

패션상품 온라인 리뷰의 분류 및 예측

Classification and Prediction of Fashion Product Online Review

김현숙*

배재대학교 의류패션학과 교수

Kim, Hyunsook*

Dept. of Clothing & Textiles, Paichai University¹⁾

Abstract

Fashion products should encompass a wide variety of styles to suit individual preferences. Therefore, online retailers should identify diverse consumer preferences and satisfaction-dissatisfaction factors early through online review analysis. This enhances consumer satisfaction and improves product planning and marketing efficiency. However, product reviews often cover multiple topics within a single document, and due to the extensive nature of unstructured data with varying expressions and lengths, it is challenging to effectively harness them for practical purposes. The objective of this study was to categorize fashion product reviews and to create sentiment prediction models for establishing systematic analysis. The specific research includes: first, classifying online fashion product reviews; and second, building a sentiment prediction model and validating it. The Amazon review dataset 2018 including 881,895 data entries was used. Fashion product reviews were categorized into three groups using the k-means clustering algorithm; "Size & Fit," "Quality & Price," and "Appearance." Validation of significance was conducted using ANOVA mean analysis, post hoc analysis, and silhouette score. To construct sentiment prediction models for fashion product online reviews, various embeddings and algorithms were combined and tested. Embeddings included Count Vectorization, TF-IDF, and Word2Vec, while algorithms comprised SVC, logistic regression, random forest classifier, bagging classifier, and multinomial NB. Comparative analysis against a bi-directional sequence LSTM model revealed that the bi-directional LSTM model achieved the highest accuracy of 0.94. In conclusion, this study proposed models for categorizing fashion product reviews and establishing a sentiment prediction model, thus providing a schema for efficient review-based product planning and marketing strategies.

Keywords: Online review, Fashion product, Sentiment prediction, K-means clustering, LSTM

I. 서론

코로나 이후 온라인 쇼핑이 급격히 증가하는 가운데 패션상품의 온라인 거래액은 총 온라인 구매의 25.8%에 달하는 50조원에 이르렀다(정정숙, 2022). 온라인 쇼핑에서 소비자는 상품을 화면상으로 봐야 하고 직접 입어보고 구

매할 수 없기 때문에 다른 소비자의 구매 리뷰를 많이 참조한다. 판매자의 관점에서도 패션상품은 상품구성이 다양하며 소비자에 따라 선호가 달라 판매예측 및 재고관리가 힘들기 때문에 온라인 리뷰 분석을 통해 사전에 고객 반응을 예측하고 이를 후속 발주, 재고관리 및 마케팅에 활용한다(McKinney & Shim, 2016). 패션상품 온라인 구매리뷰

* Corresponding author: Kim, Hyunsook
Tel: +82-42-520-5984, Fax: +82-70-4850-8468
E-mail: hskim88@pcu.ac.kr

가 소비자의 반응을 조기에 파악하고 후속 구매를 예측하는데 중요한 지표임에도 불구하고 온라인 리뷰는 표현 양식 및 길이가 다른 방대한 양의 비정형 데이터로 구성되며 하나의 리뷰에서 여러 긍·부정 요인이 동시에 언급되기 때문에 실무에 활용하는데 어려움이 있다(Almiron-chamadoira, 2018).

근래에 온라인 리뷰에 관한 연구가 활발히 이루어지고 있는데 대부분 온라인 리뷰의 분석을 통해 평점이나 감성을 예측하거나 소비자 행동이나 매출에 리뷰가 미치는 영향을 연구하였다(Chong et al., 2016; Elmurungi & Gherbi, 2008; Kim & Kim, 2018; Lee et. al., 2019; 이정원, 박철, 2020). 이들 연구에서는 온라인 리뷰의 내용과 특성, 길이, 리뷰 수 및 리뷰 작성자 등을 독립변수로 이들 변수가 소비자 감성이나 평점, 제품 및 서비스의 흥행, 매출, 인지도 및 구매의도에 미치는 영향을 연구하였다(Reich & Maglio, 2020; 이동엽 외, 2017; 이윤주 외, 2020; 이정원, 박철, 2020; 한기향, 2021).

온라인 리뷰와 평점의 관계에서 소비자마다 만족과 불만족을 느끼는 요인에 차이가 있어 평점 평균만으로 정확히 소비자 만족과 불만족의 상태를 파악하기 힘들다(김준겸 외, 2021). 또한 온라인 리뷰를 구성하는 내용은 상품 카테고리별로 다른 특성을 가지며 구매에 결정적인 영향을 미치는 요인도 다르므로 상품 카테고리별로 리뷰를 체계적으로 분류할 수 있는 시스템이 필요하다. 이에 본 연구에서는 패션상품 리뷰의 내용으로부터 텍스트 벡터를 추출하여 보다 효율적으로 리뷰 평점의 감성을 예측할 수 있는 알고리즘을 테스트하고자 한다.

본 연구의 목적은 온라인 패션상품 구매 리뷰를 의미론적 분석 및 감성분석에 의해 유형화하고 서로 다른 임베딩 방법과 알고리즘의 차이에 따른 예측 성능의 차이를 검증하는 것이다. 이를 위한 구체적인 연구문제는 다음과 같다. 첫째, 패션상품 구매 후 리뷰 텍스트를 군집분석으로 분류하고 분류의 정확도를 평가한다. 둘째, 패션상품 구매 후 리뷰의 극성을 감성 예측모델로 분류하고 분류의 정확도를 평가한다. 온라인 상품구매 리뷰 데이터가 효율적으로 사용되기 힘들다는 점을 고려할 때 본 연구에서 정확도 높은 분류체계를 제시하고 최적의 알고리즘을 구축한다면 향후 구매 리뷰의 감성 예측 시스템에서 빠른 시간에 상품 판매를 예측하여 재고관리에 활용하며 고객 불만족에 신속하게 대응할 수 있을 것이다. 또한 평점이 없는 블로그나 SNS 글에 대해서도 감성을 판별하고 이로부터 평점을 예측할 수 있다. 또 내용과 평점이 불일치하는 거짓 리뷰를 판별하여

적절히 대응할 수 있을 것이다. 결론적으로 본 연구는 패션 리뷰 연구에 기초자료 활용될 수 있을 뿐 아니라 온라인 마케팅 실무 및 상품기획, 재고관리 등에서 효율적으로 활용될 수 있을 것이다.

II. 이론적 배경

1. 온라인 상품 구매 리뷰

온라인 쇼핑물 상품 구매 리뷰는 온라인 구전의 대표적인 형태로 소비자가 제품 구매 후 남긴 글을 말한다(이호근, 박현, 2013). 구매 리뷰는 소비자의 제품 구매 의사 결정에 영향을 주므로 소비자에게나 판매자 모두에게 중요한 의미를 가진다(Almiron-Chamadoira, 2018). 이에 온라인 리뷰에 대한 연구가 활발히 이루어지고 있으며 온라인 리뷰가 평점이나 감성 및 소비자 구매행동에 미치는 영향에 대한 연구가 있다(Chong et al., 2016; Elmurungi & Gherbi, 2008; Kim & Kim, 2018; Lee et. al., 2019; Reich & Maglio, 2020; Zhang & Liu, 2018; 이동엽 외, 2017; 이윤주 외, 2020; 이정원, 박철, 2020; 한기향, 2021).

온라인 리뷰의 영향에 대한 연구결과는 다양한데 먼저 온라인 리뷰의 방향성이 제품의 품질 등 지각된 위험을 감소하는 불확실성 회피의 효과를 통해 매출액에 유의한 영향을 준다고 하였다(이정원, 박철, 2020). 한편 온라인 리뷰 자체가 구매결정에 영향을 주기 보다는 리뷰의 숫자가 인지도에 영향을 준다고 한 연구도 있다(Duan, et al., 2008). 또한 온라인 리뷰의 방향성과 양은 개별 효과보다 상호작용 효과가 크다고 하였다(Chong et al., 2016). 소비자 구매 후 반응으로 긍정보다 부정 리뷰의 중요성이 더 크기 때문에 부정 리뷰에 대한 피드백을 적절히 제공하지 못하면 소비자 신뢰와 명성에 문제가 생기게 된다(Elmurungi & Gherbi, 2018).

온라인 리뷰의 감성분석에서 중요한 것은 어떻게 부정적이거나 긍정적이거나 중립의 리뷰를 감지하느냐 하는 것이다. 리뷰 평점만으로는 소비자의 만족 및 불만족을 결정하는 핵심 요소를 파악하기 힘들며 평점과 더불어 리뷰 내용에서 언급되는 속성 및 긍·부정 감성을 포괄적으로 분석해야 한다(권혜진, 2020). 리뷰의 내용에 대한 선행연구를 살펴보면 부정 리뷰와 관련해 제품이 불만족스러울 때 리뷰의 길이가 더 길었으며 왜 불만족했는지에 대해 더 구체적으로 설명하는 경향이 있었다(조용희 외, 2018). 또 이전에 잘못 구매한 경험이 있는 리뷰어가 추천하는 제품을 더 많이 구매하

였으며(Reich & Maglio, 2020), 여러 네트워크 연결이 있는 리뷰어의 리뷰는 오히려 부정적인 영향을 미쳤다(Zhang & Liu, 2018). 때로 온라인 리뷰 내용과 평점이 불일치하는 경우도 왕왕 있으며 가짜 리뷰도 있어 신뢰도와 유용성에 의문이 제기되기도 한다(Wang & Chen, 2020; 김준겸 외, 2021).

온라인 리뷰 연구에서 많이 사용한 연구방법은 수치에 의한 모델링 연구와 리뷰의 내용에 중점을 둔 의미론적 분석이 있다. 리뷰의 의미론적 연구에서 이윤주 외(2020)는 흥행 지표를 웹툰 플랫폼 정식 연재로 보고 성공한 웹툰과 그렇지 못한 웹툰 14개의 댓글 36,194개를 TF-IDF(Term Frequency-Inverse Document Frequency)로 주요 단어를 비교 분석하였다. 또 Kim과 Kim(2018)은 신발에 대한 온라인 리뷰를 내용 분석하여 리뷰의 조절초점이 예방초점인지 향상초점인지를 분류하고 구매의 시간적 거리에 따라 리뷰가 구매의도에 미치는 영향을 분석하였다. 그 결과 가까운 미래일 때에는 예방초점의 리뷰에 대해 보다 호의적인 태도를 보였고 먼 미래의 구매에 대해서는 조절초점에 따른 차이가 나타나지 않았다. 온라인 리뷰의 모델링 연구에서는 온라인 리뷰의 감성분석 연구가 주로 이루어졌다(윤소영, 윤성대, 2020; 이동엽 외, 2017; 이윤주 외, 2020). 온라인 리뷰 연구에서 평점에만 의지해 소비자 만족도를 판단하거나 상품 정책을 전개하는 것은 불완전한 정보에 의존하게 될 위험이 크며 평점과 리뷰 내용의 일치도를 평가해야 한다. 온라인 리뷰 분석이 상품기획, 마케팅 전략 및 고객 서비스에 효율적으로 활용되기 위해서는 수치에 의한 모델링과 더불어 내용에 중점을 둔 의미론적 분석이 이루어져야 하겠다. 온라인 리뷰를 상시적으로 관리할 수 있는 시스템의 구축을 위해서는 수적 모델링 뿐 아니라 의미론적 분석에 기반한 리뷰의 유형화가 선행되어야 하겠다.

2. 온라인 패션상품 구매 리뷰

온라인 리뷰는 상품의 장점, 품질, 단점 및 기타의 내용으로 구성된다. 리뷰의 내용에는 모든 상품에 공통되는 일반적인 내용이 있고 해당 상품 카테고리별로 특정적인 부분이 있다. 온라인 리뷰의 효과를 정확히 예측하기 위해서는 상품 카테고리 별 특성을 반영한 분석이 필요하다(Almiron-Chamadoira, 2018; Lee, et al., 2019). 패션상품 구매리뷰에 대한 선행연구 역시 리뷰의 특성, 리뷰의 내용 분석, 감성 분석 예측 모델 및 리뷰의 유형화를 중심으로 이루어졌는데, 특히 패션상품은 그 성격상 소비자의 정체성

이나 자기표현과 관련이 깊기 때문에 온라인 리뷰 또한 유동적이며 비체계적이며 주관적인 요소의 개입 여지가 많아 이러한 유동성을 개념화할 수 있는 틀이 필요하다(Kawarf & Istanbuluoglu, 2019).

패션상품의 긍정적인 온라인 리뷰는 제품과 판매자에 대한 신뢰를 증가시키며 신뢰를 매개로 충동구매의 가능성을 증가시킬 뿐 아니라 브랜드 이미지에 우호적으로 작용하여 구매를 촉진한다(Hilal & Astuti, 2022; Purnamawati, et al., 2022), 그러나 온라인 리뷰는 자신의 경험을 주관적으로 전달하는 경우가 많기 때문에 긍정적 리뷰도 영향력이 낮을 수 있으며 평가내용과 방향성이 상호작용하여 주관적인 정보보다 객관적이고 긍정적인 정보가 가장 신뢰도가 높게 평가되기도 한다(서현진, 이규혜, 2013). 따라서 패션상품의 온라인 리뷰가 구매에 미치는 영향을 분석할 때에는 그 내용 및 방향성을 동시에 고려해야 할 것이며 서술형 리뷰의 내용을 유형화하고 평점과의 관계를 파악해야 할 것이다.

패션 구매 리뷰의 내용이 중요함에 따라 내용의 감성 및 토픽을 분석한 연구가 많은데 한기향(2021)은 아마존에서 크롤링한 반팔 티셔츠에 대한 구매 리뷰를 형태소 분석, 빈도분석, TF-IDF, 감성분석, 연결중심성으로 분석한 결과 반팔 티셔츠 구매 시 가장 중요하게 생각하는 요인이 사이즈임을 밝혔다. 사이즈 다음으로 심미적 요인, 소재의 품질, 세탁, 활용성이 중요하게 나타났다. 또 감성 분석 결과 긍정의 비율이 높으며 심미적 요인과 연관된 ‘예쁘다’와 ‘귀엽다’가 많이 나왔고 사이즈에 대한 긍정의 댓글도 많았다. Elmurghi와 Gherbi(2018)은 패션 리뷰에 머신러닝 알고리즘을 사용하여 감성을 분석하였는데 의복, 신발, 주얼리 리뷰, 유아용품 리뷰 및 애완동물 물품 리뷰의 세 개의 데이터 셋에 대해 4가지 머신러닝 알고리즘을 사용하여 감성 분석을 하고 모델 성능을 비교하였다.

이상에서 패션상품의 구매 리뷰에 대한 선행 연구를 살펴본 결과 패션상품 리뷰 고유의 특성을 파악하려는 연구가 있었고 패션상품의 리뷰가 다면적이고 유동적임에 따라 그 내용을 유형화하고 분류하려는 시도가 있었다. 그러나 패션 구매 리뷰에 인공지능 모델을 활용한 연구는 별로 없었으며 구매 리뷰가 방대한 양의 빅데이터로 쌓임에 따라 이를 효율적으로 관리하기 위해서는 인공지능 기반의 모델링 연구가 유형화 연구와 더불어 활성화될 필요가 있겠다.

3. 온라인 구매 리뷰의 군집화 분류

선행연구에서는 온라인 구매 리뷰를 다차원적으로 보고

이를 유형화하였다. 온라인 구매 리뷰의 유형화 연구에는 내용분석이나 연결망 분석과 같은 질적연구와 계층적 군집 분석이나 k-평균 군집화와 같은 양적 연구가 있다(Lee, et al., 2019; Korfiatis, et al., 2019; 허지욱, 2018). 패션상품 리뷰의 유형화 연구로 McKinney 와 Shin(2016)은 정장 렌탈 경험에 대한 리뷰를 내용 분석하여 의복평가기준, e-서비스 평가기준 및 기타 내용으로 분류하였다. 한편 채희주 외(2016)는 패션상품 리뷰를 긍·부정 리뷰, 도구적·쾌락적 리뷰로 분류하여 소비자의 신뢰에 미치는 영향을 분석하였다.

온라인 리뷰의 양적 유형화 연구로 Lee, et al.(2019)은 에어비앤비에 대한 온라인 리뷰 196,666개에 대해 계절 및 시간의 흐름에 따라 좋은 품질에 대한 생각이 바뀌었는지 분석하였다. 이때 분석은 상위 10개 키워드에 대해 텍스트 마이닝으로 TF-IDF를 구하고 계층적 군집분석, 시계열 분석 등으로 온라인 리뷰의 시간적 변화를 분석하였다. 그 결과 시간이 지나도 공통적으로 좋은 품질의 요건이 되는 요소가 있는 반면 특정 년도와 시즌에 독특하게 나타나는 속성도 발견되었다. 온라인 리뷰의 유형화에 인공신경망 모델을 적용한 연구에서는 주로 Word2Vec을 이용하여 단어간 유사도를 통해 임베딩하고 이를 K-평균 군집화 등의 방법으로 유형화하였다. 강장경과 박은혜(2017)는 이러닝 수강 후기를, 허지욱(2018)은 네이버 블로그의 글을 Word2Vec을 이용하여 단어 간 유사도를 구하고 임베딩하여 KNN과 같은 클러스터링 알고리즘으로 군집화하였다. 강형석과 양장훈(2019)은 Word2Vec으로 단어를 임베딩하고 군집화한 후 군집의 특성 및 t-SNE 분포로 단어간 관계 표현 능력을 정성적으로 분석하였다. 김성범(2020)은 아마존 리뷰를 군집 분석한 후 3개의 상품 유형 즉 하드웨어 제품, 서비스 상품, 하드웨어와 클라우드 서비스가 융합된 형태의 상품 유형별 리뷰가 서로 다른 군집에 속함을 확인하였다.

이같이 선행연구에서 구매 리뷰의 유형화는 주로 군집분석으로 이루어졌는데 군집분석은 비지도학습으로 이루어지며 군집 수를 연구자의 판단에 의해 결정해야 하고 분류분석이 재현될 때 분석의 정확도를 확인하기 어렵다는 단점이 있다. 특히 온라인 리뷰는 그 성격상 개인의 견해가 표현되기 때문에 유동성이 커서 이러한 유동성을 개념화하는 것이 필요하다는 점을 감안할 때 온라인 리뷰의 유형화 연구는 아직도 해결되어야 할 부분이 많은 도전적인 연구 주제이다(Kawaf & Istanbuluoglu, 2019).

4. 온라인 구매 리뷰의 감성 예측 모델

감성분석(sentiment analysis)이란 텍스트에서 의견, 정서, 감정 및 태도를 분석하여 감성지수를 계산하고, 긍정지수 및 부정지수로 구성된 감성지수를 사용하여 텍스트에 포함된 긍정, 부정의 감성을 판단하는 기법이다(윤소영, 윤성대, 2020). 온라인 구매에서 방대한 양의 사용자 리뷰에 대한 감성 분석은 사용자 만족을 효과적으로 증가시킬 수 있을 것이다(Yang, et. al., 2020). 온라인 리뷰에 감성분석을 적용한 연구에는 긍·부정의 감성에 해당하는 핵심 단어를 추출한 연구, 인공신경망 분석을 통해 감성을 예측하는 모델링 연구 등이 있으며 감성분석 결과를 활용하여 다른 변수를 예측하거나 다른 변수와 결합시켜 알고리즘을 향상시킨 연구가 있다(윤소영, 윤성대, 2020; 이동엽 외, 2017; 이지현 외, 2020; Chong et al., 2016; Elmurungi, & Gherbi, 2018; Yang, et al., 2020). 각각의 주제에 해당하는 연구를 아래에서 살펴보겠다.

온라인 패션 구매 리뷰에서 긍·부정의 감성에 해당하는 핵심 단어를 추출한 최영현과 이규혜(2020)의 연구에서는 인터넷 카페, 네이버, 다음, 유튜브에서 제공하는 웹문서를 대상으로 패션제품 리뷰에서 빈도와 TF-IDF로 핵심 단어를 추출하여 감성 사전을 구축하고 베이지안 분류기(Bayes classifier)로 긍정, 중립 및 부정의 감성을 예측하는 모델을 만들었다. 그 결과 패션 리뷰에서 구매 결정 요소로 가격, 크기, 장식, 배송, 세탁 및 소매상의 안정성, 제품구매 동기가 추출되었고, 해결해야 할 불만족 요소로 불충분한 양, 공급 배송지연, 고가가격, 불량한 장식, 퇴색한 색상, 세탁물 오염, 구매 후 불쾌한 냄새가 나왔다.

감성 예측 모델링 연구로 이동엽 외(2017)의 연구에서는 아마존 패션상품에 대한 리뷰 570만건을 Word2Vec으로 임베딩하고 SVM분류기 모델로 학습하여 사용자 감성을 예측하였고 그 결과 88.0%의 정확도를 얻었다. 한편 감성 분석 모델의 알고리즘을 개선하기 위한 연구로 Yang, et al.(2020)은 감성 어휘와 CNN, BiGRU(attention-based Bidirectional Gated Recurrent Unit)를 결합하여 새로운 감성모델을 만들었다. 또 Agarap(2018)은 감성 분류 분석에 LSTM(long-short term memory)을 사용한 RNN으로 그라디언트가 사라지는 문제를 해결하였으며 이를 추천과 비추천의 감성 분류에 적용시킨 결과 0.88과 0.93의 F-1 score를 얻었다.

감성분석 결과를 활용하여 다른 변수를 예측한 연구로 윤소영, 윤성대(2020)의 연구에서는 상품 리뷰를 분석하고 이를 가중치로 사용하여 협업 필터링에 적용하였다. 또 온라인 리뷰에서 매출을 예측한 Chong et al.(2016)의 연구에서

는 온라인 리뷰의 양과 평점, 온라인 홍보전략과 더불어 리뷰 감성이 매출에 미치는 영향을 MLP(multilayer perception)로 분석하였다. 그 결과 온라인 리뷰, 온라인 홍보전략, 온라인 감성이 제품 매출에 미치는 영향은 각각 달랐으며 온라인 리뷰 수는 감성이나 할인과 상호작용하였다.

이상에서 본 바와 같이 구매 리뷰 텍스트에서 감성 분석은 주요 키워드를 추출하거나 감성 예측 알고리즘을 구현하는 연구가 주를 이루었다. 각 연구에서 감성 분석은 구매 리뷰에서 유용한 정보를 얻는데 적합함을 알 수 있으며 선행연구에서 검증된 기법을 사용하여 보다 효율적인 리뷰 분석 시스템을 만들 수 있음을 알 수 있다.

III. 연구모델 및 연구방법

1. 연구모델

본 연구의 목적은 아마존의 패션상품 리뷰 데이터셋을 사용하여 온라인 패션 리뷰 분류 시스템을 제안하고 예측모델을 제안하는 것이다. 이를 위해 패션리뷰를 의미론적으로 분류하고 감성을 예측하여 각각의 성능을 테스트하였다. 본 연구의 범위는 아마존 패션상품 구매 후 리뷰 데이터에서 벡터를 추출하여 군집분석하고 감성을 예측하여 양적, 질적

분류 시스템을 제안하는 것으로, 연구결과는 리뷰 분석을 위한 실무적, 이론적 기초 자료로 사용될 수 있을 것이다. 연구문제는 아래와 같다.

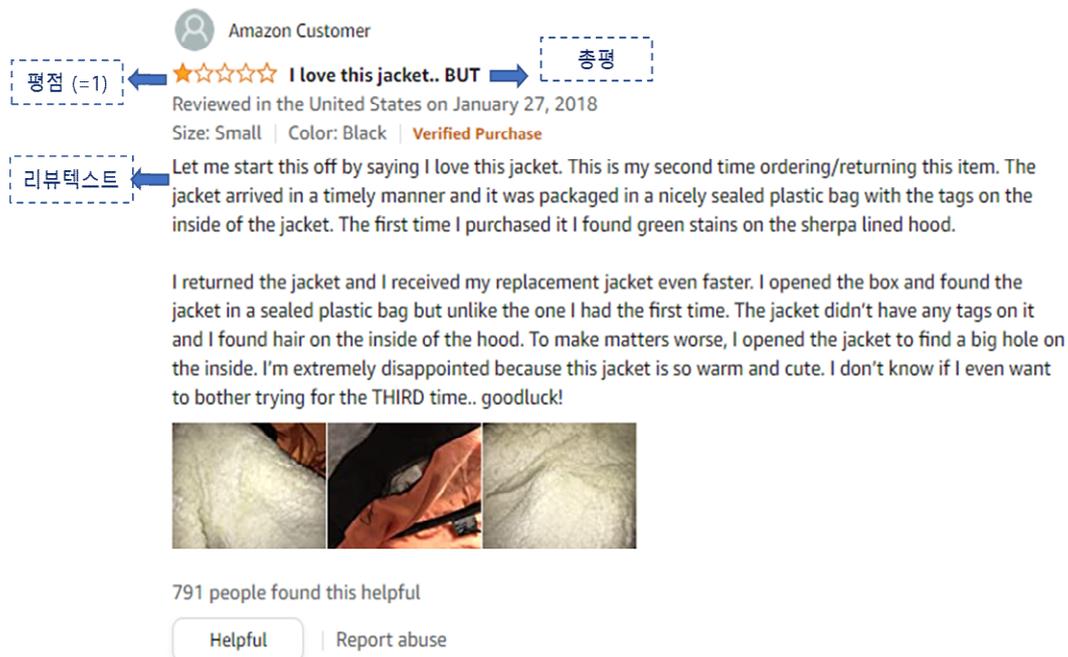
연구문제 1. 패션상품 리뷰를 유형화한다.

연구문제 2. 패션상품 리뷰를 감성 예측모델을 구하고 성능을 평가한다.

2. 데이터 수집 및 전처리

본 연구에서는 Ni(2018)의 아마존 리뷰 데이터셋에서 패션상품 리뷰데이터(Clothing Shoes and Jewelry) 883,636개를 사용하였다. 데이터는 평점, 구매확정, 리뷰 시간, 리뷰아이디, 아신코드(asin code), 리뷰어이름, 리뷰 텍스트, 총평, 표준시간, 좋아요, 상품 스타일, 상품이미지의 12개 열로 구성되었으며 이중 평점, 리뷰 텍스트 데이터 및 총평을 분석에 사용하였다. 데이터 분석환경은 Windows 11 운영체제에서 파이썬 기반의 아나콘다 쥬피터노트북(Anaconda Jupyter Notebook)을 활용하였다. 아마존리뷰 데이터셋에서 리뷰 텍스트와 총평 및 평점의 위치는 [그림 1]과 같다.

분석에서는 데이터 883,636개 중 리뷰텍스트 및 총평에서 결측치가 있는 1,741개(0.20%)를 제외하고 총 881,895개를 분석에 사용하였다. 평점 데이터 값은 최소 1점에서



[그림 1] 아마존 리뷰 데이터의 예

최대 5점까지 분포하였으며 5점이 가장 많은 464,262개 (52.6%) 였고, 4점이 149,103개(16.9%), 3점이 96,945개 (11.0%), 2점이 64,669개(7.3%), 1점이 106,916개 (12.1%)로 분포하였다. 본 연구의 리뷰 분석에서는 평점의 긍·부정을 예측의 타깃으로 하였으며 평점 5를 긍정으로 보고 평점 1점과 2점을 부정으로 보았다. 이때 평점 1점과 평점 2점 데이터의 합이 총 19.4%에 지나지 않아 52.6%를 차지하는 평점 5와는 데이터 양에 차이가 크므로 평점 1점과 평점 2점을 부정으로 코딩하고 평점 5점을 긍정으로 코딩하였으며, 데이터를 셔플(shuffle)로 섞고 위에서부터 긍정과 부정의 데이터를 같은 수로 선택하여 모델 예측에 사용하였다.

3. 연구방법

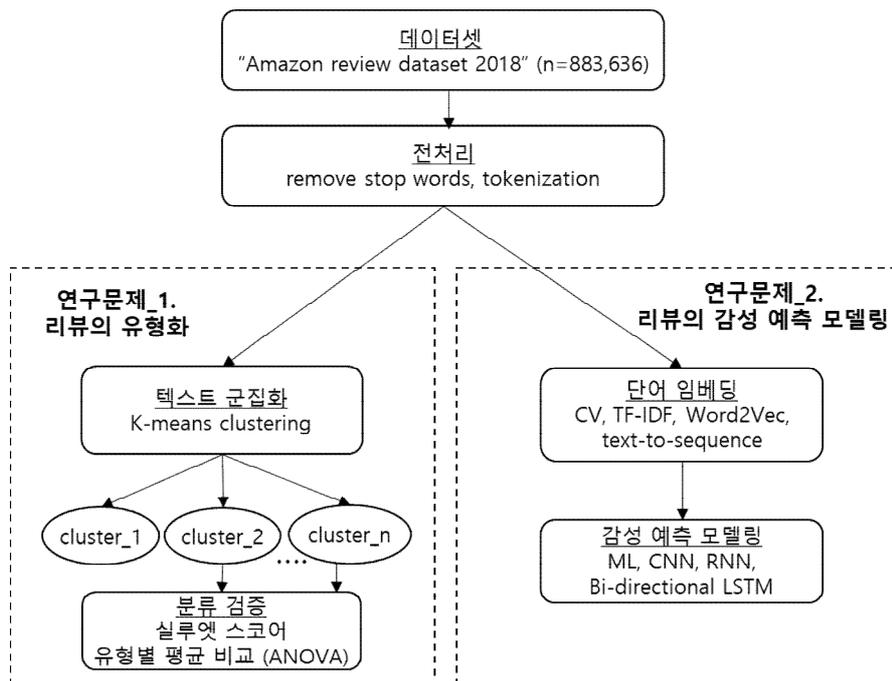
연구방법은 아마존의 패션상품 리뷰 데이터셋을 사용하여 [그림 2]의 연구 절차와 같이 텍스트 데이터를 분석하였다. 아마존 패션상품 리뷰 883,663개를 분석에 사용하였고 5점 만점의 평점에 대해 1점과 2점을 부정, 5점을 긍정으로 각각 라벨링하여 인공지능망 분석으로 감성 예측 모델을 검증하였다. 연구문제 1에서 단어 임베딩 방법으로 CV(Count Vectorization), TF-IDF, Word2Vec, text-to-sequence를 사용하였고, 감성 예측 모델링 방법으로 ML에서 SVC,

logistic regression, random forest classifier, bagging classifier, multinomial NB와 딥러닝 모델로 CNN, RNN, Bidirectional LSTM 모델을 테스트하였다. 또 리뷰 텍스트의 긍·부정 내용을 분석하기 위해 LSA, LDA로 토픽분석을 하여 주제어를 추출하였다. 연구문제 2에서는 리뷰의 유형을 k-평균 군집화와 DBSCAN을 사용하여 분류하였고 그 결과의 유효성을 검증하기 위해 실루엣 스코어와 ANOVA 검증을 사용하였다. 분류된 집단별로 평점예측 모델의 감성 예측 정확도를 검증하고, 리뷰 대신 총평을 넣어 분류 성과를 살펴보았다. 리뷰 분류 시스템을 총평으로 다시 검증한 것은 Praveenraj, et al.(2021)가 남성 패션 액세서리 귀걸이의 온라인 구매 리뷰를 대상으로 감성 분석을 한 연구에서 코멘트의 감성 점수와 타이틀의 감성 점수에서 평점과 감성 점수가 관련이 있다고 한 연구결과에 기초한다.

IV. 결과 및 논의

1. 패션상품 리뷰의 텍스트 벡터화

아마존 패션상품 리뷰 데이터인 Amazon review dataset 2018에서 리뷰 평점으로 긍정, 부정 데이터를 1:1로 추출



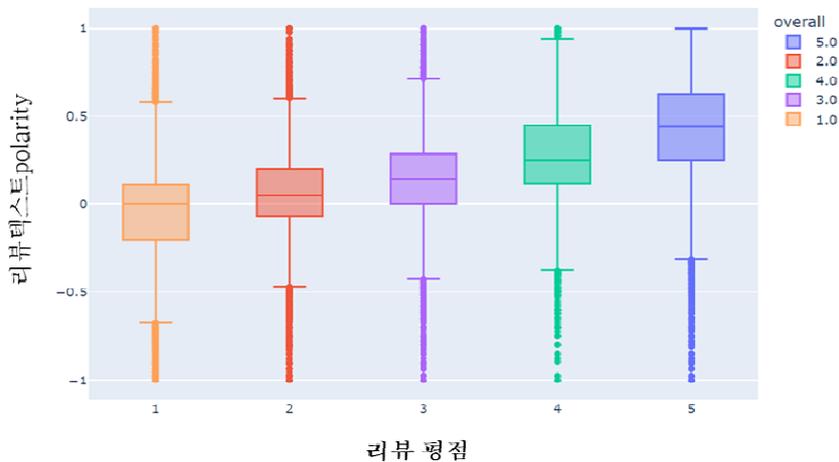
[그림 2] 연구모델

한 후 리뷰 텍스트의 임베딩을 위해 전처리를 하였다. 리뷰 텍스트의 전처리 과정에서는 먼저 모든 문자를 소문자로 바꾸고 구두점과 숫자를 제거하고 토큰화(tokenization)하여 불용어(stopwords)를 제거하였다. 또 표제어를 추출(lemmatization)하여 길이가 2이하인 단어는 제거하고 토큰을 문장으로 다시 결합하여 데이터 파일을 만들고 CV (Countvectorization), TF-IDF, Word2Vec, doc2vec 등의 임베딩 방법을 사용하여 벡터화(vectorization)하였다.

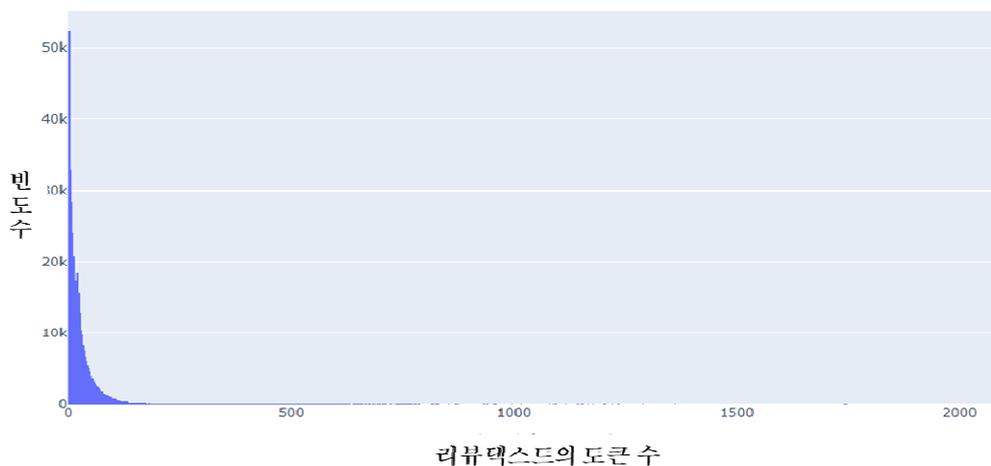
벡터화한 리뷰에 대해 텍스트블롭(TextBlob)을 사용하여 극성(polarity)을 구한 결과 [그림 3]과 같이 리뷰 평점이 높을수록 리뷰 텍스트의 극성(polarity) 평균이 높았다. 리뷰 텍스트를 토큰화 했을 때 토큰의 길이 분포는 [그림 4]와 같이 왼쪽에 극단적으로 치우쳤으며 오른쪽 꼬리가 급하게 떨어짐을 알 수 있다. 이를 통해 리뷰 길이의 분산이 크지

않으며 리뷰를 짧게 쓴 사람이 길게 쓴 사람보다 많음을 알 수 있다.

리뷰 텍스트를 긍정과 부정으로 나누기 위해 극성(polarity)이 1일 때 긍정, 0.5일 때 중립, 0이하 일 때 부정으로 정의한 결과 중립에 절반 이상(55.6%)의 데이터가 분포하였으며 긍정 대 부정 반응의 비율은 긍정 리뷰(24.2%)가 부정 리뷰(13.1%)보다 1.8배 많았다. 긍정 및 부정 리뷰 샘플을 각각 10개씩 뽑은 결과 <표 1> 및 <표 2>와 같이 긍정보다 부정 리뷰에서 문장이 긴 경우가 더 많았다. 즉 부정적인 리뷰를 쓰는 사람이 글을 더 길게 쓰는 경향이 있다고 추측할 수 있다. 이는 조용희 외(2018)의 연구에서 영화에 대해 만족스러울 경우보다 불만족스러운 경우 즉 평점이 낮을 때 영화 리뷰의 길이가 길었던 것과 일치하는 결과이다.



[그림 3] 리뷰 텍스트의 평점에 따른 극성 분포



[그림 4] 토큰화된 리뷰 텍스트의 길이 분포

<표 1> 긍정 텍스트 리뷰 샘플

 Tokenized review

- excellent
 - Beautiful!
 - Great price!
 - The best shop from I can get any color any time I want love the delivery time and material,
 - used this jacket while traveling to Dublin. Great purchase!
 - I love it!!!! Thank you!!! It looks just like the dress and it fits perfectly!!!
 - Excellent quality
 - Perfect for my newborn baby, born in the summer and who loves spending time outdoors.
 - The best price for this model! Delivered as a set, not just a frame. Excellent.
 - Beautiful! Great buy!!
-

<표 2> 부정 텍스트 리뷰 샘플

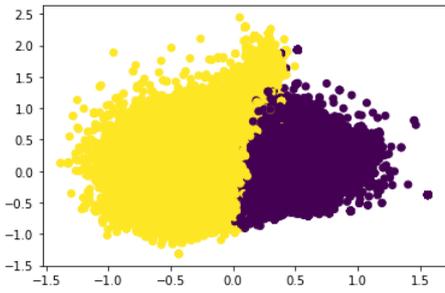
 Tokenized review

- HIDEOUS!! SO SO UGLY!!! Very boxy cut, no shape to it. It looks like a mechanics jumpsuit from a Nascar pit stop. Not denim/chambray fabric AT all; it's like a cheap blue cotton that's not even soft. Save your money
 - did not fit well.
 - Too tight, and the belt hoops were at hip level instead of normal waist level.
 - Maybe others bought from a different vendor but mine is awful. The material is not cotton like or mixed as stated and too small. I am 5'4" and it's too short also I fit in size 6 pants and the leg opening is too small. The worst is the pocket placement, might as well not have any pockets. The pockets are a good 8" from the waistband. Not a good purchase, run away.
 - These seem to do the trick in the pool, but I'm not super psyched about how small they run or how tight they seem around the waste and legs of my son. They leave deep red marks after every pool session and he outgrows them super fast because they do run so small.
 - Weird fabric, cape was smaller than in the pic, no belt
 - I bought this as a gift for my sister who's obsessed with butterflies and she loves it
 - It runs small. Is very thin. I'll likely never wear it because the quality is not what I expected.
 - Way too small. Must have been mismarked. Immediately sent it back.
 - Way to big and crazy long.
-

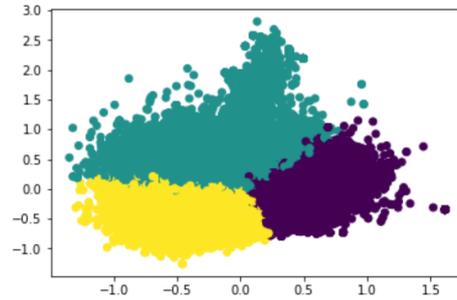
4.2 군집분석에 의한 패션상품 리뷰 분류

리뷰텍스트를 분류하기 위해 Word2Vec에서 20차원으로 데이터를 임베딩한 후 k-평균 군집화를 했다. 그 결과 [그림 5]~[그림 8]과 같이 PCA를 이용한 산점도(scatter plot)에서 군집 수가 2와 3일 때 가장 잘 분리되었으며 4와 5일 때는 잘 분리되지 않아 총 설명 분산을 높이면서 집단 간의 구별이 잘 되도록 집단수를 3개로 선택하였다. 리뷰의 텍스트벡터를 3개의 집단으로 군집화(clustering) 했을 때 각각 한 단어, 두 단어, 세 단어에 대해 n-gram을 10개씩 구해서 집단 별 특성을 살펴보았다. 그 결과 두 단어를 추출했

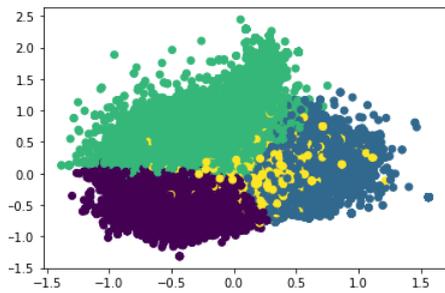
을 때 가장 해석 가능한 군집의 특성을 볼 수 있었으며 각 집단의 top-2gram을 10개씩 뽑아본 결과는 <표 3>과 같다. 집단1은 사이즈와 맞음새에 대한 내용이 많아 'size & fitting'으로 명명했다. 집단2는 품질과 가격에 대한 것이 많았는데 싼가격, 비싼가격, 우수한 품질, 가격대비 품질 등 가격과 품질 등에 대한 긍정적인 댓글과 부정적인 댓글이 다 있었으며 'quality & price'로 명명했다. 집단 3은 look well, look great, look good, like picture와 같은 외형에 대한 댓글이 많아 'appearance'로 명명했다. 군집화한 3개 집단에서 n-gram으로 뽑은 대표 단어는 [그림 9]~[그림 11]과 같다. 각 군집 내의 대표 단어들은 의미가 서로 연결



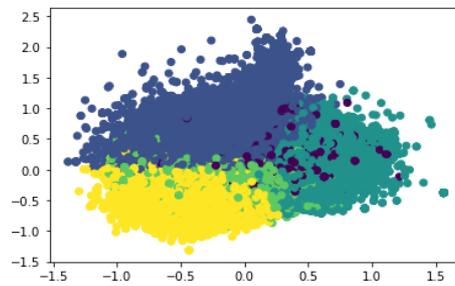
[그림 5] 군집분석 결과 (n=2)



[그림 6] 군집분석 결과 (n=3)



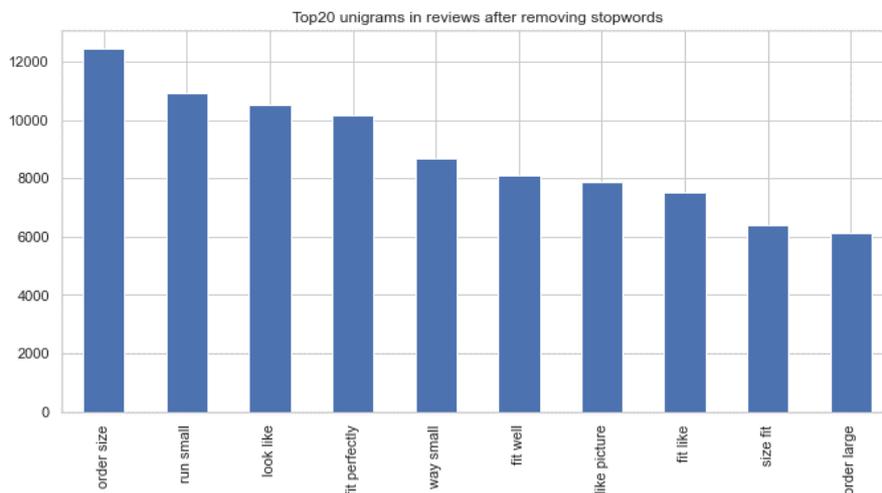
[그림 7] 군집분석 결과 (n=4)



[그림 8] 군집분석 결과 (n=5)

〈표 3〉 패션상품 리뷰 유형별 n-gram

cluster	n	Mean rank	2-gram
Group1. Size & Fit	284,245	3.33	order size, run small, look like, fit perfectly, way small, fit well, like picture, fit like, size fit, order large
Group2. Quality & Price	242,934	4.83	good quality, fit great, great quality, look great, daughter love, we'll make, great price, love love, fit perfectly, fit perfect
Group3. Appearance	353,308	3.73	look like, well make, good quality, look great, really like, year old, fit well, highly recommend, look good, like picture

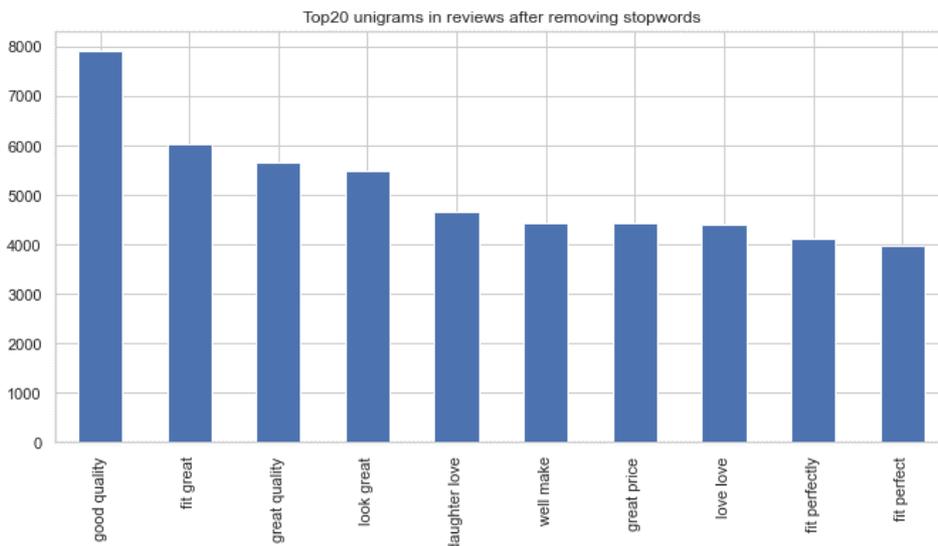


[그림 9] n-gram of the 'size & fit' group

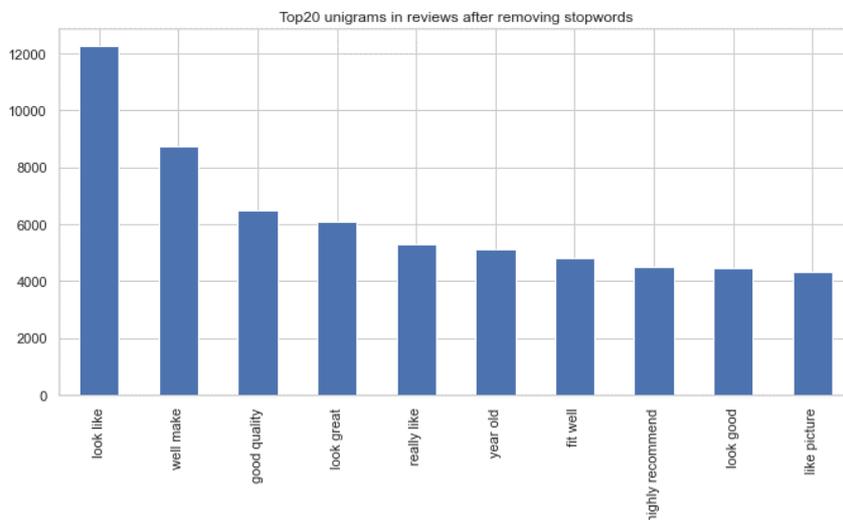
되었으며, 그렇지 않은 몇몇 단어들은 k-평균 군집화 실행 시의 노이즈로 볼 수 있다.

온라인 리뷰 분류의 유효성을 평가하기 위해 실루엣스코어(silhouette score), 일원분산분석(ANOVA) 및 사후검정(pariwise tukeyhsd)을 하였다. 실루엣스코어는 리뷰 텍스트의 k-평균 군집화 분석 결과 성능을 평가하기 위해 구하였으며 그 결과 [그림 12]~[그림 15]에서와 같이 군집이 3일 때 모든 집단이 평균을 넘고 분산이 치우치지 않았으며 실루엣스코어가 0.805로 가장 높았으며 집단 간의 크기도 균형을 이루었다.

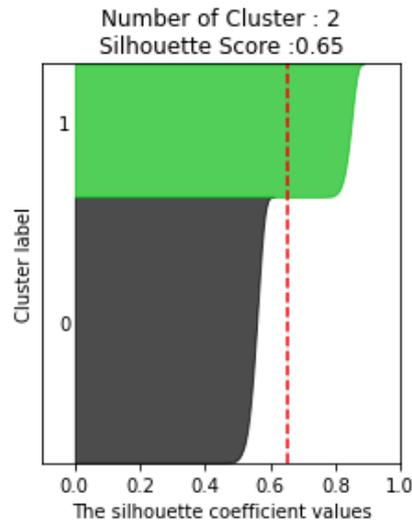
리뷰 텍스트를 군집화한 결과 나온 3개의 유형별로 평점에 차이가 있는지를 보기 위해 ANOVA 분석을 하였다. 각 유형별 평균은 집단1(‘Size & Fit’ group)에 속하는 284,245개 리뷰의 평점 평균은 3.33 이었고, 집단2(‘Quality & Price’ group)에 속하는 242,934개 리뷰의 평점 평균은 4.83 이었으며, 집단3(‘Appearance’ group)에 속하는 355,308개 리뷰의 평점 평균은 3.73로 나왔다. 이들 평균 점수의 유형 간 차이를 보기 위해 ANOVA 분석을 실시한 결과 F값이 94166.24(p=0.0)으로 유의하게 나와 3개 집단의 평점 평균에 유의한 차이가 있음을 알 수 있었다. 또 두 집



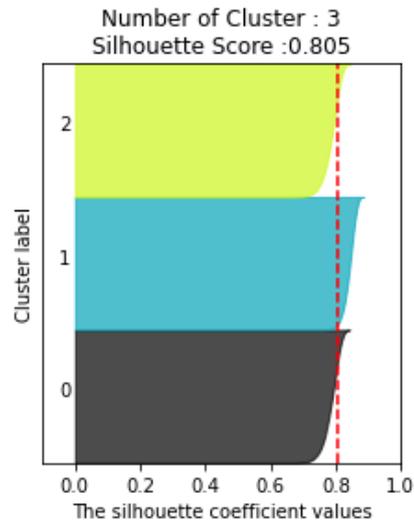
[그림 10] n-gram of the ‘quality & price’ group



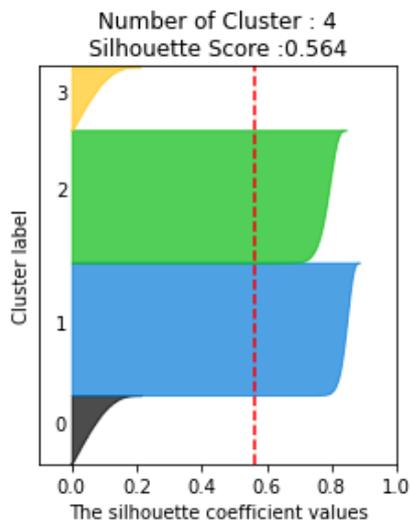
[그림 11] n-gram of the ‘appearance’ group



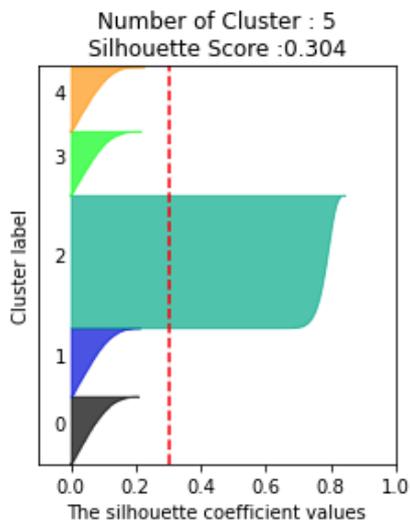
[그림 12] 실루엣스코어 (n=2)



[그림 13] 실루엣스코어(n=3)



[그림 14] 실루엣스코어(n=4)



[그림 15] 실루엣스코어(n=5)

단 간 평점 평균 값의 차이가 유의한지 검정하기 위해 사후검정으로 pairwise tukeyhsd 사후검정을 실시하였으며 그 결과 두 집단 간의 차이가 모두 유의하게 나왔다 ($\alpha=0.05$).

3. 패션상품 리뷰의 감성 예측 모델링

패션상품 리뷰의 긍·부정을 예측하기 위해 머신러닝과 딥러닝 모델을 활용하여 감성예측모델을 구축하고 각 알고리즘의 성능을 비교 분석하였다. 이를 위해 먼저 아마존 리뷰 텍스트 데이터의 임베딩을 위해 전처리를 하였다. 모든

문자를 소문자로 바꾸고 구두점과 숫자를 제거하여 토큰화(tokenization)하고 불용어(stop words)를 제거하였다. 또 표제어 추출(lemmatization)을 하였고 길이가 2이하인 단어는 제거하고 나서 마지막으로 토큰을 문장으로 다시 결합하여 데이터 파일을 만들고 CV(Countverization), TF-IDF, Word2Vec, post sequence 기법을 사용하여 임베딩 하였다.

이렇게 임베딩한 텍스트 벡터로부터 머신러닝과 딥러닝의 다양한 모델을 활용하여 감성 예측모델을 구축하고 성능을 비교하여 최종 모델을 선정하였다. 리뷰 텍스트로부터 평점을 예측하는 모델을 구축하기 위해 먼저 학습데이터와 테스트데이터의 비율은 8:2로 하였다. 온라인 리뷰로부터

평점을 예측하는 알고리즘을 구축하고 성능을 평가하기 위해 CV, TF-IDF, Word2Vec, post-sequence 로 각각 임베딩한 값을 SVC, 로지스틱회귀(logistic regression), 랜덤 포레스트구분자(random forest classifier), 배깅구분자(bagging classifier), multinomial NB 등의 머신러닝 알고리즘을 사용하여 모델링하였으며 또한 CNN과 bi-directional LSTM와 같은 딥러닝 알고리즘을 사용하여 모델링하였다. 그 결과 모델 정확도는 <표 4>와 같이 SVC와 logistic regression에서는 TF-IDF 임베딩 방법에서 정확도가 0.924로 가장 높았다. 한편 random forest classifier와

multinomial NB 모델에서는 CV 임베딩 했을 때 정확도가 0.910과 0.93으로 가장 높았다. 또한 bagging classifier에서는 Word2Vec으로 임베딩했을 때 정확도가 0.890으로 가장 높았다. 전체적으로 보면 머신러닝 알고리즘 모델에서는 CV와 TF-IDF 임베딩 방법이 Word2Vec보다 우수한 것으로 보였으며 각 알고리즘에 따라 적합한 임베딩 방법에 차이가 있음을 알 수 있다. 전체적으로는 post-sequence 임베딩의 bidirectional LSTM이 정확도 0.94로 가장 높게 나타나 온라인 리뷰의 감성 예측모델로 LSTM의 우수성을 보여주었다.

<표 4> 감성 예측 모델 정확도

algorithm		embedding			
		CV	TF-IDF	Word2Vec	post- sequence
Machine learning	SVC	0.909	0.924	0.886	-
	logistic regression	0.921	0.924	0.886	-
	random forest classifier	0.910	0.909	0.901	-
	bagging classifier	0.875	0.874	0.890	-
	multinomial NB	0.93	0.92	-	-
Deep learning	CNN	-	-	0.696	-
	Bidirectional LSTM	-	-	-	0.94

<표 5> LSTM 모델 개요

Feature Extraction	review text	pad-sequence
	sentiment(positive, negative) index	text-to-sequence
Data split	train data : test data	8:2
Embedding	sequential	Dimension 64
Modeling	model	bidirectional LSTM (units = 100)
	embedding	input_dim: 5000(MY VOCAB) embedding dim: 64 (dropout rate=0.5)
	dense	dimension 3
	activation	Softmax
	optimizer	adam
	loss	sparse-categorical_crossentropy
	epochs	10

<표 6> LSTM sequential 모델의 파라미터

Layer	Output shape	Number of parameters
Embedding	(None, None, 64)	320,000
Dropout	(None, None, 64)	0
Bidirectional	(None, 200)	132,000
Dense	(None, 6)	1206

LSTM모델의 개요는 <표 5>와 같다. 데이터는 리뷰평점이 3,4로 중간인 것과 결측치를 제외하고 긍정과 부정을 각각 70,000개씩 총 70,000개가 사용되었다. 모델학습과 테스트를 위해 데이터를 훈련데이터와 평가 데이터를 8:2로 나누어 112,000개의 훈련데이터와 28,000개의 평가데이터를 분석에 사용하였다. 리뷰 텍스트를 불용어 처리한 후 pad-sequence로 토큰화 하였다. 제일 긴 리뷰는 624, 제일 짧은 리뷰는 1개의 단어로 이루어졌으며, 사용된 총 단어 수는 22,290개였다. 이때 padding을 사용하여 데이터 길이를 200으로 균등하게 길이를 맞추었다. 또 긍정 부정의 인덱스를 text-to-sequence로 토큰화 하였다. 모델은 bi-directional sequential LSTM을 사용하였으며 하이퍼 파라미터는 <표 5>와 같이 단어 수는 5000, 임베딩 차원은 64, LSTM 셀의 규모는 100, 기사의 길이는 200, 학습용 데이터의 비율은 0.8, 반복 학습 수(epochs)는 10으로 하였다. 모델은 Dencs=6, activation은 softmax, optimizer는 adam, loss함수는 sparse-categorical crossentropy를 사용하였다. 모델 요약은 <표 6>과 같다. 모델의 정확도는 0.94 이고 loss는 0.093으로 나왔다. 혼동행렬은 $\begin{bmatrix} 13246 & 755 \\ 850 & 13149 \end{bmatrix}$ 으로 나왔다.

각 세분 집단 별 모델 평가를 위해 아마존 패션 제품 리뷰 8개를 랜덤으로 가져와서 성능을 테스트한 결과 세 집단에서 모두 8개 샘플 다 긍·부정을 정확하게 예측하였다. 또한 본 연구에서 사용한 아마존 패션 데이터셋에 리뷰 텍스트 대신 총평을 이용하여 리뷰 텍스트의 유형분류의 재현성을 테스트 하였다. 총평을 임베딩하여 LSTM 알고리즘으로 유형 예측 모델을 적합시킨 결과 0.79의 정확도가 나왔다.

V. 결론 및 제언

본 연구의 목적은 온라인 쇼핑물에서 패션상품 구매 후 고객이 작성한 리뷰 텍스트를 군집분석으로 유형화하고 리뷰 텍스트로부터 감성을 예측하는 모델을 제안하고 성능을 평가하는 것이다. 이를 위해 먼저 아마존 온라인 패션상품 리뷰 데이터에서 리뷰 텍스트를 임베딩하여 군집화함으로 리뷰를 의미론적으로 분석하였고, 리뷰에서 추출한 벡터로 고객 평점의 긍·부정을 예측하는 감성 예측 모델을 만들었다. 패션상품 구매 리뷰를 의미론적으로 분석하기 위해 리뷰 텍스트의 임베딩 벡터를 k-평균 군집분석으로 세 개의 유형으로 분류하였다. 패션상품 구매리뷰는 “Size & Fit”,

“Quality & Price”, “Appearance”의 세 개 유형으로 분류할 수 있었으며 각 군집 별로 고객평점의 평균은 “Quality & Price” 그룹이 가장 높았고, “Size & Fit”와 “Appearance” 그룹은 낮게 나왔다. 이는 온라인에서 패션상품을 구매할 때 가성비와 이점이 크지만 상대적으로 사이즈나 맞춤새 및 착용 후 외모에서 불만족할 수 있는 여지가 많음을 반영한다고 볼 수 있어 온라인에서 패션상품을 직접 보거나 착용하지 못하고 구매해야 하는 것의 한계점을 말해주며 동시에 가성비를 증시하는 아마존 쇼핑물의 특성 및 고객의 특성을 반영한다고 볼 수 있겠다. 패션상품 구매 리뷰 분류의 유효성을 평가하기 위해 실루엣 스코어, ANOVA, pairwise tukeyhsd 사후검정을 구하고, 분류 집단 별로 평점 예측 모델의 성능을 평가하였다. 결론적으로 패션상품 구매 리뷰는 bi-directional LSTM 알고리즘으로 분석할 때 긍·부정의 감성을 효율적으로 예측할 수 있으며, 리뷰의 의미론적 분석에 따라 “Size & Fit”, “Quality & Price”, “Appearance”의 3가지 유형으로 분류할 수 있다. 한편 감성 예측 모델을 구성함에 있어 임베딩 방법으로 CV, TF-IDF, Word2Vec와 SVC, logistic regression, random forest classifier, bagging classifier, multinomial NB, CNN, bi-directional sequential LSTM 알고리즘을 결합하여 성능을 테스트하였다. 성능 평가 결과 bi-directional sequential LSTM 모델의 성능이 가장 우수한 것으로 나타났으며 군집분석으로 유형화한 리뷰 분류 시스템에 적용할 때 만족할만한 결과를 얻었다. 본 연구에서 증명된 LSTM알고리즘의 우수성은 Agarap(2018)의 연구에서 LSTM이 RNN에서 성분이 사라진다는 단점을 보완하기 때문에 리뷰 텍스트에서 단어의 문맥을 보다 잘 파악하는데 좋은 성능을 낸다고 한 것과 일치하는 결과이다.

본 연구는 온라인 패션상품 구매 리뷰의 의미 및 감성 분석에 의한 분류 모델을 제시함으로 온라인 패션상품 구매와 관련된 연구에 필요한 기초자료를 제공하고자 하였다. 온라인 상품구매 리뷰에는 다양한 내용이 섞여 있는 경우가 많아 자세히 읽어 보기 전에는 대표적인 내용을 추출하거나 긍·부정의 감성을 파악하기 힘들며 때때로 거짓 평점으로 인해 리뷰의 감성과 평점이 일치하지 않는 경우도 있다. 이러한 이유로 풍부한 양의 온라인 상품구매 리뷰 데이터가 유용하게 사용되기 힘들었던 점을 고려할 때 본 연구와 같이 리뷰를 유형화하고 bi-directional sequence LSTM 알고리즘을 사용하여 구매 리뷰에서 높은 정확도로 감성 예측이 가능함을 보여 준 것은 상품 기획 및 마케팅에서 실무적 의의가 크다고 하겠다. 패션상품 리뷰의 감성 예측 모델을 통한 실시

간 모니터링으로 해당 상품에 대한 긍·부정 감성을 조기에 예측하는 것이 가능하며 이에 따라 효율적인 수요예측 및 재고 관리가 가능하다. 또 평점과 댓글의 감성이 일치하지 않을 때 평점만으로 정확히 파악할 수 없는 소비자 만족도를 댓글을 통해 직접 예측함으로써 효율적인 대응이 가능하다.

본 연구는 패션 구매 리뷰를 의미론적 분석에 의해 유형화함으로써 패션상품 구매와 관련된 소비자 관심 영역 및 만족 불만족의 결정요인을 파악할 수 있었다. 패션상품 리뷰의 의미론적 유형화 결과 “Size & Fit”, “Quality & Price”, “Appearance”로 분류되었다. 이는 온라인 패션상품 구매 후 평가 기준으로 사이즈와 핏, 품질, 가격, 외형 등이 중요하다는 것을 말해준다. 특히 사이즈와 맞춤새에 대해 언급한 사람들의 평점평균이 가장 낮게 나온 것은 온라인 패션 쇼핑에서 피팅 부분에 아직도 해결되어야 문제가 많다는 것을 의미한다. 또한 패션 구매 리뷰가 이렇게 세 영역으로 집단화 할 수 있다는 것은 온라인 구매 후 평가기준에 따라 가성비를 중시하는 소비자, 사이즈와 피팅을 중시하는 소비자, 미적인 부분을 중시하는 소비자로 유형화할 수 있다는 의미가 되기도 한다.

소비자 구매의사결정 시 평가 기준에 대한 연구는 많으나 본 연구에서는 구매 후 평가 기준을 제시했다는 데 의미가 있다. 온라인 리뷰는 그 성격상 개인의 견해가 표현되기 때문에 유동성이 커서 이러한 유동성을 개념화하는 것이 필요하다는 점을 감안할 때 온라인 리뷰의 유형화 연구는 아직도 해결해야 할 부분이 많은 도전적인 연구 주제다(Kawaf & Istanbulluoglu, 2019). 이에 본 연구에서는 온라인 리뷰 유형화의 재현가능성을 보완하기 위해 실루엣 스코어, ANOVA, 사후검정과 같은 통계수치를 구해 군집 수의 적정성을 검증하고자 하였고 토픽분석을 통해 유형화한 군집의 유동성을 개념화하고자 하였다. 패션상품 영역에서의 온라인 리뷰 유형화를 위한 본 연구의 시도는 실무에서 유동성이 높은 패션 구매 리뷰를 조기에 파악하고 관리하는데 유용하게 쓰일 수 있을 것이다.

본 연구에서는 아마존 데이터셋을 이용하여 영어로 된 구매 리뷰를 분석하였는데 향후 한국어로 된 리뷰를 대상으로 본 연구 주제를 적용시켜 볼 필요가 있다. 본 연구에서 k-평균 군집화에 의해 패션상품 리뷰의 유형을 도출했음에도 불구하고, 개별 데이터의 유형을 분류하는 정답이 없기 때문에 분류의 효율성을 직접적으로 증명하기 힘들기 때문에 향후 연구에서 패션상품 리뷰의 유형을 라벨링 할 수 있는 리뷰 프로토콜을 개발할 필요가 있겠다. 또한 본 연구에서는 리뷰텍스트를 Word2Vec으로 임베딩하여 k-평균 군집화

로 군집을 추출하였는데 품사별로 단어를 추출하여 군집화하거나 주제별 솔림을 균등화 하는 등 다양한 군집화 방안을 모색할 필요가 있겠다. 또 구매 리뷰 텍스트와 더불어 한 줄 총평을 이용한 연구를 상품 정보 및 소비자 정보와 연동시켜 연구한다면 온라인 소비자 행동을 이해하는데 도움이 될 것이다. 본 연구에서는 감성 예측 모델의 검증을 위해 다양한 임베딩 방법과 알고리즘을 사용하여 우수한 성능을 증명하였는데 향후 연구에서는 여러 임베딩 방법을 합성하여 가중치를 주던가 전이학습 등의 방법을 사용할 수 있을 것이며 분류 모델의 성능을 평가할 수 있는 다양한 방안을 개발할 수 있을 것이다. 또 다양한 질적, 양적 연구방법을 보완한 리뷰 분류 시스템과 실무에서의 빠르고 보다 세분화된 리뷰 분류를 위한 프로토콜을 만드는 것이 필요하겠다.

주제어: 온라인리뷰, 패션상품, 감성 예측, k-평균군집화, LSTM

REFERENCES

- 강형석, 양장훈(2019). Word2Vec 모델로 학습된 단어 벡터의 의미 관계 분석. *정보과학논문지*, 46(10), 1088~1093.
- 권혜진(2020). *숙박공유서비스의 온라인 리뷰를 활용한 전반적 만족도 결정 요인 분석*. 부경대학교 박사학위논문.
- 김성범(2020). 제품, 서비스, 융합제품서비스의 소비자 니즈 비교 분석: 아마존 온라인 리뷰를 중심으로. *한국콘텐츠학회논문지*, 20(7), 316-300.
- 김장영, 박은혜(2017). 빅데이터 분석을 이용한 이리닝 수장 후기 분석. *한국정보통신학회논문지*, 21(2), 423-428.
- 김준겸, 최은솔, 윤수현, 이유빈, 김동환 (2021). 치킨 리뷰의 이면: 텍스트 마이닝을 통한 리뷰의 탐색적 분석을 중심으로. *한국콘텐츠학회논문지*, 21(11), 30-40.
- 서현진, 이규혜(2013). 온라인 구매후기의 방향성과 평가 내용이 패션상품에 대한 소비자 태도에 미치는 영향. *복식문화연구*, 21(3), 440-451.
- 윤소영, 윤성대(2020). 상품 리뷰 감성분석을 이용한 아이템 기반 협업 필터링 추천 기법. *한국정보통신학회논문지*, 24(8), 970-977.

- 이동엽, 조재춘, 임희석(2017). 워드 임베딩을 이용한 아마존 패션 상품 리뷰의 사용자 감성 분석. *한국융합학회논문지*, 8(4), 1-8.
- 이윤주, 소현정, 광기영, 안현철(2020). 텍스트 마이닝 기법을 활용한 웹툰 댓글 분석: 네이버 베스트 도전 웹툰을 중심으로. *한국컴퓨터정보학회 학술발표논문집*, 28(2), 219-222.
- 이정원, 박철(2020). 온라인 리뷰 특성이 영화 매출액에 미치는 영향: 한국과 미국의 비교. *경영학연구*, 49(6), 1575-1594.
- 이지현, 정상형, 김준호, 민은주, ... 김종우(2020). 온라인 리뷰 분석을 통한 상품 평가 기준 추출: LDA 및 k-최근접 이웃 접근법을 활용하여. *지능정보연구*, 26(1), 97-117.
- 이호근, 광현(2013). 온라인 소비자 리뷰의 효과에 영향을 미치는 요인에 대한 고찰. *정보화정책*, 20(3), 3-17.
- 정정숙(2022). 통계청, 온라인 패션 거래액 9.2% 성장한 49조7192억원, <https://www.ktnews.com/news/articleView.html?idxno=122454>에서 인출
- 조용희, 박이슬, 김혜진(2018). 빅데이터를 활용한 영화 흥행에 따른 리뷰길이 변화. *한국콘텐츠문화학회논문지*, 18(5), 367-375.
- 최영현, 이규혜(2020). 겨울왕국2의 콜라보레이션 패션제품에 대한 소비자 리뷰 -의미 네트워크와 감성분석-. *복식문화연구*, 28(2), 265-284.
- 채희주, 박수현, 고은주(2016). 서포터즈의 온라인 리뷰 유형에 따른 패션 브랜드의 온라인 인상형성과 구전효과에 대한 연구. *한국의류산업학회지*, 18(1), 15-26.
- 한기향(2021). 반팔 티셔츠의 온라인 리뷰 분석에 관한 융합적 분석 연구. *한국과학예술통합학회*, 39(4), 541-555.
- 허지욱(2018). Word2Vec을 이용한 한국어 단어 군집화 기법. *한국인터넷방송통신학회 논문지*, 18(5), 25-30.
- Agarap, A. F.(2018). Statistical analysis on E-commerce review, with sentiment classification using bidirectional recurrent neural network(RNN). *arXiv*, 2018; arXiv:1805.03687
- Almiron-chamadoira, P.(2018). Online reviews as a genre: A semiotic analysis of Amazon.com 2010-2014 review on the categories 'Clothing' and 'Electronics'. *DTUC'18: Proceedings of the 1st International Conference on Digital Tools & Uses Congress*(p.1-4), Paris, France.
- Chong, A. Y. L., Li, B., Ngai, E. W. T., Ch'NG E., & Lee, F.(2016). Predicting online product sales via online reviews, sentiments, and promotion strategies: a big data architecture and neural network approach. *International Journal of operations & Production Management*, 36(4), 358-383.
- Duan, W., Gu, B., & Whinston, A. B.(2008). Do online reviews matter? -An empirical investigation of panel data. *Decision Support Systems*, 45(4), 1007-1016.
- Elmurngi, E. I., & Gherbi, A.(2018). Unfair review detection on Amazon reviews using sentiment analysis with supervised learning techniques. *Journal of Computer Science*, 14(5), 714-726.
- Hilal, A. K., & Astuti, S. P.(2022). The role of online customer reviews in increasing impulsive purchase of fashion products online with customer trust as a mediator. *Journal of Management and Islamic Finance*, 2(2), 310-323.
- Kawaf, F., & Istanbuluoglu, D.(2019). Online fashion shopping paradox: the role of customer reviews and facebook marketing. *Journal of Retailing and Consumer Studies*, 48, 144-153.
- Kim, S., & Kim, Y. K.(2018). The interplay of regulatory focus and temporal distance on consumer evaluation of online reviews. *Clothing and Textiles Research Journal*, 36(3), 135-150.
- Korfiatis, N., Stamolampros, P., Kourouthanassis, P., & Sagiadinos, V.(2019). Measuring service quality from unstructured data: a topic modeling application on airline passengers' online reviews. *Expert Systems with Applications*, 116, 472-486.
- Lee, C. K. H., Tse, Y. K., Zhang, M., & Ma, J.(2019). Analysing online reviews to investigate customer behavior in the sharing economy: the case of Airbnb. *Information Technology & People*, 33(3), 945-961.
- McKinney, E., & Shin, E.(2016). Exploiting criteria consumers use in evaluating their online formal wear rental experience: A content analysis of online reviews. *Clothing and Textiles Research*

- Journal*, 34(4), 272-286.
- Ni, J.(2018). Amazon review data, Clothing shoes and jewelry, UCSD. Retrieved from <https://nijianmo.github.io/amazon/index.html>
- Praveenraj, D. D. W., Subramani, K., Murugan, R. T., & Manikandan, S.(2021), Exploring consumer reviews for men's fashion accessories in online purchase platforms using sentiment analysis. *Revista GEINTEC*, 11(2), 676-685.
- Pumamawati, I. P. S., Putra, I M. W., & Wahyuni, N. M.(2022). The effect of brand image on online purchase decision with mediation word of mouth on fashion products in Denpasar-Bali Province. *Journal Ekonomi & Bisnis JAGADITHA*, 9(1), 90-98.
- Reich, T., & Maglio, S. J.(2020). Featuring mistakes: the persuasive impact of purchase mistakes in online reviews. *Journal of Marketing*, 84(1), 52-65
- Wang, Z., & Chen, Q.(2020). Monitoring online reviews for reputation fraud campaigns. *Knowledge-Based Systems*, 195, 105685.
- Yang, L., Li, Y., Wang, J., & Sherratt, R. S.(2020). Sentiment analysis for e-commerce product review in Chinese based on sentiment lexicon and deep learning. *IEEE Access*, 8, 23522-23530.
- Zhang, J., & Liu, R. R.(2018). The more the better? Exploring the effects of reviewer social networks on online review. *Journal of Marketing Management*, 35(17-18), 1667-1688.

Received 15 August 2023;

1st Revised 1 October 2023;

Accepted 4 October 2023