

## 리뷰 데이터 기반 개인 맞춤형 음식점 추천 시스템

정승혜, 임예빈, 최가연, 장혜원, 김현희

동덕여자대학교 정보통계학과

[20191071@dongduk.ac.kr](mailto:20191071@dongduk.ac.kr), [20191068@dongduk.ac.kr](mailto:20191068@dongduk.ac.kr), [20201082@dongduk.ac.kr](mailto:20201082@dongduk.ac.kr),[20201071@dongduk.ac.kr](mailto:20201071@dongduk.ac.kr), [heekim@dongduk.ac.kr](mailto:heekim@dongduk.ac.kr)Personalized restaurant recommendation system  
based on customer's review data

Jeong Seung Hye, Lim Yea Bin, Choi Ga Yeon, Chang Hye Won,

Kim Hyon Hee

Dept. of Information Statistics, Dongduk Women's University

## 요 약

사람들은 각자 원하는 조건에 부합한 식당과 카페를 찾곤 한다. 그러나 개인별로 원하는 조건들이 다양하고 그 조건들이 모두 부합하는 음식점을 찾기에는 적지 않은 시간과 노력이 필요한 일이다. 이 불편함을 해소하고자, 사용자가 원하는 조건을 입력하면 그 조건에 부합하는 몇 개의 음식점들을 추천해 주고, 지도상으로 위치를 표시해 주는 개인 맞춤형 음식점 추천 시스템을 개발하였다. 본 연구에서 제안하는 추천 시스템은 사용자가 입력한 우선순위에 따라 차별화된 음식점 추천을 받을 수 있으므로, 시간과 노력을 투자하지 않고도 자신이 원하는 음식점을 쉽게 찾을 수 있을 것으로 예상된다.

## 1. 서론

고객들은 식당과 카페를 방문하기 전, 포털 사이트의 검색을 통해 음식점의 메뉴, 가격 리뷰를 보고 방문 여부를 결정한다. 그러나 포털 사이트에 검색하면 광고 매장들이 최상단에 위치할 뿐만 아니라, 상단에 떠 있는 리뷰 몇 개만 읽고 방문한 경우 괴리가 있어 실망하는 경우가 대다수이다.

이러한 문제점을 해결하기 위해 많은 리뷰를 분석하여 키워드를 추출한 뒤 그 음식점의 맛·분위기·서비스를 좀 더 객관화하여 알려주고 사용자가 원하는 조건과 부합하는 음식점을 추천해 주는 추천 시스템 [1]을 개발하였다. 먼저 교통카드 사용량 데이터에서 서울시 2호선 중 유동 인구가 많은 20개 역을 선정하였다. 매장 후기를 최대 600개까지 추출하고, 음식점들의 종류를 총 9가지로 분류하였다. 리뷰의 유사도를 TF-IDF 가중치의 합으로 산출한 뒤 매장별 긍·부정 분석을 실시하였다. 맛·분위기·서비스와 매장별 유사도를 도출하여 사용자에게 적합한 음식점 3개를 선정하여 추천하고 지도상으로 위치를 표시해 준다.

이 추천 시스템을 통해 사람들은 시간과 노력을

들이지 않아도 개인별 맞춤 음식점을 편리하게 알 수 있을 것으로 보인다.

본 논문은 다음과 같이 구성된다. 제2장에서는 데이터 수집 및 전처리를 나타내고, 제3장에서는 추천 모델에 대해 구체적으로 서술한다. 마지막으로 제4장에서 결론을 맺는다.

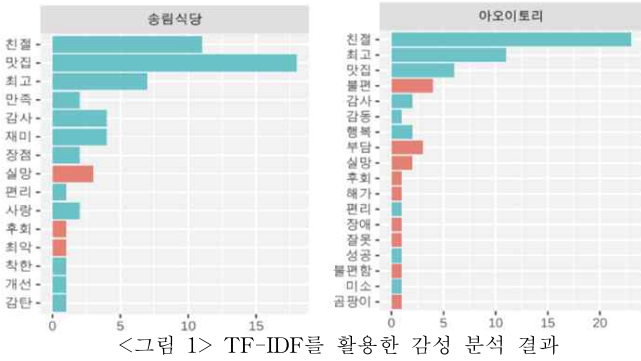
## 2. 데이터 수집 및 전처리

본 연구에서는 티머니에서 제공한 교통카드 사용량 데이터 중 접근성이 좋은 서울시 2호선에서 출·퇴근 시간과 새벽 시간을 제외한 시간에 유동 인구가 많은 20개 역을 선정하였다. 역별 인기순 상위 40개 매장을 추출하였고, 총 716개의 음식점들을 대상으로 리뷰 182,790개, 메뉴 4,433개, 거리를 수집했다. 거리는 지도에 나와 있는 지하철역을 기준으로 걸리는 도보 시간으로 지정했다. 매장당 리뷰는 최소 50개부터 최대 600개까지 추출하고, 음식점별 메뉴 카테고리는 9가지로 분류했다.

추천 시스템에 사용자 리뷰를 적용하기 위하여 Konlpy[2]를 이용하여 TF-IDF 기반의 감성분석[3]을 실시하였다. <그림 1>에서 보는 바와 같이 매장별로 긍·부정 단어 20개씩 추출한 뒤 유사도 분석을 진행한 후, 맛·분위기·서비스별로 3가지에 대한 연관

\* 본 연구는 2023학년도 중소기업벤처부의 SW·콘텐츠 인재 양성사업(벤처스타트업아카데미) 지원에 의한 연구임

성 점수를 계산하여 상위 20개 단어를 추출했다.



<그림 1> TF-IDF를 활용한 감성 분석 결과

### 3. 개인 맞춤형 추천 모델

사용자가 원하는 역, 역으로부터의 도보 시간, 음식 종류, 메뉴당 최소·최대 금액, 맛·분위기·서비스의 우선순위를 입력받는다. 이 정보를 바탕으로 전체 데이터에서 조건에 부합하는 데이터들을 필터링한다. 역, 음식 카테고리에 일치하는 데이터를 뽑고, 거리는 입력받은 값보다 낮은 수치에 해당하는 데이터로 필터링한다. 그 후, 사용자로부터 입력받은 맛·분위기·서비스의 우선순위를 고려하여 각 음식점들의 점수를 계산한다. 음식점 리뷰와 맛·분위기·서비스의 유사도 분석에서 뽑은 20개 단어가 각 기준의 단어에 포함되면 1순위로 지정한 것은 0.5, 2순위는 0.3, 3순위는 0.1의 가중치를 부여하여 계산했다. 이때, 유사도 분석으로 추출한 단어들의 순위에서도 높은 순위의 TF-IDF 키워드에는 더 큰 가중치를 주어 점수를 합산했다.

원하는 역 이름을 입력해주세요 :  
강남역

- 1 : 도보 5분 이하
- 2 : 도보 10분 이하
- 3 : 도보 20분 이하
- 4 : 상관없음
- 2

원하는 카테고리를 입력해주세요 :

- 1 : 한식
- 2 : 아시안
- 3 : 육류
- 4 : 일식
- 5 : 중식
- 6 : 분식
- 7 : 양식
- 8 : 해산물
- 9 : 카페
- 0 : 모두포함
- 9

원하는 메뉴의 가격대를 입력해주세요

- 최소 : 1000
- 최대 : 9500

맛, 분위기, 서비스의 우선순위를 입력해주세요!

- ex) 1 3 2
- 2 1 3

<그림 2> 사용자 입력 화면

<그림 2>는 사용자 입력을 나타내는 화면이다.

먼저 원하는 역이름으로 강남역을 입력하고, 거리는 도보 10분 이하, 원하는 카테고리는 카페를 선택하였으며, 원하는 메뉴의 가격대도 입력하였다. 또한 분위기 - 맛 - 서비스로 우선순위를 정하였다. 해당 입력에 따라 추천한 음식점을 지도 시각화 한 결과는 <그림 3>에서 볼 수 있다.



<그림 3> 추천된 음식점 지도 시각화 결과

### 4. 결론

본 연구에서는 음식점 방문자들이 남긴 리뷰 데이터를 이용하여 사용자의 원하는 장소, 가격, 음식 카테고리에 맞는 식당, 카페를 추천해 준 후 지도에 위치를 표시해 주는 추천 시스템을 제시했다. 제안한 모델은 소비자들이 시간과 노력을 들이지 않고 현명한 선택을 할 수 있도록 도와준다. 더 나아가 기업에서는 소비자의 니즈를 파악하여 마케팅 전략을 수립할 수 있고, 자영업자들에게는 제품 개발 전략을 수립할 수 있을 것으로도 기대하는 바이다.

추후, 회원가입을 통해 해당 서비스를 이용하고 서비스 결과 통계를 제공한다면, 해당 결과를 통해 소비자별 데이터 분석이 가능하고 제품 개발 분석도 가능할 것이라고 예상된다. 또한, 서울시 전체로 범위로 확장하여 추천 시스템을 구현하고 포털 사이트에 등록되지 않은 음식점의 모든 메뉴까지 추출하는 방안을 모색하고 있다.

### 참고문헌

[1] 장예화 외 (2021). 리뷰 데이터 마이닝을 이용한 하이브리드 추천시스템 개발: Amazon Kindle Store 데이터 분석사례. 경영정보학연구, 23(1), 155-172.  
 [2] 박은정, 조성준, "KoNLPy: Korean natural language processing in Python", 제 26회 한글 및 한국어 정보처리 학술대회 논문집, 2014년, pp.133 - 136  
 [3] 감성사전 : KNU 한국어 감성사전