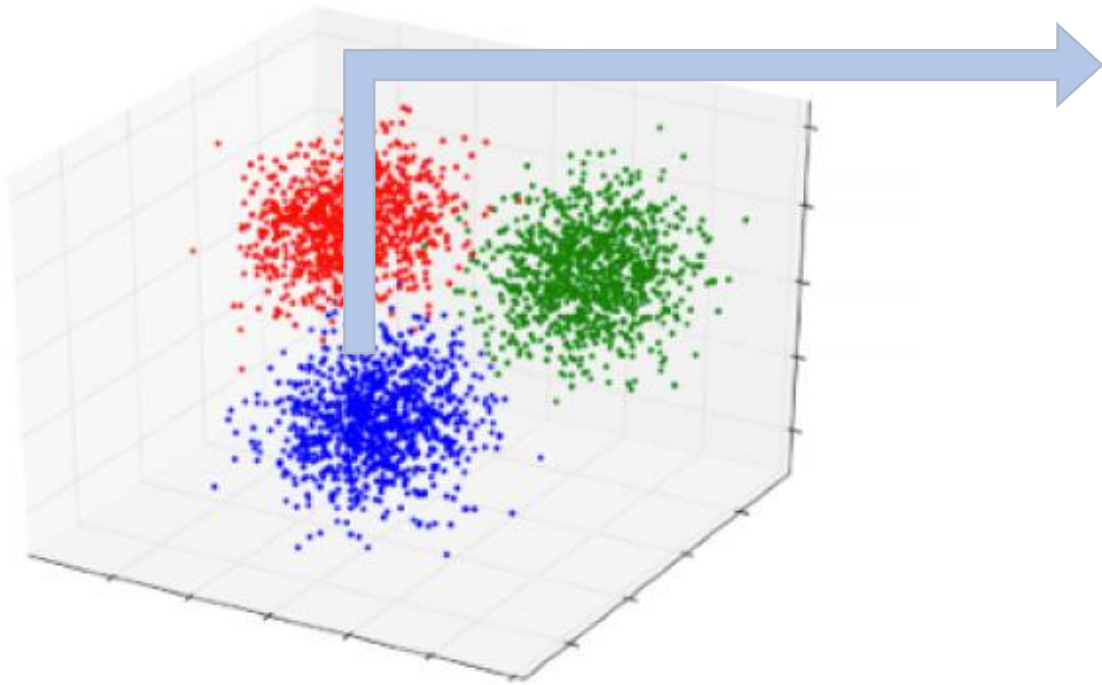
그룹 1



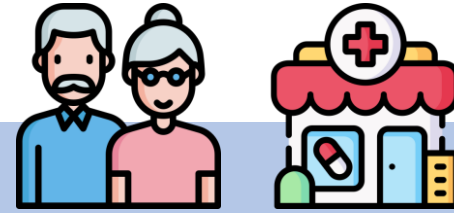
- MZ세대를 중심으로 세 그룹 중
저연령대가 가장 많이 분포
- 초등자녀부모가 가장 많음
- 그룹 2에 비해 '온라인쇼핑', '요식', '쇼핑',
'생활', '교육'의 매출금액합계 평균이 더 높음

02 군집분석

군집분석 결과



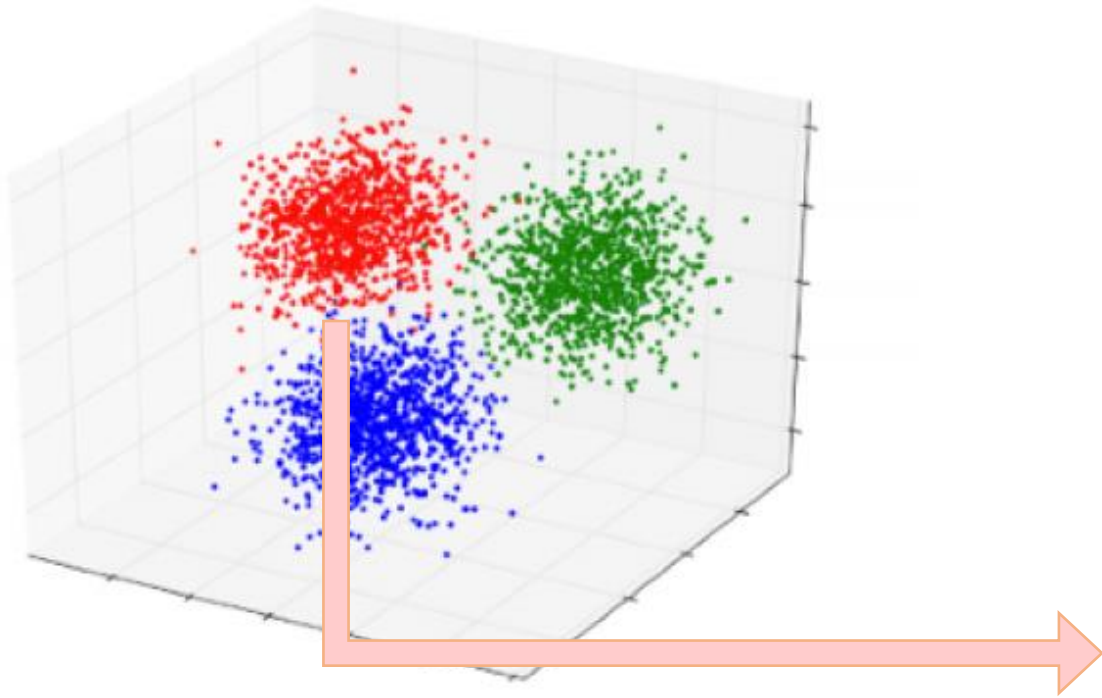
그룹 2



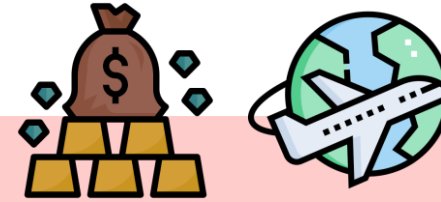
- 세 그룹 중 고연령대가 가장 많이 분포
- 성인자녀부모가 가장 많음
- 그룹 1에 비해 '건강보조식품', '건강', '양약국'의 매출금액합계 평균이 높음

02 군집분석

군집분석 결과



그룹 3



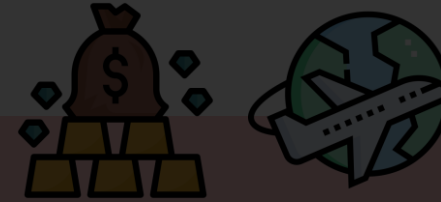
- 연령대가 고르게 분포
- 소비수준이 높고 고소득자의 비율이 높음
- 다른 두 그룹에 비해 반려동물을 보유하지 않을 가능성이 높고 해외여행에 대한 관심이 높음

02 군집분석

군집분석 결과



그룹 3



매출 관련 변수를 활용한 군집분석으로 도출된 결과에서
단순히 변수별로 기초통계량을 확인했을 때보다 훨씬 많은 정보를 파악

- 연령대가 고르게 분포
- 소마누인이 높고 고소득자의 비율이 높음
- 다른 두 그룹에 비해 반려동물을 보유하지 않을 가능성이 높고 해외여행에 대한 관심이 높음

03

자동차보험
사고율 예측

03 자동차보험 사고율 예측을 통한 결측치 처리

자동차보험 결측치 처리

자동차보험의 보험료가 모두 **결측치**로 사고율(손해액÷보험료)의 **직접적 계산 불가**

장기보험 사고율을 통해 자동차보험 사고율 **추정**

장기보험 : 자동차보험과 동일한 손해보험 범주, 사고율 계산식이 자동차보험과 동일(손해액÷보험료)



대용량 데이터에 적합

빠른 속도 및 높은 성능



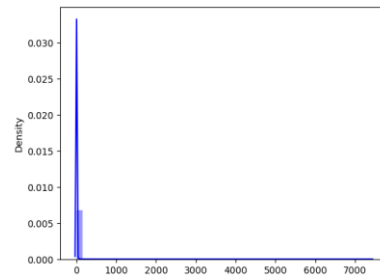
제한된 분석환경 내에서 신속한 최적화 및 모형 적합 가능

03 자동차보험 사고율 예측을 통한 결측치 처리

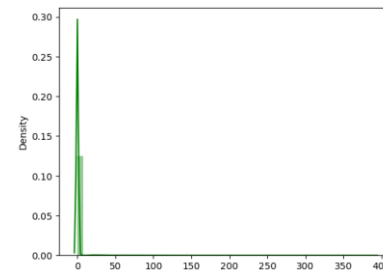
예측(대체) 모델링

모델의 최적화를 위해 gridsearchCV를 사용한 파라미터 튜닝 진행
최종 모델 : LGBMRegressor(max_depth=4, n_estimators=50)
통계적 모형에 비해 높은 예측 정확도

분포의 유사성을 통해 결측 사고율 대체가
적절하게 되었음을 확인



장기보험의 사고율 분포



자동차보험의 추정 사고율 분포

04
보험종목별
사고율 분석

04 보험 종목별 사고율 분석

1. 사고율 분석 모델

GLM (일반화 선형 모형)

목적변수(y)가 **정규분포를 따르지 않을때** 사용하는 회귀분석의 확장 모형

분석 데이터 특성 상 log 함수를 사용한 형태의 모형 사용

$$\log(y) = \alpha + \beta X$$



데이터 & 분석 목적에 대한 최적의 적합성



- 변수 영향력 설명에 있어서 **해석이 용이**하며 **불균형한 데이터**에 사용 가능
- 보험종목별 사고율에 대한 각 변수의 **영향력의 방향과 크기**를 개별적으로 파악 가능

04 보험 종목별 사고율 분석

1. 사고율 분석 모델

GLM (일반화 선형 모형)

목적변수(y)의 형태에 따른 다양한 세부 모형 존재

EX. 포아송 회귀, 음이항 회귀, 로지스틱 회귀, 감마 회귀

① 생명보험

사고건수 : 포아송 회귀모형

경과계약건수 : 음이항 회귀모형

$$\log(\text{사고율}) = \log(\text{사고건수}) - \log(\text{경과계약건수})$$

생명보험의 사고빈도는 **포아송 분포**를 따른다는
선행연구 기반

② 장기보험

로그 감마 모형

③ 자동차보험

로그 감마 모형

동일하게 **손해보험**으로 분류
손해보험의 사고율은 **감마분포**를 따른다는 선행 연구에 기반

차재형 외(2000), 기승도 외(2009)

04 보험 종목별 사고율 분석

2. 소비패턴에 따른 보험종목별 사고율 분석

① 생명보험

전반적으로 **건강** 관련 소비변수가 유의한 영향력 행사
매출금액에 비해 **매출건수**의 영향력이 상대적으로 크게 나타남

사고율을 높이는 요인(+)	사고율을 낮추는 요인(-)
매출건수합계_건강보조_Sum(3.64%)	매출건수합계_건강_Sum(-3.27%)
매출건수합계_병원_Sum(2.05%)	매출건수합계_주차_Sum(-2.68%)
매출금액합계_건강_Sum(1.23%)	매출건수합계_양약국_Sum(-2.04%)
매출금액합계_주차_Sum(1.18%)	매출금액합계_건강보조_Sum(-1.29%)
매출금액합계_양약국_Sum(0.76%)	

설명변수가 1% 증가할수록, 목적변수는 괄호 안의 수치% 만큼 증가 · 감소

04 보험 종목별 사고율 분석

2. 소비패턴에 따른 보험종목별 사고율 분석

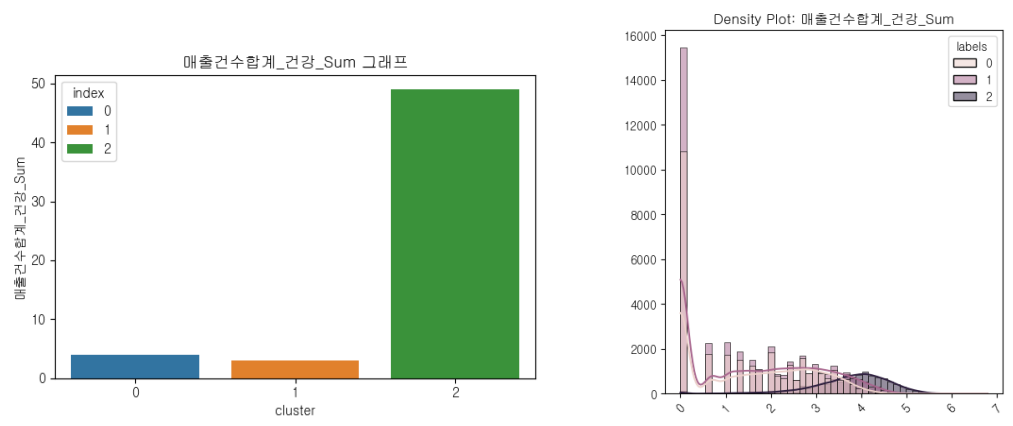
① 생명보험

성별에 따른 영향력 추이 : 남성 > 여성
 사고율 경향성 : 그룹 2 > 그룹 1 >>> 그룹 3

(By 건강보조 매출건수 등) (By 건강 매출건수 등)

사고율을 높이는 요인(+)	사고율을 낮추는 요인(-)
매출건수합계_건강보조_Sum(3.64%)	매출건수합계_건강_Sum(-3.27%)
매출건수합계_병원_Sum(2.05%)	매출건수합계_주차_Sum(-2.68%)
매출금액합계_건강_Sum(1.23%)	매출건수합계_양약국_Sum(-2.04%)
매출금액합계_주차_Sum(1.18%)	매출금액합계_건강보조_Sum(-1.29%)
매출금액합계_양약국_Sum(0.76%)	

변수에 대한 그룹 간 차이가 클수록 사고율 차이 극대화



건강 매출건수에 대한 그룹 별 중앙값 및 분포 시각화

04 보험 종목별 사고율 분석

2. 소비패턴에 따른 보험종목별 사고율 분석

② 장기보험

매출금액에 비해 **매출건수**의 영향력이 상대적으로 크게 나타남

사고율을 높이는 요인(+)	사고율을 낮추는 요인(-)
매출건수합계_유흥 (1.3302)	매출금액합계_유흥(0.9999)
초등학생자녀부모스코어 (1.04204)	매출건수합계_온라인쇼핑(0.9943)
전업주부스코어 (1.035651)	매출건수합계_여가(0.9736)
매출건수합계_쇼핑 (1.004934)	매출건수합계_건강보조식품(0.6044)
매출금액합계_건강보조식품 (1.00002)	매출건수합계_요양원(0.5360)

그룹 1

건강보조식품 매출건수의
차이로 낮은 사고율 추정

그룹 2 > 그룹 1 >>> 그룹 3

그룹 2

낮은 온라인쇼핑 매출건수
높은 전업주부스코어

그룹 3

주요 유효 변수들의
높은 매출건수 · 금액

온라인쇼핑 매출건수, 유흥 매출금액 등

04 보험 종목별 사고율 분석

2. 소비패턴에 따른 보험종목별 사고율 분석

③ 자동차보험

관심성향에 대한 등급 변수의 영향력이 두드러짐

등급(1~20)이 1에 가까울수록 사고율이 높아짐

야외활동관심성향 & 얼리어답터성향

사고율과 음의 관계를 보임

관심도 : 그룹 3 > 그룹 1 > 그룹 2

사고율 : 그룹 2 > 그룹 1 > 그룹 3

사고율을 높이는 요인(+)	사고율을 낮추는 요인(-)
프리미어성향등급(0.9996)	야외활동관심성향등급(1.01801)
반려동물보유성향등급(0.9967)	얼리어답터성향등급(1.0161)
인테리어스타일링관심성향등급(0.9899)	온라인소비선호성향등급(1.004701)
배달식품관심성향등급(0.9886)	
인테리어모델링관심성향등급(0.9798)	

등급(1~20)이 1에 가까울수록 사고율이 낮아짐

04 보험 종목별 사고율 분석

2. 소비패턴에 따른 보험종목별 사고율 분석

③ 자동차보험

관심성향에 대한 등급 변수의 영향력이 두드러짐

배달식품관심성향 & 반려동물보유성향

그룹 3에서 가장 낮게 나타남

➡ 그룹 3의 낮은 사고율에 영향!

등급(1~20)이 1에 가까울수록 사고율이 높아짐

사고율을 높이는 요인(+)	사고율을 낮추는 요인(-)
프리미어성향등급(0.9996)	야외활동관심성향등급(1.01801)
반려동물보유성향등급(0.9967)	얼리어답터성향등급(1.0161)
인테리어스타일링관심성향등급(0.9899)	온라인소비선호성향등급(1.004701)
배달식품관심성향등급(0.9886)	
인테리어모델링관심성향등급(0.9798)	

등급(1~20)이 1에 가까울수록 사고율이 낮아짐

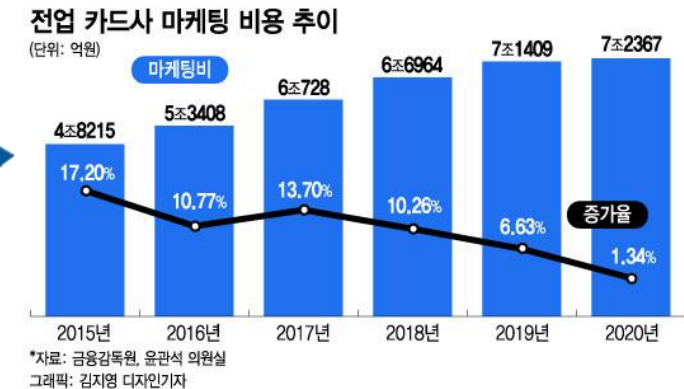
05
아이디어 및
기대효과

05 아이디어 및 기대효과

삼성카드 ① : 고객 맞춤 카드 혜택 필터링 시스템

아이디어 배경

전년대비 카드사 이익 감소로 여러 카드사가 할인 혜택 축소하고 있는 실정
카드 혜택 시행은 **고객 문의 응대, 마케팅 제휴** 등의 측면에서 카드사에 **비용적 부담**을 초래



05 아이디어 및 기대효과

삼성카드 ① : 고객 맞춤 카드 혜택 필터링 시스템

아이디어 배경

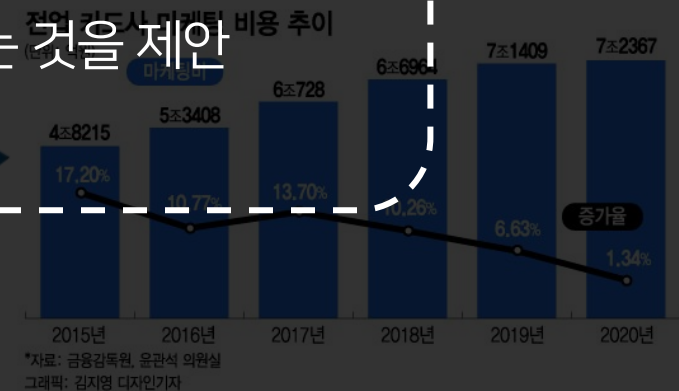
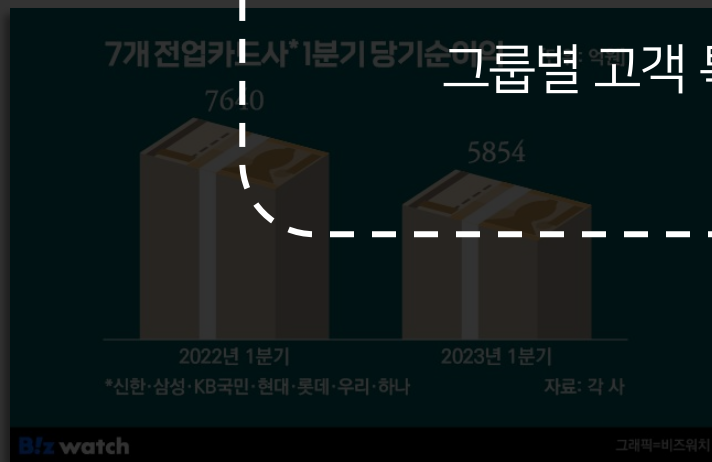


전년대비 카드사 이익 감소로 여러 카드사가 할인 혜택 축소하고 있는 실정

카드 혜택 시행은 **고객 문의 응대, 마케팅 제휴** 등의 측면에서 카드사에 **비용적 부담**을 초래

고객 맞춤 카드 혜택 필터링 시스템을 통해

그룹별 고객 특성에 기반한 카드 혜택을 추천하는 것을 제안

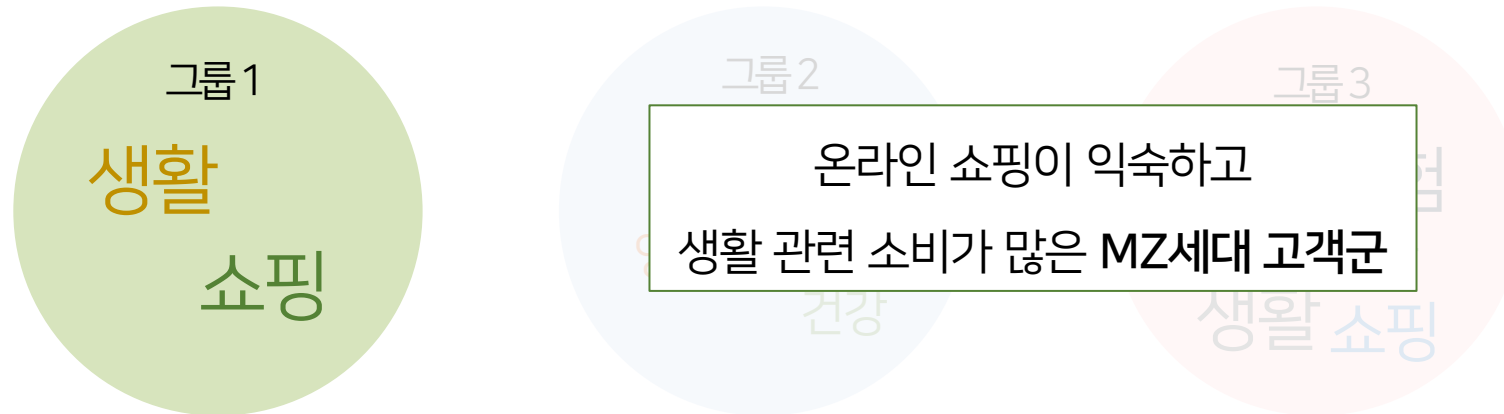


05 아이디어 및 기대효과

삼성카드 ① : 고객 맞춤 카드 혜택 필터링 시스템

아이디어 세부내용

군집분석 기법을 적용해 변화하는 고객 데이터에 적합한 카드 혜택 추천 및 필터링
'삼성 iD VITA 카드' 할인 혜택을 분석 결과에 기반해 재구성

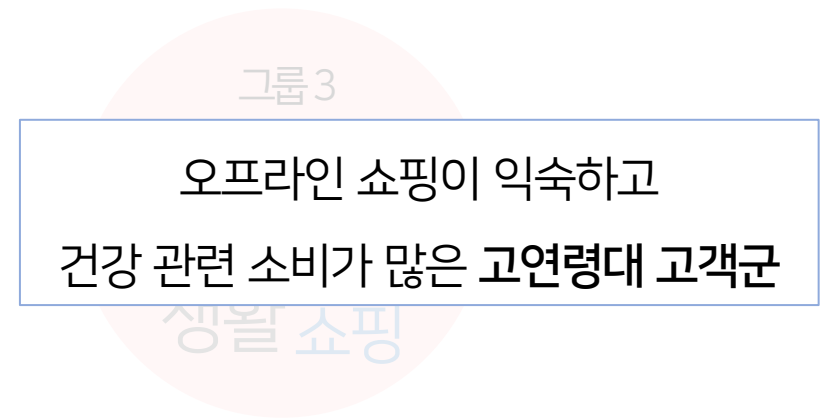


05 아이디어 및 기대효과

삼성카드 ① : 고객 맞춤 카드 혜택 필터링 시스템

아이디어 세부내용

군집분석 기법을 적용해 변화하는 고객 데이터에 적합한 카드 혜택 추천 및 필터링
'삼성 iD VITA 카드' 할인 혜택을 분석 결과에 기반해 재구성

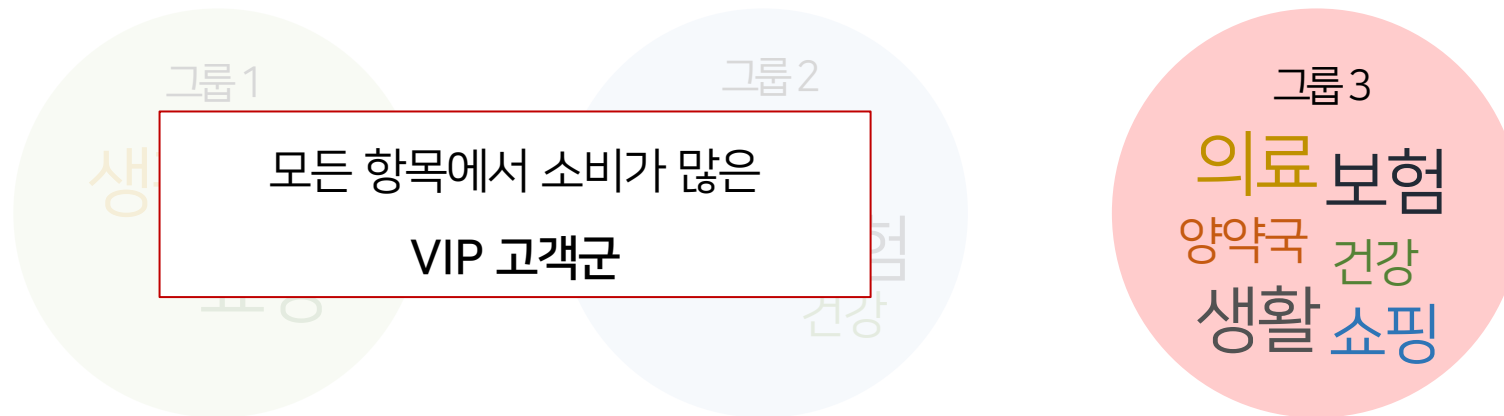


05 아이디어 및 기대효과

삼성카드 ① : 고객 맞춤 카드 혜택 필터링 시스템

아이디어 세부내용

군집분석 기법을 적용해 변화하는 고객 데이터에 적합한 카드 혜택 추천 및 필터링
'삼성 iD VITA 카드' 할인 혜택을 분석 결과에 기반해 재구성



05 아이디어 및 기대효과

삼성카드 ① : 고객 맞춤 카드 혜택 필터링 시스템

아이디어 세부내용

군집분석 기법을 적용해 변화하는 고객 데이터에 적합한 카드 혜택 추천 및 필터링
'삼성 iD VITA 카드' 할인 혜택을 분석 결과에 기반해 재구성

오프라인 쇼핑이 익숙하고
건강 관련 소비가 많은 고연령대 고객군



의료, 보험 할인 혜택이 많은
'삼성 iD VITA' 카드 추천

이동통신/렌탈/멤버
십 10% 결제일 할인

헬스/뷰티 20% 결제
일 할인 中 온라인 아
모레몰 혜택

05 아이디어 및 기대효과

삼성카드 ① : 고객 맞춤 카드 혜택 필터링 시스템

아이디어 세부내용

군집분석 기법을 적용해 변화하는 고객 데이터에 적합한 카드 혜택 추천 및 필터링
'삼성 iD VITA 카드' 할인 혜택을 분석 결과에 기반해 재구성

오프라인 쇼핑이 익숙하고
건강 관련 소비가 많은 고연령대 고객군

그룹2

추천 고객의 특성을 고려해
iD VITA 카드 혜택 필터링



혜택



이동통신/렌탈/멤버
십 10% 결제일 할인

헬스/뷰티 20% 결
제일 할인 중 온라인
아모레몰 혜택

05 아이디어 및 기대효과

삼성카드 ① : 고객 맞춤 카드 혜택 필터링 시스템

아이디어 세부내용

군집분석 기법을 적용해 변화하는 고객 데이터에 적합한 카드 혜택 추천 및 필터링
'삼성 iD VITA 카드' 할인 혜택을 분석 결과에 기반해 재구성

오프라인 쇼핑이 익숙하고
건강 관련 소비가 많은 고연령대 고객군

그룹2

'생활 영역의 소비가 적다'는 점과 '온라인 소비가 적다'는 점을 고려해
관련 기존 혜택을 줄일 수 있음

혜택

이동통신/렌탈/멤버
십 10% 결제일 할인

헬스/뷰티 20% 결
제일 할인 중 온라인
아모레몰 혜택

05 아이디어 및 기대효과

삼성카드 ① : 고객 맞춤 카드 혜택 필터링 시스템

아이디어 기대효과

고객이 선호할만한 혜택들만 담아 만든 카드를 통해 소비자의 마음을 사로잡을 수 있음

할인 혜택의 필터링을 통한 비용 절감

(혜택 문의 비용, 할인 혜택 제공을 위한 제휴 비용 등)



05 아이디어 및 기대효과

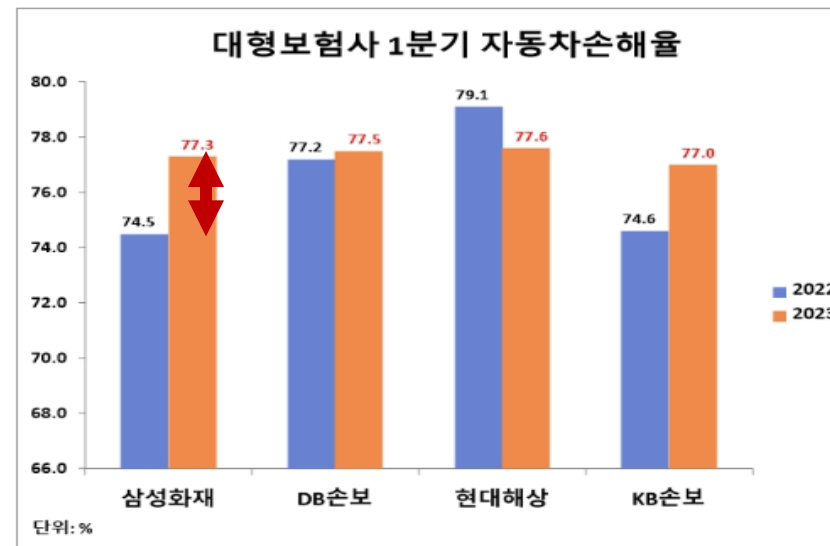
보험개발원 ① : 카드 소비변수 활용 보험손해율 예측모형 고도화

아이디어배경

보험 손해율의 변화를 예측한다면 대비책 마련 가능

언더라이팅에 보편적인 지표(ex : 사고 건수) + 카드 소비 변수 **종합적**으로 고려한 손해율 예측

→ **우량물건 정교하게 선별할 필요성**



출처: 손해보험업계

05 아이디어 및 기대효과

보험개발원 ①: 카드 소비변수 활용 보험손해율 예측모형 고도화

아이디어 세부 내용 1) 통계적 모델링 기법 활용 변수 선택

GLM을 통해 사고율 예측에 유의한 영향을 미치는 카드매출 관련 변수 도출
해당 변수를 활용해 보험손해율 예측 모형을 개발할 수 있음

보험	방향성	변수명
생명보험	양의 방향 (+)	건강보조, 병원 등의 매출 건수 합계
	음의 방향 (-)	건강, 주차 등의 매출 건수 합계
장기보험	양의 방향 (+)	쇼핑, 유흥 등의 매출 건수 합계
	음의 방향 (-)	건강보조식품, 여가 등의 매출건수합계
자동차보험	양의 방향 (+)	배달식품관심성향, 프리미어성향
	음의 방향 (-)	야외활동관심성향, 얼리어답터성향

05 아이디어 및 기대효과

보험개발원 ①: 카드 소비변수 활용 보험손해율 예측모형 고도화

아이디어 세부 내용 2) 머신러닝 모델 & XAI 활용 예측

현업의 대규모 데이터 처리에는 머신러닝 모델링 기법이 효과적일 것으로 판단해 LGBM 활용 예측 모델링 제안
XAI(설명가능한 인공지능) 기법을 이용해 머신러닝 예측 근거 제공



“위 아이디어의 실현가능성을 아래와 같이 확인”

장기보험 손해율 AI 예측 모델링을 적용
XAI 기법으로 변수 중요도를 통해 예측에 유의미한 영향을 준 요인을 추출

05 아이디어 및 기대효과

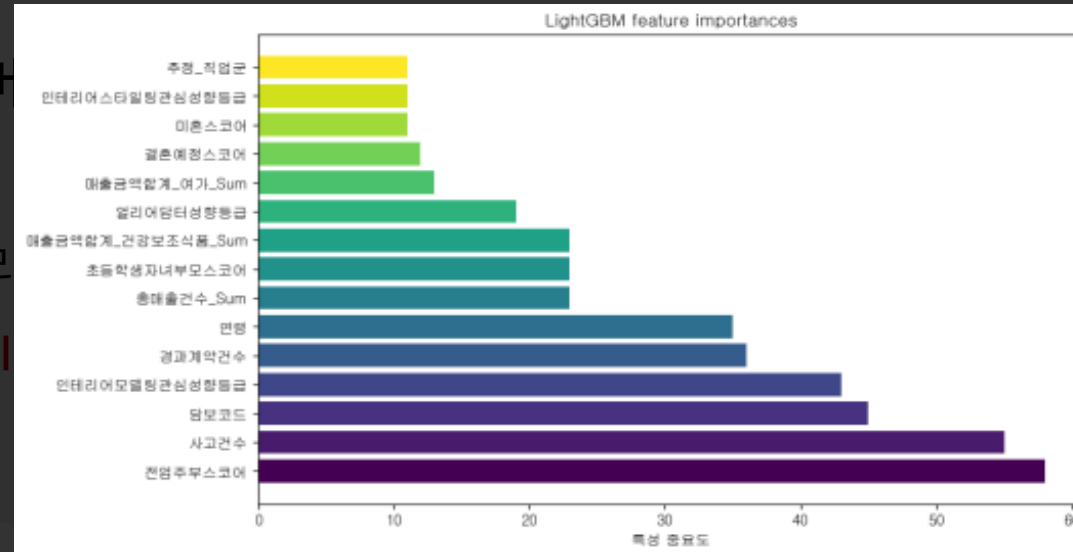
보험개발원 ①: 카드 소비빈수 등등 보험손에볼 예측모형 고도화

XAI(설명가능한 인공지능)기법 적용 결과

아이디어 세부 내용 2)마

현업의 대규모 데이터 처리

XAI



BM 활용 예측 모델링 제안

제공

“위 아이디어의 실현가능성을 아래와 같이 확인”

전업주부스코어, 사고건수, 담보코드, 인테리어모델링관심성향 등급, 경과계약건수 등이

장기보험 손해율 AI 예측 모델링을 적용

AI 손해율 예측 결과의 영향을 많이 미친 것으로 추정

XAI 기법으로 Feature Importance를 사용해 예측에 유의미한 영향을 준 요인을 추출

05 아이디어 및 기대효과

보험개발원 ①: 카드 소비변수 활용 보험손해율 예측모형 고도화

아이디어 기대효과

- ➡ • 언더라이팅 심사 기준을 카드 소비 관련 변수로 **확대**한 엄밀한 심사 → **손해율 완화**
- 손해율을 낮추는 데에 기여하는 카드 소비 변수를 통한 보험료 할인 등의 **혜택** 적용 → **고객 만족도 상승** 기대
- 통계적 모델과 머신러닝 기법을 동시에 고려 → 계약 심사의 **정확성**과 **속도**를 모두 제고

05 아이디어 및 기대효과

보험개발원 ① : 카드 소비변수 활용 보험손해율 예측모형 고도화

아이디어 기대효과

- 언더라이팅 심사 기준을 카드 소비 관련 변수로 **확대**한 엄밀한 심사 → **손해율 완화**

- ➡
- 손해율을 낮추는 데에 기여하는 카드 소비 변수를 통한 보험료 할인 등의 **혜택** 적용 → **고객 만족도 상승** 기대

- 통계
- ▶ 계약 심사의 **정확성**과 **속도**를 모두 제고

생명보험 사고율을 낮추는 요인
매출건수합계_건강_Sum(-3.27%)
매출건수합계_주차_Sum(-2.68%)
매출건수합계_양약국_Sum(-2.04%)
매출금액합계_건강보조_Sum(-1.29%)

건강, 양약국 매출건수합계, 건강보조 매출금액합계

증가하면 보험료 할인혜택 제공



05 아이디어 및 기대효과

보험개발원 ① : 카드 소비변수 활용 보험손해율 예측모형 고도화

아이디어 기대효과

- 언더라이팅 심사 기준을 카드 소비 관련 변수로 확대한 엄밀한 심사 → 손해율 완화
- 손해율을 낮추는 데에 기여하는 카드 소비 변수를 통한 보험료 할인 등의 혜택 적용 → 고객 만족도 상승 기대
- ➔ • 통계적 모델과 머신러닝 기법을 동시에 고려 → 계약 심사의 정확성과 속도를 모두 제고

05 아이디어 및 기대효과

보험개발원 ② : 고객 특성을 고려한 맞춤형 보험 상품 추천

아이디어 배경

보험개발원 : "자체 보유 보험 정보 + 타 기관 데이터 결합" → 새로운 마케팅 포인트를 발굴하고자 함

자신에게 필요한 것을 파악한 합리적 선택 중시하는 MZ세대 : 고객에게 필요한 보험 상품 추천이 MZ세대 타겟팅에 효과적

허창언 보험개발원장은 빅데이터와 연계해 이용량 연계형 보험 상품(UBI)을 개발하는 등 디지털 중심 보험 시장을 구축하겠다고 14일 밝혔다.

허창언 원장은 이날 취임 100일을 맞아 개최한 기자회견에서 "대내외 복합 위기에 직면한 보험 산업이 새롭게 도약하기 위해서는 고객 맞춤형 보험 상품 개발이 필수적이다. **보험개발원은 올해 초 자체 보유한 보험 정보와 다른 기관의 금융 및 비금융 데이터 결합을 통한 마케팅 포인트 제시 계획을 밝혔다**

이를 위해 보험개발원은 올해 자체 보유한 보험 정보와 다른 기관의 금융·비금융 데이터 결합을 통해 새로운 마케팅 포인트를 제시할 계획이다.

예를 들어 보험 정보와 은행 등의 금융 정보를 결합해 소득 수준별 금융상품 가입 현황을 분석하고 소비자의 노후 소득 부족시 장수 위험에 대비할 수 있는 보험 상품을 안내하는 방식이다.

나에게 적합한 상품 추천,
저렴한 가격, 합리적 추천,
유익한 정보 제공 등

보험소비에서 MZ세대가 원하는 경험

합리

신뢰

편리

신속

자기주도

출처: 보험연구원

05 아이디어 및 기대효과

보험개발원 ② : 고객 특성을 고려한 맞춤형 보험 상품 추천

아이디어 세부 내용

군집분석을 통해 도출한 고객 그룹별 특성 도출
→ 해당 그룹 고객을 위한 보험 상품 추천



그룹 1

MZ세대의 비율이 높음
가성비 & 보험 가입의 편의성 중시
여행 관련 소비 多



보험 상품 추천

온라인으로 가입 가능한 보험 상품

미니보험상품

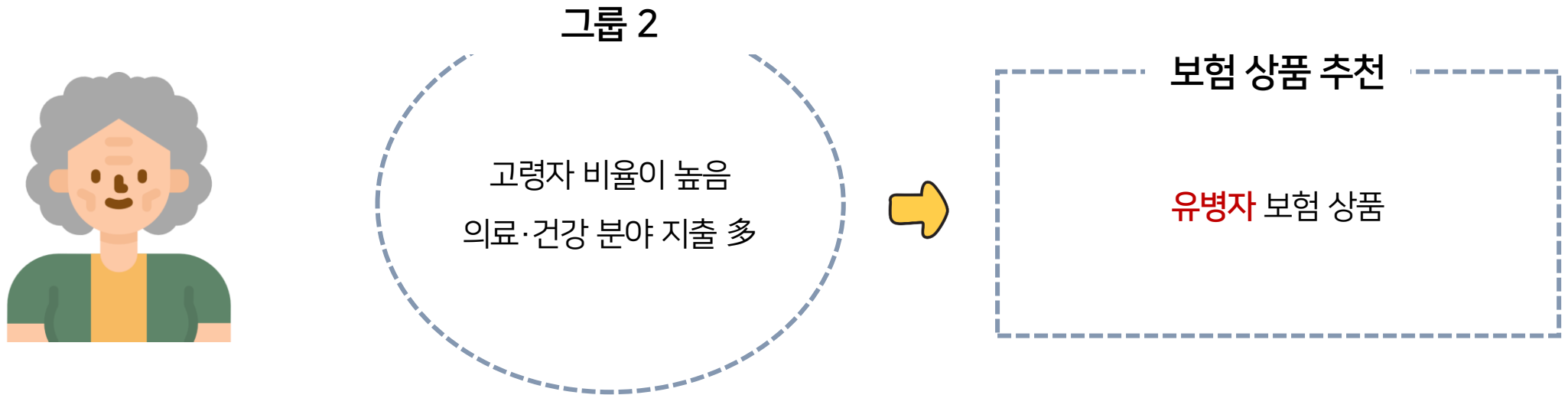
해외여행자보험

05 아이디어 및 기대효과

보험개발원 ② : 고객 특성을 고려한 맞춤형 보험 상품 추천

아이디어 세부 내용

군집분석을 통해 도출한 고객 그룹별 특성 도출
→ 해당 그룹 고객을 위한 보험 상품 추천



05 아이디어 및 기대효과

보험개발원 ②: 고객 특성을 고려한 맞춤형 보험 상품 추천

아이디어 세부 내용

군집분석을 통해 도출한 고객 그룹별 특성 도출
→ 해당 그룹 고객을 위한 보험 상품 추천



그룹 3

소비 수준이 높은 집단
프리미엄 소비성향이 높음
타 그룹 비해 높은 전문직 비율



보험 상품 추천

연금저축보험 상품 등의 자산관리 상품

05 아이디어 및 기대효과

보험개발원 ② : 고객 특성을 고려한 맞춤형 보험 상품 추천

아이디어 기대효과

맞춤형 마케팅 시행 → 보험가입에 대한 관심이 적은 **MZ세대의 보험 가입률 개선** 효과 기대

보험상품 탐색 시간 단축 → **고객 만족도 제고** 효과 기대



감사합니다