# 자연어처리와 대형언어모델을 활용한 소비자 리뷰 기반

# 착한가격 업소

## 1. 분석배경 및 인터뷰



#### 착한가격 업소의 중묘성과 한계

외식 물가 상승으로 가성비 높은 업소에 대한 **수요가 증가**하는 가운데, 정부는 '**착한가격업소**' 제도로 이를 장려하고 있으나 **여전히 현실적인 한계가 존재**함

착한가격업소 신규 발굴, 선정 이후 관리 어려움



### 발굴과 관리가

방문객의 실제 리뷰, 키워드, 가격 정보를 기반으로 한 데이터 자동화 시스템을 통해 착한가격업소의 효율적인 관리와 신규 업소의 체계적 발굴을 지향

#### 현재 기준표에 의한 평과과정 중 느끼는 어려움은 무엇인가요? 위생 평가는 비정량적인 부분이 많아 일관성 등 많은 어려움이 있습니다

#### 실사를 통한 평가자의 정성평가

(비정량적 묘소와 주관성으로 인한 일관성 부족 문제)



가는지 잘 모르겠어요 ↑HITH AMI 점해진 후보만을 대상으로 현장 조사

(전수 조사 불가능)

착한가격 업소 선점 기준이

자주 가는 가격도 저렴하고 서비스도 좋은

가게가 있는데 왜 착한가격 업소에 만들어

소비자 입장에서 잘 체감 되시나요?

크롤링 및 자동 분석 시스템 활용 (전수 조사를 통한 후보군 선점 가능)

# 2. 착한가격점수 도출

**착한가격점수**는 착한가격업소를 신규 발굴하기 위해 **정량적으로 평가**할 수 있도록아래와 같은 공식으로 산출하였다

기존 평가 방식 🖒 착한가격 점수 = ( 0.5 x 가격 + 0.5 x 위생 ) x 100

#### 소비자 체감 착한가격 점수 = (0.358 x 가격 + 0.289 x 위생 + 0.353 x 서비스) x 100

- 기존의 평가 항목을 보다 실제 이용자 경험을 반영하는 방향으로 재해석하고자 하였으며 기존 의 형식적인 기준을 넘어 **현실적이고 체감 가능한 평가 지표**를 구축하였다
- 실제 이용자 네이버 키워드 데이터를 이를 **가격, 위생, <u>서비스</u> 세** 가지 카테고리로 분류 후 각 - 키워드 빈도수를 기반으로 **가중치(W)**를 설정하였다

저렴하게 안심하고 이용할 수 있는 착한가격업소 효율적인 경영을 통하여 소비지에게 저렴한 가격으로 양질의 서비스를 제공하는 정부와 지방자치단체가 지정, 관리하는 입소를 말합니다.				
₩ <u>₩</u> 저렴한 가격 안전한 재료	™			
	<b>○</b> 행정안전부 <mark>◆한가격업</mark> ▲			

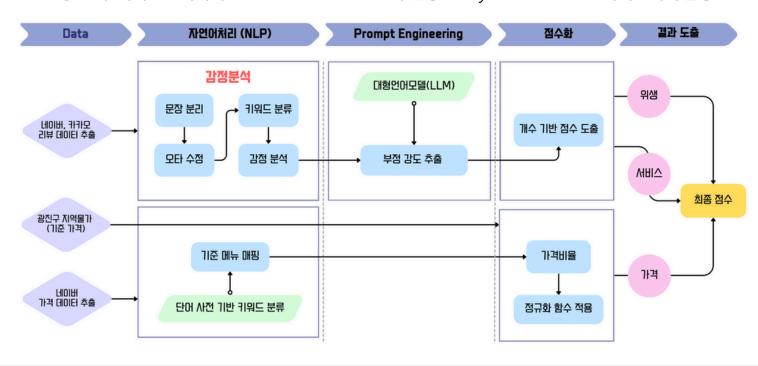
:	카테고리	키워드 목록	비율 (%)
	가격	가격이 합리적이에요, 가성비가 좋아요	35.8%
	위생	재료가 신선해요, 위생적이에요	28.9%
	서비스	친절해요, 직접 잘 구워줘요, 맞춤 지도를 잘해줘요	35.3%

## 3. 데이터 및 분석과정

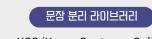
• 사용 데이터 출처: 네이버, 카카오 리뷰 데이터

광진구 관계자

- 분석 대상 지역: 광진구
- 업종 범위: 외식업 + 비외식업
- 선정 기준: 착한가격 선정 기준 메뉴 포함 업소 전체
- 기간: 1년 이내 리뷰 (24년~25년)
- 수집방법 : Python Selenium 라이브러리 활용



## 4. 리뷰 데이터 정제



KSS (Korea Sentence Splitter) 문장 분리 방법

#### 1.Kss 라이브러리 사용 문장 단위 분리 2. 5자 미만 <u>짧은 문장</u>은 다음 문장과 <u>병합</u>

3. 불필요한 공백 제거 의미 없는 문장 필터링

1.**한글 자모/영어 단일** 문장 제거 (예시: "ㄱ", "hi") 2. 특수기호/숫자만 존재하는 문장 제거 (예시: "!!", "12345")

## 모타 교정 모델 T5 기반 한국어 오타 교정 특화 모델

모타 교정 방법 1.오타만 고쳐줘 → 프롬프트 작성 2. 문맥을 고려해 오타 수정

#### 모타 교정 파이프라인

1.**반복 문자** 제거 및 문장 정제 2.**오타 교정** 프롬포트 생성 3.유사도 검증 후 <u>원본과 70% 유사 시</u> 원문 유지

#### 문장 분류 시스템 설계

- 착한가격 업소 지정 평가표 기반
- 리뷰 내 위생 및 서비스 요소 자동 분류 • 항목별 키워드 사전 + 공간 기반 분류 로직

#### 사전 구축 기준

기준표 항목 단어 반영 주방 위생, 매장 청결, 화장실 상태

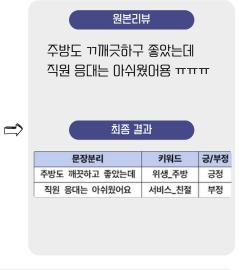
추가항목 설정 :

위생 기타 (기준표에 없는 위생표현) 서비스 관련 항목 (친절도, 정보제공, 편의성)

# 감정 분석 모델 한국어 감정 분석 모델 감정분석 방법

1. 텍스트 입력 2. 각 감정에 대한 확률 출력 3. 가장 높은 확률 가진 감정 기준 분류 4. "0" = Negative / "1" = Positive

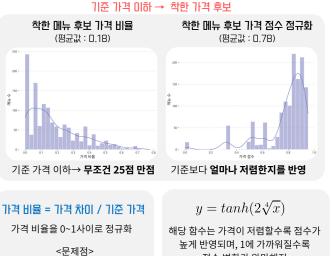
#### 카테고리 리뷰개수 서비스\_친절도 12,275 위생\_매장 9,349 위생\_공통 4,061 서비스 정보제공 3,413 서비스\_편의성 2,988 위생\_주방 2,571 위생\_기타 701 257 위생\_화장실



# 5. 서비스/위생 점수화



# 6. 가격 점수화



1) 1점 만점 체계에 직접 대응할 경우, 대부분의 점수가 낮게 몰리며 분포가 왜곡되는 경향

2) 위생 및 서비스 점수 분포와

일관되지 않는다는 문제

점수 변화가 완만해져

위생, 서비스 항목과 동일한 구조이면서 가격 차이가 작더라도 업소 간 점수 차이를 정밀하게 구분 가능

# 7. 최종 착한 가격 점수



착한가격 점수 & 소비자 체감 착한가격 점수



현실적이고 변별력있는 후보군 선정 가능

# 8. 프로토타입



# 9. 활용방안 및 한계점

#### □1 비신청 업소 자동 후보화

1,536개의 잠재적 착한가격 업소 발굴 신청하지 않아도, 소비자 체감과 리뷰 데이터를 기반으로 잠재적 착한가격 업소를 행정이 먼저 발굴하는 시스템 디지털 취약계층의 업소가 신청하기 힘든 부분을 해결 ➡️ 행정 사각지대 해소, 정책 포용성 강화,능동형 행정

#### 02소비자 체감형 점수를 통한 소통체계 활성화 소비자의 의견을 적극 반영한 행정 판단을 통해

착한가격업소 선정의 신뢰도를 높이고, 소비자와 업소 간 **공감대를** 형성함 (업소 - 행정 - 소비자) 간 신뢰 형성

#### 03 인증업소 효율적인 모니터링

인증 이후 업소에 대해 정기적인 사후 모니터림 체계를 구축함 분기별 리뷰 기반 모니터링 보고서를 자동 생성 소비자 체감 데이터를 바탕으로 업소의 개선점 객관적 파악

행정체계 📥 단순화, 비용절감

업소 ፫ 지속적인 문명 개선

#### 01 리뷰 데이터 부족 업소의 확보 방안

일부 업소는 온라인 리뷰 데이터가 부족해 분석 및 체감지표 반영에 한계 특히 기존 광진구는 기존 인증 업소 중심 리뷰 수집 방식이므로 사전 발굴이 어려운 구조임

데이터 불균형 해소 + 주민 자발적 리뷰 참여 유도 필요

"소비자 참여형 리뷰 인센티브 제도"

홈페이지를 통해 방문 업소 리뷰 작성 시 소량의 지역화폐 마일리지 제공

<sup>\*</sup>데이터 기반 행정 신뢰성 확보 \*공감형 정책 설계 가능

#### □ 작한 가격 메뉴 기준의 구조적 한계

"더 넓은 메뉴 분류 가능

광진구 착한가격 기준 메뉴가 수량 제한적이고 세분화 부족으로 인해 신규 업소나 다양한 음식군 반염 어려움

★ 저렴하고 만족도 높은 업소가 후보에서 제외 경우 발생

"LLM 기반 메뉴 분류 시스템' 메뉴 데이터를 분석해 기준표에 없는 메뉴는 분류 후 신규제안 항목으로 추천

현실 적합성 향상