

치킨 리뷰의 이면: 텍스트 마이닝을 통한 리뷰의 탐색적 분석을 중심으로

BEHIND CHICKEN RATINGS: An Exploratory Analysis of Yogiyo Reviews Through Text Mining

김준겸*, 최은솔*, 윤수현*, 이유빈*, 김동환**

연세대학교 언더우드국제대학*, 연세대학교 커뮤니케이션대학원**

Jungyeom Kim(jghg2724@yonsei.ac.kr)*, Eunsol Choi(jupiteraca@yonsei.ac.kr)*,
Soohyun Yoon(syooun071@yonsei.ac.kr)*, Youbeen Lee(youbeenlee@yonsei.ac.kr)*,
Dongwhan Kim(dongwhan@yonsei.ac.kr)**

요약

코로나 19의 영향으로 배달앱 시장이 빠르게 성장하며, 리뷰와 평점이 더욱 중요해지고 있다. 그러나 급격하게 늘어난 리뷰와 평점의 신뢰도에 의문이 제기되고 있다. 따라서 본 연구는 리뷰와 평점을 탐색적으로 분석하여 배달앱 리뷰의 신뢰도와 유용성을 파악하고, 이를 높일 방법을 탐구하였다. 텍스트 마이닝 기법을 사용하여 도출한 결과는 다음과 같다. 첫째, 요기요와 네이버 지도, 구글 지도의 음식점 평점을 분석한 결과, 요기요는 가장 우편향된 평점 분포를 보여주었다. 둘째, 요기요의 세부 평점 요인(맛, 양, 배달)간에는 모두 강한 양의 상관관계가 있었고, 이는 부정 리뷰의 단어 분석에서도 드러났다. 셋째, 리뷰의 극성에 따라 사용되는 품사와 형태소의 비율이 달랐다. 넷째, 전체 리뷰 데이터에서 367개의 부정어를 선별한 후, 이를 분류하여 치킨 전용 부정어 사전을 제작하였다. 본 연구는 치킨 리뷰의 탐색적 분석을 통해 앞으로 배달앱 리뷰에 대한 연구가 나아가야 할 방향을 제시하였다.

■ 중심어 : | 배달앱 | 텍스트 마이닝 | 온라인 리뷰 | 리뷰 분석 | 감성 사전 |

Abstract

Ratings and reviews, despite their growing influence on restaurants' sales and reputation, entail a few limitations due to the burgeoning of reviews and inaccuracies in rating systems. This study explores the texts in reviews and ratings of a delivery application and discovers ways to elevate review credibility and usefulness. Through a text mining method, we concluded that the delivery application 'Yogiyo' has (1) a five-star oriented rating dispersion, (2) a strong positive correlation between rating factors (taste, quantity, and delivery) and (3) distinct part of speech and morpheme proportions depending on review polarity. We created a chicken-specialized negative word dictionary under four main topics and 20 sub-topic classifications after extracting a total of 367 negative words. We provide insights on how the research on delivery app reviews should progress, centered on fried chicken reviews.

■ keyword : | Delivery Application | Text Mining | Online Review | Review Analysis | Sentiment Dictionary |

* 본 연구는 2021년 산업자원통상부와 한국디자인진흥원의 디자인 주도 다학제 교육과정 지원 사업의 지원과 2021년도 연세대학교 연구비의 지원을 받아 수행된 것임(2021-22-0320)."

접수일자 : 2021년 07월 09일

수정일자 : 2021년 07월 29일

심사완료일 : 2021년 07월 29일

교신저자 : 김동환, e-mail : dongwhan@yonsei.ac.kr

태소, 숫자의 사용 빈도 등이 부정적인 리뷰를 나타내는 지표가 될 수 있음을 확인하였다. 그러나 이러한 문법적인 특징 외에도, 부정적인 리뷰에서 특정 단어들 자주 등장하는 것을 발견하였고, 이 단어를 바탕으로 부정어 사전을 만들어 부정 리뷰를 판별하는 데 사용할 수 있으리라 판단하였다. 단순히 부정적인 의미를 띤 단어만을 부정어로 분류하기보다, 부정적인 리뷰에 반복적으로 등장하는 단어를 부정어 사전에 포함했고, 이 단어를 기준에 따라 분류하여 부정어 사전을 제작하였다.

부정어 사전을 제작하기 위해 프랜차이즈 음식점의 리뷰 5,000개에서 부정 리뷰에 자주 등장하는 부정어 키워드 440개를 수동으로 추출한 후, NLP를 사용하여 정규화하였다. 이후 해당 키워드가 포함된 리뷰의 평가 항목별 평균 평점을 계산하여, 맛, 양, 배달 점수가 모두 4점 이상이면 해당 키워드를 부정어 목록에서 제외했다. 그 결과 총 367개의 키워드가 남았고, 이들을 평가 항목별 평점(맛, 양, 배달) 분포에 따라 1차 분류한 후, 수동으로 [Table 6]의 분류 기준에 따라 재배치하였다.

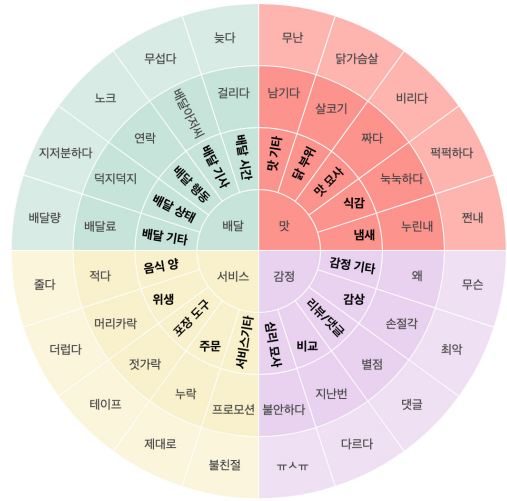


Figure 3. Example of Negative Words

묘사를 정리하였다. 닭 부위는 맛에 관한 묘사 중 '살코기', '닭가슴살', 등을 정리하였다.

배달의 경우 '배달 시간', '배달 기사', '배달 행동', '배달 상태', '배달 기타'로 나누었다. 배달 시간에는 '걸리다', '늦다' 등의 키워드가 포함되었고, 배달 기사의 경우 '배달아저씨', '무섭다' 등 배달 기사를 부정적으로 평가하는 단어들로 구성하였다. 배달 행동에는 '연락', '노크' 등 배달 기사의 행동 중 부정적인 것들이 포함되었다. 배달 상태는 배달된 음식물의 상태를 묘사하는 '떡지덕지', '지저분하다' 등으로 구성되었다.

서비스의 경우 '양', '위생', '포장 도구', '주문', '서비스 기타' 항목으로 나누었다. 양 카테고리에는 양에 관련된 부정어 키워드인 '적다', '줄다' 등이 포함되었다. 위생의 경우 '머리카락', '더럽다' 등 매장과 음식의 위생 상태에 대한 부정어 키워드들로 구성되었으며, '젓가락', '테이프' 등 포장 시 매장에서 함께 챙겨주어야 하는 도구들이 누락되어 부정적으로 사용된 단어들은 포장 도구 항목에 포함되었다. 특히나 서비스 경험에 큰 영향을 미치는 '주문' 소분류의 경우 '누락', '환불' 등의 키워드로 구성되었다.

Table 6. Categories of Negative Words

대분류	맛	배달	서비스	감정·심리
소분류	냄새	배달 시간	양	심리 묘사
	식감 및 상태	배달 기사	위생	비교
	맛 묘사	배달 행동	포장 도구	리뷰/댓글
	닭 부위	배달 상태	주문	감상
	맛 기타	배달 기타	서비스 기타	감정 기타

전체 부정 키워드를 크게 <맛>, <배달>, <서비스>, <감정 심리>로 분류하였다. 4개의 대분류는 각각 기타를 포함하여 5개의 소분류를 포함하고 있으며, 기타의 경우 기존의 소분류에는 해당하지 않지만 명백하게 해당 카테고리에 속하는 키워드들을 정리했다. 그 예시는 [Figure 3]과 같다.

먼저 맛의 경우 세부 항목을 '냄새', '식감 및 상태', '맛 묘사', '닭 부위', '맛 기타'로 나누었다. 냄새의 경우 '누린내', '쩨내' 등 치킨의 냄새를 묘사하는 부정어가 포함되어 있다. 식감 및 상태의 경우 '눅눅하다', '마르다' 등의 키워드를 포함하여 분류하였으며, 맛 묘사의 경우 '짜다', '비리다' 등 미각으로 느껴지는 맛 자체에 대한