

CNN 기반 리뷰 유용성 점수 예측을 통한 개인화 추천 서비스 성능 향상에 관한 연구*

이청용

경희대학교 빅데이터응용학과
(leecy@khu.ac.kr)

이흥철

경희대학교 빅데이터응용학과
(lixz@khu.ac.kr)

이병현

경희대학교 빅데이터응용학과
(leebh1129@khu.ac.kr)

김재경

경희대학교 경영대학 & 빅데이터응용학과
(jaek@khu.ac.kr)

전자상거래 시장이 빠르게 성장하면서 다양한 유형의 제품이 출시되고 있으며, 이로 인해 사용자들은 구매 의사결정 과정에 많은 시간이 소요되는 정보 과부하 문제에 직면하고 있다. 따라서 사용자에게 맞춤형 제품 및 서비스를 제공해줄 수 있는 개인화 추천 서비스의 중요성이 대두되고 있다. 대표적으로 Netflix, Amazon, Google 등 세계적 기업은 개인화 추천 서비스를 도입하여 사용자의 구매 의사결정을 지원하고 있다. 이에 따라 사용자의 정보탐색 비용이 감소하는 효과가 나타났고, 기업의 매출 상승에도 긍정적인 영향을 끼치고 있다. 기존 개인화 추천 서비스 관련 연구에서 주로 사용된 협업 필터링(Collaborative Filtering, CF) 기법은 정량화된 정보를 활용하여 사용자의 선호도를 예측하였다. 그러나 정량화된 정보만을 활용하면 사용자의 구매 의도는 고려하지 못하므로 추천 성능이 저하될 수 있다는 문제점이 제기되고 있다. 이와 같은 기존 연구의 문제점을 개선하기 위해 최근에는 사용자가 작성한 리뷰를 활용한 개인화 추천 서비스 연구가 활발히 진행되고 있다. 그러나 리뷰에는 광고성 내용, 거짓 후기, 의미를 전혀 파악할 수 없거나 제품과 관련 없는 내용 등 구매 의사결정을 저해하는 요소들이 포함되어 있다. 이러한 요소들이 포함된 리뷰를 활용하여 추천 서비스를 제공하게 되면, 추천 성능이 저하되는 문제가 발생할 수 있다. 따라서 본 연구에서는 이러한 문제점을 개선하기 위해 Convolutional Neural Network(CNN) 기반 리뷰 유용성 점수 예측을 통한 새로운 추천 방법론을 제안하였다. 본 연구에서 제안하는 유용한 리뷰를 포함하는 방법론과 기존 모든 선호도 평점을 고려하는 추천 방법론을 비교한 결과, 본 연구에서 제안한 방법론이 더 우수한 예측 성능을 나타내고 있음을 확인할 수 있었다. 또한 본 연구의 결과는 리뷰 유용성에 대한 정보를 개인화 추천 서비스에 반영하면 전통적인 CF의 성능을 향상할 수 있음을 시사한다.

주제어 : 개인화 추천 서비스, 리뷰 유용성, 협업 필터링, 딥러닝, CNN

논문접수일 : 2021년 5월 21일 논문수정일 : 2021년 6월 29일 게재확정일 : 2021년 7월 12일

원고유형 : 일반논문 교신저자 : 김재경

1. 서론

최근 정보통신기술의 발전과 모바일 기기의 대중화로 인해 전 세계적으로 전자상거래 시장 규모가 급속하게 커지면서 새로운 제품들이 계

속 출시되고 있으며, 이에 따라 다양한 유형의 제품이 활발하게 거래되고 있다. 그러나 사용자들은 선호하는 제품을 선택하는 데 많은 시간이 소모되며 구매 의사결정 과정에서 어려움을 겪는 정보 과부하 문제에 직면하고 있다(Park et

* 본 논문은 교육부 및 한국연구재단의 4단계 두뇌한국21 사업(4단계 BK21 사업)으로 지원된 연구임.

al., 2012; Su and Khoshgoftaar, 2009). 특히, 사용자들은 제품을 온라인에서 간접적으로 확인하기 때문에 정보 과부하 문제는 더 크게 드러날 수 있다. 이러한 배경하에 개인화 추천 서비스의 중요성이 대두되고 있으며, Netflix, Amazon, Google 등 세계적인 전자상거래 기업들은 개인화 추천 서비스를 도입하여 사용자의 구매 의사결정을 지원하고 있다(Bennett and Lanning, 2007; Das et al., 2007; Linden et al., 2003). 이와 같은 개인화 추천 서비스 제공으로 인해 사용자의 정보탐색 비용이 감소하는 효과가 나타났으며, 더 나아가 기업의 매출 증대에도 긍정적인 영향을 끼치고 있다. 예를 들어, Netflix는 사용자가 시청한 전체 동영상의 75%가 개인화 추천 서비스를 통해 제공되고 있으며, Amazon은 전체 매출의 35%가 개인화 추천 서비스를 통해 사용자에게 추천된 제품에 의해 발생하고 있다(Lee and Hosanagar, 2019; Na and Nam, 2020).

협업 필터링(Collaborative Filtering, CF)은 개인화 추천 서비스 관련 연구에서 현재까지 가장 우수한 추천 성능을 나타내고 있는 기법으로, 사용자와 제품 간의 상호관계를 파악하여 사용자에게 맞춤형 제품을 추천한다(Kim et al., 2009; Su and Khoshgoftaar, 2009). 기존 CF 관련 연구에서는 주로 사용자가 제품에 부여한 선호도 평점, 구매 및 방문 여부 등 정량화된 정보를 바탕으로 사용자의 선호도를 예측하였다(Li et al., 2014; Park et al., 2012; Ricci et al., 2011). 그러나 정량화된 정보만을 고려하면 제품 구매 동기, 이유 등 정성적인 선호도는 고려하지 못하므로 추천 성능이 저하될 수 있다는 문제점이 꾸준히 제기되고 있다(Lu et al., 2015; Qiu et al., 2016). 최근에는 이러한 문제점을 개선하기 위해 다양한 추가 정보를 사용하는 연구들이 진행되고 있으

며, 대표적으로 사용자가 작성한 리뷰를 사용한 연구가 활발히 진행되고 있다(Cheng et al., 2018; Lei et al., 2016; Zheng et al., 2017). 대부분의 전자상거래 사이트에서는 사용자가 구매한 제품에 대해 직접 리뷰를 작성할 수 있는 기능을 제공하고 있다. Moore (2015)에 따르면, 사용자의 88%는 제품을 구매할 때 다른 사용자가 작성한 리뷰를 참고하여 구매 의사결정을 내리고 있다. 리뷰에는 사용자가 해당 제품을 구매한 이유, 제품에 대한 평가 등 구체적이고 신뢰할 수 있는 정성적인 정보가 포함되기 때문에 개인화 추천 서비스를 제공할 때 유용하게 사용될 수 있다(Qiu et al., 2016). 그러나 기존의 리뷰를 사용한 개인화 추천 서비스 관련 연구는 주로 리뷰에 대한 감성 특성을 추출하거나 리뷰에 포함된 다양한 속성을 파악하여 이를 CF에 반영하는 연구에 중점을 두었다(Srifi et al., 2020; Yun et al., 2018). 그러나 사용자가 작성한 리뷰에는 광고성 내용, 거짓 후기, 의미를 전혀 파악할 수 없거나 제품과 관련 없는 내용 등 구매 의사결정을 저해하는 내용이 포함되어 있다(Ge et al., 2019). 따라서 사용자가 작성한 리뷰가 신뢰할 수 있는지 또는 다른 사용자에게 유용한 정보를 제공할 수 있는지 등에 대한 품질을 고려하지 않고 추천 서비스를 제공하게 되면 추천 성능이 저하되는 문제가 발생할 수 있다(Hu et al., 2017).

따라서 본 연구는 정량화된 선호도 평점만을 사용하는 기존 CF 기법의 추천 성능을 개선하기 위해, 사용자의 구매 의사결정에 영향을 줄 수 있는 리뷰 유용성(Helpfulness) 정보를 개인화 추천 서비스에 반영하는 새로운 방법론을 제안하고자 한다. 최근 전자상거래 사이트를 통해 제품을 구매하는 사용자들이 증가하면서 제품에 대한 리뷰 수가 증가하고 있다. <Table 1>에서 사

〈Table 1〉 Number of Reviews Received by Amazon Best Sellers Items

Product Category	Product Name	Number of Reviews
Electronics	Amazon Fire TV Stick 4K	613,424
Books	The Midnight Library by Matt Haig	70,460
Kindle Store	These Tangled Vines by Julianne MacLean	4,460
Sports & Outdoors	Hanes Women's Jersey Short	21,897
Video Games	\$10 PlayStation Store Gift Card	224,887

용자가 리뷰를 통해 제품의 특성을 파악하기 위해 구매 의사결정 과정에 많은 양의 정보를 사용할 수 있음을 확인할 수 있다. 그러나 구매 의사결정 과정에서 사용자가 모든 리뷰를 고려하는 것은 불가능하다. 따라서 사용자는 제품 구매과정에서 참고할 수 있는 유용한 리뷰를 탐색하는데, 어려움을 겪고 있다. 이러한 문제를 해결하기 위해 Amazon은 2007년에 사용자가 작성한 리뷰가 의사결정에 도움이 되었는지 확인할 수 있는 리뷰 유용성 투표 기능을 도입하였다. 리뷰 유용성 투표수를 기반으로 리뷰의 순위가 정렬되며 가장 투표수를 많이 받은 리뷰가 목록의 상단에 표시된다(Kaushik et al., 2018). Castelli et al. (2017)의 연구에 따르면, 이와 같은 리뷰 유용성 정보는 사용자의 구매 의사결정 과정에 큰 영향을 끼치므로 개인화 추천 서비스 제공에 있어 중요한 요소로 볼 수 있다.

사용자의 구매 의사결정에 큰 영향을 끼치는 리뷰의 유용성 정보를 개인화 추천 서비스에 반영하기 위해 본 연구에서는 다음과 같은 방법론을 제안한다. 먼저, 최근 자연어 처리 분야에서 우수한 성능을 보이는 Convolutional Neural Network(CNN)를 활용하여 리뷰의 유용성 점수를 정교하게 예측할 수 있는 CHSC(CNN-based Helpfulness Score Classifier) 모델을 구축한다

(Johnson and Zhang, 2014; Wu et al., 2020). 다음으로 사용자가 작성한 리뷰에 CHSC 모델을 적용하여 유용성 점수를 정교하게 예측하고, 유용한 리뷰만을 포함하는 새로운 사용자 프로파일을 구축한다. 마지막으로 개인화 추천 서비스 관련 연구에서 가장 우수한 성능을 나타낸 CF 기법을 활용하여 본 연구에서 제안한 모델의 추천 성능을 평가한다. 본 연구에서는 리뷰의 유용성 정보를 개인화 추천 서비스에 직접 적용한다. 따라서 제품을 구매할 때 리뷰를 고려하는 사용자들의 구매 행동을 반영하므로, 개인화 추천 서비스의 추천 성능 향상에 기여할 수 있다. 또한, 개인화 추천 서비스를 제공하거나 도입을 검토하고 있는 전자상거래 사이트에 효율적인 개인화 추천 서비스 방법론에 대한 다양한 시사점을 제공할 수 있다.

본 연구의 나머지 구성은 다음과 같다. 제2장에서는 개인화 추천 서비스와 리뷰 기반의 개인화 추천 서비스에 대한 이론적 배경을 설명한다. 제3장에서는 본 연구에서 제안하는 추천 방법론에 대해 설명한다. 제4장에서는 실험 데이터, 평가 지표 및 결과에 대해 설명한다. 마지막으로 제5장에서는 결론, 연구의 한계점 그리고 향후 연구 계획에 대해 설명한다.

2. 이론적 배경

2.1. 협업 필터링(Collaborative Filtering)

개인화 추천 서비스는 사용자 선호도 평점, 구매 기록, 검색 기록 등 데이터를 사용하여 사용자에게 맞춤형 제품이나 서비스를 제공하는 것을 목표로 한다(Bennett and Lanning, 2007). 이와 같은 개인화 추천 서비스는 다양한 유형의 제품 및 서비스 속에서 구매 의사결정의 어려움을 겪는 사용자에게 편의성을 제공할 수 있다(Wang et al., 2019). 대표적으로 Netflix, Amazon, Google 등의 전자상거래 기업은 개인화 추천 서비스 도입을 통해 효율적인 이익 성과를 거두고 있다(Lee and Hosanagar, 2019; Na and Nam, 2020). 따라서 개인화 추천 서비스는 다양한 분야의 산업에서 활용되고 있으며 이와 관련된 연구들 또한 꾸준히 진행되고 있다(Lu et al., 2015; Paradarami et al., 2017). 현재 학계와 산업계에서 널리 사용되고 있는 방법으로는 CF 기법이 있다(Kim et al., 2010; Park et al., 2012; Zhang et al., 2013).

CF 기법은 Goldberg et al. (1992)에 의해 처음으로 제안된 방식이며, 특정 제품에 대해 선호도를 보인 사용자는 다른 제품에 대해서도 비슷한 선호도를 보인다는 것을 가정하고 사용자 또는 제품 간의 유사도를 기반으로 추천하는 방법이다(Lee et al., 2020; Su and Khoshgoftaar, 2009). 협업 필터링은 메모리 기반 협업 필터링(Memory-based Collaborative Filtering)과 모델 기반 협업 필터링(Model-based Collaborative Filtering)으로 분류할 수 있다. 메모리 기반 협업 필터링 방식은 사용자 기반 협업 필터링(User-based Collaborative Filtering; UBCF)과 아이템 기반 협업 필터링(Item-based

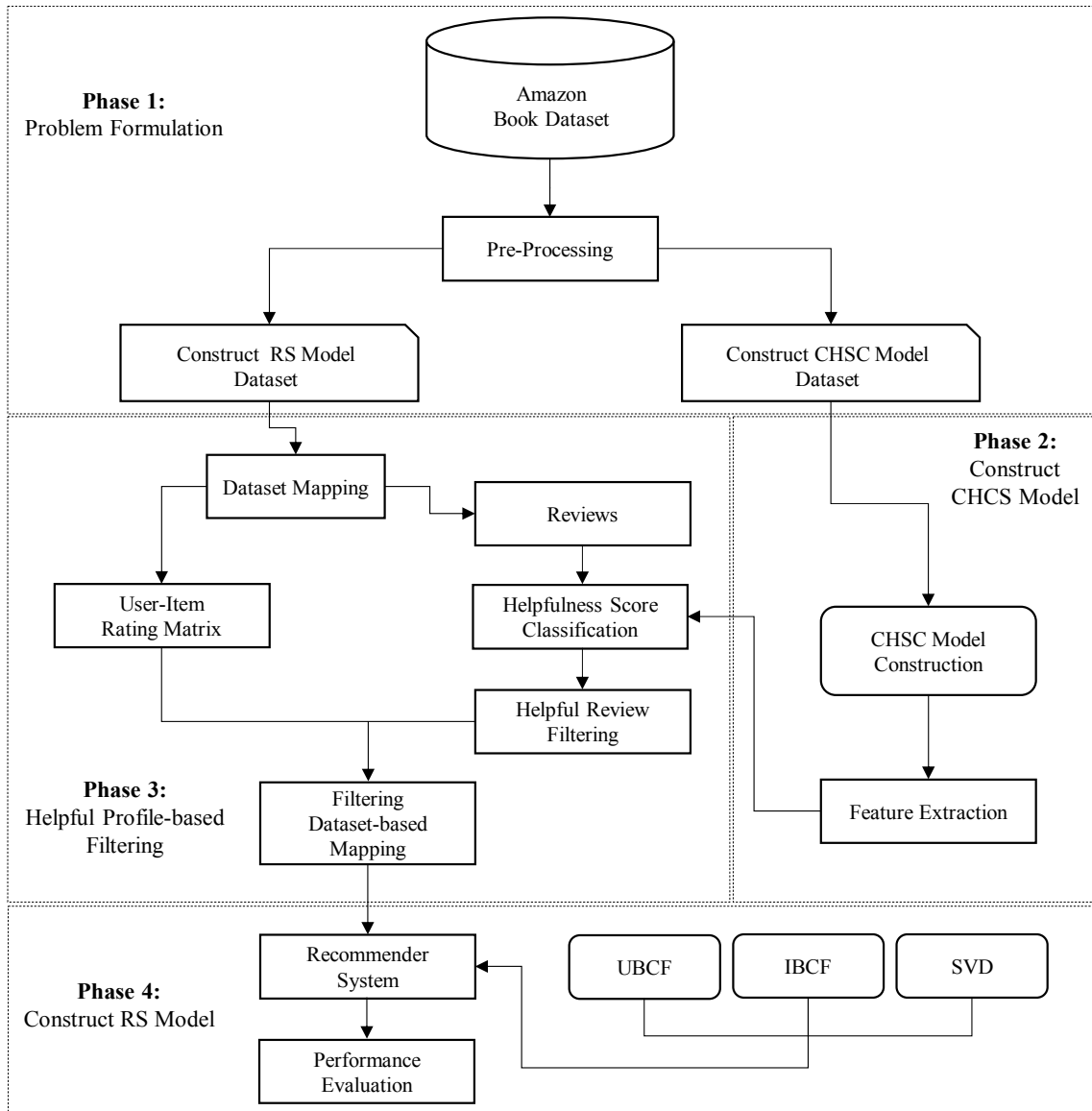
Collaborative Filtering; IBCF) 방식으로 분류된다(Al-Bashiri et al., 2018). UBCF는 추천 대상 사용자와 유사한 성향을 지닌 사용자가 구매한 제품을 추천하는 방식으로 먼저 사용자 간의 유사도를 측정하여 추천 대상 사용자와 유사한 이웃 사용자를 선정한다. 다음으로 추천 대상 사용자를 위한 제품별 선호도 예측값을 산출하여 최종적으로 선호도 예측값이 큰 제품을 사용자에게 추천한다(Elahi et al., 2016). IBCF는 사용자가 과거에 구매했던 제품과 유사한 제품을 선호할 것이라는 가정하에 추천이 이루어진다. 즉, 추천 대상 사용자가 과거에 구매한 제품을 기준으로 가장 유사도가 높은 제품을 추천하는 방식이다. 모델 기반 협업 필터링은 사용자의 선호도 평점을 예측하기 위해 CF 기법을 기반으로 하며 분류, 군집화 그리고 예측 과정에서는 데이터 마이닝 기법을 사용하여 사용자가 선호할 만한 제품을 예측하여 추천하는 방식이다(Fu et al., 2018). 즉, 과거 사용자가 제품에 부여한 평점을 가지고 기계학습이나 데이터 마이닝 모델을 활용하여 사용자가 평점을 부여하지 않은 제품에 부여할 평점을 예측하는 방식이다. 평점을 예측하는 방법으로는 대표적으로 행렬 분해(Matrix Factorization) 방식이 있다(Abdollahi and Nasraoui, 2017). 행렬 분해는 사용자와 제품에 대한 평점으로 구성된 행렬을 두 개의 행렬로 분해하는 방법이다. 가장 널리 사용되는 행렬 분해 방법으로는 특이값 분해(Singular Value Decomposition; SVD)가 있다(Bang et al., 2015). SVD는 고차원의 행렬을 저차원의 행렬로 축소하는 차원 축소의 개념이다(Mishra et al., 2015). 따라서 행렬을 차원 축소 후, 다시 원래의 행렬로 복원하면 결측치 값들이 채워지게 되며, 이를 기반으로 추천을 하는 방식이다.

기존 CF 기법 관련 개인화 추천 서비스 연구에서는 주로 정량화된 구매 및 거래 내역, 평점 정보 등의 데이터를 사용하여 추천 대상 사용자와 유사한 이웃 사용자를 선택하고, 이를 바탕으로 추천 대상 사용자의 선호도를 예측하였다. 그러나 이러한 정량적 정보는 사용자의 선호도를 구체적으로 반영하지 못하여 추천 성능이 저하되는 문제가 발생한다(Lu et al., 2015; Qiu et al., 2016). 따라서 본 연구에서는 리뷰 유용성 반영 여부에 따른 CF 기법의 성능 결과를 확인하고자 하며, 이를 통해 기존 CF 기법의 성능을 개선할 수 있을 것으로 기대한다.

2.2. 사용자 리뷰를 고려한 개인화 추천 서비스

사용자가 작성한 리뷰는 제품에 대한 정보와 사용 후기 등을 텍스트로 표현한 정성적 정보이다. 이러한 리뷰는 사용자가 제품에 대해 상세한 의견을 표현할 수 있는 중요한 매개체이다(Guy et al., 2017). 따라서 정량적 정보만을 활용한 기존 추천시스템의 한계를 보완하기 위해, 리뷰를 활용한 다양한 추천시스템이 개발되고 있다. 리뷰를 활용한 추천 시스템 연구는 2000년대 중반 학계에서 처음으로 소개되었으며, Leung et al. (2006)의 연구는 추천시스템에 사용자 리뷰를 반영한 최초의 시도라 할 수 있다. 이 연구에서는 IMDb 영화 리뷰에 감성분석을 적용하여, 사용자가 작성한 리뷰가 긍정 또는 부정인지에 대한 감성 방향과 해당 리뷰의 강도를 추정할 수 있는 확률적 평가 추론 모형을 개발하였다. 그리고 모형의 산출 값을 CF 기법에 반영한 추천시스템을 제안하였다. 따라서 추천시스템에 정성적 정보인 사용자 리뷰를 적용한 최초의 연구라는 점에

서 의의가 있다. 그러나 정량적 정보인 평점과 정성적 정보인 리뷰를 동시에 고려했다면 더 우수한 성능의 추천 알고리즘을 제안할 수 있음에도 불구하고 정량적 정보인 평점을 전혀 고려하지 않았다. 그리고 리뷰에 대한 전처리 과정에서 감성 특성만을 고려하였으며, 리뷰의 품질은 파악하지 못한 한계점이 존재한다. Garcia-Cumberas et al. (2013)은 사용자가 작성한 리뷰를 감성 분석하여 해당 사용자가 주로 긍정적인 평가 또는 부정적인 평가를 하는지에 따라 직관주의자와 비관주의자로 분류하고, 집단별로 CF 기법을 적용한 방법을 시도하였다. 따라서 IMDB 영화 리뷰를 활용한 실증 분석에서 전체 사용자 대상으로 CF 기법을 적용한 결과보다 사용자를 직관주의자랑 비관주의자로 분류하여 CF 기법을 적용하였을 때의 선호도 예측 성능이 더 높게 나타나는 것을 확인하였다. 따라서 리뷰를 집단으로 분류한 점에서 기존 연구들과 차별성이 있다. 그러나 리뷰의 내용을 추천시스템에 직접적으로 반영하지 못하였고, 이로 인해 정보 손실이 발생한다는 한계점이 있다. Zhang et al. (2014)은 리뷰를 반영한 추천 방법론을 제안하였다. 이 연구에서는 Zhou and Chaovalit (2008)의 영화 리뷰 온톨로지에 있는 총 32개의 영화에 대한 리뷰의 특성을 TF-IDF와 유사한 개념인 FF-IRF를 활용하여 도출하였다. 각 특성에 사용자의 감성 극성을 반영하여 사용자 간의 유사도 값을 산출하고, 이를 기반으로 한 CF 기법을 제안하였다. 그리고 제안한 방법론을 Yahoo Movies 데이터에 적용한 결과, 전통적인 CF 기법보다 6.18~8.24% 정도 예측 정확도가 향상됨을 확인하였다. 따라서 예측 성능이 우수하고, 메모리 기반 협업 필터링 방식을 활용하여 구현이 쉽다는 장점이 있다. 그러나 리뷰의 내용을 고려하지 않기 때문에 잘못



〈Figure 1〉 Research Flow Chart

된 극성 파악이 이루어질 수 있다는 한계점이 있다. Jeon and Ahn (2015)은 전통적인 CF 기법의 성능을 개선하기 위해, 사용자가 작성한 리뷰를 고려한 방법론을 제안하였다. 그리고 이를 스마

트폰 앱 추천시스템에 적용하여 제안한 방법론의 효과성을 검증하였다. 정성적 정보인 사용자의 리뷰는 텍스트 마이닝을 활용하여 정량화하였으며, 이를 기반으로 사용자 간의 유사도를 산

출하였다. 이때, 유사도를 산출하는 방법으로 중복으로 사용된 단어의 빈도로 도출하는 방법과 TF-IDF 가중치 합으로 도출하는 2가지 방법을 제안하고 성능을 비교하였다. 실험 결과, CF 기법에 사용자 간의 리뷰 유사도를 추가로 반영한 경우가 평점만을 고려한 전통적인 CF 기법보다 더 우수한 예측 성능을 나타냈다. 또한 TF-IDF 가중치 합으로 도출하는 방법이 중복으로 사용된 단어의 빈도만을 고려했을 때보다 더 우수한 예측 성능을 나타냈다. Hyun et al. (2019)은 사용자가 작성한 리뷰와 평점을 결합하여 CF 기법에 반영한 추천 알고리즘을 제안하였다. 이를 위해 영화 리뷰 데이터를 활용하여 도메인 맞춤형 감성 사전을 구축하고, 이를 활용하여 리뷰의 감성 수치를 도출하였다. 그리고 도출한 감성 수치를 평점과 결합하여 생성한 새로운 평점을 CF 기법에 반영하였다. 연구 결과, 제안한 방법론이 평점만을 고려하는 전통적인 CF 기법보다 더 우수한 성능을 나타내는 것을 확인하였다. 따라서 기존 연구와는 달리 리뷰와 평점을 결합한 새로운 방법론을 제안하였고, 기존 CF 기법보다 우수한 성능을 보였다는 점에서 의의가 있다. 그러나 기존 연구와 마찬가지로 리뷰의 내용 및 품질은 고려하지 않고 긍정과 부정에 대한 특성만을 반영하였다.

따라서 기존 연구에서는 리뷰의 감성 특성을 정량화하는 데 초점을 두었다. 그러나 리뷰의 내용 및 유용성을 고려하지 않은 리뷰가 추천 알고리즘에 포함되면 예측 성능을 저하하는 문제점이 있다(Cheng et al., 2018). 따라서 본 연구에서는 리뷰의 유용성 반영 여부에 따른 알고리즘 예측 성능 결과를 비교하고자 하며, 이를 통해 기존의 사용자 리뷰를 반영한 추천 서비스의 성능을 개선할 수 있을 것으로 기대한다.

3. 제안 방법론

본 연구에서는 기존 CF 기법의 성능을 개선하기 위해 사용자가 작성한 리뷰에서 유용성 점수 예측을 통해 유용한 리뷰를 포함하는 사용자 프로파일을 추천시스템에 반영하는 새로운 방법론을 제안하고자 한다. 제안한 방법론의 추천 성능을 검증하기 위해, 먼저 리뷰에 대한 유용성 점수를 예측하는 CHSC 모델을 구축한다. 이를 기반으로 사용자가 작성한 리뷰에서 유용성 점수를 예측하고, 유용한 리뷰만을 포함하는 새로운 사용자 프로파일을 구축한다. 마지막으로 개인화 추천 서비스 관련 연구에서 가장 우수한 성능을 나타낸 CF 기법을 활용하여 본 연구에서 제안한 추천 방법론의 추천 성능을 평가한다. 본 연구의 절차는 <Figure 1>과 같이 총 4단계로 구성된다.

3.1. Phase 1: Problem Formulation

첫 번째 단계에서는 제안한 방법론의 추천 성능을 평가하기 위해 리뷰 유용성 정보가 있는 데이터는 CHSC 모델 구축에 사용하며, 유용성 정보가 없는 데이터 세트는 추천 성능 평가에 사용하였다(Baek et al., 2012). CHSC 모델을 학습하고 성능을 평가하는 모델 구축용 데이터 세트에서 리뷰를 $R_1 = \{r_1, r_2, \dots, r_{n-1}, r_n\}$ 이라고 정의할 때, 각 리뷰 r_i 는 5개 특성 $[P, U, C, M, H]$ 을 포함하는 튜플(Tuple) 형태로 표현할 수 있다. 이때 P 는 제품 특성, U 는 사용자 특성, C 는 리뷰 원문(Textual), M 은 메타 데이터 특성 그리고 H 는 총 투표수(Total Votes)에 대한 유용한 투표수(Helpful Votes)의 비율로 계산되는 유용성 점수를 의미하며, H 는 0과 1 사이의 값으로 도출된다

(Liu et al., 2008; Zhang and Varadarajan, 2006). F 를 CHSC 모델 구축에 사용되는 리뷰의 특성 행렬($n \times m$)이라고 정의할 때 n 은 리뷰의 개수, m 은 리뷰 특성의 개수를 의미한다. Z 는 CHSC 모델 구축에 사용되는 모든 리뷰에 대한 분류 값으로 구성된 벡터이며, Z_i 는 리뷰 r_i 가 유용하거나 유용하지 않음을 의미한다. 따라서 유용한 리뷰와 유용하지 않은 리뷰를 분류하는 임계값(Threshold)을 각각 θ_1 과 θ_2 로 표현하며, 최종적으로 Z_i 는 수식 (1)과 같이 정의된다.

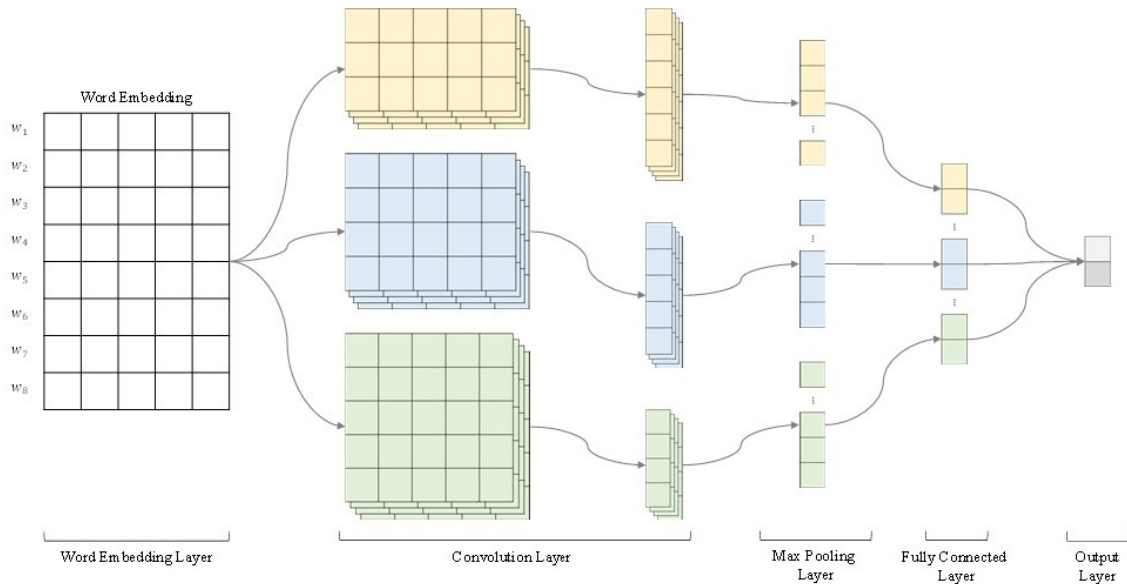
$$Z_i = \begin{cases} 1, & \text{if } H_i > \theta_1 \\ 0, & \text{if } H_i < \theta_2 \end{cases} \quad (1)$$

다음으로 추천 성능을 평가하는 데이터 세트에서 리뷰를 $R_2 = \{R_1, R_2, \dots, R_{n-1}, R_n\}$ 이라고 정의할 때, 각 리뷰 R_i 는 4개 특성 [I, B, T, S]을 포함하는 튜플(Tuple) 형태로 표현할 수 있

다. 이때 I 는 제품 특성, B 는 사용자 특성, T 는 리뷰 원문(Textual) 특성 그리고 S 는 메타 데이터 특성을 의미한다. Q 를 추천 성능 평가에 사용되는 리뷰의 특성 행렬($v \times u$)이라고 정의할 때 v 는 리뷰의 개수, u 는 리뷰 특성의 개수를 의미한다. 또한, W 는 추천 성능 평가에 사용된 모든 리뷰의 유용성에 대한 여부를 분류한 값으로 구성된 벡터이며, W_i 는 리뷰가 유용하거나 유용하지 않는지를 나타내고 있다. 본 연구는 먼저 F 가 주어지면 Z 의 예측 오류를 최소화하는 모델을 구축하는 것이다. 그런 다음 학습된 모델을 사용하여 새로운 리뷰 R_i 에 대한 유용한지를 나타내는 값 W 를 정교하게 예측하고자 한다.

3.2. Phase 2: CHSC Model

최근 리뷰에서 감성 특성을 정교하게 추출하기 위해 CNN, RNN 등 딥러닝 알고리즘을 적용하는 연구가 활발히 수행되고 있다. 본 연구에서



<Figure 2> The Architecture of CHSC Model

는 자연어 처리 연구에서 우수한 분류 성능을 나타내고 있는 CNN을 활용하여 CHSC 모델을 구축하였다(Johnson and Zhang, 2014; Wu et al., 2020). CNN은 연산 속도가 RNN보다 빠르므로 리뷰에 포함되는 의미론적 특성을 정교하게 추출하여 리뷰 유용성 여부를 분류한다. 또한, 집계된 리뷰 대신 각 리뷰를 입력으로 사용하여 텍스트의 길이를 줄이고 CNN의 연산 속도를 높일 수 있으며, 모델 학습에 필요한 연산 비용을 감소시킬 수 있다(Cao et al., 2019). 본 연구에서 구축하고자 하는 CHSC 모델은 <Figure 2>와 같이 Word Embedding Layer, Convolution Layer, Max Pooling Layer, Full Connected Layer로 구성된다.

3.2.1. Word Embedding Layer

딥러닝 알고리즘에서는 텍스트 데이터를 학습 시키기 위해 각 단어를 정형화된 벡터 형태로 변환하는 작업이 필요하다. 예를 들어, 사용자 u 가 특정 아이템 i 에 남긴 리뷰 R_{ui} 를 $\{w_1, w_2, \dots, w_n\}$ 로 정의했을 때, n 은 리뷰에 포함되는 단어를 개수를 나타내고 있다. 기존 텍스트 마이닝 연구에서는 주로 One-hot Encoding 기법을 적용하여 각 단어를 벡터로 변환하였다. 그러나 이와 같은 방법은 행렬의 차원이 지나치게 커지고 벡터의 대부분 값이 0으로 채워지는 데이터 희소성 문제가 발생한다. 이와 같은 문제점을 보완하기 위해 본 연구에서는 Word Embedding Layer를 통해 리뷰에 포함되어 있는 각 단어를 벡터 형태로 변환하였다(Yoo et al., 2018). 이를 위해 리뷰의 각 단어에 워드 임베딩 $f: w_n \rightarrow R^D$ 를 적용하면 각 단어는 정형화된 밀집 벡터(Dense Vector) 형태로 변환한다. 따라서 리뷰는 행렬 $E \in R^{V \times D}$ 형태로 표현되며, D 는 벡터의 차원을 나타내고 V 는 리

뷰에 포함되는 단어의 개수를 의미한다.

3.2.2. Convolution Layer

두 번째 단계에서 컨볼루션 레이어는 첫 번째 단계의 Word Embedding Layer를 통해 입력된 리뷰에서 맥락 특성을 추출할 때 사용된다. 이때 특성을 추출하기 위해 다양한 크기의 필터 K_j 를 통해 리뷰의 슬라이딩 윈도우(Sliding Window)에서 수식 (2)와 같이 컨볼루션 연산을 수행한다.

$$c_j = \Phi(E * K_j + b_j) \quad (2)$$

식(2)에서 $*$ 는 컨볼루션 연산자를 나타내고, $K_j \in R^{k \times m}$ 는 필터 커널의 파라미터(Parameter)를 나타내며, $k \times m$ 은 커널 크기를 의미한다. b_j 는 편향(Bias)을 나타내고 Φ 는 활성화 함수를 의미하며 수식 (3)과 같이 정의할 수 있다.

$$\text{relu}(x) = \max(0, x) \quad (3)$$

3.2.3 Max Pooling Layer

세 번째 단계의 Max Pooling Layer는 앞서 컨볼루션 연산을 통해 추출한 특성에서 중요한 특성은 유지하고 노이즈(Noise) 특성을 제거하기 위해 사용되며, 수식 (4)와 같이 정의할 수 있다.

$$O_j = \max([c_1, c_2, \dots, c_{l-t+1}]) \quad (4)$$

리뷰에 포함되는 다양한 특성을 추출하기 위해 본 연구에서는 다양한 크기의 필터 커널을 적용하였다. 마지막으로 Max Pooling Layer의 출력은 수식 (5)와 같이 정의할 수 있다.

〈Table 2〉 Examples of Review Helpfulness Classification Results on Evaluation Performance Dataset

Users	Items	Reviews	Helpfulness Score	Ratings
52481	0	book provide reflection apply life ... short time plane	1	5
163431	367934	gift complimentary copy book ... enjoyable read	1	3
101585	38	great hop disappoint book none ... recommend book	1	2
36188	367803	tire nico disrespecting tania turn... get next one	0	2
255387	367972	omfg hell read good way serious ... nut book know	1	5
551	0	spiritually mentally inspire book ... help discover really	0	5
230414	366809	book find love kind sordid ... author book seem	1	1

$$O = [o_1, o_2, \dots, o_n] \quad (5)$$

3.2.4. Full Connected Layer

마지막으로 Full Connected Layer는 리뷰에서 추출한 특성을 바탕으로 유용성 점수를 예측하여 결과를 분류하며, 수식 (6)과 같이 정의할 수 있다.

$$Y = \theta(W_s \cdot O + b_s) \quad (6)$$

식 (6)에서 θ 는 Sigmoid 활성화 함수, W_s 는 가중치 행렬 그리고 b_s 는 편향을 나타낸다. Full Connected Layer는 입력 데이터 특성을 0 또는 1로 분류하여 출력한다. 출력된 값이 0이면 리뷰 r_i 는 유용하지 않음을 나타내며, 1이면 유용한 리뷰를 의미한다.

3.3. Phase 3: Helpful Profile-based Filtering

세 번째 단계에서는 먼저, 앞서 구축한 CHSC 모델을 적용하여 추천 성능 평가용 데이터 세트

에서 유용성 정보가 없는 리뷰에 적용하여, 사용자가 작성한 리뷰가 유용한지에 대한 여부를 분류한다. 〈Table 2〉는 사용자가 작성한 리뷰가 유용한지에 대해 분류한 결과 예시이며 Users, Items, Reviews, Helpfulness Score 그리고 5점 척도로 평가된 Ratings로 구성되어 있다. Helpfulness Score는 리뷰 유용성에 대해 분류한 결과를 나타내며 유용성 여부에 따라 0과 1로 분류하였다. 0은 해당 리뷰가 유용하지 않음을 의미하고 1은 해당 리뷰가 유용하다는 것을 의미한다. 그리고 앞서 분류한 결과를 바탕으로 〈Table 3〉과 같이 Helpfulness가 1인 유용한 리뷰만을 포함하는 새로운 사용자 프로파일을 구축한다.

3.4. Phase 4: Recommender System Model

본 연구에서 제안한 방법론의 추천 성능을 검증하기 위해 개인화 추천 서비스 관련 연구에서 대표적으로 사용되고 있는 UBCF, IBCF 그리고 SVD 모델을 적용하여 선호도 평점을 예측하였다(Postmus and Bhulai, 2018). 이를 통해 리뷰의 유용성에 대한 여부를 고려하지 않고 추천 서비

〈Table 3〉 Examples of Filtering Helpful Review on Evaluation Performance Dataset

Users	Items	Reviews	Helpfulness Score	Ratings
52481	0	book provide reflection apply life ... short time plane	1	5
163431	367934	gift complimentary copy book ... enjoyable read	1	3
101585	38	great hop disappoint book none ... recommend book	1	2
255387	367972	omfg hell read good way serious ... nut book know	1	5
230414	366809	book find love kind sordid ... author book seem	1	1

스를 제공할 때의 예측 성능과 유용한 리뷰를 별도로 고려하여 추천 서비스를 제공할 때의 예측 성능을 비교하고자 한다.

UBCF 기법은 추천 대상 사용자에게 적합한 제품을 추천해주기 위해 먼저 선호도 평점 정보를 기반으로 추천 대상 사용자와 유사한 이웃 사용자를 선정한다. 다음으로 선정된 이웃 사용자들이 공통으로 선호하는 제품을 추천 대상 사용자에게 최종적으로 추천한다(Elahi et al., 2016). UBCF 기법을 사용하여 사용자의 선호도 평점을 예측하기 위해서는 먼저 사용자 간의 유사도를 계산한다. 사용자 a 와 사용자 b 간의 유사도는 수식 (7)과 같이 정의된다(Isinkaye et al., 2015; Sarwar et al., 2001).

$$sim(a,b) = \frac{\sum_{i \in I} (R_{a,i} - \bar{R}_a)(R_{b,i} - \bar{R}_b)}{\sqrt{\sum_{i \in I} (R_{a,i} - \bar{R}_a)^2} \sqrt{\sum_{i \in I} (R_{b,i} - \bar{R}_b)^2}} \quad (7)$$

수식 (7)에서 I 는 전체 제품의 집합을 나타내며, $R_{a,i}$ 와 $R_{b,i}$ 는 사용자 a , 사용자 b 가 각각 제품 i 에 부여한 선호도 평점이다. \bar{R}_a 와 \bar{R}_b 는 사용자 a , 사용자 b 가 각각 전체 제품에 부여한 선호도 평점의 평균값을 의미한다. 이처럼 먼저, 유사도 계산을 통해 유사한 이웃 사용자를 선정하고, 다

음으로 이웃 사용자들의 구매 내역을 기반으로 가중 합(Weighted Sum)을 산출하여 선호도 평점을 예측한다(Choi et al., 2016; Polatidis and Georgiadis, 2016). 추천 대상 사용자 u 가 특정 제품 i 에 대한 예측 선호도 평점은 수식 (8)과 같이 계산된다(Ar and Bostanci, 2016; Bobadilla et al., 2011).

$$\hat{R}_{u,i} = \bar{R}_u + \frac{\sum_{v \in N_i^k(u)} sim(u,v) \cdot (R_{v,i} - \bar{R}_v)}{\sum_{v \in N_i^k(u)} |sim(u,v)|} \quad (8)$$

수식 (8)에서 \bar{R}_u 는 추천 대상 사용자가 모든 제품에 부여한 선호도 평점의 평균값을 나타내고, $R_{v,i}$ 는 다른 사용자 v 가 특정 제품 i 에 부여한 선호도 평점, \bar{R}_v 는 다른 사용자 v 가 모든 제품에 부여한 선호도 평점 평균값을 의미한다. 그리고 $sim(u, v)$ 는 추천 대상 사용자 u 와 다른 사용자 v 간의 유사도 값을 나타낸다.

IBCF 기법은 먼저 특정 제품을 기준으로 선택하고, 사용자가 선호도 평점을 부여한 제품과 유사한 제품을 이웃 제품으로 선정한다. 다음으로 제품에 부여한 선호도 평점을 기반으로 추천 대상 사용자가 특정 제품에 대한 선호도 평점을 예

측한다. 제품 i 와 제품 j 간의 유사도 $sim(i, j)$ 는 아래 수식 (9)와 같이 계산된다(Chung et al., 2014; Wei et al., 2012).

$$sim(i, j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_i)(R_{u,j} - \bar{R}_j)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_i)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_j)^2}} \quad (9)$$

수식 (9)에서 U 는 제품 i 와 j 에 모두 선호도 평점을 부여한 전체 사용자 집합을 의미한다. $R_{u,i}$ 와 $R_{u,j}$ 는 사용자 u 가 제품 i , 제품 j 에 각각 부여한 선호도 평점이며, \bar{R}_i 와 \bar{R}_j 는 전체 사용자가 각각 제품 i , 제품 j 에 부여한 평점의 평균이다. IBCF 기법은 가중 평균(Weighted Average)을 계산해 선호도 등급을 수식(10)과 같이 계산할 수 있다.

$$\hat{R}_{a,i} = \frac{\sum_{n \in N} R_{a,n} \cdot sim(i, n)}{\sum_{n \in N} sim(i, n)} \quad (10)$$

수식 (10)에서 각 선호도 평점은 제품 i 와 제품 n 간의 유사도 $sim(i, n)$ 에 따라 가중치가 부여된다. 이는 예측하려는 제품과 유사한 제품의 점수를 크게 반영하여 선호도 평점을 계산하는 방식이다.

모델 기반 SVD(Singular Value Decomposition)는 고차원의 행렬을 저차원의 행렬로 축소하는 기법으로 행렬 인수분해(Matrix Factorization) 관련 연구에서 널리 사용되고 있는 방법의 하나다 (Bokde et al., 2015; Koren and Bell, 2015). 모든 사용자와 제품에 대한 $m \times n$ 크기의 행렬 $M = U \Sigma V^T$ 을 분해하면 $U_{m \times m}$, $\Sigma_{m \times n}$, $V_{n \times n}^T$ 3개의 행렬로 나타낼 수 있다. $U_{m \times m}$ 는 사용자 행렬을 나타

내고, $\Sigma_{m \times n}$ 은 대각 행렬을 나타내며 $V_{n \times n}^T$ 는 제품 행렬을 나타낸다. SVD 기법에서 사용자 선호도 평점은 수식 (11)과 같이 계산된다.

$$\hat{R} = \mu + b_u + b_i + q_i^T p_u \quad (11)$$

수식 (11)에서 μ 는 선호도 평점 평균을 나타내며, b_u 와 p_u 는 각각 사용자와 제품에 대한 편향 값이다. q_i 와 p_u 는 저차원의 공간에서 사용자와 제품에 대한 잠재적 요인을 의미하고 있다. 행렬 $M = U \Sigma V^T$ 은 사용자가 부여한 선호도 평점 정보가 부족해 대부분 결측값(Missing Value)을 나타내는 데이터 희소성 문제가 존재한다. 따라서 사용자가 기존에 부여한 선호도 평점 정보를 사용하여 제품 오차를 최소화하는 잠재적 요인을 찾기 위해 수식 (12)와 같이 정규화 과정을 수행한다.

$$\min_{q, p} \sum_{(u,i) \in M} (R_{ui} - \hat{R}_{ui}) + \lambda (\|q_i\|^2 + \|p_u\|^2) \quad (12)$$

$$\begin{aligned} b_u &\leftarrow b_u + \gamma(e_{ui} - \lambda b_u) \\ b_i &\leftarrow b_i + \gamma(e_{ui} - \lambda b_i) \\ p_u &\leftarrow p_u + \gamma(e_{ui} \cdot q_i - \lambda p_u) \\ p_i &\leftarrow q_i + \gamma(e_{ui} \cdot p_u - \lambda q_i) \end{aligned} \quad (13)$$

제품 오차를 최소화하는 방법은 수식 (13)과 같이 SGD(Stochastic Gradient Descent)를 활용하여 단계마다 가중치를 수정하면서 사용자의 선호도 평점 \hat{R}_{ui} 을 예측한다. 여기서 e_{ui} 는 실제 선호도 평점 R_{ui} 와 예측 선호도 평점 \hat{R}_{ui} 간의 차이로 계산된다. 또 사용자 또는 제품의 잠재적 요인은 정규분포에 따라 무작위로 초기화되어 계산되

(Table 4) Statistics of Amazon Book Dataset

Users	Items	Ratings & Reviews	Helpfulness
603,556	367,975	8,872,495	4,745,574

며, 이를 바탕으로 사용자와 제품의 특성을 표현할 수 있는 최적의 요인 수를 설정한다.

($\Theta_2 < 0.2$)를 각각 35,000개를 활용하여 모델을 학습시키고 성능을 평가했다(Ge et al., 2019).

4. 실험

4.1. 데이터 및 평가지표

본 연구에서 제안한 방법론의 추천 성능을 평가하기 위해 세계적인 전자상거래 사이트 Amazon.com에서 1996년 5월부터 2014년 7월까지 수집된 도서 관련 데이터¹⁾를 사용하였으며, 해당 데이터 세트에는 리뷰를 5개 이상 작성한 사용자와 5개 이상의 리뷰가 작성된 제품만 포함되어 있다. 데이터 세트에 대한 기초통계는 <Table 4>와 같다.

본 연구에서 제안한 방법론의 추천 성능을 평가하기 위해 먼저 CHSC 모델 구축용 데이터 세트와 추천시스템 구축용 데이터 세트로 분할한다. 본 연구에서 제안한 모델을 효과적으로 학습시키기 위해 기존 연구의 방법을 참고하여 유용성 투표수가 최소 10개 이상의 리뷰만을 학습 데이터 세트로 사용한다(Ge et al., 2019; Liu et al., 2008). 여기에서 리뷰 유용성 점수는 수식 (14)와 같이 총투표수에 대한 유용한 리뷰 투표수의 백분율로 계산된다(Krishnamoorthy, 2015; Zhang and Varadarajan, 2006). CHSC 모델이 유용성 점수를 정교하게 예측할 수 있도록 학습시키기 위해 유용한 리뷰($\Theta_1 > 0.9$)와 유용하지 않은 리뷰

$$\text{Helpfulness Score} = \frac{\# \text{Helpful Vote}}{\# \text{Total Vote}} \quad (14)$$

다음으로 추천시스템을 구축하기 위해 수집한 데이터 세트에서 유용성 정보가 반영되지 않은 리뷰만을 활용하여 사용자 프로파일을 구축하고 실험에 사용했다. 해당 사용자 프로파일은 사용자 정보, 제품 정보, 평점 및 리뷰 등 4가지 속성으로 구성되어 있다. 또한, 추천시스템 구축 과정에서 데이터 희소성 문제를 방지하기 위해 기존 연구의 방법을 따라 최소 20개 이상의 평점을 부여한 사용자만을 선택하여 실험에 사용했다(He et al., 2017; Moore, 2015). 따라서 최종적으로 사용한 데이터 세트는 30,745명의 사용자, 40,405개의 제품 그리고 916,067개의 평점 및 리뷰를 포함하고 있다.

CHSC 모델을 적용한 유용성 점수 분류 결과의 성능은 <Table 5>와 같이 Confusion Matrix를 사용하여 평가하였다(Nassirtoussi et al., 2014; Siering et al., 2018; Zafari et al., 2020). 이때 Accuracy는 분류 성능을 측정할 때 가장 일반적으로 사용되는 평가지표이며, 수식 (15)와 같이 전체 분류 결과 중에서 유용한 리뷰와 유용하지 않은 리뷰를 정확하게 분류한 개수의 비율을 나타내고 있다. Precision은 수식 (16)과 같이 모델이 유용하다고 분류한 리뷰 중에서 실제 유용한

1) <http://jmcauley.ucsd.edu/data/amazon/links.html>

〈Table 5〉 Confusion Matrix

		Predicted Class	
		Helpful	Unhelpful
Actual Class	Helpful	TP	FN
	Unhelpful	FP	TN

리뷰에 속하는 비율을 나타내며, Recall은 수식 (17)과 같이 실제 유용한 리뷰 중에서 모델이 유용하다고 분류한 리뷰 개수의 비율을 나타낸다. F1-Score는 Precision과 Recall의 평가 결과를 조화 평균하여 모델의 성능을 평가할 수 있으며 식 (18)과 같이 정의한다.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (15)$$

$$Precision = \frac{TP}{TP + FP} \quad (16)$$

$$Recall = \frac{TP}{TP + FN} \quad (17)$$

$$F1-Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (18)$$

개인화 추천 서비스의 추천 성능을 평가하기 위해 다양한 평가 방법이 제안되었다(Isinkaye et al., 2015). 본 연구에서는 사용자의 선호도 평점을 예측할 때 널리 사용되는 MAE(Mean Average Error), RMSE(Root Mean Squared Error) 평가지표를 적용하여 예측 성능을 측정하였다(Herlocker et al., 2004; Sarwar et al., 2001). 먼저, 예측 평점과 실제 평점 간의 차이를 비교할 때 가장 널리 사용되는 MAE는 수식 (19)와 같이 정의한다(Hammou and Lahcen, 2017; Jeong et al., 2010).

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (19)$$

수식 (19)에서 N 은 전체 실험 데이터 세트 개수를 나타내며, y_i 는 실제 평점, \hat{y}_i 는 예측 평점을 나타내고 있다. 즉, 실제 선호도 평점과 예측 선호도 평점 간의 오차를 절대값 형태로 나타내며, 최종적으로 이를 합치고 실험 데이터 세트 개수만큼 나누는 방식이다. MAE는 실제 평점과 예측 평점 간의 오차 크기와 관계없이 같은 가중치를 부여한다(Chai and Draxler, 2014). 다음으로, RMSE는 Netflix Prize에서 사용된 평가지표로 수식 (20)과 같이 정의한다.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|^2} \quad (20)$$

수식 (20)에서, y_i 는 실제 평점, \hat{y}_i 는 예측 평점을 나타내며, N 은 전체 실험 데이터 세트 개수를 나타내고 있다. RMSE는 실제 선호도 평점과 예측 선호도 평점 간의 오차를 제곱한 형태로 나타내며, 최종적으로 이를 합치고 실험 데이터 세트 개수만큼 나누고 제곱근을 계산하는 방식이다(Ngo-Ye and Sinha, 2014; Sánchez-Moreno et al., 2016). 또한, 실제 평점과 예측 평점 간의 오차가 큰 값에 대해 상대적으로 높은 가중치를 부여한다(Chai and Draxler, 2014).

4.2. 실험설계

본 연구의 CHSC 실험에서는 리뷰를 효과적으로 분석하기 위해 NLTK 패키지를 적용하여 모델 학습에 크게 중요하지 않은 불용어, 특수 문자, 기호 및 숫자 등을 제거하였다(Moon et al., 2019; Song et al., 2020). 또한, 제안한 방법론의 성능을 평가하기 위해 데이터 세트를 8:2 비율의 학습과 검증 데이터로 분할하였다. 모델 학습을 위해 단어 임베딩 차원을 300으로 설정하고, 필터 크기는 100으로 설정하였으며, 윈도우 크기(Window Size)는 각각 3, 4, 5로 설정하였다(Khan and Niu, 2020). 리뷰의 길이는 최대 길이로 고정하고, 단어 크기(Vocabulary Size)는 다양하게 설정하고 분류 성능에 따라 파라미터를 설정했다(Zhang and Wallace, 2015). 실험 과정에서 과적합 문제를 개선하기 위해 각 Convolutional Layer와 Dense Layer에 드롭아웃(Dropout) 방법을 적용하였으며, 드롭아웃 비율은 0.5로 설정하였다. 최적화 알고리즘은 기존 연구에서 널리 사용되는 Adam을 적용하고 배치 크기(Batch Size)는 50으로 설정하였으며, Epoch는 다양한 크기로 설정하고 분류 성능에 따라 최적의 파라미터를 설정했다(Janke et al., 2019; Ullah et al., 2018). 추천시스템 구축 실험에서 CF 기법은 피어슨 상관계수(Pearson Correlation Coefficient)를 활용하여 유사도를 측정하고, UBCF 기법과 IBCF 기법은 각각 가중 합과 가중 평균을 계산하여 사용자 선호도 평점을 예측하였다(Barragáns-Martínez et al., 2010; Moon et al., 2017). 또한, 이웃의 크기를 1부터 100까지 다양하게 설정하고 예측 성능을 확인하였다. SVD 기법은 잠재적 벡터 크기를 8, 16, 32, 64, 128로 설정하고, 학습률(Learning Rate)은 0.005로 설정하였다(Cui and Fearn, 2018;

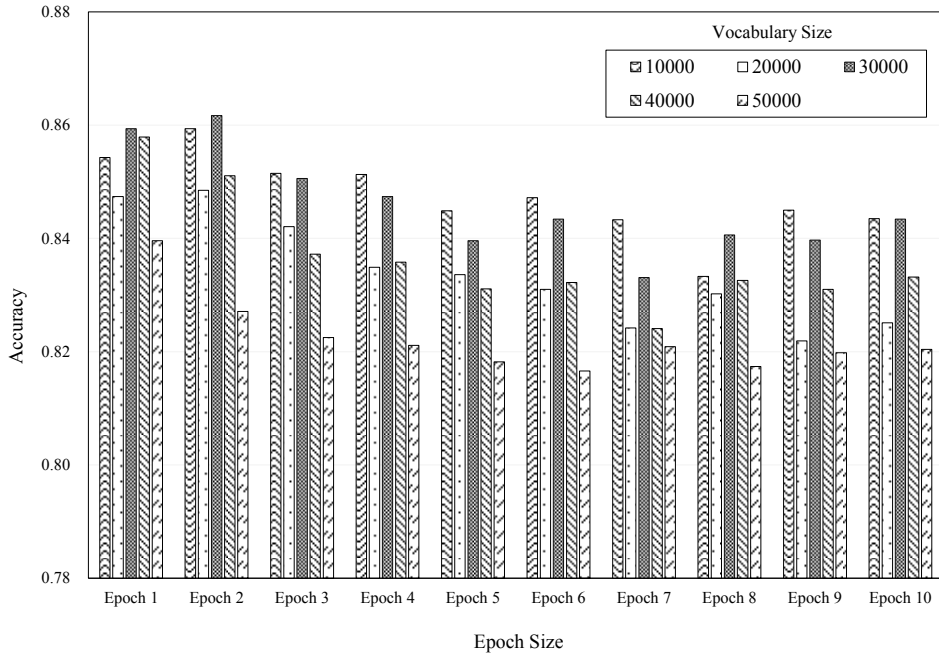
Hu et al., 2017; Knees et al., 2014). 본 연구에서는 TensorFlow, Keras 및 Surprise 패키지를 사용하여 CPU Intel Core i9-9900KF, 64GB RAM, GeForce RTX 2080 Ti 환경에서 실험을 수행하였다.

4.3. 실험결과

4.3.1. 리뷰 유용성 분류 성능 평가

본 연구에서 제안한 방법론의 성능을 평가하기 위해 먼저, 리뷰의 유용성 여부를 분류하는 CHSC 모델을 구축하였다. 기존 리뷰 관련 연구에서는 주로 평점 정보를 기반으로 리뷰의 감성 특성을 긍정 및 부정으로 분류하였으나, 본 연구에서는 해당 리뷰가 사용자에게 실제로 유용한지에 대한 여부를 정교하게 분류하였다. 또한, 모델의 분류 정확도를 높이기 위해 실험 데이터에 따라 등장하는 단어 빈도가 다른 점을 고려하여 다양한 단어 크기를 설정하고 실험을 수행하였다(Yoo et al., 2018). <Figure 3>은 Epoch 크기를 1에서 10까지 설정하였을 때 단어 크기에 따른 정확도를 나타내고 있다. 실험 결과에서 단어 크기를 각각 10,000개, 20,000개, 30,000개로 설정하였을 때는 Epoch 크기가 2일 때 분류 정확도가 가장 높게 나타났고, 이후에는 정확도가 다소 낮아지고 있다. 반면, 단어 크기를 각각 40,000개, 50,000개로 설정하였을 때는 Epoch 크기가 1일 때 가장 높게 나타났다.

<Table 6>은 각 단어 크기에서 최적의 Epoch를 설정하였을 때의 분류 성능을 나타내고 있다. 다양한 실험을 수행한 결과에 따라 Epoch가 2, 단어 크기가 30,000개일 때 최적의 분류 성능을 보여주고 있다. 따라서 본 연구에서는 앞서 언급한 최적의 파라미터를 CHSC 모델에 적용하여



〈Figure 3〉 Classification Performance on the Different Sizes of Vocabulary for the CHSC Model

모델을 학습시키고, 이를 활용하여 새로운 리뷰의 유용성 점수를 정교하게 예측하였다.

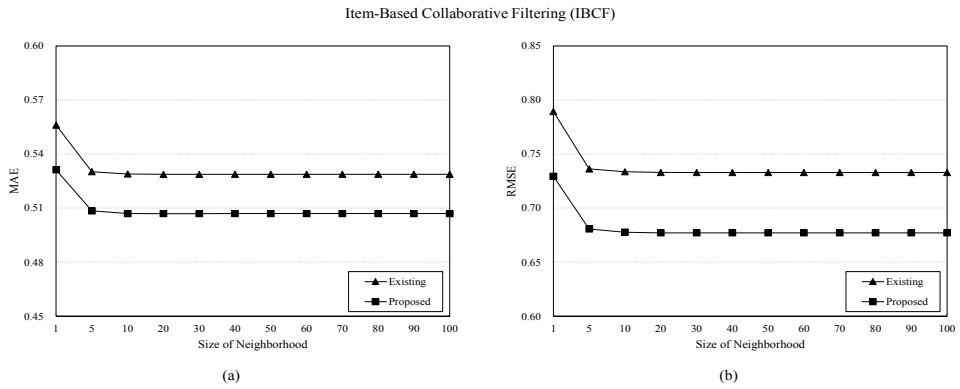
4.3.2. 선호도 예측 성능 평가

추천 성능 평가 데이터를 사용하여 사용자가 작성한 리뷰의 유용성 점수를 예측하여 리뷰가 유용한지에 대한 여부를 분류하였다. 그리고 유용한 리뷰만을 활용하여 새로운 사용자 프로파

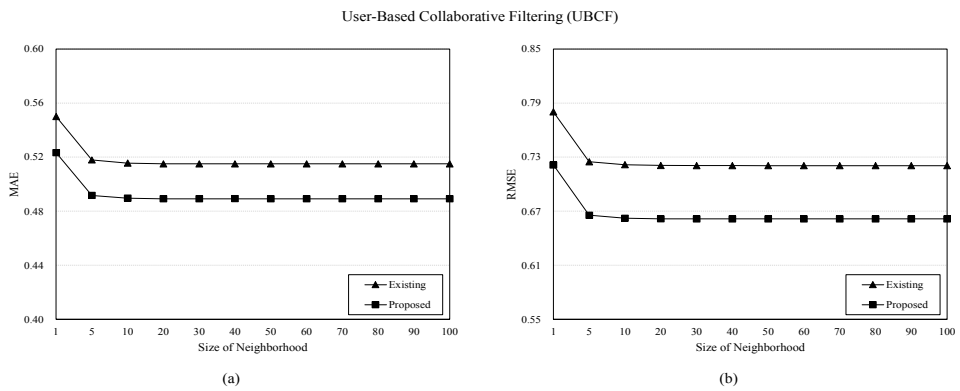
일을 구축하였다. 이때 “Existing”은 모든 리뷰를 포함하는 기존의 전통적인 추천 방법론이다. “Proposal”은 본 연구에서 제안하는 유용한 리뷰만을 포함한 사용자 프로파일을 사용한 추천 방법론을 의미한다. 본 연구에서는 UBCF, IBCF 기법에 대해 이웃 사용자의 크기를 다양하게 설정했을 때의 예측 성능을 확인하기 위해, 이웃 사용자의 크기를 1부터 100까지 설정하였다(Wang et al., 2021; Zhu et al., 2014). 모델 기반 협업 필

〈Table 6〉 Classification Performance Comparison Results at Optimization Epoch Size

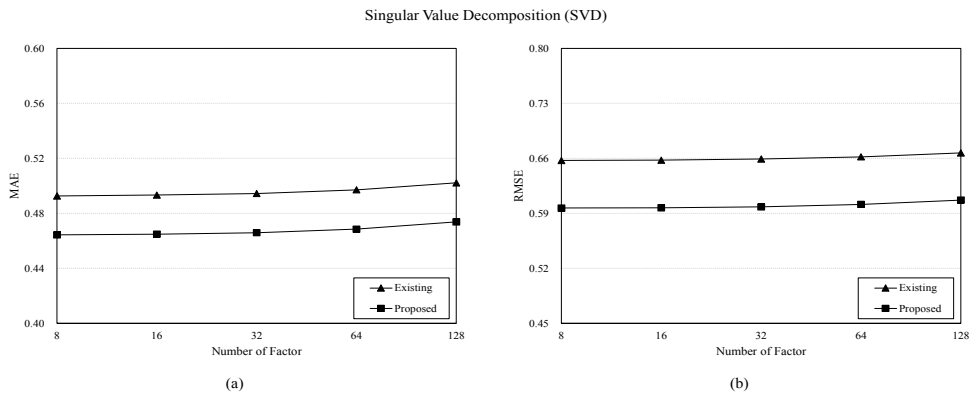
Vocabulary Size	Accuracy	Precision	Recall	F1-Score
10000	0.859	0.849	0.876	0.860
20000	0.849	0.861	0.839	0.847
30000	0.862	0.833	0.888	0.864
40000	0.858	0.862	0.851	0.854



〈Figure 4〉 Prediction Performance of MAE (a) and RMSE (b) on the Different Sizes of the Neighborhood for the IBCF Method



〈Figure 5〉 Prediction Performance of MAE (a) and RMSE (b) on the Different Sizes of the Neighborhood for the UBCF Method.



〈Figure 6〉 Prediction Performance of MAE (a) and RMSE (b) on the Different Number of the Factor for the SVD Method

터링의 SVD 기법은 기존 행렬 분해 관련 연구 방법을 참고하여 잠재적 요인 크기를 8, 16, 32, 64, 128로 설정하여 예측 성능을 비교하였다 (Mandal and Maiti, 2021). 또한, MAE와 RMSE 평가지표를 사용하여 예측 평점과 실제 평점 간의 오차에 대해 다른 가중치를 부여했을 때의 예측 성능 변화를 확인했다(Chai and Draxler, 2014). 기존 추천 방법론과 본 연구에서 제안한 추천 방법론을 UBCF, IBCF 그리고 SVD 기법을 통해 예측 성능을 비교한 결과는 <Figure 4>, <Figure 5> 그리고 <Figure 6>과 같다. 실험 결과, 이웃 사용자의 크기와 잠재적 요인 크기와 관계 없이 본 연구에서 제안한 추천 방법론이 기존 추천 방법론보다 전체적으로 예측 성능이 개선되었음을 확인할 수 있다. UBCF 기법에 MAE와 RMSE 평가지표를 적용한 결과, 이웃 사용자 크기가 각각 10, 20일 때 가장 우수한 예측 성능이 나타났고 이후에는 이웃 사용자 크기와 관계없이 일관된 예측 성능을 나타냈다. 마찬가지로 IBCF 기법에도 MAE와 RMSE 평가지표를 적용한 결과, 이웃 사용자의 크기가 각각 10, 30일 때 가장 우수한 예측 성능이 나타났고 이후에는 일관된 예측 성능을 나타내고 있다. SVD 기법에서 잠재적 요인의 크기가 모두 8일 때 가장 우수한 예측 성능을 나타내고 있다. 따라서 본 연구에서 제안한 방법론을 기존 방법론과 비교한 결과, MAE 평가지표에서는 예측 성능이 3.39%(UBCF), 5.00%(IBCF), 5.53%(SVD)만큼 개선되었고, RMSE 평가지표에서는 예측 성능이 8.16%(UBCF), 8.15%(IBCF), 8.88%(SVD)만큼 개선되었다.

실험 결과에서 나타내는 바와 같이, 사용자의 선호도 평점을 예측할 때 본 연구에서 제안하고 있는 유용한 리뷰만 포함한 사용자 프로파일을 사용하는 경우가 기존 방법론보다 더 우수한 예

측 성능을 나타내고 있음을 확인할 수 있다. 이는 기존 개인화 추천 서비스 연구에서 주로 사용되던 정량화 정보를 통해 확인할 수 없었던 리뷰 유용성에 대한 정보를 개인화 추천 서비스에 반영하면 예측 성능을 개선할 수 있음을 시사한다.

5. 결론

5.1. 연구 결과 토의

본 연구는 사용자의 구매 의사결정에 큰 영향을 미치는 리뷰에서 유용성 정보를 정교하게 예측하여 추천시스템에 반영하는 새로운 추천 방법론을 제안하였으며, 이를 위해 전자상거래 사이트 Amazon.com에서 수집한 데이터를 사용하여 추천 성능을 평가하였다. 실험 결과, 본 연구에서 제안한 CHSC 모델을 통해 분류된 유용한 리뷰만을 포함하는 사용자 프로파일을 사용했을 때의 예측 성능이 기존 모든 리뷰를 포함할 때의 예측 성능보다 우수함을 확인할 수 있었다. 그중에서도 모델 기반 협업 필터링의 SVD 기법이 기존 방법보다 7.5%만큼 개선되어 가장 우수한 성능을 보여주었다. 이는 선호도 평점을 바탕으로 사용자 간의 유사도를 측정할 때 리뷰의 품질이 선호도 평점에 유의미한 영향을 미칠 수 있음을 시사한다. 즉, 사용자가 작성한 리뷰의 품질이 높으면 사용자가 부여한 선호도 평점 정보가 더 높은 신뢰성을 제공할 수 있음을 보여준다(Qiu et al., 2016). 또한, 본 연구를 통해 리뷰의 유용성을 예측하는 CHSC 모델의 분류 정확도가 모델 학습에 사용되는 단어 개수에 따라 다양하게 나타나는 것을 확인할 수 있었다. 이는 학습 데이터에 전체 단어를 포함하면 분석에 크게 중요

하지 않은 노이즈 특성을 포함할 수 있으며, 전체적으로 연산 비용과 시간이 증가하고 정확도는 오히려 감소할 수 있음을 보여준다(Ge et al., 2019). 즉, 전체 리뷰에서 등장 빈도가 낮은 단어들을 분석과정에서 제외하면 중요한 단어는 유지되고, 분석에 크게 중요하지 않은 단어는 제거해도 실제 모델 학습에는 큰 영향이 없음을 확인할 수 있다.

5.2. 연구의 한계점 및 향후 방향

본 연구는 CNN을 적용한 CHSC 모델을 통해 리뷰의 유용한지 아닌지를 분류하고, 이를 바탕으로 새로운 추천 방법론을 제안하고 성능을 평가하였다. 본 연구의 한계점 및 향후 연구 방향은 다음과 같다. 첫째, 본 연구에서는 Amazon 온라인 도서 관련 데이터를 수집하여 제안한 방법론의 추천 성능을 평가하였다. 본 연구에서 제안한 방법론을 다른 분야에 맞게 적용하기 위해서는 다양한 분야의 데이터를 추가로 수집하여 성능을 평가하는 연구가 필요하다. 둘째, 본 연구에서는 리뷰 유용성 점수 예측 모델을 구축하기 위해 자연어 처리에서 우수한 성능을 보여준 CNN을 적용하였다. 최근에는 BERT(Bidirectional Encoder Representations from Transformers)와 같은 딥러닝 모델이 자연어 처리 분야에서 우수한 성능을 보인다. 따라서 향후 연구에서는 다양한 딥러닝 모델을 적용하였을 때의 추천 성능을 비교할 수 있다. 셋째, 본 연구에서는 리뷰의 유용성 여부를 예측하고 이를 바탕으로 유용한 리뷰를 작성한 사용자 프로파일을 구축하여 추천 서비스를 제공하는 방법론을 제안했다. 그러나 리뷰 유용성 정보 외에도 사용자의 선호도를 파악할 수 있는 제품 특성, 구매 시기 등 다양한 정보

를 고려한다면 제안한 방법론의 추천 성능을 더 향상할 수 있다고 본다. 마지막으로, 새로운 제품이 출시한 초기에 작성된 리뷰는 일반적으로 나중에 게시된 리뷰보다 많은 유용성 평가를 받게 되고, 이로 인해 사용자의 과도한 관심을 받게 되는 순차적 편향 문제가 일어난다. 본 연구에서는 리뷰 유용성 예측 모델 구축 시, 이러한 순차적 편향은 고려하지 않았다. 따라서 앞으로는 이러한 문제점을 해결하기 위해, 리뷰의 작성 일자를 고려하여 유용성 점수를 예측할 필요가 있다고 본다.

5.3. 학술적 의의와 실무적 시사점

본 연구는 대표적인 비정형 데이터인 리뷰를 딥러닝 기법을 통해 분석하고 이를 추천시스템에 반영하였고, 추천 성능을 향상했다. 이와 같은 본 연구의 학술적 의의를 살펴보면 다음과 같다. 첫째, 리뷰를 활용한 기존 개인화 추천 서비스 관련 연구에서는 제품에 작성된 모든 리뷰를 활용하여 리뷰에 포함된 감성 특성을 추출하여 이를 추천시스템에 반영하였다. 그러나 사용자가 작성한 리뷰에는 광고성 내용 또는 의미를 전혀 알 수 없거나 거짓 리뷰 등이 포함되어 있다(Saumya and Singh, 2018). 즉 사용자에게 유용하지 않고 제품과 관련 없는 리뷰 또한 포함되어 있으므로 이는 추천시스템의 성능을 저하하는 문제가 있다. 따라서 본 연구에서는 리뷰의 유용성 여부를 정교하게 예측하여 추천시스템에 반영하는 방법론을 제안하여 추천 성능을 향상함으로써 개인화 추천 서비스 연구의 확장에 기여할 수 있다. 둘째, 본 연구에서는 리뷰 유용성 정보를 고려한 개인화 추천 서비스의 추천 성능을 검증하기 위해 리뷰 유용성을 고려한 결과와 고

려하지 않은 결과를 비교하였다. 연구 결과, 리뷰 유용성을 고려하였을 때의 예측 성능이 더 높은 것으로 나타났다. 따라서 제품에 대한 특성, 가격 그리고 사용자의 감정 이외에도 리뷰의 유용성 정보가 구매 의사결정에 있어 중요한 요인이라는 것을 확인할 수 있었으며, 이는 제품에 대한 주관적 리뷰보다는 사용자가 작성한 리뷰의 유용성 투표수와 같이 객관적인 정보가 선호도에 영향을 끼치는 요인임을 시사한다.

본 연구의 실무적 시사점은 다음과 같다. 첫째, 본 연구에서는 사용자가 작성한 리뷰에서 유용성 점수를 정교하게 예측하여 이를 추천시스템에 반영하여 추천 서비스를 제공하는 방법론을 제안하였으며, 기존 방법론보다 높은 추천 성능을 확인하였다. 대부분의 전자상거래 사이트는 사용자가 제품에 대한 리뷰를 남길 수 있지만, 리뷰의 유용성 정보를 추천 서비스에 반영한 경우는 많지 않은 실정이다. 따라서 기업은 리뷰에 대한 유용성을 평가할 수 있는 서비스를 제공할 필요가 있다. 또한 유용성 지수가 높게 평가된 리뷰의 작성자에게 마일리지 또는 쿠폰으로 보상을 한다면 제품에 대한 리뷰의 가치를 높일 수 있다고 본다. 둘째, 기존에는 제품에 대한 리뷰 개수에 중점을 두고 사용자들이 제품에 리뷰를 남길 수 있도록 유도하였다. 그러나 리뷰의 개수보다는 사용자들이 양질의 리뷰를 작성할 수 있도록 유도하는 전략을 고려할 필요가 있다고 본다. 셋째 본 연구에서 제안한 방법론은 리뷰 유용성 정보를 제공하는 다양한 분야의 개인화 추천 서비스에 활용 가능할 것으로 본다. 마지막으로, 제안한 방법론을 실제 전자상거래 사이트에 적용하여 사용자에게 서비스를 제공한다면 추천 목록에 따른 사용자들의 요구를 도출할 수 있다. 또한 이를 활용하여 더욱 정교한 추천

서비스를 구축하고 마케팅, 고객 관리 등 기업 내 여러 분야에서 유용하게 활용할 수 있다. 따라서 기업은 사용자의 편의성과 만족도를 개선할 수 있고 매출 상승 또한 기대할 수 있을 것으로 본다.

참고문헌(References)

- Abdollahi, B., and O. Nasraoui, "Using explainability for constrained matrix factorization", *Proceedings of the Eleventh ACM Conference on Recommender Systems*, (2017), 79~83.
- Al-Bashiri, H., M. A. Abdulgaber, A. Romli, and H. Kahtan, "An improved memory-based collaborative filtering method based on the TOPSIS technique", *PloS one*, Vol. 13, No.10(2018), e0204434.
- Ar, Y., and E. Bostanci, "A genetic algorithm solution to the collaborative filtering problem", *Expert Systems with Applications*, Vol.61 (2016), 122~128.
- Baek, H., J. Ahn, and Y. Choi, "Helpfulness of online consumer reviews: Readers' objectives and review cues", *International Journal of Electronic Commerce*, Vol.17, No.2(2012), 99~126.
- Bang, H. B., H. W. Lee, and J. H. Lee, "TV Program Recommender System Using Viewing Time Patterns", *Journal of Korean Institute of Intelligent Systems*, Vol.25, No.5(2015), 431~436.
- Barragáns-Martínez, A. B., E. Costa-Montenegro, J. C. Burguillo, M. Rey-López, F. A. Mikic-Fonte, and A. Peleteiro, "A hybrid content-based and item-based collaborative

- filtering approach to recommend TV programs enhanced with singular value decomposition”, *Information Sciences*, Vol.180, No.22(2010), 4290~4311.
- Bennett, J., and S. Lanning, “The netflix prize”, *Proceedings of KDD Cup and Workshop*, Vol.2007, (2007), 35.
- Bobadilla, J., F. Ortega, A. Hernando, and J. Alcalá, “Improving collaborative filtering recommender system results and performance using genetic algorithms”, *Knowledge-Based Systems*, Vol.24, No.8(2011), 1310~1316.
- Bokde, D., S. Girase, and D. Mukhopadhyay, “Matrix factorization model in collaborative filtering algorithms: A survey”, *Procedia Computer Science*, Vol.49, (2015), 136~146.
- Cao, R., X. Zhang, and H. Wang, “A Review Semantics Based Model for Rating Prediction”, *IEEE Access*, Vol.8, (2019), 4714~4723.
- Castelli, M., L. Manzoni, L. Vanneschi, and A. Popovič, “An expert system for extracting knowledge from customers’ reviews: The case of Amazon. com, Inc”, *Expert Systems with Applications*, Vol.84, (2017), 117~126.
- Chai, T., and R. R. Draxler, “Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature”, *Geoscientific Model Development*, Vol.7, No.3(2014), 1247~1250.
- Cheng, Z., Y. Ding, L. Zhu, and M. Kankanhalli, “Aspect-aware latent factor model: Rating prediction with ratings and reviews”, *Proceedings of the World Wide Web Conference*, (2018), 639~648.
- Choi, I. Y., M. G. Oh, J. K. Kim, and Y. U. Ryu, “Collaborative filtering with facial expressions for online video recommendation”, *International Journal of Information Management*, Vol.35, No.3(2016), 397~402.
- Chung, K. Y., D. Lee, and K. J. Kim, “Categorization for grouping associative items using data mining in item-based collaborative filtering”, *Multimedia Tools and Applications*, Vol.71, No.2(2014), 889~904.
- Cui, C., and T. Fearn, “Modern practical convolutional neural networks for multivariate regression: Applications to NIR calibration”, *Chemometrics and Intelligent Laboratory Systems*, Vol.182, (2018), 9~20.
- Das, A. S., M. Datar, A. Garg, and S. Rajaram, “Google news personalization: scalable online collaborative filtering”, *Proceedings of the 16th International Conference on World Wide Web*, (2007), 271~280.
- Elahi, M., F. Ricci, and N. Rubens, “A survey of active learning in collaborative filtering recommender systems”, *Computer Science Review*, Vol.20, (2016), 29~50.
- Fu, M., H. Qu, D. Moges, and L. Lu, “Attention based collaborative filtering”, *Neurocomputing*, Vol.311, (2018), 88~98.
- Garcia-Cumbreras, M. A., A. Montejo-Raez, and M. C. Diaz-Galiano, “Pessimists and optimists: Improving collaborative filtering through sentiment analysis”, *Expert Systems with Applications*, Vol.40, No.17(2013), 6758~6765.
- Ge, S., T. Qi, C. Wu, F. Wu, X. Xie, and Y. Huang, “Helpfulness-aware review based neural recommendation”, *CCF Transactions on Pervasive Computing and Interaction*, Vol.1, No.4(2019), 285~295.

- Goldberg, D., D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry", *Communications of the ACM*, (1992), 61~70.
- Guy, I., M. Avihai, A. Nus, and F. Raiber, "Extracting and Ranking Travel Tips from User-Generated Reviews", *Proceedings of the 26th International Conference on World Wide Web*, (2017), 987~996.
- Hammou, B. A., and A. A. Lahcen, "FRAIPA: A fast recommendation approach with improved prediction accuracy", *Expert Systems with Applications*, Vol.87, (2017), 90~97.
- He, X., L. Liao, H. Zhang, L. Nie, X. Hu, and T.S. Chua, "Neural collaborative filtering", *Proceedings of the 26th International Conference on World Wide Web*, (2017), 173~182.
- Herlocker, J. L., J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems", *ACM Transactions on Information Systems*, Vol.22, No.1(2004), 5~53.
- Hu, Y. H., Y. L. Chen, and H. L. Chou, "Opinion mining from online hotel reviews—a text summarization approach", *Information Processing & Management*, Vol.53, No.2(2017), 436~449.
- Hyun, J., S. Ryu, and S. Y. Lee, "How to improve the accuracy of recommendation systems: Combining ratings and review texts sentiment score", *Journal of Intelligence and Information Systems*, Vol.25, No.1(2019), 219~239.
- Isinkaye, F. O., Y. O. Folajimi, and B. A. Ojokoh, "Recommendation systems: Principles, methods and evaluation", *Egyptian Informatics Journal*, Vol.16, No.3(2015), 261~273.
- Janke, J., M. Castelli, and A. Popovič, "Analysis of the proficiency of fully connected neural networks in the process of classifying digital images. Benchmark of different classification algorithms on high-level image features from convolutional layers", *Expert Systems with Applications*, Vol.135, (2019), 12~38.
- Jeon, B. K., and H. Ahn, "A Collaborative Filtering System Combined with Users' Review Mining: Application to the Recommendation of Smartphone Apps", *Journal of Intelligence and Information Systems*, Vol.21, No.2(2015), 1~18.
- Jeong, B., J. Lee, and H. Cho, "Improving memory-based collaborative filtering via similarity updating and prediction modulation", *Information Sciences*, Vol.180, No.5(2010), 602~612.
- Johnson, R., and T. Zhang, "Effective use of word order for text categorization with convolutional neural networks", *arXiv preprint arXiv:1412.1058*, (2014).
- Kaushik, K., R. Mishra, N. P. Rana, and Y. K. Dwivedi, "Exploring reviews and review sequences on e-commerce platform: A study of helpful reviews on Amazon", *Journal of Retailing and Consumer Services*, Vol.45, (2018), 21~32.
- Khan, Z. Y., and Z. Niu, "CNN with Depthwise Separable Convolutions and Combined Kernels for Rating Prediction", *Expert Systems with Applications*, (2020), 114528.
- Kim, H. K., J. K. Kim., and Y. U. Ryu, "Personalized recommendation over a customer network for ubiquitous shopping", *IEEE Transactions on Services Computing*, Vol.2, No.2(2009), 140~151.

- Kim, J. K., H. K. Kim, H. Y. Oh, and Y. U. Ryu, "A group recommendation system for online communities", *International Journal of Information Management*, Vol.30, No.3(2010), 212~219.
- Knees, P., D. Schnitzer, and A. Flexer, "Improving neighborhood-based collaborative filtering by reducing hubness", *Proceedings of International Conference on Multimedia Retrieval*, (2014), 161~168.
- Koren, Y., and R. Bell, *Recommender systems handbook*, Springer, New York, USA, 2015.
- Krishnamoorthy, S., "Linguistic features for review helpfulness prediction", *Expert Systems with Applications*, Vol.42, No.7(2015), 3751~3759.
- Lee, D., and K. Hosanagar, "How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment", *Information Systems Research*, Vol.30, No.1(2019), 239~259.
- Lee, Y., H. Won, J. Shim, and H. Ahn, "A Hybrid Collaborative Filtering-based Product Recommender System using Search Keywords", *Journal of Intelligence and Information Systems*, Vol.26, No.1(2020), 151~166.
- Lei, X., X. Qian, and G. Zhao, "Rating prediction based on social sentiment from textual reviews", *IEEE Transactions on Multimedia*, Vol.18, No.9(2016), 1910~1921.
- Leung, C. W., S. C. Chan, and F. Chung, "Integrating Collaborative Filtering and Sentiment Analysis: A Rating Inference Approach", *Proceedings of the ECAI Workshop on Recommender Systems*, (2006), 62~68.
- Li, X., M. Wang, and T. P. Liang, "A multi-theoretical kernel-based approach to social network-based recommendation", *Decision Support Systems*, Vol.65, (2014), 95~104.
- Linden, G., B. Smith, and J. York, "Amazon.com recommendations: Item-to-item collaborative filtering", *IEEE Internet Computing*, Vol.7, No.1(2003), 76~80.
- Liu, Y., X. Huang, A. An, and X. Yu, "Modeling and predicting the helpfulness of online reviews", *8th IEEE International Conference on Data Mining*, (2008), 443~452.
- Lu, J., D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: a survey", *Decision Support Systems*, Vol.74, (2015), 12~32.
- Mandal, S., and A. Maiti, "Deep collaborative filtering with social promoter score-based user-item interaction: a new perspective in recommendation", *Applied Intelligence*, (2021), 1~26.
- Mishra, R., P. Kumar, and B. Bhasker, "A web recommendation system considering sequential information", *Decision Support Systems*, Vol.75, (2015), 1~10.
- Moon, H. S., D. Sung, and J. K. Kim, "An Analytical Approach Using Topic Mining for Improving the Service Quality of Hotels", *Journal of Intelligence and Information Systems*, Vol.25, No.1(2019), 21~41.
- Moon, H. S., J. H. Yoon, I. Y. Choi, and J. K. Kim, "An Exploratory Study of Collaborative Filtering Techniques to Analyze the Effect of Information Amount", *Asia Pacific Journal of Information Systems*, Vol.27, No.2(2017), 126~138.
- Moore, S. G., "Attitude predictability and helpfulness in online reviews: The role of explained actions and reactions", *Journal*

- of Consumer Research*, Vol.42, No.1(2015), 30~44.
- Na, H., and K. Nam, “Application of diversity of recommender system according to user preference change”, *Journal of Intelligence and Information Systems*, Vol.26, No.4(2020), 67~86.
- Nassirtoussi, A. K., S. Aghabozorgi, T. Y. Wah, and D. C. L. Ngo, “Text mining for market prediction: A systematic review”, *Expert Systems with Applications*, Vol.41, No.16(2014), 7653~7670.
- Ngo-Ye, T. L., and A. P. Sinha, “The influence of reviewer engagement characteristics on online review helpfulness: A text regression model”, *Decision Support Systems*, Vol.61, (2014), 47~58.
- Paradarami, T. K., N. D. Bastian, and J. L. Wightman, “A hybrid recommender system using artificial neural networks”, *Expert Systems with Applications*, Vol.83, (2017), 300~313.
- Park, D. H., H. K. Kim, I. Y. Choi and J. K. Kim, “A literature review and classification of recommender systems research”, *Expert Systems with Applications*, Vol.39, No.11(2012), 10059~10072.
- Polatidis, N., and C. K. Georgiadis “A multi-level collaborative filtering method that improves recommendations”, *Expert Systems with Applications*, Vol.48, (2016), 100~110.
- Postmus, S., and S. Bhulai, “Recommender system techniques applied to Netflix movie data”, *Research Paper Business Analytics*, Vrije Universiteit Amsterdam, Netherlands, 2018.
- Qiu, L., S. Gao, W. Cheng, and J. Guo, “Aspect-based latent factor model by integrating ratings and reviews for recommender system”, *Knowledge-Based Systems*, Vol.110, (2016), 233~243.
- Ricci, F., L. Rokach and B. Shapira, *Introduction to recommender systems handbook*, Springer, Boston, USA, 2011.
- Sánchez-Moreno, D., A. B. G. González, M. D. M. Vicente, V. F. L. Batista, and M. N. M. Garcia, “A collaborative filtering method for music recommendation using playing coefficients for artists and users”, *Expert Systems with Applications*, Vol.66, (2016), 234~244.
- Sarwar, B., G. Karypis, J. Konstan, and J. Riedl, “Item-based collaborative filtering recommendation algorithms”, *Proceedings of the 10th International Conference on World Wide Web*, (2001), 285~295.
- Saumya, S., and J. P. Singh, “Detection of spam reviews: a sentiment analysis approach”, *CSI Transactions on ICT*, Vol.6, No.2(2018), 137~148.
- Siering, M., A. V. Deokar, and C. Janze, “Disentangling consumer recommendations: Explaining and predicting airline recommendations based on online reviews”, *Decision Support Systems*, Vol.107, (2018), 52~63.
- Song, C., X. K. Wang, P. F. Cheng, J. Q. Wang, and L. Li, “SACPC: A framework based on probabilistic linguistic terms for short text sentiment analysis”, *Knowledge-Based Systems*, Vol.194, (2020), 105572.
- Srifi, M., A. Oussous, A. A. Lahcen, and S. Mouline, “Recommender Systems Based on Collaborative Filtering Using Review Texts—A

- Survey”, *Information*, Vol.11, No.6(2020), 317.
- Su, X., and T. M. Khoshgoftaar, “A survey of collaborative filtering techniques”, *Advances in Artificial Intelligence*, (2009).
- Ullah, I., M. Hussain, and H. Aboalsamh, “An automated system for epilepsy detection using EEG brain signals based on deep learning approach”, *Expert Systems with Applications*, Vol.107, (2018), 61~71.
- Wang, X., X. Lin, and M. K. Spencer, “Exploring the effects of extrinsic motivation on consumer behaviors in social commerce: Revealing consumers’ perceptions of social commerce benefits”, *International Journal of Information Management*, Vol.45, (2019), 163~175.
- Wang, X., Z. Dai., H. Li, and J. Yang, “Research on Hybrid Collaborative Filtering Recommendation Algorithm Based on the Time Effect and Sentiment Analysis”, *Complexity*, (2021).
- Wei, S., N. Ye, S. Zhang, X. Huang, and J. Zhu, “Item-based collaborative filtering recommendation algorithm combining item category with interestingness measure”, *International Conference on Computer Science and Service System*, (2012), 2038~2041.
- Wu, P., X. Li., S. Shen, and D. He, “Social media opinion summarization using emotion cognition and convolutional neural networks”, *International Journal of Information Management*, Vol.51, (2020), 101978.
- Yoo, S., J. Song, and O. Jeong, “Social media contents based sentiment analysis and prediction system”, *Expert Systems with Applications*, Vol.105, (2018), 102~111.
- Yun, Y., D. Hooshyar, J. Jo, and H. Lim, “Developing a hybrid collaborative filtering recommendation system with opinion mining on purchase review”, *Journal of Information Science*, Vol.44, No.3(2018), 331~344.
- Zafari, F., I. Moser, and T. Sellis, “ReEx: An integrated architecture for preference model representation and explanation”, *Expert Systems with Applications*, Vol.161, (2020), 113706.
- Zhang, Y., and B. Wallace, “A sensitivity analysis of (and practitioners’ guide to) convolutional neural networks for sentence classification”, *arXiv preprint arXiv:1510.03820*, (2015).
- Zhang, Z., and B. Varadarajan, “Utility scoring of product reviews”, *Proceedings of the 15th ACM International Conference on Information and Knowledge Management*, (2006), 51~57.
- Zhang, Z., D. Zhang, and J. Lai, “urCF: User Review Enhanced Collaborative Filtering”, *Proceedings of the 20th Americas Conference on Information Systems*, (2014).
- Zhang, Z., H. Lin, K. Liu, D. Wu, G. Zhang, and J. Lu, “A hybrid fuzzy-based personalized recommender system for telecom products/services”, *Information Sciences*, Vol.235, (2013), 117~129.
- Zheng, L., V. Noroozi, and S. Yu, “Joint deep modeling of users and items using reviews for recommendation”, *Proceedings of the 10th ACM International Conference on Web Search and Data Mining*, (2017), 425~434.
- Zhou, L., and P. Chaovalit, “Ontology-Supported Polarity Mining”, *Journal of the American*

Society for Information Science and Technology,
Vol.59, No.1(2008), 98~110.

Zhu, T., Y. Ren, W. Zhou, J. Rong, and P.
Xiong, “An effective privacy preserving

algorithm for neighborhood-based collaborative
filtering”, *Future Generation Computer
Systems*, Vol.36, (2014), 142~155.

Abstract

A Study on Enhancing Personalization Recommendation Service Performance with CNN-based Review Helpfulness Score Prediction

Qinglong Li* · Byunghyun Lee* · Xinzhe Li* · Jae Kyeong Kim**

Recently, various types of products have been launched with the rapid growth of the e-commerce market. As a result, many users face information overload problems, which is time-consuming in the purchasing decision-making process. Therefore, the importance of a personalized recommendation service that can provide customized products and services to users is emerging. For example, global companies such as Netflix, Amazon, and Google have introduced personalized recommendation services to support users' purchasing decisions. Accordingly, the user's information search cost can reduce which can positively affect the company's sales increase. The existing personalized recommendation service research applied Collaborative Filtering (CF) technique predicts user preference mainly use quantified information. However, the recommendation performance may have decreased if only use quantitative information. To improve the problems of such existing studies, many studies using reviews to enhance recommendation performance. However, reviews contain factors that hinder purchasing decisions, such as advertising content, false comments, meaningless or irrelevant content. When providing recommendation service uses a review that includes these factors can lead to decrease recommendation performance. Therefore, we proposed a novel recommendation methodology through CNN-based review usefulness score prediction to improve these problems. The results show that the proposed methodology has better prediction performance than the recommendation method considering all existing preference ratings. In addition, the results suggest that can enhance the performance of traditional CF when the information on review usefulness reflects in the personalized recommendation service.

Key Words : Personalized Recommendation Service, Review Helpfulness, CF, Deep Learning, CNN

Received : May 21, 2021 Revised : June 29, 2021 Accepted : July 12, 2021

Corresponding Author : Jaekyeong Kim

* Department of Big Data Analytics, Kyung Hee University

** Corresponding Author: Jaekyeong Kim

School of Management & Department of Big Data Analytics, Kyung Hee University

26, Kyunghedae-ro, Dongdaemun-gu, Seoul, 02447, Korea

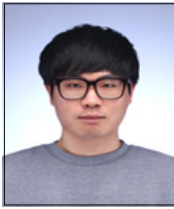
Tel: +82-2-961-9355, Fax: +82-2-961-9355, E-mail: jaek@khu.ac.kr

저 자 소개



이청용

경희대학교 경영학과에서 경영학 학사학위를 취득하고, 동 대학원 빅데이터응용학과에서 공학 석사학위를 취득하였다. 현재 빅데이터응용학과 박사과정에 재학 중이며, 주요 관심 분야로는 개인화 추천 서비스, 자연어 처리, 데이터 마이닝, 빅데이터 분석, 딥러닝 등이다.



이병현

경희대학교 대학원 빅데이터응용학과에서 경영학 석사학위를 취득하고, 현재 빅데이터응용학과 박사과정에 재학 중이다. 주요 관심 분야는 개인화 추천 서비스, 딥러닝, 자연어 처리, 데이터 마이닝, 소셜 네트워크 분석 등에 관심을 두고 연구하고 있다.



이흠철

경희대학교 대학원 빅데이터응용학과 석사과정에 재학 중이며, 주요 연구 관심 분야는 추천시스템, 딥러닝, 자연어 처리 및 빅데이터에 분석 등이다.



김재경

서울대학교에서 산업공학학사, 한국과학기술원에서 경영정보시스템 전공으로 석사 및 박사학위를 취득하였으며 현재 경희대학교 경영대학 및 빅데이터응용학과 교수로 재직하고 있다. 미국 미네소타 주립대학교 그리고 텍사스 주립대학교(달라스)에서 교환교수를 역임하였다. 주요 관심 분야로는 개인화 서비스, 추천시스템, 빅데이터 및 딥러닝 등이다. IEEE Transaction on Services Computing, IEEE Transaction on SMC-A, International Journal of Human Computer Studies, International Journal of Information Management, Information and Management, Expert Systems with Applications, Applied Artificial

Intelligence, 등 다수의 학술지에 논문을 게재하였다. 현재 4단계 BK21사업 연구단장 (빅데이터 분야) 및 AI 비즈니스 연구센터 센터장을 맡고 있다.