

제13회 2025 빅콘테스트 AI데이터 분석분야

우리 동네 가맹점, 위기 신호를 미리 잡아라!

배경

우리동네의 작은 가게, 즉 소상공인과 자영업자가 운영하는 가맹점들은 우리 경제의 근간을 이루는 중요한 축입니다. 하지만, 급변하는 환경 속에서 다양한 위기에 쉽게 노출됩니다. 대부분의 위기는 매출, 고객, 상권 등 데이터 속에 미세한 신호로 먼저 나타나지만, 개인의 경험만으로는 이를 미리 감지하기 어렵습니다.

2025 빅콘테스트에서 AI기술을 활용해 이러한 위기 징후를 선제적으로 포착하는 'AI 조기 경보 시스템' 예측 모델을 제안해 주세요. 여러분의 아이디어가 소상공인의 안정적인 성장과 건강한 지역 상권을 만드는데 기여 할 수 있습니다.

주제

미래를 예측하는 데이터 챌린지 : 우리 동네 가맹점, 위기 신호를 미리 잡아라!
(영세/중소 요식 가맹점 경영 위기 조기 경보 시스템 제안)

설명

- 제공된 다양한 데이터를 활용해 우리 주변의 영세/중소 가맹점이 갑자기 어려워지는 순간(매출 급락, 폐업 위험 등)을 미리 감지할 수 있는 'AI 조기 경보 시스템' 제안
- 모델 개발뿐 아니라, 데이터에서 어떤 신호가 위기를 미리 알려주는지 직접 찾아내고, 이 결과를 실제로 소상공인이 쉽게 활용할 수 있는 아이디어도 함께 제안

제공 데이터-1

[데이터셋1] 가맹점 개요정보

[상세조건]

- 신한카드 요식업종 서울 성동구 지역 영세/중소 가맹점
- 데이터 총 건수 4,244건

구분	컬럼명	컬럼한글명	타입	NULL	항목 설명
1	ENCODED_MCT	가맹점구분번호	VARCHAR	N	
2	GDS_AF_MCT_AR	가맹점주소	VARCHAR	N	
3	MCT_NM	가맹점명	VARCHAR	N	
4	MCT_BRD_NUM	브랜드구분코드	VARCHAR	Y	동일 브랜드 매장 식별을 위한 구분코드
5	MCT_SIGUNGU_NM	가맹점지역	VARCHAR	N	
6	HPSN_MCT_ZCD_NM	업종	VARCHAR	N	
7	HPSN_MCT_BZN_CD_NM	상권	VARCHAR	Y	
8	ARE_D	개설일	VARCHAR	N	
9	MCT_ME_D	폐업일	VARCHAR	Y	

제공 데이터-2

[데이터셋2] 가맹점 월별 이용 정보

[상세조건]

- 신한카드 요식업종 서울 성동구 지역 영세/중소 가맹점
- 기간 : 2023년 01월 ~ 2024년 12월
- 데이터 총 건수 87,826건

구분	컬럼명	컬럼한글명	타입	NULL	항목 설명
1	ENCODED_MCT	가맹점구분번호	VARCHAR	N	
2	TA_YM	기준년월	VARCHAR	N	
3	MCT_OPE_MS_CN	가맹점 운영개월수 구간	VARCHAR	N	6개 구간(10%이하, 10-25%, 25-50%, 50-75%, 75-90%, 90%초과) 0%에 가까울 수록 상위 (운영개월 수 상위)
4	RC_M1_SAA	매출금액 구간	VARCHAR	N	6개 구간(10%이하, 10-25%, 25-50%, 50-75%, 75-90%, 90%초과) 0%에 가까울 수록 상위 (매출금액 상위)
5	RC_M1_TO_UE_CT	매출건수 구간	VARCHAR	N	6개 구간(10%이하, 10-25%, 25-50%, 50-75%, 75-90%, 90%초과) 0%에 가까울 수록 상위 (매출건수 상위)
6	RC_M1_UE_CUS_CN	유니크 고객 수 구간	VARCHAR	N	6개 구간(10%이하, 10-25%, 25-50%, 50-75%, 75-90%, 90%초과) 0%에 가까울 수록 상위 (고객 수 상위)
7	RC_M1_AV_NP_AT	객단가 구간	VARCHAR	N	6개 구간(10%이하, 10-25%, 25-50%, 50-75%, 75-90%, 90%초과) 0%에 가까울 수록 상위 (객단가 상위)
8	APV_CE_RAT	취소율 구간	VARCHAR	Y	6개 구간, 1구간에 가까울 수록 상위 (취소율 낮음)
9	DLV_SAA_RAT	배달매출금액 비율	NUMBER	N	배달매출 미존재시 SV (-999999.9)
10	M1_SME_RY_SAA_RAT	동일 업종 매출금액 비율	NUMBER	N	동일 업종 매출 금액 평균 대비 해당 가맹점 매출 금액 비율 (평균과 동일 : 100%)
11	M1_SME_RY_CNT_RAT	동일 업종 매출건수 비율	NUMBER	N	동일 업종 매출 건수 평균 대비 해당 가맹점 매출 건수 비율 (평균과 동일 : 100%)
12	M12_SME_RY_SAA_PCE_RT	동일 업종 내 매출 순위 비율	NUMBER	N	업종 내 순위 / 업종 내 전체 가맹점 * 100 (0에 가까울수록 상위)
13	M12_SME_BZN_SAA_PCE_RT	동일 상권 내 매출 순위 비율	NUMBER	N	상권 내 순위 / 상권 내 전체 가맹점 * 100 (0에 가까울수록 상위)
14	M12_SME_RY_ME_MCT_RAT	동일 업종 내 해지 가맹점 비중	NUMBER	N	
15	M12_SME_BZN_ME_MCT_RAT	동일 상권 내 해지 가맹점 비중	NUMBER	N	상권 미존재시 SV (-999999.9)

제공 데이터-3

[데이터셋3] 가맹점 월별 이용 고객정보

[상세요건]

- 신한카드 요식업종 서울 성동구 지역 영세/중소 가맹점
- 기간 : 2023년 01월 ~ 2024년 12월
- 데이터 총 건수 87,826건

구분	컬럼명	컬럼한글명	타입	NULL	항목 설명
1	ENCODED_MCT	가맹점구분번호	VARCHAR	N	
2	TA_YM	기준년월	VARCHAR	N	
3	M12_MAL_1020_RAT	남성 20대이하 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
4	M12_MAL_30_RAT	남성 30대 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
5	M12_MAL_40_RAT	남성 40대 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
6	M12_MAL_50_RAT	남성 50대 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
7	M12_MAL_60_RAT	남성 60대이상 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
8	M12_FME_1020_RAT	여성 20대이하 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
9	M12_FME_30_RAT	여성 30대 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
10	M12_FME_40_RAT	여성 40대 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
11	M12_FME_50_RAT	여성 50대 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
12	M12_FME_60_RAT	여성 60대이상 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
13	MCT_UE_CLN_REU_RAT	재방문 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
14	MCT_UE_CLN_NEW_RAT	신규 고객 비중	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
15	RC_M1_SHC_RSD_UE_CLN_RAT	거주 이용 고객 비율	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
16	RC_M1_SHC_WP_UE_CLN_RAT	직장 이용 고객 비율	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)
17	RC_M1_SHC_FLP_UE_CLN_RAT	유동인구 이용 고객 비율	NUMBER	N	고객 정보 미존재시 SV (-9999999.9)

평가

[1차 서류 평가 기준]

- 모델링의 적합성/완성도(25점)
 - : 전처리, 변수선정, 적절한 모델 선택 및 적용, 모델 평가, 해석(25점)
- 데이터 활용/분석력(25점)
 - : 데이터 수집/정제/해석, 탐색적분석&시각화, 논리성
- 인사이트/실효성(20점)
 - : 도출된 인사이트의 참신성, 현실적 타당성, 실질적 활용 가능성
- 맞춤형 금융상품/서비스 제안(20점)
 - : 아이디어 창의성, 구체성 및 실현 가능성, 시장 경쟁력
- 완성도(10점)
 - : 논리적 서술, 시각적 퀄리티, 보고서 구조, 프레젠테이션 효과(10점)

[2차 발표 평가 기준]

- 발표 내용 및 구성(40점)
 - : 핵심 메시지 전달력, 논리적 스토리텔링
- 분석 결과 설명력(25점)
 - : 데이터 해석 능력, 인사이트 근거 제시
- 제안의 설득력 (25점)
 - : 비즈니스 임팩트, 실현가능성, 차별화전략
- 발표 스킬 및 소통(10점)
 - : 시각 자료 활용, 질의응답 대응

데이터 활용 규칙

- 제공한 데이터 중 최소 1개 데이터셋은 필수로 사용
- 외부 데이터 사용 가능 (무료 공개 데이터에 한함). 출처·수집방법 명시 및 원본파일 제출 필수.
유료 데이터 사용 시 수상 취소
- LLM 사용 여부는 가점 또는 감점 요인으로 작용하지 않음.
단, LLM 사용시 구체적인 활용 방식과 분석 과정에 대한 설명자료를 제출해야 함.

결과물 제출

- 핵심 인사이트 및 비즈니스 제안(구제책 제안, 실행 전략 및 기대효과 분석)
- 데이터 분석 근거 자료(탐색적 데이터 분석 및 세대별 비교분석 결과, 분석 방법론 및 과정 상세 기술)
- 예측/분류 모델링 결과(모델 개발 과정 및 성능 평가, 예측 결과의 비즈니스 해석, 재현 가능한 코드)