

2025 경영경제대학 학술제 참가보고서

| | | | | | |
|----------------------------------|--|----------|---------------|---------------|---------------------|
| 제 목 | 리뷰의 신뢰도 : 플랫폼별 리뷰 특성에 관한 연구 (팀명 : 리뷰이벤트) | | | | |
| 신청팀 인적사항 | 성명 | 학번 | 생년월일 | 연락처 | E-mail |
| | 정형준 | 20212132 | 2002. 09. 01. | 010-8338-8077 | cloud0901@naver.com |
| | 이상민 | 20212344 | 2001. 04. 11. | 010-6422-1060 | lsm041111@naver.com |
| 선정주제 | 플랫폼별(네이버지도, 카카오맵) 리뷰의 특성에 관한 연구 | | | | |
| 주제 선정 이유 (사회적 현상과 관련지어 서술) | <p>온라인 리뷰가 현대 소비자들에게 중요한 의사결정요소가 되면서, 점점 다양한 플랫폼에서 리뷰를 제공하고 있다. 상품 리뷰에서 일상생활 속에 방문하는 다양한 장소(음식점, 카페, 헤어샵, 관광지 등)에 대한 리뷰까지 리뷰의 범위가 점점 확대되고 있다.</p> <p>하지만 동일한 업체라도, 각 플랫폼별 다양한 요인(마케팅 리뷰, 플랫폼별 이용자의 특성 등)에 따라 별점이 차이나 나는 것처럼 보이는 경우가 있다. 이런 상황에서 리뷰의 신뢰도에 대해 의심을 갖게 되었고, 리뷰의 신뢰도에 관한 연구를 진행하기로 하였다.</p> | | | | |
| 주제 분석 및 연구 | <p>I. 서론</p> <p>최근 온라인 쇼핑이 확산되면서 리뷰의 기능은 점점 확대되고 있다. 물건을 직접 보지 못하게 된 사람들은 다른 소비자들의 리뷰를 통해 의사결정을 하게 됐고, 기업들은 소비자들을 유혹하기 위해 리뷰를 이용한 마케팅을 펼치고 있다. 이에 따라 실제 소비자들의 ‘실제 리뷰’가 아닌 마케팅을 위한 ‘허위 리뷰’가 만연하게 되었고 소비자들은 점점 객관적인 판단에 어려움을 겪게 되고 있다.</p> <p>이런 문제를 해결하기 위해 플랫폼들은 자체 리뷰 필터링 시스템을 개발하여 적용하고 있다. 포털 서비스 “네이버”는 자체 플레이스 서비스인 “네이버 지도”에 있는 리뷰에 대하여 “네이버 리뷰 클렌징 시스템”을 적용하여 필터링하고 있다. 또한 배달음식 플랫폼인 “배달의 민족”에서도 리뷰에 대한 자체 필터링 시스템을 적용하여 허위 리뷰를 단속하고 있다.</p> <p>앞서 언급한 업체들의 리뷰 필터링 시스템은 실제 소비자 영수증을 통해 실제 방문 여부와 위치 정보 등을 통해 허위 리뷰를 판별하는 방식이다. 하지만 실제 리뷰임에도 허위 리뷰로 판별되어 필터링 되거나 허위 리뷰임에도 실제 리뷰로 판별되는 경우가 빈번히 발생한다. 특히 업체에서 진행하는 “리뷰 이벤트”의 경우에는 실제로 방문한 사람이 영수증 인증까지 했기 때문에 필터링 시스템이 판별하는 데에는 어려움이 있다.</p> <p>또한, 사용자, 리뷰 방식, 노출 알고리즘 등에 따라 같은 업체라도 별점과 평가가 다른 경우가 많고, 이는 사용자들에게 더 많은 혼란을 안겨줄 수 있다.</p> <p>그래서 리뷰의 신뢰도에 대해 의심을 갖게 되었고, 리뷰의 신뢰도를 판단할 수 있는 새로운 기준에 대한 연구를 진행하게 되었다. 이 연구는 네이버와 카카오 플랫폼에서 동일한 업체의 별점과 리뷰를 비교분석 함으로써 플랫폼별 리뷰와 별점의 특성을 파악하고자 한다.</p> <p>II. 본론</p> <p>1. 데이터 수집</p> <p>플랫폼별(네이버지도, 카카오맵) 별점 비교를 위해 서울 시내 업체 300개의(카페, 음식점, 주점 각각 100개씩) 별점 데이터를 수집하였다. 리뷰 수가 극히 적을 경우 소수의 사용자의 별점에 의해 업체의 별점이 크게 치우치는 경우가 있는데, 신뢰성을 확보하기 위해 리뷰가 30회 이상인 업체를 대상으로 하였다.</p> <p>또한, 리뷰 내용 분석을 위해 앞서 수집한 300개의 업체 중 임의로 180개를 추출하여(카페, 음식점, 주점 각각 60개씩) 다른 별도의 업체 리스트를 만들었다. 웹 크롤링을 통해 업체별로 100개씩 리뷰 데이터를 수집하였고, 네이버 지도 리뷰 데이터 18,677개와 카카오맵 리뷰 데이터 4,619개를 수집하여 텍스트 분석을 시행하였다.</p> <p>2. 데이터 분석</p> <p>2-1. 플랫폼별 별점 비교</p> <p>전체적인 분석을 진행하기 전, 플랫폼별 별점이 통계적으로 유의한 차이가 있는지 확인하기 위해 앞서 수집한 업체 300개를 대상으로 업체별 네이버 지도와 카카오 맵의 별점 차이를 비교하기 위한 검정을 시행하였다.</p> <p>검정에 앞서 변점의 diff(네이버 지도 별점 - 카카오 맵 별점) 대한 정규성</p> | | | | |

검정(Shapiro-Wilk test) 시행 결과, 별점 분포가 정규성을 따르지 않는다는 결과가 나왔다.

- 표1. Shapiro-Wilk test, Wilcoxon Signed-Rank test의 검정통계량 & p-value (소수 둘째 자리까지 표시)

| Shapiro-Wilk test | | Wilcoxon Signed-Rank test | |
|-------------------|---------|---------------------------|---------|
| 검정통계량 | p-value | 검정통계량 | p-value |
| 0.92 | 0.00 | 694 | 0.00 |

따라서 기존에 고려했던 Paired t-test 대신 비모수적 검정인 Wilcoxon Signed-Rank Test를 시행했다. 검정 결과 p-value가 모두 0에 매우 가까운 수치가 나왔고, 두 플랫폼의 별점 차이가 상당히 유의함을 알 수 있다.

2-2. 데이터 전처리

본 연구에서는 리뷰 텍스트 데이터를 활용하여 자연어 처리 기반의 분석을 수행하였다. 텍스트 데이터의 특성상 수치형 데이터와는 다른 전처리 방식을 통해 데이터 전처리 과정을 진행하였다.

2-2-1. 형태소 분석기 사용

먼저 리뷰 텍스트 데이터를 단어 수준으로 나누었다. 온라인 리뷰의 특성상 신조어, 줄임말, 이모티콘의 사용 빈도가 높는데, 이를 잘 처리할 수 있도록 Okt 형태소 분석기를 사용하였다. 형태소 분석을 통해 문장을 어절 단위가 아닌 의미 단위로 분리하고, 주요 명사 및 형용사, 동사를 추출하였다.

2-2-2. 불용어 제거 및 품사 기반 필터링

형태소 분석기를 사용해 분리한 단어들 중 의미 없는 단어들을 제거하기 위해 불용어를 제거하여 텍스트의 잡음을 줄였다. ‘입니다’, ‘하다’, ‘예요’ 같은 등 분석에 영향을 주지 않는 요소들을 불용어로 제거하였다.

2-3. 리뷰속 단어 등장 빈도 분석

플랫폼별 리뷰의 특성을 확인하기 위해 각 리뷰에서 자주 사용된 단어들을 파악해 보았다.

2-3-1. 플랫폼별 워드 클라우드 맵

- 그림1. 네이버지도 리뷰(왼쪽), 카카오맵 리뷰(오른쪽) 워드 클라우드 맵

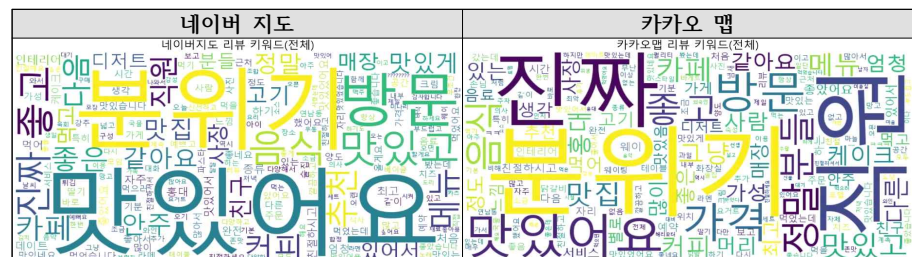


그림1을 보면 두 플랫폼 모두 긍정적인 단어(맛있어요, 분위기, 좋고) 위주로 리뷰가 작성되어 있는 것을 확인할 수 있다. 큰 차이가 나타나지는 않지만, ‘직원’이 카카오맵 리뷰에 더 빈번히 등장한 것을 알 수 있다. 이를 통해 카카오맵의 리뷰 작성자들은 네이버 지도 리뷰 작성자들에 비해 직원의 태도나 서비스를 중요하게 생각한다는 것을 알 수 있다.

2-3-2. TF-IDF 분석

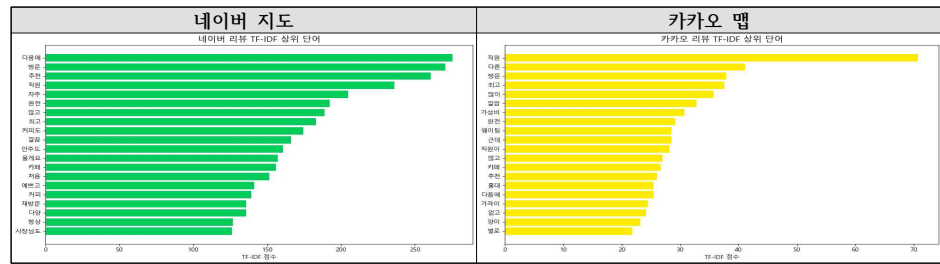
TF-IDF 분석은 텍스트 데이터를 숫자로 표현하기 위해 쓰이는 대표적인 방법으로 특정 단어가 해당 문서에서 얼마나 중요한가를 수치로 계산해주는 방식이다. TF(Term Frequency)는 특정 문서에서 단어가 얼마나 자주 등장했는지, IDF(Inverse Document Frequency)는 다른 문서들에서는 얼마나 드물게 등장했는지를 측정하는 방식이다.

$$\overline{TF} = (1 + \log_2 TF) / n_i \quad n_i: \text{리뷰 } i \text{ 에서 총 단어의 빈도수}$$

$$IDF = \log_2 (N / DF) \quad N: \text{전체 리뷰수, } DF: \text{문서 빈도수}$$

\overline{TF} 와 IDF 를 이용하여 정규화 $TF-IDF$ 를 정의할 수 있고 이 수치를 통해 개별 문서에서 각 단어의 상대적 중요도를 표현할 수 있어서 개별 문서에서 존재하는 키워드를 추출하는 데 활용되고 있다. 따라서 정규화 $TF-IDF$ 를 활용하여 플랫폼별 키워드의 중요도를 파악하고자 한다.

- 그림2. 네이버지도(왼쪽), 카카오맵(오른쪽) TF-IDF 점수 상위 20개



TF-IDF 점수를 보면 네이버 지도의 경우 ‘추천’, ‘최고’, ‘예쁘다’ 등 긍정적인 단어들이 상단에 위치했다. 카카오 맵의 경우에는 ‘직원’의 점수가 압도적으로 높았고, ‘별로’, ‘없고’, ‘웨이팅’ 같은 부정적인 단어들이 눈에 띄었다.

2-4. 리뷰 길이 분석

- 그림3. 네이버지도(왼쪽), 카카오맵(오른쪽) 리뷰 길이 분포

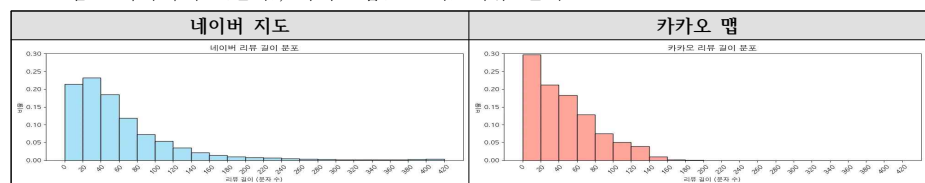


그림3을 보면, 두 플랫폼 모두 오른쪽으로 꼬리가 긴(right-skewed) 분포를 보인다. 네이버 리뷰는 주로 20~40자 구간에, 카카오 리뷰는 20자 미만에 집중되어 있으며, 이는 간단한 형식의 리뷰가 대부분이라는 점을 시사한다. 한편, 글자 수가 많은 리뷰는 적은 빈도로 존재하지만, 이는 상대적으로 풍부한 의견을 담거나 불만 사항을 구체적으로 기술했을 가능성이 있다.

2-4-1. 요약 통계량 비교

- 표2. 플랫폼 별 리뷰 길이에 대한 요약통계량

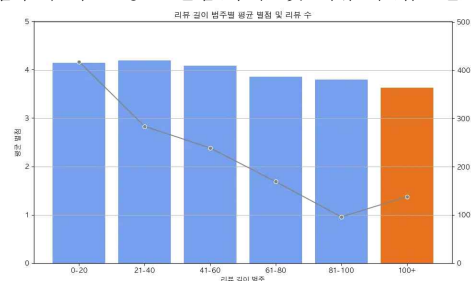
| 플랫폼 | 네이버 | 카카오 |
|---------|--------|-------|
| Count | 18,677 | 4,619 |
| Mean | 61.47 | 45.94 |
| Std | 60.96 | 36.65 |
| Min | 1 | 2 |
| Q1(25%) | 22 | 16 |
| Q2(50%) | 44 | 39 |
| Q3(75%) | 79 | 68 |
| Max | 420 | 181 |

표2에 따르면, 평균, 중앙값, 최댓값 모두 네이버가 카카오보다 높게 나타났으며, 네이버는 리뷰 수가 많을 뿐 아니라, 리뷰의 분포와 길이 면에서도 더 다양하고 긴 형태의 리뷰가 포함되어 있음을 보여준다.

2-4-2. 리뷰 길이와 별점 간의 관계 분석 (카카오 리뷰)

네이버 리뷰는 개별 별점이 없어, 카카오 리뷰에 대해서만 각각 리뷰와 별점 간의 관계 분석을 수행하였다. 리뷰 길이에 따라 20자 단위로 범주를 나누고 각 범주별 평균 별점을 확인한 결과, 리뷰 길이가 길어질수록 평균 별점이 낮아지는 경향을 보였다.

- 그림4. 리뷰 길이 범주에 따른 평균 별점(막대) 및 리뷰 수(꺾은선)



이를 정량적으로 검증하기 위해 일원분산분석(ANOVA)를 실시한 결과, p-value 는 0.0001로 모든 범주의 평균 별점이 동일하지 않음이 확인되었다.

또한, 사후 검정(쌍별 비교) 결과 다음 비교 쌍들에서 유의한 차이가 나타났다.

100자 이상 vs 20자 미만 / 100자 이상 vs 20~40자 / 100자 이상 vs 40~60자
 이로써, 100자 이상의 긴 리뷰일수록 상대적으로 낮은 별점이 부여되는 경향이 있다고 해석할 수 있다.

2-5. 감성분석

감성분석은 글에 내재되어 있는 사람들의 주관적인 태도나 감성을 추출해 내는 분석 기법을 말한다. Nsmc(네이버 영화 리뷰 공개 데이터셋)을 기반으로 사전을 학습시킨 뒤, 각 플랫폼에 맞게 감성사전을 구축하였고, 이를 활용하여 플랫폼별 리뷰에 대한 감성분석을 진행하였다.

- 표3. 플랫폼 별 리뷰 감성 분석 결과 (긍정, 중립, 부정)

| 플랫폼 | 네이버 | 카카오 |
|-----|--------|-------|
| 긍정 | 3,141 | 471 |
| 중립 | 13,356 | 2,751 |
| 부정 | 1,478 | 692 |

표3에 따르면, 대부분의 리뷰들이 중립성을 띄는 것으로 나타났다. 리뷰가 형식적, 간결한 문구 중심으로 작성되어 감성 사전에 강하게 반응하지 않는 경우가 많았을 수 있다. 플랫폼간 긍정-부정 비율이 다르게 나타났는데, 긍정비율의 경우, 네이버가 약 14%로 약 10%인 카카오보다 높게 나타났다. 이는 표현방식의 차이일 수도 있으나, 플랫폼간 리뷰문화의 차이로 인해 나타날 수 있다.(네이버는 이벤트, 후기 블로그 작성처럼 긍정 리뷰가 많고 카카오는 불만을 토로하거나 단순 불친절 경험 공유가 많은 가능성)

2-5-1. 감성분석과 별점의 관계 분석

- 그림5. 별점에 따른 감성 점수의 평균

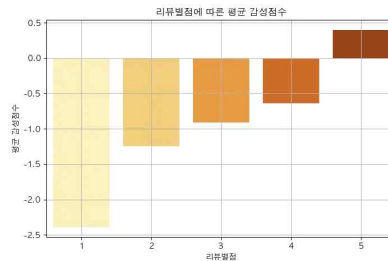
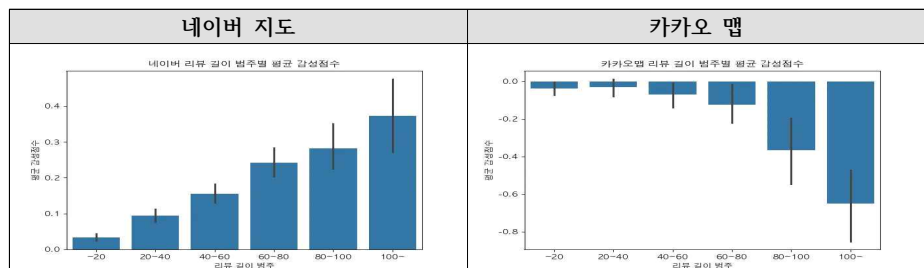


그림5에 따르면, 별점이 높아질수록 평균 감성점수가 증가하는 경향이 나타난다. 또한, 2, 3, 4점 구간의 평균 감성점수도 음수로 나타나는 것으로 보아, 전체적으로 감성점수가 보수적으로 평가되었거나 리뷰가 부정적인 니앙스를 담고 있는 경우가 많았을 수도 있음을 나타내며, 이는 카카오 플랫폼의 성향과 일치함을 나타낸다. 이러한 결과는 감성분석이 별점을 평가하는 데에 있어서 일관되게 작동하며, 추후에 불일치 리뷰 탐지나, 신뢰도 산출 등 추가적인 연구에도 사용될 수 있다는 것을 시사한다.

2-5-2. 감성분석과 리뷰 길이의 관계 분석

- 그림6. 리뷰 길이 범주에 따른 플랫폼별 감성 점수의 평균



리뷰 길이를 일정 구간으로 나누어서 분석한 결과, 카카오와 네이버 플랫폼간 명확한 차이가 나타났다.

카카오의 경우, 전반적으로 리뷰가 길어질수록 감성 점수가 낮아지는 경향을 보였다. 이는 불만이나 부정적인 감정을 자세히 서술한 긴 리뷰가 많다는 점을 시사한다.

네이버의 경우, 카카오와는 반대로 리뷰가 길어질수록 감성점수가 높아지는 경향을 보였다. 이는 이벤트성 리뷰가 많아서 리뷰가 길어질수록 긍정적인

| | | |
|-------------------|-------------------------------|---|
| | | <p>표현을 풍부하게 담는 경향이 있음을 시사한다. 이러한 결과는 신뢰도 분석이나 필터링 시스템 설계 시 플랫폼별 특징을 반영해야 한다는 점을 보여준다.</p> <p>Ⅲ. 결론</p> <p>본 연구는 동일한 업체에 대한 네이버 지도와 카카오 맵의 리뷰 데이터를 비교분석 함으로써 플랫폼별 리뷰 특성과 별점 체계의 차이를 정량적으로 분석하고자 하였다. 별점 비교, 텍스트 빈도 비교, 리뷰 길이 분석, 감성분석 등 총 네 가지 분석을 시행하였고, 플랫폼별로 리뷰의 특성(길이, 평점에 따른 감성 점수 등)이 다르다는 결과를 얻을 수 있었다.</p> <p>이러한 분석 결과는 단순히 리뷰를 수집하여 활용하는 것을 넘어, 구조적 특성을 이해하고, 리뷰 신뢰도 판단 및 허위 리뷰 탐지 기준 설계 등에 기여할 수 있다. 특히 플랫폼별 필터링 시스템이 리뷰 형식과 내용에 영향을 줄 수 있음을 고려할 때, 리뷰 데이터를 연구할 때 플랫폼 특성을 반영한 분석이 필수적임을 시사한다.</p> <p>다만, 본 연구는 서울 지역 업체에 한정하였고, 수집 시점에 따라 데이터 편향이 존재할 수 있다는 한계가 있다. 또한 리뷰 감성분석은 사전 학습된 감성사전에 기반하기 때문에 특정 문맥이나 맥락에 대한 이해가 부족할 수 있다. 향후에는 더 다양한 지역과 업종, 시계열 데이터를 활용한 확장 연구 및 리뷰 신뢰도 측정을 위한 모델링이 병행된다면 보다 실용적이고 정교한 인사이트를 도출할 수 있을 것이다.</p> |
| 현실성 및 지속가능성 | 주제의 현실 적용 가능성 | <p>소비자들은 리뷰를 선택을 위한 주요 척도로 사용하고 있으나, 마케팅 목적의 허위 리뷰가 증가하면서 신뢰성 문제가 발생한다. 이 연구는 여러 가지 분석을 통해 리뷰의 신뢰성에 대한 의문을 제기하고 있고, 더 나은 신뢰성 있는 리뷰의 필요성을 시사하고 있으므로 리뷰 서비스 개선을 위한 문제제기 측면에서 현실 적용 가능성이 높을 것으로 기대된다.</p> |
| | 주제의 발전 가능성 및 타당성 | <p>본 연구는 단순히 텍스트의 정성적인 분석이 아닌 감성점수, 리뷰 길이 등 정량지표를 결합하여 리뷰의 신뢰도를 분석했다. 특히 분포 비교, 범주형 변수 분석 등을 통해 통계적으로 타당성을 검증하였다.</p> <p>또한 단순한 통계적 설명을 넘어, 향후 지도학습 기반의 모델로서의 확장 가능성도 있다. 추후에는 기업이 보유한 메타 데이터를 결합한다면 더 높은 정확도를 확보할 수 있을 것으로 기대된다.</p> |
| 창의성 및 전문성 | 주제에 대한 분석의 차별성 및 독창성 | <p>기존의 선행연구에서는 단순히 리뷰를 긍정/부정으로 나누거나, 텍스트 자체의 분석을 통한 접근방식을 지녔다. 하지만 본 연구에서는 단순 분류를 넘어 감성점수와 별점 간의 불일치, 짧지만 과도하게 긍정/부정적인 리뷰 등의 패턴을 문제 삼아 신뢰도 판단의 지표로 사용하였다.</p> <p>또한, 플랫폼별 비교를 통해 동일 업체에 대한 소비자들의 반응 차이를 확인하였고, 이를 기반으로 신뢰성 있는 리뷰 판단지표의 필요성을 제시한다.</p> |
| | 의견 도출의 논리 및 과정 | <p>본 연구는 리뷰의 신뢰성에 의문을 갖고 플랫폼별 리뷰 특성의 차이를 비교하기 위해 네이버 지도와 카카오 맵의 리뷰와 별점 차이를 정량적으로 비교하였고, 플랫폼별로 유의미한 차이를 보였다. 이어진 텍스트 분석을 통해 네이버 리뷰는 더 긍정적인, 카카오 리뷰는 아쉬운 점에 대한 언급이 많음을 확인하였다. 특히, 리뷰의 길이가 길어질수록 이런 경향이 커지는 패턴이 나타났고 플랫폼 간 리뷰 문화 및 구조적 차이가 존재한다는 결론과 함께 신뢰할 수 있는 리뷰 시스템이 필요하다는 의견을 제시한다.</p> |
| | 전공융합 적절성 | <p>감성점수와 별점의 분포 등은 통계학의 정량분석기법을 기반으로 하여 데이터 기반의 합리적인 판단의 근거가 되었다. 동시에, 이 연구는 소비자 행동분석과, 마케팅 전략과 관련된 연구로 경영경제학적 접근과도 융합될 수 있다. 기업의 입장에서 신뢰도는 기업의 이미지와 연결되며, 이를 통해 소비자의 행동에 영향을 미칠 수 있다.</p> <p>결론적으로, 통계학적 기법을 활용하여 실질적인 경제적/경영적 가치로 확장이 가능하다.</p> |
| 의의 및 기대효과 | | <p>해당 연구는 여러 플랫폼의 리뷰 데이터를 기반으로 감성분석과 정량분석을 통해 신뢰성 있는 리뷰 판단지표의 필요성을 제시한다. 리뷰를 통한 마케팅으로 왜곡될 수 있는 소비자 경험을 개선하고, 최종적으로는 실제 이용자들의 정보만 제공하는 것을 목표로 한다.</p> <p>이를 통해 소비자들의 선택에 실질적인 도움을 줄 수 있으며, 플랫폼 운영자들에게는 더 발전된 실질적인 허위 리뷰 탐지 시스템의 필요성과 방향성을 제시한다. 또한, 리뷰 기반의 마케팅 연구에도 응용될 수 있을 것으로 기대된다.</p> |