

부산 해운대구 한식당 리뷰 분석과 다국어 AI 추천 시스템

카카오맵 리뷰 크롤링부터 EDA, 감성분류 머신러닝, RAG 기반 다국어 LLM 추천 챗봇까지 — 실제 데이터로 완성한 엔드투엔드 프로젝트.

분석 대상 — 해운대구 · 한식업종 · **영업연수 10년 이상** 매장 (전체 1,518개 중 791개, 52.1%)

01 분석 범위와 데이터 파이프라인

해운대구 한식 매장 1,518개 가운데 영업연수 10년 이상 매장만 골라 카카오맵 리뷰를 크롤링했다. 매장 실패와 리뷰 0건 매장을 제외하고 579개 매장·15,309건을 모았고, 중복 제거 등 전처리를 거쳐 최종적으로 14,878건을 분석에 썼다.

- 수집 — 카카오 로컬 검색 API와 카카오맵 리뷰 API를 requests로 직접 호출했다. 네이버지도는 Selenium 크롤링 중 캡차에 막혀 중단했다.
- 전처리 — 중복 431건 제거, 결측치는 삭제 대신 플래그로 구분, review_length-rating_group 등 파생변수 생성.

분석 범위의 한계 — 영업연수 10년 미만 727개(47.9%) 매장은 대상에서 빠졌고, 매장 실패율도 25.2%에 달한다. 이 결과는 "**해운대구 한식당 전체**"가 아니라 **장기 영업 매장 중심의 결과**로 읽어야 한다.

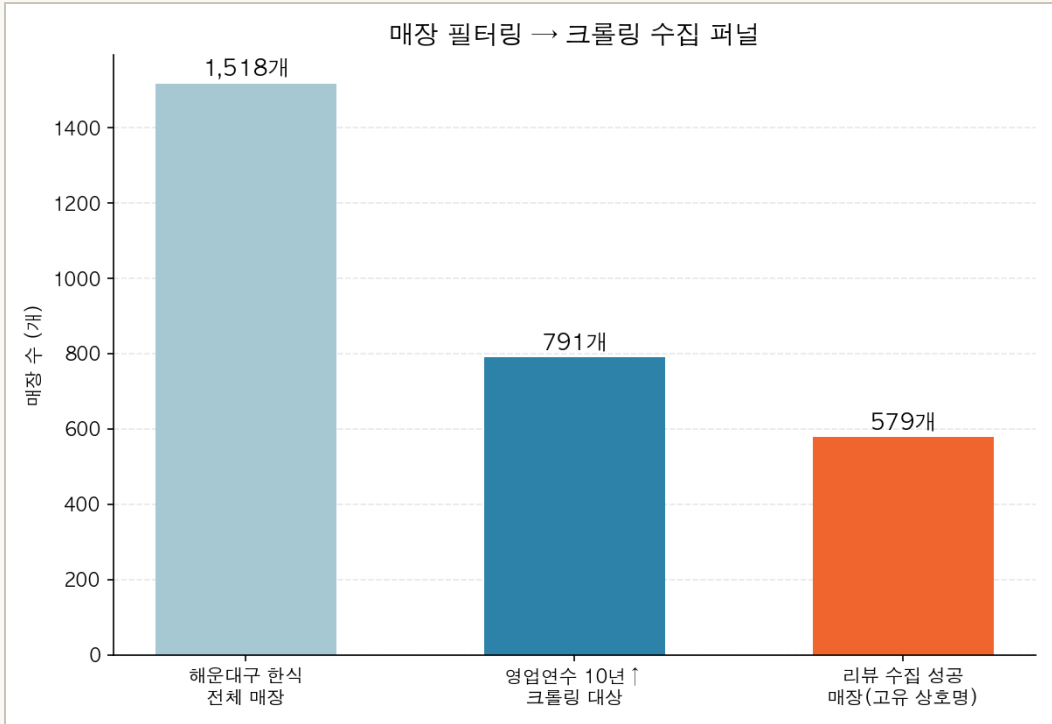


그림 1. 1,518개 전체 매장 → 영업연수 10년+ 필터 791개(52.1%) → 매장실패 199개 제외 → 최종 579개 매장·14,878건.

- 579개 매장 · **14,878건** 정제 리뷰 — 원본 15,309건에서 중복 431건 제거
- 매장 실패 199개 — 크롤링 대상 791개 중 25.2%가 상호명 표기 차이·폐업 등으로 카카오맵 매장 실패

02 감성 분류 머신러닝 모델

평점을 가리고 리뷰 본문만으로 긍정·중립·부정을 맞추는 모델이다. 12,216건(80/20 분할)으로 베이שראל인 → Naive Bayes → 로지스틱회귀 → 로지스틱회귀(balanced) 네 가지를 비교했다.

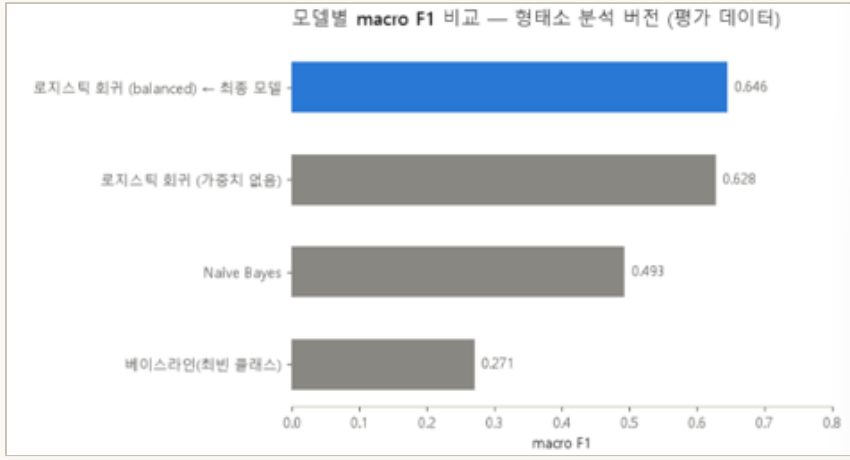


그림 2. 모델별 macro F1 — 베이שראל인 0.271 → Naive Bayes 0.493 → 로지스틱회귀 0.628 → balanced 0.646.

최종 모델(로지스틱회귀 balanced) — 정확도 80.3%, macro F1 0.646. character n-gram(0.670)이 형태소 분석(0.646)보다 근소하게 우세하다는 결과도 그대로 남겼다 — "형태소 분석이 항상 정답은 아니다."

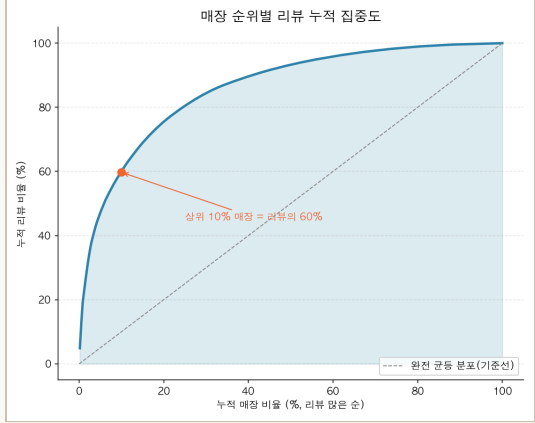
	precision	recall	f1-score	support
긍정	0.903	0.880	0.891	1669
부정	0.741	0.779	0.760	548
중립	0.277	0.295	0.286	227
accuracy			0.803	2444
macro avg	0.640	0.651	0.646	2444
weighted avg	0.808	0.803	0.805	2444

그림 3. 최종 모델 classification report (평가 데이터 2,444건).

- 긍정 — 표본 1,669건(68.3%), f1 0.891 — 데이터가 많아 가장 안정적으로 잡아낸다.
- 부정 — 표본 548건(22.4%), f1 0.760 — 양호한 수준.
- 중립 — 표본 227건(9.3%)뿐이라 f1 0.286 — "애매함"은 클래스 자체의 성격이자 데이터 부족의 결과이기도 하다.

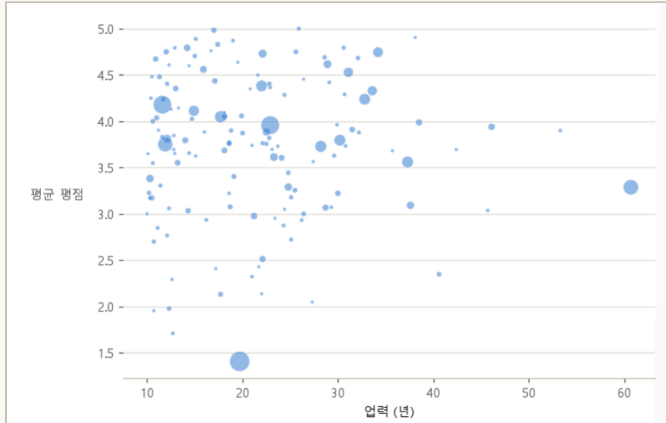
03 핵심 인사이트 — 탐색적 데이터 분석

① 리뷰는 소수 매장에 집중된다



매장당 평균 26.4건(중앙값 8건) — 상위 10%가 리뷰의 60%를 차지하는 롱테일 구조.

③ 업력과 평점은 사실상 무관하다



매장 단위 $r \approx 0.02$, 리뷰 단위 $r = 0.088$ — 분석 단위를 바꿔도 **같은 결론**이다.

04 다국어 LLM 추천 챗봇 (RAG)

- 매장 프로필 생성 — 평점·업력·형태소 기반 특징 태그·대표 리뷰를 텍스트로 요약한다.
- 임베딩 — OpenAI text-embedding-3-small로 매장 프로필을 벡터화한다.
- 검색 — 질문을 같은 모델로 임베딩해 코사인 유사도 상위 5개 후보를 찾는다.
- 생성 — GPT-4o-mini가 후보 정보만 근거로 추천을 쓴다. 질문 언어와 100% 동일한 언어로만 답하도록 시스템 프롬프트 최우선 규칙으로 못박았다.

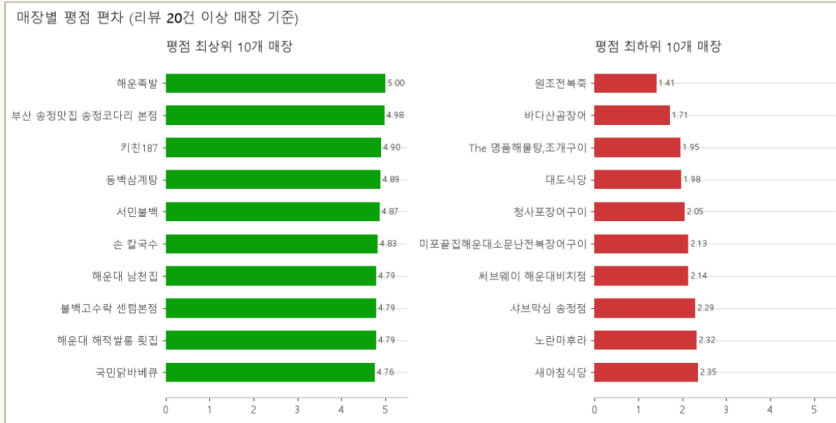
"불친절하지 않고 친절한 식당 추천해줘"

해운대구에서 친절한 식당을 찾으신다면 "술내음한정식"을 추천드립니다. 평균 평점 3.44, 친절 관련 긍정 리뷰 다수...

"Recommend a friendly, good-value pork soup restaurant with no long wait"

I recommend "오복식당" — average rating 3.88 (16 reviews). Friendly staff, tender fresh pork, unlimited rice/noodles...

④ 인기와 평점은 다르다



리뷰 수 1위 "원조전복죽"(747건)이 평점은 **최하위(1.41점)** — 인기는 품질의 증가가 아니다.

05 데이터·모델 한계 및 고찰

- 롱테일 편중 — 상위 10% 매장이 리뷰 60%를 차지해, LLM 추천도 그쪽으로 쏠릴 수 있다.
- 범위 편중 — 10년 이상 매장만 대상으로 삼아 신생 매장 727개(47.9%)를 배제했다.
- 생존편향 — 인허가일자 기준 데이터라 이미 폐업한 매장은 반영되지 않는다.

- 중립 클래스 저조 — 데이터가 9.3%로 적어 f1이 0.286에 머문다.

- 형태소 분석의 한계 — character n-gram이 근소하게 더 나아 트레이드오프가 남는다.

- LLM 근거 제한 — 리뷰 텍스트에만 근거해 답하고, 실시간 정보는 반영하지 못한다.

579개	14,878건	80.3%	RAG · 다국어
분석 매장	정제 리뷰	ML 정확도	LLM 추천 챗봇

크롤링부터 EDA, 감성분류 ML, 다국어 LLM 추천 챗봇까지 **엔드투엔드**로 구현했다. 다음 단계는 중립 클래스 성능 보완, 다른 구·신생 매장으로의 확장, 실사용자 피드백을 통한 개선이다.