

1. 프로젝트 개요

소속 회사명	콘센트릭스 서비스 코리아(유)
프로젝트 진행일	2020.07 - 2020.11
주사용 기술	Machine Learning, Python, SQL
프로젝트 개요	충성고객을 정의하고 이탈 원인을 분석하여 이탈 확률과 이탈 원인을 예측하는 모델 개발

2. 주요업무

충성고객 정의

- 구매횟수, 구매금액, 객단가 등의 지표를 이용하여 고객 유형 분리
- 구매횟수와 멤버십 등급 기준으로 충성고객 정의

이탈 기준 정의

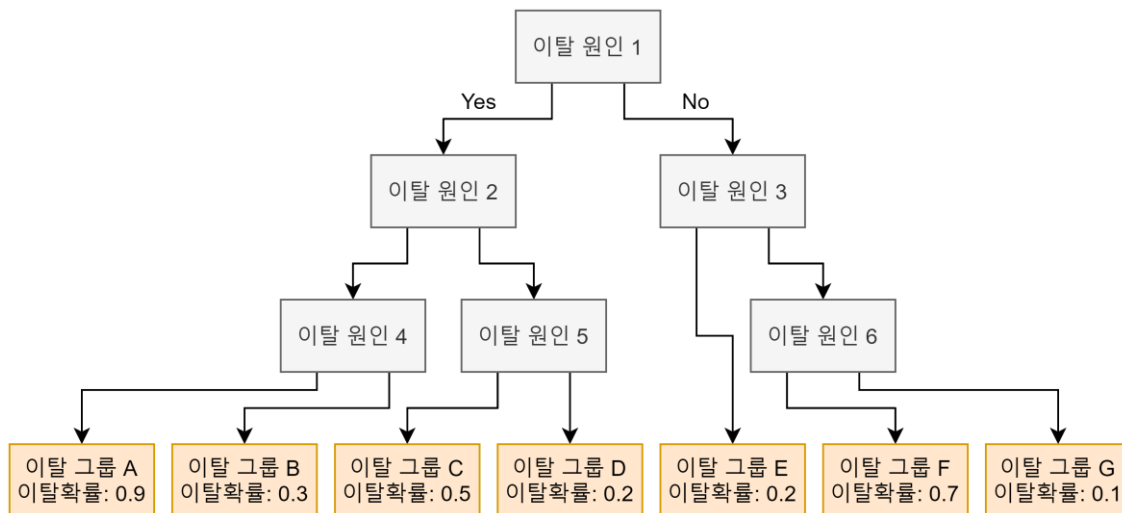
- 미구매 기간에 따른 미구매 기간 1년 이상인 고객의 비율 비교
- 급격히 비율이 상승하는 구간을 참고하여 이탈 기준 정의

이탈 원인 분석

- 이탈 유형 정의
- 이탈 모델의 성능 향상을 위한 이탈 원인 Feature 분석

이탈 원인 해석 모델 개발

- 이탈 원인 Feature를 모델(Decision Tree)에 적용
- 이탈 원인에 따른 이탈 고객 유형화
- [그림 2.1]에서 이탈 그룹 A의 경우, 이탈 확률이 90%이며 '이탈 원인은 이탈 원인 1,2,4이다'라고 해석
- [그림 2.1]에서 이탈 그룹 G의 경우, 이탈 확률이 10%이며 '이탈 원인을 파악할 수 없다'라고 해석



[그림 2.1] 이탈 원인 해석 모델

3. 주요 이슈 및 해결

활성 고객 전체를 대상으로 모델링 진행

- 1회성 고객 또는 구매력이 낮은 고객의 경우 데이터 정보가 부족하여 이탈 원인이 해석되지 않음
- 구매 활동이 활발한 충성고객을 정의하고 충성고객을 대상으로 모델링 진행

딥러닝으로 모델링 후, 딥러닝 해석 모델인 라임(LIME)으로 이탈 원인을 해석하고자 함

- 시간이 많이 소비됨
- 딥러닝과 성능이 비슷하고 해석이 용이한 Decision Tree 사용
(실험결과 Decision Tree의 성능이 더 우수했음)

이탈 원인이 해석되지 않은 고객

- [그림 2.1]의 이탈 그룹 G와 같이 이탈 원인을 파악할 수 없는 고객이 많았음
- 이탈 원인을 파악할 수 없는 고객의 신규 고객 비율, 미구매 경과월 등 추가 분석 진행
이탈 불분명 고객 중 미구매 경과 3개월 이상인 고객이 약 30%이며 실제 이탈 고객 비중은 약 60%에 달했음
- '미구매 경과 3개월 이상 여부'를 모델에 적용

4. 성과

분석을 통해 충성고객 정의

충성고객 중 예상 이탈 고객의 이탈 원인에 따른 이탈 방지 프로모션 지원

머신러닝 기반 CRM 운영 시스템의 충성고객 이탈방지 모델 추가