

온라인 호텔 리뷰와 평점 불일치 문제 해결을 위한 딥러닝 기반 개인화 추천 서비스 연구

A Study of Deep Learning-based Personalized Recommendation Service for Solving Online Hotel Review and Rating Mismatch Problem

이 청 용 (Qinglong Li) 경희대학교 대학원 빅데이터응용학과 박사과정
최 사 박 (Shibo Cui) 경희대학교 대학원 빅데이터응용학과 석사
신 병 규 (Byunggyu Shin) 경희대학교 대학원 경영학과 박사수료, 교신저자
김 재 경 (Jaekyeong Kim) 경희대학교 경영대학/빅데이터응용학과 교수

요 약

세계적인 전자상거래 기업들은 지속 가능한 경쟁력을 확보하기 위해 사용자 맞춤형 추천 서비스를 제공하고 있다. 기존 관련 연구에서는 주로 평점, 구매 여부 등 정량적 선호도 정보를 사용하여 개인화 추천 서비스를 제공하였다. 하지만 이와 같은 정량적 선호도 정보를 사용하여 개인화 추천 서비스를 제공하면 추천 성능이 저하될 수 있다는 문제점이 제기되고 있다. 호텔을 이용한 사용자가 호텔 서비스, 청결 상태 등에 대하여 만족하지 못한다고 리뷰를 작성하였으나 선호도 평점 5점을 부여했을 때 정량적 선호도(평점)와 정성적 선호도(리뷰)가 불일치한 문제가 발생할 수 있다. 따라서 본 연구에서는 정량적 선호도 정보와 정성적 선호도 정보가 일치하는지를 확인하고 이를 바탕으로 선호도 정보가 일치하는 사용자를 바탕으로 새로운 프로파일을 구축하여 개인화 추천 서비스를 제공하고자 한다. 리뷰에서 정성적 선호도를 추출하기 위해 자연어 처리 관련 연구에서 널리 사용되고 있는 CNN, LSTM, CNN + LSTM 등 딥러닝 기법을 사용하여 감성분석 모델을 구축하였다. 이를 통해 사용자가 작성한 리뷰에서 정성적 선호도 정보를 정교하게 추출하여 정량적 선호도 정보와 비교하였다. 본 연구에서 제안한 추천 방법론의 성능을 평가하기 위해 세계 최대 여행 플랫폼 TripAdvisor에서 실제 호텔을 이용한 사용자 선호도 정보를 수집하여 사용하였다. 실험 결과 본 연구에서 제안한 추천 방법론이 기존의 정량적 선호도만을 고려하는 추천 방법론보다 우수한 추천 성능을 나타냄을 확인할 수 있었다.

키워드 : 개인화 추천 서비스, 협업 필터링, 감성분석, 딥러닝

I. 서 론

최근 정보통신기술의 발전과 스마트폰의 대중

화로 인해 온라인 전자상거래 시장은 폭발적으로 성장하고 있으며, 이에 따라 온라인 쇼핑 거래 규모도 꾸준히 성장하고 있다. 이로 인해 매년 수많은

제품과 서비스가 새롭게 출시되면서 사용자들의 접근성과 편의성은 향상되었지만, 구매 의사결정 과정에서 정보탐색 비용이 증가하는 문제점이 발생하고 있다. 즉, 수많은 제품 중에서 자신의 선호도에 적합한 제품을 선택할 때 많은 시간이 걸리고 어려움을 겪는 정보 과부하(Information Overload) 문제에 직면하고 있다(Park *et al.*, 2012; Su and Khoshgoftaar, 2009). 특히, 최근에는 온라인 쇼핑에 대한 수요가 급증하고 있지만, 사용자들은 선호하는 제품이나 서비스를 직접 확인하고 체험하는 데 한계를 가지고 있으므로 정보 과부하 문제가 더 크게 드러날 수 있다. 또한, 기업 측면에서도 자사의 제품이나 서비스를 선호하고 구매할 가능성이 큰 사용자에게 홍보하고 노출할 기회가 감소하여 수익 창출에 어려움을 겪고 있다. 이러한 배경하에 사용자에게 맞춤형 제품이나 서비스를 효과적으로 추천하는 개인화 추천 서비스의 중요성이 제기되고 있다. 대표적으로 Amazon, Netflix, Google 등 세계적인 전자상거래 기업들은 지속 가능한 기업 경쟁력을 강화하기 위해 개인화 추천 서비스를 제공하고 있다(Bennett and Lanning, 2007; Das *et al.*, 2007; Linden *et al.*, 2003). 예를 들어, Amazon은 기업 매출의 35%가 개인화 추천 서비스가 제공하는 제품이나 서비스를 통해 발생하고 있으며, Netflix는 사용자가 시청한 전체 동영상의 75%가 개인화 추천 서비스를 통해 제공되고 있다(Nguyen *et al.*, 2014). 이처럼 개인화 추천 서비스는 사용자의 정보탐색 비용을 줄일 수 있으며 기업의 수익 창출에도 긍정적인 영향을 주고 있다(Li *et al.*, 2014).

기존 개인화 추천 서비스 관련 연구는 주로 평점 등 사용자가 제품에 대한 선호도를 정량적으로 표현하는 명시적 피드백(Explicit Feedback) 데이터와 구매 혹은 방문 여부 등을 나타내는 암묵적 피드백(Implicit Feedback) 데이터를 사용하여 사용자의 선호도를 예측하였다(Moon *et al.*, 2017; Su and Khoshgoftaar, 2009). 그중에서도 많은 연구는 주로 사용자의 선호도 정보를 정량적으로 나타낼

수 있는 선호도 평점을 활용하였다(Choi *et al.*, 2016; Li *et al.*, 2014; Ricci *et al.*, 2011). 이처럼 많은 연구에서 사용되는 평점 정보는 추천 시스템을 구축할 때 선호도 정보를 활용하기 편리하고 수리적으로 분석하고 처리하기가 쉽다는 장점이 있다(전병국, 안현철, 2015; 현지연 등, 2019). 그러나 정량적 평점만을 사용하면 사용자의 선호도를 제대로 고려하지 않기 때문에 추천 성능을 저하할 수 있다는 문제점이 제기되고 있다(Seo *et al.*, 2017; Zhang *et al.*, 2014; 조승연 등, 2017).

최근에는 정량적 평점만을 사용하는 개인화 추천 서비스 관련 연구의 한계를 개선하기 위해 다양한 유형의 데이터를 추가로 고려하는 연구들이 꾸준히 증가하고 있다(Zhang *et al.*, 2019). 대표적으로 제품이나 서비스에 대한 구매 동기 혹은 선호 이유 등을 확인할 수 있는 리뷰 데이터를 많이 사용하고 있다(Cheng *et al.*, 2018; Lei *et al.*, 2016; Zheng *et al.*, 2017). 사용자가 작성한 리뷰에는 구매 동기, 제품을 선호하는 이유 등 구체적이고 신뢰할 수 있는 정성적 선호도 정보가 포함되기에 유용하게 사용할 수 있다(Lei *et al.*, 2016; Tang *et al.*, 2013). 기존 리뷰를 활용한 개인화 추천 서비스 관련 연구에서는 주로 리뷰의 감성 특성을 분석하여 추천 시스템에 반영하는 연구에 중점을 두었다(Paradarami *et al.*, 2017; 전병국, 안현철, 2015; 현지연 등, 2019). 그러나 정량적 선호도 정보와 리뷰에 포함된 정성적 선호도 정보가 일치하지 않으면 추천 성능이 저하될 수 있는 문제가 발생할 수 있다. 예를 들어, 사용자 A는 특정 제품을 구매하고 선호도 평점 5점을 부여하고, 리뷰를 통해 해당 제품의 포장 상태, 배송 속도, 디자인 등을 포함하여 모두 만족한다고 작성했다. 반대로 사용자 B는 선호도 평점 5점을 부여했으나 해당 제품의 디자인, 제품 구성 등에 대하여 만족스럽지 못하다고 리뷰를 작성했다. 이처럼 사용자가 특정 제품이나 서비스를 구매하고 부여한 정량적 선호도 정보와 사용자가 작성한 리뷰에 포함되는 정성적 선호도 정보가 서로 일치하지 않다는 것을 알 수 있다. 이와

같은 문제점은 개인화 추천 서비스의 성능을 저하하는 중요한 요인이므로 개선이 필요하다.

따라서 본 연구의 목적은 리뷰에 포함된 정성적 선호도 정보와 정량적 선호도 정보가 일치하지 않는 문제를 해결하여 개인화 추천 서비스의 추천 성능을 높이는 것이다. 이를 위해 정성적 선호도 정보와 정량적 선호도 정보가 일치하는 사용자를 바탕으로 새롭게 프로파일을 구축하여 사용자 맞춤형 추천 서비스를 제공하는 새로운 추천 방법론을 제안하고자 한다. 먼저, 사용자가 작성한 리뷰에 포함된 정성적 선호도의 의미론적(Semantic) 특성을 효과적으로 추출하기 위해 자연어 처리 분야에서 우수한 성능을 보이는 CNN(Convolution Neural Network), LSTM(Long Short-Term Memory) 및 CNN + LSTM 결합 모델을 적용하여 감성분석 모델을 구축하였다(Zhang *et al.*, 2018). 이를 통해 사용자가 작성한 리뷰에서 정성적 선호도 정보의 의미론적 감성 특성을 정교하게 추출하여 정량적 선호도 정보와 비교하고자 한다. 다음으로 선호도 정보가 일치하는 사용자 프로파일을 바탕으로 개인화 추천 서비스 관련 연구에서 널리 사용되는 협업 필터링 알고리즘을 적용하여 사용자의 선호도를 예측한다(Park *et al.*, 2012). 본 연구에서는 기존 개인화 추천 서비스 관련 연구에서 시도하지 않은 사용자가 작성한 리뷰에 포함된 정성적 선호도 정보와 정량적 선호도 정보가 일치하는지를 판단하여 추천 성능을 높이는 새로운 추천 방법론을 제안하였다. 이를 통해 개인화 추천 서비스를 제공하고 있거나 도입을 계획하고 있는 전자상거래 기업에 추천 시스템 알고리즘과 방법론을 제공하여 사용자 맞춤형 추천 서비스 전략에 대한 다양한 시사점을 제공할 수 있을 것으로 기대한다.

본 연구의 나머지 구성은 다음과 같다. 제Ⅱ장에서는 개인화 추천 서비스, 감성분석 및 사용자 리뷰를 반영한 개인화 추천 서비스에 대한 이론적 배경과 관련 연구에 관하여 기술한다. 제Ⅲ장에서는 본 연구에서 제안하는 추천 방법론에 대하여 구체적으로 기술한다. 제Ⅳ장에서는 실험 데이터,

추천 성능 평가 방법과 실험 결과를 서술한다. 마지막으로 제Ⅴ장에서는 연구 결과 토의, 한계점과 추후 연구 계획 및 시사점에 대하여 구체적으로 기술한다.

Ⅱ. 관련 연구

2.1 개인화 추천 서비스

개인화 추천 서비스의 관련 연구는 주로 사용자가 직접 부여한 선호도 평점, 구매 기록, 행동 패턴, 인구 통계학적 정보 등을 사용하여 사용자 맞춤형 제품이나 서비스를 제공하는 것을 목적으로 한다(Kim *et al.*, 2010; Park *et al.*, 2012). 대표적으로 Netflix, Amazon, Google 등 세계적인 전자상거래 기업들은 기업의 지속 가능한 경쟁력 강화를 위해 개인화 추천 서비스를 도입하여 사용자에게 맞춤형 서비스를 제공하고 있다(Bennett and Lanning, 2007; Das *et al.*, 2007; Linden *et al.*, 2003). 국내에서 네이버는 AiRS 알고리즘을 기반으로 사용자 맞춤형 뉴스를 추천하고 있으며, 카카오는 루빅스(RUBICS) 알고리즘을 기반으로 사용자 반응을 분석하여 맞춤형 뉴스 콘텐츠를 실시간으로 제공하고 있다(Park *et al.*, 2017; 박승택 등, 2017). 이처럼 국내외 다양한 분야에서 사용자에게 맞춤형 정보를 제공하기 위한 개인화 추천 서비스 관련 연구가 꾸준히 진행되고 있다.

개인화 추천 서비스 관련 연구에서 널리 사용되고 있는 추천 알고리즘은 협업 필터링(Collaborative Filtering, CF) 기법이다. CF 기법은 1990년대에 학계에 처음으로 제안된 이후 현재까지 우수한 성능을 보여주면서 다양한 분야의 연구에서 널리 사용되고 있다(Konstan *et al.*, 1997). CF 기법은 추천 대상으로 선택된 사용자와 선호도가 가장 유사한 사용자를 찾고 그중에서 선호도가 높은 제품이나 서비스를 추천하는 것을 목표로 한다(Kim *et al.*, 2009; Su and Khoshgoftaar, 2009). 개인화 추천 서비스 연구에서 CF 기법은 주로 메모리 기반 방법

과 모델 기반 방법으로 나눌 수 있다(Park *et al.*, 2012). 메모리 기반 방법은 크게 사용자 기반 협업 필터링(User-based Collaborative Filtering, UBCF)과 아이템 기반 협업 필터링(Item-based Collaborative Filtering, IBCF)으로 구성된다. UBCF는 사용자 간의 선호도 평점 유사도를 계산하여 선호도가 유사한 다른 사용자가 평가한 제품이나 서비스를 기반으로 추천 대상 사용자의 선호도 평점을 예측한다. 반대로 IBCF는 사용자가 선호도 평점을 부여한 기존 제품과 추천하고자 하는 제품 간의 유사도를 계산하여 추천 대상 사용자가 선호하는 제품을 예측한다. 메모리 기반 방법은 개인화 추천 서비스를 제공할 때 실제 모델을 구축하지 않고 추천을 제공할 때 연산을 수행하는 휴리스틱 기법이다(Adomavicius and Tuzhilin, 2005). CF 기법은 기본적으로 추천 대상 사용자의 구매 내역, 선호도 평점 등 정보를 기반으로 사용자 혹은 제품 간의 유사도를 계산하여 맞춤형 추천 서비스를 제공한다. 하지만 새로운 사용자의 경우 유사도를 계산할 수 있는 구매 내역이나 선호도 평점 등 정보가 존재하지 않을 수 있다. CF 기법은 새로운 사용자의 경우 선호도를 예측하기 어려운 “Cold Start” 문제점과 특정 제품이 아직 구매가 이루어지지 않으면 추천 목록을 생성할 수 없는 “First Start” 문제점이 존재한다. 이처럼 사용자의 선호도 평점, 구매 내역 등 정보가 부족하여 발생하는 문제를 희소성(Sparsity) 문제점이라고 한다(Herlocker *et al.*, 2000). 또한, 사용자의 구매 내역 혹은 선호도 평점 정보가 꾸준히 증가하게 되면 추천 알고리즘의 연산 규모가 커지게 되어 연산 처리에 많은 시간과 비용이 발생하는 확장성(Scalability) 문제점도 존재하고 있다. 이와 같은 문제점을 개선하기 위해 다양한 연구가 진행되고 있으며, 그중에서도 데이터 학습을 통해 모델을 구축하여 추천 서비스를 제공하는 모델 기반 방법이 우수한 성능을 나타내고 있다. 이 방법은 추천 서비스 과정에서 데이터 마이닝이나 기계학습 기법을 적용하여 기존 메모리 기반 방법이 가지고 있는 다양한 문제점을 일부

개선하였다(Hofmann, 2004). 모델 기반 방법에서 널리 사용되고 있는 대표적인 기법은 데이터의 차원을 축소하는 SVD(Singular Value Decomposition)이다. SVD 기법은 사용자 선호도 행렬에서 중요하지 않은 사용자나 제품을 제외하고 행렬의 차원을 축소하는 방법이다. 이처럼 데이터의 차원을 축소하면 기존 데이터보다 연산할 때 정보 손실을 감소시켜 사용자 선호도 예측에 우수한 성능을 보여주고 있다(Sarwar *et al.*, 2000).

기존 CF 기법을 활용한 개인화 추천 서비스 관련 연구들은 주로 선호도 평점, 구매 내역 등 데이터를 사용하여 먼저 유사한 이웃 사용자를 선택하고 이를 통해 추천 대상 사용자의 선호도를 예측하였다(Park *et al.*, 2012). 그러나 이와 같은 방법은 주로 정량적인 데이터를 사용하므로 사용자 선호도를 제대로 반영하지 못하여 추천 성능이 저하되는 문제가 발생한다(Seo *et al.*, 2017; Zhang *et al.*, 2014; 조승연 등, 2017). 본 연구에서는 먼저 사용자가 부여한 정량적 선호도 정보와 리뷰의 정성적 선호도 정보를 비교하여 일치하는지를 확인한다. 이를 통해 선호도 정보가 일치한 사용자 프로파일을 구축하여 추천 서비스를 제공하여 정량적 선호도 정보와 정성적 선호도 정보가 일치하지 않는 문제로 인해 추천 성능이 저하되는 문제점을 해결할 수 있을 것을 기대한다.

2.2 감성분석

감성분석은 자연어 처리 기법의 일부분으로 주로 리뷰에서 사용자의 감성과 주관적인 의견을 종합적으로 분류하는 것을 목표로 하고 있다(Christopher Frey and Patil, 2002). 감성분석 연구는 주로 사전 기반 방법(Lexicon-based Approach)과 기계학습 기반 방법(Machine Learning Approach)으로 나누고 있다(Borgonovo and Plischke, 2016). 사전 기반 방법은 긍정 혹은 부정 단어로 구성된 감성 사전을 기반으로 리뷰에 포함되는 감성 단어 빈도 등을 종합적으로 파악하여 감성을 분류한다.

반대로 기계학습 기반 방법은 리뷰에 포함된 감성 특성이나 맥락을 정교하게 학습시켜 이를 기반으로 감성을 분류하고 있다. 특히, 기계학습 기반 방법은 많은 양의 리뷰를 대상으로 감성분석을 수행할 때 우수한 성능을 나타내어 최근 많은 연구에서 사용되고 있다(Borgonovo and Plischke, 2016).

기계학습 기반 방법에서 대표적으로 Decision Tree, Support Vector Machines(SVM), Naive Bayes, Maximum Entropy, Artificial Neural Network 등 기법들이 널리 사용되고 있다. Pang *et al.*(2002)은 대표적인 SVM, Naive Bayes, Maximum Entropy 등 기계학습 기법을 사용하여 긍정 리뷰와 부정 리뷰를 각 700개를 포함하는 영화 리뷰에 적용하여 감성분석을 수행했다. 실험 결과는 연구에서 제안한 기계학습 기반 방법이 우수한 분류 정확도를 보여주었으며, 그중에서도 SVM 기법이 가장 우수한 성능을 나타냈다. 이외에도 다양한 연구에서 SVM 기법을 감성분석에 적용하여 우수한 성능을 입증하였다(Annett and Kondrak, 2008; Kennedy and Inkpen, 2006; König and Brill, 2006). Ferguson *et al.*(2009)은 Naive Bayes 기법을 적용한 다중 분류 방법론을 제안하여 금융 관련 리뷰를 감성분석을 수행한 결과 기존 SVD 기법을 적용하는 방법보다 우수한 성능을 나타냄을 확인하였다. 해당 연구에서는 감성 특성을 긍정 혹은 부정으로 분류할 뿐만 아니라 긍정, 부정 및 중립으로 분류를 시도하였으며, 실험 결과는 단순히 긍정과 부정으로 분류할 때 분류 성능이 우수하게 나타났다. Pak and Paroubek(2010)가 제안한 방법론에서도 Naive Bayes 기법을 사용한 다중 분류 방법론이 SVM 기법보다 우수한 성능을 나타냈음을 확인했다. 또한, 해당 연구에서는 온라인 트위터 리뷰를 사용하여 실험을 수행하였으며 실험 결과로는 n-gram 과 POS tagging 기법을 적용했을 때 더 우수한 성능을 나타냈다. 최근에는 기존의 감성분석 기법의 한계를 보완하기 위해 딥러닝 기법을 적용하는 연구가 꾸준히 증가하고 있다(Zhang *et al.*, 2018). 감성분석에 딥러닝 기법을 적용하면 기존의 분석

방법에서 입력 특성들의 차원 수가 증가할수록 분류 성능이 저하되는 문제점을 개선할 수 있다. 자연어 처리 관련 연구에서 대표적으로 CNN을 사용하여 감성분석을 진행하면 우수한 분류 성능을 나타내고 있다(Zhang and Wallace, 2015). Kim(2014)은 Word2Vec 기법을 적용하여 텍스트를 단어 벡터로 변환하고 이를 사용하여 CNN을 통해 학습한 결과 우수한 분류 성능을 나타낼 수 있음을 확인하였다. Lai *et al.*(2015)은 Word2Vec 기법을 통해 텍스트를 단어 벡터로 변환하고 RNN(Recurrent Neural Networks)을 적용하여 텍스트를 분류한 결과 우수한 분류 정확도를 나타냄을 확인하였다. 또한, Tai *et al.*(2015)은 기존 RNN의 장기 의존성 문제점을 개선한 LSTM을 적용하여 감성분석을 수행하여 우수한 성능을 확인하였다.

본 연구에서는 감성분석 관련 연구에서 우수한 성능을 나타낸 여러 유형의 딥러닝 기법을 적용하여 사용자가 작성한 리뷰에서 정성적 선호도의 의미론적 감성 특성을 정교하게 추출하고자 한다. 사용자가 작성한 리뷰에는 제품을 선호하는 이유 등과 같이 구체적이고 신뢰할 수 있는 정보를 포함하고 있으므로 유용하게 사용할 수 있다(Lei *et al.*, 2016; Tang *et al.*, 2013). 따라서 본 연구에서는 딥러닝 기반 감성분석 모델을 구축하고 이를 통해 리뷰의 정성적 선호도를 정교하게 예측하고자 한다. 이를 통해 정량적 선호도 정보와 정성적 선호도 정보가 일치하는 사용자를 바탕으로 프로파일을 구축하고 이를 통해 개인화 추천 서비스를 제공하여 추천 성능을 향상하고자 한다.

2.3 사용자 리뷰를 반영하는 개인화 추천 서비스

기존 개인화 추천 서비스 연구는 주로 선호도 평점 등 사용자가 직접 제품에 대한 선호도를 정량적으로 표현하는 명시적 피드백 데이터와 구매 혹은 방문 여부 등 나타내는 암묵적 피드백 데이터를 사용하여 사용자의 선호도를 예측했다. 하지

만 사용자가 직접 작성한 리뷰 등 정성적 선호도를 고려하지 않고 평점, 구매 여부 등 정량적 선호도 정보만을 고려하면 추천 성능을 저하하는 문제점이 꾸준히 제기되었다(Seo *et al.*, 2017; Zhang *et al.*, 2014). 최근에는 이러한 기존 연구의 한계를 개선하기 위해 리뷰에 포함된 정성적 선호도 정보를 정교하게 분석하여 추천 시스템에 반영하는 연구가 활발히 진행되고 있다(Cheng *et al.*, 2018; Lei *et al.*, 2016; Zheng *et al.*, 2017).

개인화 추천 서비스에 사용자가 작성한 리뷰를 고려한 연구는 Leung *et al.*(2006)이 처음으로 제안하였다. 이 연구에서는 영화 리뷰를 사용하여 사용자가 작성한 리뷰가 긍정 혹은 부정인지 분류하고 감성의 강도를 측정할 수 있는 모형을 개발하고 이를 기반으로 추천 서비스를 제공하는 방법을 제안했다. 해당 연구는 처음으로 사용자의 정성적 선호도를 개인화 추천 서비스에 반영했다는 측면에서는 시사하는 바가 크다. 하지만 사용자의 정량적 선호도와 정성적 선호도를 모두 고려하면 우수한 추천 성능을 나타낼 수 있었음에도 정량적 선호도는 반영하지 못한 한계점을 가지고 있다. Jakob *et al.*(2009)은 사용자가 작성한 리뷰를 사용하여 하이브리드 추천 방법론을 제안하였다. 이 연구는 사용자가 작성한 리뷰에 포함된 상품 특성을 추출하여 새로운 사용자 프로파일을 구축하고 이를 기반으로 맞춤형 추천 서비스를 제공하는 새로운 방법론을 제안했다. 실험 결과 해당 연구에서 제안한 방법론이 기존의 방법론에 비해 통계적으로 유의하게 성능이 개선되었음을 확인했다. Moshfeghi *et al.*(2011)과 Levi *et al.*(2012)은 CF 기법이 가지고 있는 한계점을 개선하기 위해 사용자 리뷰 마이닝 기법을 적용하는 방법론을 제안했다. 이 연구에서는 제안 방법론을 호텔과 영화 분야 개인화 추천 서비스에 적용하고 실증 분석을 수행하여 제안한 방법론이 새로운 사용자에게 만족할 수 있는 추천 서비스를 제공하였음을 확인했다. García-Cumbreras *et al.*(2013)은 사용자가 작성한 리뷰를 사용하여 감성분석을 수행하여 해당 사용

자가 낙관주의자 혹은 비관주의자인지 분류하고 각 집단별로 CF 기법을 적용하여 집단을 분류하는 기준으로 리뷰를 사용하는 새로운 접근 방식을 제안했다. Zheng *et al.*(2017)은 기존 연구의 데이터 희소성 등 문제를 해결하기 위해 리뷰에 포함되는 제품 속성과 사용자 행동을 정교하게 학습하기 위해 DeepCoNN 방법론을 제안했다. 이 연구는 사용자가 작성한 리뷰와 제품에 대하여 작성한 리뷰에 각각 CNN을 적용하여 사용자와 제품의 상호관계를 학습시키고 최종적으로 사용자 선호도 평점을 예측하였다. 실험 결과 연구에서 제안한 방법론이 기존에 추천 방법론에 비해 우수한 성능을 나타냄을 확인하였다.

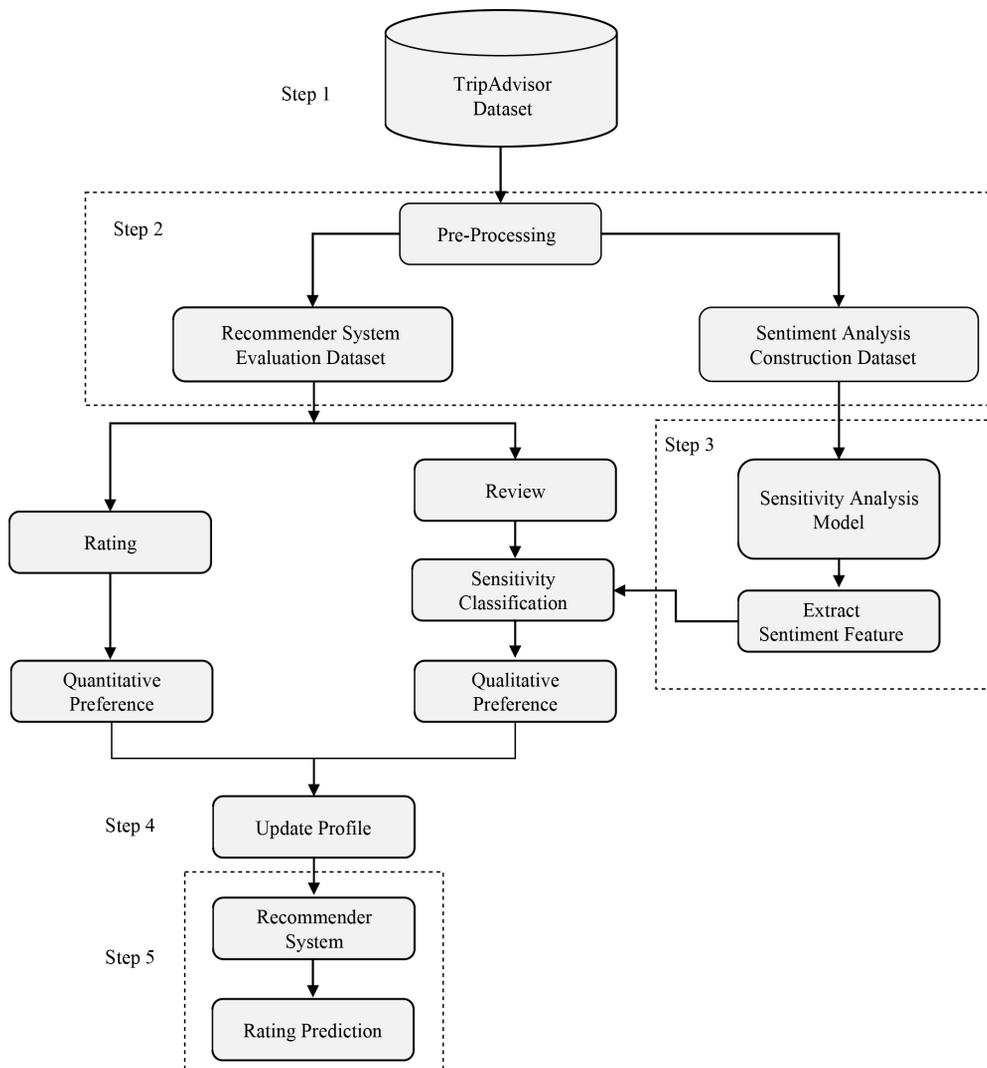
기존 리뷰를 활용한 개인화 추천 서비스 관련 연구는 주로 리뷰의 감성 특성을 분석하여 추천 시스템에 반영하는 연구에 중점을 두었다. 그러나 정량적 선호도 정보와 정성적 선호도 정보가 불일치한 선호도 정보를 활용하여 개인화 추천 서비스를 제공하는 것은 추천 성능을 저하할 수 있는 요소가 될 수 있다. 본 연구는 여러 유형의 딥러닝 기법을 적용한 감성분석 모델을 통해 정량적 선호도와 리뷰의 정성적 선호도 정보의 일치하는지를 먼저 판단하고자 한다. 다음으로 선호도 정보가 일치하는 사용자 프로파일을 생성하고 이를 통해 최종적으로 사용자의 선호도를 예측하는 추천 시스템을 구축하고자 한다.

III. 추천 방법론

본 연구에서는 사용자가 작성한 리뷰에 포함된 정성적 선호도 정보와 정량적 선호도 정보가 일치하는지를 확인하고 이를 통해 선호도 정보가 일치하는 사용자를 바탕으로 새롭게 프로파일을 구축하여 추천 서비스를 제공하는 새로운 추천 방법론을 제안하고자 한다. 본 연구에서 제안하는 방법론은 <그림 1>과 같이 데이터 수집, 데이터 정제 및 분할, 감성분석 모델 구축, 사용자 프로파일 구축, 추천 시스템 구축 등 5단계로 구성된다. 먼저

제안한 방법론의 성능을 평가하기 위해 세계 최대 여행 플랫폼 TripAdvisor에서 사용자, 호텔, 리뷰 및 선호도 평점을 수집하였다. 수집된 데이터에서 리뷰는 딥러닝 모델 학습을 위한 입력 데이터로 사용하기 위해 먼저 정제하는 작업을 수행하고 감성분석 모델 구축을 위한 데이터와 추천 성능 평가를 위한 데이터로 나누었다. 사용자가 작성한 리뷰에서 정성적 선호도 특성을 효과적으로 추출

하기 위해 자연어 처리 분야에서 우수한 성능을 나타내는 CNN, LSTM 및 CNN + LSTM을 적용하여 감성분석 모델을 구축하였다. 이를 통해 사용자가 작성한 리뷰에 포함된 정성적 선호도와 정량적 평점 선호도와 비교하여 선호도 정보가 일치하는 사용자를 바탕으로 새롭게 프로파일을 구축하였다. 마지막으로 개인화 추천 서비스 연구에서 널리 사용되고 있는 CF 기법을 적용하여 최종적



〈그림 1〉 추천 방법론 프로세스

으로 사용자의 선호도를 예측하였다. 구체적인 연구의 절차는 다음과 같다.

3.1 1단계: 데이터 수집

먼저 본 연구에서 제안하는 추천 방법론의 성능을 평가하기 위해 세계 최대 여행 플랫폼 TripAdvisor 사이트에서 미국 뉴욕 호텔을 방문한 사용자들의 선호도 데이터를 수집하였다. 이를 위해 웹 크롤링(Web Crawling) 기법을 사용하여 사용자, 호텔, 선호도 평점, 리뷰 등 정보를 수집하였으며, 선호도 평점은 1점부터 5점 사이의 값을 나타내고 있다. 감성분석을 위해 선호도 평점이 1점과 2점의 경우 부정(Negative), 3점의 경우 중립(Neutral), 4점과 5점의 경우에는 긍정(Positive)으로 분류한다(Borgonovo and Plischke, 2016). 전체 호텔 중에서 사용자가 선호도 평점을 부여한 호텔보다 부여하지 않은 호텔의 개수가 많으면 희소성 문제점이 발생한다. 본 연구에서는 이러한 희소성 문제점을 방지하기 위해 적어도 5개 이상의 호텔에 대하여 선호도 평점을 부여한 사용자를 선택하여 실험에 사용했다.

3.2 2단계: 텍스트 데이터 정제 및 분할

온라인에서 수집한 텍스트 데이터를 바로 분석에 사용할 수 있는 것은 아니므로 분석 목적에 적합하게 정제하는 작업이 필요하다(Kazmaier and van Vuuren, 2020). 특히, 사용자가 작성한 리뷰에는 모든 단어가 의미를 나타내는 것은 아니다. 예를 들어 “service”, “perfect”, “good” 등 단어는 사용자의 감성 특성을 정확하게 나타내고 있지만 “I”, “you” “are” 등은 의미가 있는 감성 특성을 나타내지 못한다. 리뷰에서 감성 특성을 추출할 수 있는 단어를 사용하여 분석하면 데이터의 크기를 줄이고 연산 처리 시간 및 분석 효율을 높일 수 있다. 온라인에서 수집한 리뷰는 HTML 태그, 특수 부호, 기호 등 특수 문자를 포함하고 있으나 실제

분석에서는 모델의 분류 성능을 저하하므로 제거해야 한다. 본 연구에서는 정규 표현식을 정의하여 해당 특수 문자를 제거하였다. 또한, “I”, “my”, “you”, “he” 등 불용어는 리뷰에 포함되는 빈도수는 많지만 실제 분석에서는 큰 의미가 없으므로 제거하였다. 또한, 영어 리뷰에서 대문자와 소문자를 통합하거나 같은 의미를 나타내는 단어를 통일시키면 데이터의 크기를 줄일 수 있다. 따라서 어간 추출(Stemming) 기법을 통해 리뷰에서 같은 의미를 나타내는 단어를 통일시켰다. 단어의 등장 빈도가 낮으면 모델을 학습하는 과정에서 정교하게 감성 특성을 추출하는 데 한계가 있다. 이와 같은 문제를 개선하기 위해 등장 빈도가 2회 이상의 단어만을 선택하여 사용했다. 마지막으로 리뷰에 포함되는 단어를 벡터값으로 변환하기 위해 인코딩 작업을 수행하여 각 단어에 대한 고유한 벡터값을 할당하였다. 이와 같은 전처리 작업이 끝나면 감성분석 모델 구축을 위한 데이터와 개인화 추천 서비스 구축을 위한 데이터로 나누었다.

3.3 3단계: 딥러닝 기반 감성분석 모델 구축

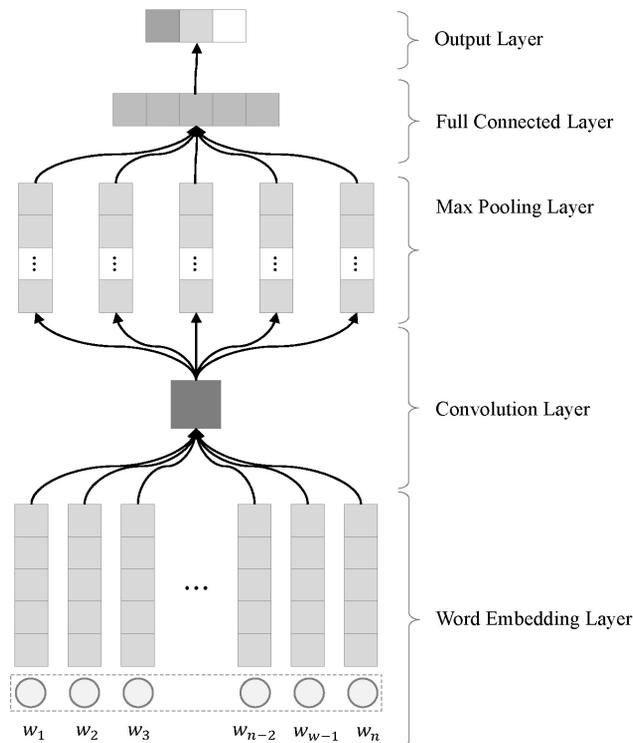
본 연구에서는 리뷰에 포함되는 정성적 선호도 감성 특성을 정교하게 추출하기 위해 CNN, LSTM, CNN + LSTM을 적용한 감성분석 모델을 구축하였다.

첫 번째는 CNN을 적용한 감성분석 모델은 Word Embedding Layer, Convolution Layer, Max Pooling Layer, Full Connected Layer로 구성된다. CNN은 이미지나 영상 처리 관련 연구에서 주로 활용하였으나 최근에는 자연어 처리 관련 연구에서 우수한 성능을 나타내기 시작하면서 다양한 연구에서 널리 사용되고 있다(Kim, 2014). CNN의 특징은 Convolution Layer와 Max Pooling Layer를 통해 리뷰에 포함된 감성 특성을 정교하게 추출하는 것이다. 또한, 고차원의 입력 데이터에서 중요한 특성들만 추출하여 이를 저차원의 데이터로 변환하여 분류하는 방법으로 감성분석에서 우수한 성능을 나타내고 있다. CNN을 적용한 감성분석 모델의 구축

과정은 <그림 2>와 같다. 먼저, CNN을 적용하여 감성분석 모델을 구축할 때 리뷰는 모델에 학습시키기 위해 벡터 형태로 변환해야 한다. 이를 위해 리뷰에 포함된 각 단어는 Embedding Layer를 통과하여 벡터 형태로 변환된다. 다음으로 Convolution Layer는 앞서 벡터 형태로 변환되어 입력된 행렬에서 필터 커널을 통해 감성 특성을 정교하게 추출한다. 이때 필터 커널의 너비는 리뷰의 벡터 행렬에서 임베딩 차원의 크기와 같이 설정한다. 또한, 특성을 추출하기 위해 일정한 크기의 필터 커널을 적용하여 컨볼루션 연산을 수행한다. Convolution Layer를 통해 출력된 벡터 행렬은 Max Pooling Layer를 통해 각 컨볼루션 연산의 결과에서 가장 큰 스칼라값을 선택하여 적용한다. 마지막으로 이와 같은 스칼라값을 Full Connected Layer를 통과하여 출력층에서 소프트맥스(Softmax) 함수를 통해 최종적으로 리

뷰의 의미론적 감성 특성에 대한 다중 분류를 수행한다(Severyn and Moschitti, 2015).

두 번째는 LSTM은 RNN의 장기 의존성 문제점을 개선하기 위해 제안되었다(Zhang *et al.*, 2019; 배상현, 최병구, 2021). LSTM은 은닉층의 메모리 셀에 입력 게이트, 망각 게이트, 출력 게이트를 추가하여 장기 의존성 문제점을 개선하였다. 이를 통해 LSTM은 리뷰의 길이와 관계없이 감성분석 분류에 우수한 성능을 나타내고 있다(Tai *et al.*, 2015; Zhang *et al.*, 2019). LSTM을 적용한 감성분석 모델의 구축 과정은 <그림 3>과 같이 Word Embedding Layer, LSTM Layer, Full Connected Layer로 구성된다. 앞서 CNN과 같이 리뷰에 포함되는 각 단어는 Embedding Layer를 통과하여 벡터 형태로 변환된다. 다음으로 LSTM Layer는 앞서 벡터 형태로 변환되어 입력된 행렬에서 리뷰의 맥락 특



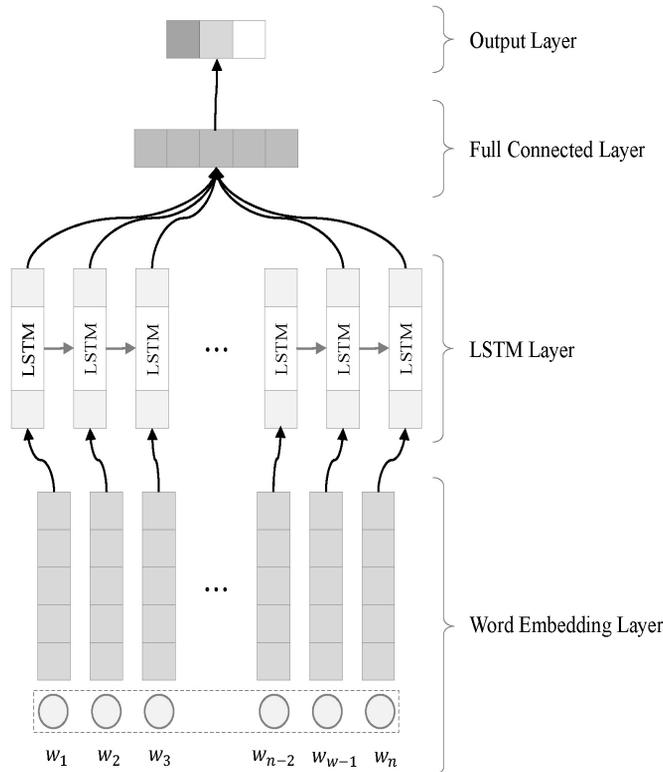
<그림 2> CNN 기반 감성분석 모델 아키텍처

성을 파악하여 학습을 진행한다. 또한, 학습하는 과정에서 과적합을 방지하기 위해 Dropout Layer를 적용하여 효율적으로 학습이 진행되도록 설정했다. 마지막으로 Full Connected Layer를 통과하여 출력층에서 소프트맥스 함수를 통해 최종적으로 감성 특성에 대한 다중 분류를 수행한다(Severyn and Moschitti, 2015).

세 번째는 CNN + LSTM을 적용한 감성분석 모델이며 구축 과정은 <그림 4>와 같이 Word Embedding Layer, Convolution Layer, Max Pooling Layer, LSTM Layer, Full Connected Layer로 구성된다. CNN + LSTM은 앞서 구축한 CNN과 LSTM의 장점을 조합하여 구축했다. 앞서 과정과 마찬가지로 리뷰에 포함되는 각 단어는 Embedding Layer를 통과하여 벡터 형태로 변환하고 Convolution Layer에서 필터 커널을 적용하여 중요한 의미론적 감성

특성을 추출한다. 다음으로 LSTM Layer는 앞서 벡터 형태로 변환되어 입력된 행렬에서 리뷰의 시퀀스 맥락 특성을 파악하여 학습을 진행한다. 또한, Dropout Layer를 적용하여 과적합을 방지하고 효율적으로 학습이 진행되도록 설정했다. 마지막으로 Full Connected Layer를 통과하여 출력층에서 소프트맥스 함수를 통해 최종적으로 감성 특성에 대한 다중 분류를 수행한다(Severyn and Moschitti, 2015).

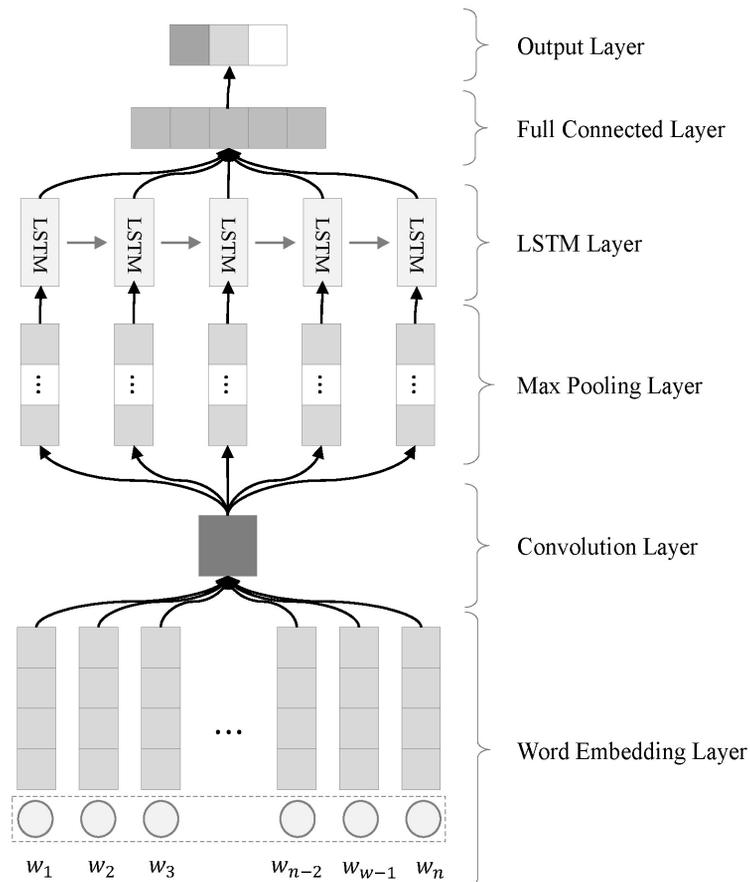
본 연구에서는 효과적으로 모델을 비교하기 위해 기존 감성분석 관련 연구의 파라미터 조합을 참고하여 <표 1>과 같이 하이퍼 파라미터(Hyper Parameter)를 적용하였다. 리뷰의 길이가 지나치게 길거나 짧으면 정보 손실이 발생하므로 모델의 성능이 저하되는 문제가 발생한다(Wen et al., 2016). 따라서 중요한 감성 특성은 유지하고 분석에 크게



<그림 3> LSTM 기반 감성분석 모델 아키텍처

중요하지 않은 노이즈(Noise) 특성은 제거하기 위해 리뷰의 길이의 다양하게 설정하여 실험을 진행했다. 실험을 통해 리뷰의 길이를 250으로 설정하였을 때 우수한 분류 성능을 보여주었으며, 이는 전체 리뷰의 97%를 포함하고 있다(Shen *et al.*, 2018; Wen *et al.*, 2016). 본 연구에서는 등장 빈도가 2회 이상인 45,686개 단어를 선택하여 모델 학습에 사용하였으며 효율적인 학습을 위해 각 단어의 임베딩 차원을 300으로 설정하였다. 또한, 임베딩 벡터 행렬에서 지역적 특성을 추출하기 위해 필터 커널의 크기는 3으로 설정하고, 모델의 학습 과정에서 과적합 문제를 방지하기 위해 드롭아웃(Dropout) 비율을 0.5로 설정하였다(Hassan and Mahmood,

2017; Jangid *et al.*, 2018). Convolution Layer에서 커널의 개수는 기존 연구의 전략을 따라서 300으로 설정하였으며, LSTM Layer에서 뉴런의 개수는 관련 연구를 참고하여 128개를 설정하고 실험을 수행하였다(Wielgosz *et al.*, 2017; Yenter and Verma, 2017). 모델을 학습할 때 최적화 알고리즘은 Adam을 적용하였으며 파라미터마다 다른 크기의 값을 반영하여 정확도를 개선할 수 있도록 하였다(Alhuzali *et al.*, 2018; Kingma and Ba, 2015). 학습된 모델에서 분류 정확도를 평가할 때 Epoch와 Batch Size는 다양한 크기를 적용하여 실험을 진행하였으며, 최종적으로 크기가 각각 30, 32로 설정했을 때 가장 높은 분류 정확도를 나타냈다.



〈그림 4〉 CNN + LSTM 기반 감성분석 모델 아키텍처

〈표 1〉 감성분석 모델 하이퍼 파라미터

Parameter	CNN	LSTM	CNN +LSTM
Review Length		250	
Embedding Dimension		300	
Vocabulary Size		45686	
Filter Kernel Size	3×3	-	3×3
Convolution Layer Units	300	-	300
LSTM Layer Units	-	128	128
Dropout		0.5	
Batch size		32	
Epoch		30	

3.4 4단계: 새로운 사용자 프로파일 구축

이 단계에서는 딥러닝 기반 감성분석 모델을 적용하여 리뷰에 포함되는 감성 특성을 추출하여 긍정, 중립 그리고 부정으로 분류한다. 이를 통해 정성적 선호도 정보와 정량적 선호도 정보가 일치하는 사용자를 바탕으로 새롭게 프로파일을 구축한다. 예를 들어, 사용자가 C가 부여한 선호도 평점이 5점이고, 리뷰의 감성을 분류한 결과가 긍정이면 정량적 선호도 정보와 정성적 선호도 정보가 일치한다고 판단한다. 반대로 사용자가 D가 부여한 선호도 평점이 2점이고, 리뷰의 감성을 분류한 결과가 긍정이면 정량적 선호도 정보와 정성적 선호도 정보가 불일치하다고 판단한다. 따라서 기존의 사용자 프로파일은 정성적 선호도 정보와 정량적 선호도 정보가 불일치하여 추천 성능이 저하되는 문제점이 존재하지만, 제안된 방법론은 정성적 선호도 정보와 정량적 선호도 정보를 동시에 고려하기에 기존 연구의 한계를 개선할 수 있다.

3.5 5단계: 개인화 추천 서비스 구축

본 연구에서 제안하는 추천 방법론의 추천 성능을 평가하기 위해 개인화 추천 서비스 관련 연구에서 널리 사용되는 메모리 기반 UBCF, IBCF 기법과 모델 기반 SVD 기법을 적용하였다(Park et

al., 2012).

UBCF 기법은 사용자 간의 선호도 유사도를 계산하여 선호도가 유사한 이웃 사용자를 선정하고 이웃 사용자들이 평가한 호텔을 기반으로 추천 대상 사용자의 선호도를 예측한다(Ricci et al., 2011). UBCF 기법으로 사용자의 선호도 평점을 예측할 때 먼저 사용자 간의 유사도를 계산한다. 사용자 a와 사용자 b 사이의 유사도는 다음 식 (1)과 같이 계산된다(손지은 등, 2015).

$$sim(a, b) = \frac{\sum_{i \in I} (r_{a,i} - \bar{r}_a)(r_{b,i} - \bar{r}_b)}{\sqrt{\sum_{i \in I} (r_{a,i} - \bar{r}_a)^2} \sqrt{\sum_{i \in I} (r_{b,i} - \bar{r}_b)^2}} \quad (1)$$

여기서 I 는 호텔 전체의 집합을 나타내며, $r_{a,i}$ 는 사용자 a가 특정 호텔 i에 부여한 선호도 평점을 나타내고 \bar{r}_a 는 사용자 a가 전체 호텔에 부여한 선호도 평점 평균을 나타내고 있다.

사용자 간의 유사도 계산을 통해 추천 대상 사용자와 선호도가 유사한 사용자를 선택하고 이 사용자들의 과거 내역을 바탕으로 추천 대상 사용자의 선호도를 예측할 수 있다. 추천 대상 사용자 a가 특정 호텔 i에 대한 선호도 평점을 식 (2)와 같이 가중합 방식으로 계산할 수 있다(Su and Khoshgoftaar, 2009; 손지은 등, 2015).

$$p(a,i) = \bar{r}_a + \frac{\sum_{u \in U} (r_{u,i} - \bar{r}_u) \cdot \text{sim}(a,u)}{\sum_{u \in U} |\text{sim}(a,u)|} \quad (2)$$

여기서 \bar{r}_a 는 추천 대상 사용자가 모든 호텔에 부여한 선호도 평점 평균을 나타내고, \bar{r}_u 는 다른 사용자가 모든 호텔에 부여한 선호도 평점 평균을 나타낸다. $\text{sim}(a,u)$ 는 추천 대상 사용자 a 와 사용자 u 사이의 유사도 계산에 따른 가중치를 의미하며 유사도가 높을수록 큰 가중치를 부여하고 있다.

IBCF 기법은 사용자가 평점을 부여한 기존 호텔과 추천하고자 하는 호텔 간의 유사도를 계산하여 추천 대상 사용자가 선호하는 호텔을 예측한다(Linden *et al.*, 2003). UBCF 기법과 마찬가지로 먼저 유사도 계산을 수행해야 하며, 여기서 호텔 i 와 호텔 j 사이의 유사도는 식 (3)과 같이 계산한다(Su and Khoshgoftaar, 2009).

$$\text{sim}(i,j) = \frac{\sum_{u \in U} (r_{u,i} - \bar{r}_i)(r_{u,j} - \bar{r}_j)}{\sqrt{\sum_{u \in U} (r_{u,i} - \bar{r}_i)^2} \sqrt{\sum_{u \in U} (r_{u,j} - \bar{r}_j)^2}} \quad (3)$$

여기서 U 는 호텔 i 와 j 에 모두 선호도 평점을 부여한 사용자 전체의 집합을 나타내고, $r_{u,i}$ 는 사용자 u 가 호텔 i 에 부여한 선호도 평점을 나타내고 \bar{r}_u 는 전체 사용자가 특정 호텔 i 에 부여한 선호도 평점 평균을 나타내고 있다.

IBCF 기법은 앞서 UBCF 기법과 달리 가중평균 방식을 통해 선호도 평점을 식 (4)와 같이 계산할 수 있다(손지은 등, 2015).

$$p(a,i) = \frac{\sum_{n \in N} r_{a,n} \cdot \text{sim}(i,n)}{\sum_{n \in N} |\text{sim}(i,n)|} \quad (4)$$

여기서 \bar{r}_a 는 추천 대상 사용자가 모든 호텔에 부여한 선호도 평점 평균을 나타내고, \bar{r}_u 는 다른

사용자가 모든 호텔에 부여한 선호도 평점 평균을 나타낸다. $\text{sim}(i,n)$ 는 선호도 평점을 예측하려는 호텔과 다른 호텔 간의 유사도에 가중치로 부여하여 예측하려는 호텔과 유사한 호텔의 선호도 평점을 반영하여 최종 선호도 평점을 계산한다.

모델 기반 SVD 기법은 고차원의 행렬을 저차원의 행렬로 축소하는 방법으로 개인화 추천 서비스 연구에서 널리 사용되고 있다(Ricci *et al.*, 2011). 모든 사용자와 호텔에 대한 $m \times n$ 크기의 행렬 $M = U \Sigma V^T$ 을 분해하면 $U_{m \times m}$, $\Sigma_{m \times n}$, $V_{n \times n}^T$ 로 구성된다. 여기서 $U_{m \times m}$ 는 사용자 행렬을 나타내고, $V_{n \times n}^T$ 는 호텔 행렬을 나타내며 $\Sigma_{m \times n}$ 은 대각행렬을 나타내고 있다. SVD 기법에서 사용자 선호도 평점은 식 (5)와 같이 계산된다(Su and Khoshgoftaar, 2009).

$$\hat{r}_{ui} = \mu + b_u + b_i + q_i^T p_u \quad (5)$$

여기서 μ 는 선호도 평점 평균을 나타내며, b_u 와 b_i 는 각각 사용자와 호텔에 대한 편향 값을 나타내고 있다. q_i 와 p_u 는 저차원의 공간에서 사용자와 호텔의 잠재적 요인을 나타내고 있다. 특히, 행렬 $M = U \Sigma V^T$ 은 희소성을 나타내고 있으므로 사용자가 기존에 부여한 선호도 평점 정보를 사용하여 제곱 오차를 최소화하는 잠재적 요인을 찾기 위해 식 (6)과 같이 정규화 작업이 필요하다.

$$\min \sum_{(u,i) \in M} (r_{ui} - \hat{r}_{ui}) = \lambda(b_i^2 + b_u^2 + \|q_i\|^2 + \|p_u\|^2) \quad (6)$$

최소화는 확률적 경사 하강법(Stochastic Gradient Descent)을 적용하여 아래와 같이 가중치를 수정하면서 최종적으로 사용자의 선호도 평점 \hat{r}_{ui} 을 예측한다.

$$b_u \leftarrow b_u + \gamma(e_{ui} - \lambda b_u) \quad (7)$$

$$b_i \leftarrow b_i + \gamma(e_{ui} - \lambda b_i) \quad (8)$$

$$p_u \leftarrow p_u + \gamma(e_{ui} \cdot q_i - \lambda p_u) \quad (9)$$

$$q_i \leftarrow q_i + \gamma(e_{ui} \cdot p_u - \lambda q_i) \quad (10)$$

여기서 e_{ui} 는 실제 평점 r_{ui} 와 예측 평점 \hat{r}_{ui} 의 차이로 계산된다. 또한, 사용자 및 호텔의 잠재적 요인은 정규분포에 따라 무작위로 초기화되어 계산되며, 이를 통해 사용자와 호텔의 특성을 표현할 수 있는 최적의 요인 수를 설정한다.

IV. 실험

4.1 데이터 및 평가지표

본 연구는 세계 최대 여행 플랫폼 TripAdvisor (www.tripadvisor.com)에서 미국 뉴욕 지역의 호텔 평가 데이터를 수집하여 제안한 추천 방법론의 성능을 평가하였다. 수집한 데이터에는 사용자, 호텔, 평점 및 리뷰 정보를 포함하고 있다. 감성분석 모델 학습을 위해 수집한 데이터에는 39,945명의 사용자 정보, 69,593개의 호텔 정보 및 161,057개 리뷰와 선호도 평점으로 구성되어 있으며, 선호도 평점이 1점과 2점의 경우 부정, 3점의 경우 중립, 4점과 5점의 경우에는 긍정으로 분류하였다. 여기서 긍정, 중립 및 부정에 포함되는 데이터를 각각 5,000개를 샘플링(Sampling)하여 감성분석 모델 학습에 사용하였다. 또한, 개인화 추천 서비스의 성능 평가를 위해 수집한 데이터에는 103,024명의 사용자와 61개의 호텔에 대한 132,669개의 리뷰와 선호도 평점이 포함되어 있다. 본 연구에서는 딥러닝 기법을 적용한 감성분석 모델을 학습할 때 편향을 줄이고, 추천 성능을 효율적으로 측정하기 위해 K-겹 교차 검증(K-fold Cross Validation)을 수행하였다. K-겹 교차 검증은 데이터를 K개로 분할하고 그중에서 1개는 성능 평가를 위해 사용하고 나머지는 모델 학습을 위해 사용한다. 여기에서 K값을 다양하게 설정할 수 있으나 큰 값을 설정하면 교차 검증 시간이 길어지고 연산 비용이 커지는 문제가 발생한다(Cakir and Dogdu, 2018). 본 연

구에서는 전체 데이터를 5개로 분할하고 그중에서 1개 데이터는 성능 평가를 위해 사용하고 나머지는 모델 학습을 위해 사용하였다(Cakir and Dogdu, 2018; Faker and Dogdu, 2019). 감성분석 모델의 분류 성능 평가는 <표 2>의 혼동 행렬(Confusion Matrix) 기준에 따라 식 (11)과 같이 정확하게 분류한 결과에서 전체 분류 결과를 나눈 백분율로 계산된다(Park et al., 2012). 본 연구에서는 5-겹 교차 검증을 사용하므로 총 5회를 실험하여 계산된 분류 정확도를 평균하여 최종적으로 모델의 분류 정확도를 계산했다.

<표 2> 혼동 행렬(Confusion Matrix)

Actual Class \ Predicted Class	Positive	Negative
	Positive	TP
Negative	FP	TN

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

개인화 추천 서비스의 예측 성능 평가는 주로 예측 선호도 등급과 실제 선호도 등급의 차이를 비교하여 측정한다. 본 연구에서는 개인화 추천 서비스 연구에서 주로 사용되고 있는 MAE(Mean Absolute Error) 평가지표를 사용하여 예측 성능을 평가하였다(Park et al., 2012; Su and Khoshgoftaar, 2009). MAE 평가지표는 실제 선호도 평점과 예측 선호도 평점 간의 오차를 절댓값으로 나타내며 이를 모두 합하여 예측 데이터의 개수만큼 나누는 방식으로 사용된다. 여기서 MAE 값이 작을수록 개인화 추천 서비스의 예측 성능이 더 우수하다고 평가할 수 있다. MAE는 아래 식 (12)와 같이 정의된다. 여기서 N 은 예측 평점의 개수를 나타내고, $p_{u,i}$ 와 $r_{u,i}$ 는 각각 예측 평점과 실제 평점을 나타내고 있다.

$$MAE = \frac{1}{N} \sum_{i=1}^N |p_{u,i} - r_{u,i}| \quad (12)$$

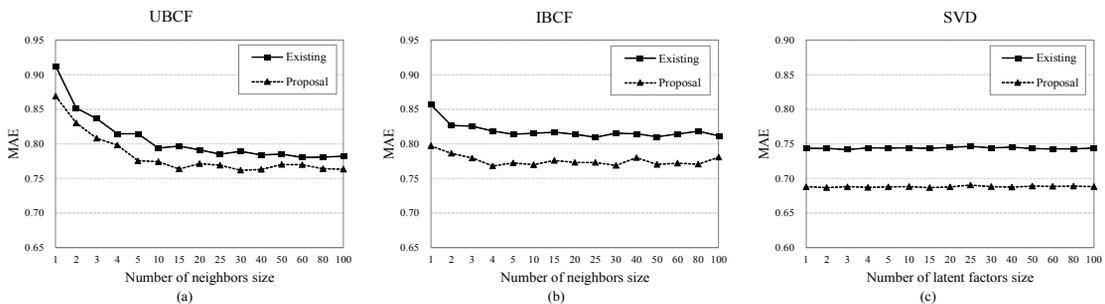
4.2 실험 결과

본 연구에서는 구매 내역, 평점 등 정량적 선호도만을 고려하는 기존 개인화 추천 서비스 관련 연구의 한계를 개선하기 위해 사용자의 정량적 선호도와 정성적 선호도를 동시에 고려하는 새로운 추천 방법론을 제안했다. 본 연구는 i5-9400F CPU, 16GB RAM, RTX 2060 개발환경에서 수행하였다. 먼저 리뷰에 포함된 감성 특성을 추출하기 위해 CNN, LSTM, CNN + LSTM을 적용한 감성분석 모델을 구축하고 각 모델의 분류 성능을 평가했다. 본 연구에서 구축한 3가지 유형의 감성분석 모델의 분류 정확도는 <표 3>과 같다. 실험에서 사용된 감성분석 모델 중에서 CNN + LSTM을 적용하였을 때 분류 정확도가 0.918로 가장 우수한 분류 성능을 나타내고 있다. CNN을 적용하였을 때 학습 시간이 다른 모델에 비해서 가장 빠르게 나타나고 있으나, 분류 성능이 0.8836을 나타내며 다른 모델에 비해서 다소 떨어진다는 단점을 가지고 있다. 반면 LSTM을 적용하였을 때 분류 성능은 0.905로 우수하게 나타나고 있으나, 모델 학습 시간이 오랫동안 지속된다는 문제점을 확인할 수 있다. 본 연구에서 구축한 3가지 유형의 감성분석 모델을 사용하여 정성적 선호도 정보와 정량적 선호도 정보가 일치하는 사용자를 바탕으로 새롭게 프로파일을 구축하고 기존의 모든 선호도 정보를 고려하는 추천 방법론과 예측 성능을 비교하였다.

<표 3> 감성분석 모델 분류 정확도

Method	Training Time (Sec)	Accuracy
CNN	259	0.8836
LSTM	1647	0.9050
CNN + LSTM	423	0.9180

첫 번째는 CNN을 적용하여 사용자의 감성 특성을 분류하고 이를 통해 정량적 선호도 정보와 일치하는지를 확인하고 선호도 정보가 일치한 사용자를 바탕으로 새로운 프로파일을 구축하였다. 여기서 “Existing”은 모든 선호도 평점을 포함하는 기존 추천 방법론을 나타내고, “Proposal”은 본 연구에서 제안하는 정량적 선호도 정보와 정성적 선호도 정보가 일치하는 선호도 정보만을 고려하는 추천 방법론을 나타낸다. UBCF와 IBCF 기법에서 이웃 사용자의 크기가 다양할 때 예측 정확도를 평가하기 위해 이웃 사용자의 크기를 1부터 100까지 설정하였다. SVD 기법은 잠재적 요인(Latent Factor) 크기를 1부터 100까지 설정하였으며, 구체적인 실험 결과는 <그림 5>와 같다. 본 연구에서 제안한 추천 방법론의 경우 기존 추천 방법론과 비교했을 때, 이웃 사용자와 잠재적 요인 크기와 상관없이 전체적으로 우수한 예측 성능을 보여주고 있다. 구체적으로 UBCF(a)와 IBCF(b) 기법에서 이웃 사용자의 크기가 각각 30, 4일 때 가장 우수한 예측을 보여주고 있으며, SVD(c) 기법은 잠재적 요인 크기가 15일 때 가장 우수한 예측을 성능



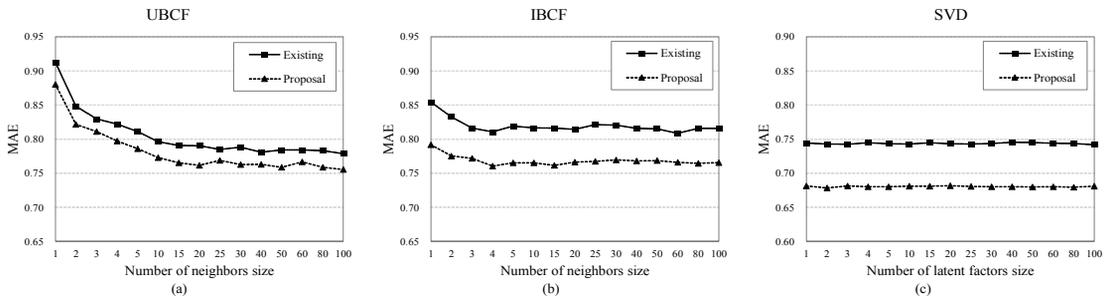
<그림 5> 기법별 CNN 기반 감성분석 모델 적용에 따른 예측 성능 비교

을 나타내고 있다. 모든 선호도 평점을 포함하는 기존 추천 방법론과 비교했을 때 각각 평균적으로 2.8%(UBCF), 5.2%(IBCF), 7.5%(SVD)만큼 예측 정확도가 개선되었다.

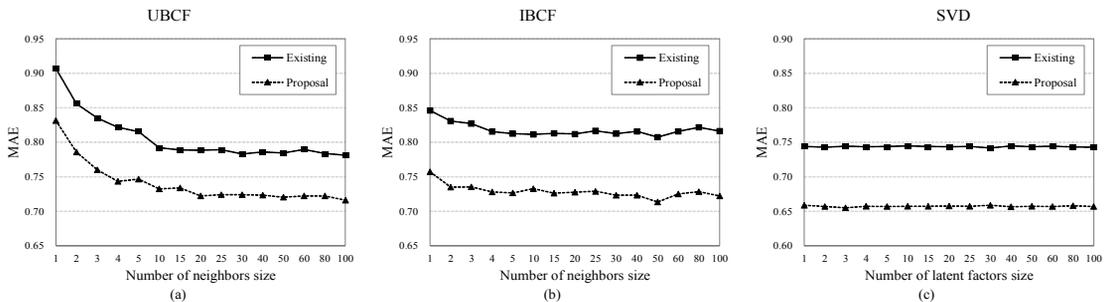
두 번째는 LSTM을 사용하여 사용자의 감성 특성을 분류하고 이를 통해 정량적 선호도 정보와 일치하는지를 확인하고 선호도 정보가 일치한 사용자를 바탕으로 새로운 프로파일을 구축하였다. 마찬가지로 UBCF와 IBCF 기법은 이웃 사용자의 크기를 1부터 100까지 설정하고 예측 정확도를 평가하였다. SVD 기법은 잠재적 요인의 크기를 1부터 100까지 설정하고 예측 성능을 평가하였으며, 구체적인 실험 결과는 <그림 6>과 같다. 전통적인 추천 방법론과 비교했을 때 본 연구에서 제안한 추천 방법론이 전체적으로 우수한 예측 성능을 보여주고 있다. 구체적으로 UBCF(a)와 IBCF(b) 기법에서 이웃 사용자의 크기가 각각 100, 4일 때 가장 우수한 예측을 보여주고 있으며, SVD(c) 기법은

잠재적 요인 크기가 2일 때 가장 우수한 예측 성능을 나타내고 있다. 모든 선호도 평점을 포함하는 기존 추천 방법론과 비교했을 때 각각 평균적으로 2.9%(UBCF), 6.2%(IBCF), 8.5%(SVD)만큼 예측 정확도가 개선되었다.

세 번째는 CNN + LSTM을 적용하여 사용자의 감성 특성을 분류하고 이를 통해 정량적 선호도와 일치하는지를 확인하고 선호도가 일치한 사용자를 선별하여 새로운 프로파일을 구축하였다. 다양한 이웃 사용자 크기에 따른 예측 정확도를 평가하기 위해 UBCF와 IBCF 기법은 이웃 사용자 크기 범위를 1부터 100까지 설정하였다. SVD 기법은 잠재적 요인의 크기를 1부터 100까지 설정하고 예측 성능을 평가하였으며, 구체적인 실험 결과는 <그림 7>과 같다. 앞서 실험과 같이 기존 추천 방법론과 비교했을 때 본 연구에서 제안한 추천 방법론이 전체적으로 우수한 예측 성능을 보여주고 있다. 구체적으로 UBCF(a)와 IBCF(b) 기법에



<그림 6> 기법별 LSTM 기반 감성분석 모델 적용에 따른 예측 성능 비교



<그림 7> 기법별 CNN + LSTM 기반 감성분석 모델 적용에 따른 예측 성능 비교

〈표 4〉 제안 방법론 성능평가 결과 요약

Method		CNN	LSTM	CNN + LSTM
Sentiment Analysis	Training Time(Sec)	259	1647	423
	Accuracy	0.883	0.905	0.918
Performance Improvement Rate	UBCF	2.8%	2.9%	8.2%
	IBCF	5.2%	6.2%	10.9%
	SVD	7.5%	8.5%	11.6%

서 이웃 사용자의 크기가 각각 100, 50일 때 가장 우수한 예측을 보여주고 있으며, SVD(c) 기법은 잠재적 요인 크기가 3일 때 가장 우수한 예측 성능을 나타내고 있다. 모든 선호도 평점을 포함하는 기존 추천 방법론과 비교했을 때 각각 평균적으로 8.2%(UBCF), 10.9%(IBCF), 11.6%(SVD)만큼 예측 정확도가 개선되었다.

본 연구에서 제안한 추천 방법론에 대한 성능 평가 결과를 요약하면 <표 4>와 같다. 실험 결과, 전체적으로 본 연구에서 제안한 방법론에 따라 평가된 예측 성능이 기존의 모든 선호도 평점을 포함하는 추천 방법론에 의한 예측 성능보다 우수함을 확인할 수 있었다. 그중에서도 CNN + LSTM을 적용하였을 때 분류 정확도가 0.918로 가장 우수하게 나타남을 확인할 수 있다. 또한, 본 연구에서 제안한 추천 방법론을 통해 예측 성능을 평가했을 때 기존 추천 방법론보다 전체적으로 우수하게 나타나고 있으며, 특히 CNN + LSTM을 적용할 때는 예측 정확도 증가율이 뚜렷하게 드러나고 있다. 이와 같은 결과는 리뷰에 포함되는 감성 특성을 분류할 때 단어 특성 추출과 맥락(Context)을 동시에 고려할 때 분류 성능이 우수하게 나타날 수 있음을 시사한다(Borgonovo and Plischke, 2016).

V. 결 론

5.1 연구결과 토의

본 연구에서는 정량적 선호도 정보와 리뷰의 정성적 선호도 정보가 일치하지 않는 문제를 개선

하기 위해 새로운 추천 방법론을 제안하였다. 제안한 방법론의 추천 성능을 평가하기 위해 세계 최대 여행 플랫폼 TripAdvisor에서 데이터를 수집하였다. 실험 결과, 본 연구에서 제안하는 정성적 선호도와 정량적 선호도가 일치하는 정보만을 고려하는 방법론이 기존의 정량적 평점 선호도 정보만을 고려하는 방법론보다 우수한 예측 성능을 보여주고 있음을 확인할 수 있었다. 이는 선호도 평점을 바탕으로 사용자 간의 유사도를 측정할 때 정성적 선호도 정보와 정량적 선호도 정보의 일치하는지가 선호도 평점에 유의미한 영향을 미칠 수 있음을 시사한다. 즉, 사용자가 작성한 리뷰에 포함된 정성적 선호도 정보와 사용자가 부여한 평점에 포함되는 정량적 선호도 정보가 일치하면 더 높은 신뢰성을 제공할 수 있음을 보여준다. 또한, CNN + LSTM을 적용한 감성분석 모델을 사용할 때 가장 우수한 예측 성능을 나타내고 있음을 확인할 수 있었다. 이는 리뷰의 정성적 선호도 특성을 추출할 때 감성 특성과 맥락을 동시에 고려할 때 분류 성능이 우수하게 나타날 수 있음을 시사한다(Borgonovo and Plischke, 2016).

5.2 연구 한계점 및 추후 연구 계획

본 연구에서는 사용자가 작성한 리뷰에서 정성적 선호도를 예측하고 이를 정량적 평점 선호도와 비교하여 최종적으로 선호도 정보가 일치한 사용자를 바탕으로 추천 서비스를 제공하는 방법론을 제안하고 성능을 평가하였다. 본 연구의 한계점과 추후 연구 계획은 다음과 같다. 첫째, 본 연구에서는

제안한 방법론의 성능을 평가하기 위해 TripAdvisor에서 호텔 관련 데이터를 직접 수집하였다. 그러나 본 연구에 사용된 제품 카테고리의 범위가 제한적인 문제가 존재한다. Amazon 사이트에는 수십 개로 구성된 다양한 유형의 제품 및 서비스 카테고리가 존재한다. 따라서 본 연구에서 제안한 방법론을 다른 분야에도 쉽게 적용하기 위해서는 다양한 카테고리를 포괄적으로 적용할 수 있는 데이터를 수집하여 추가로 검증하는 과정이 필요하다.

둘째, 본 연구에서 수집한 리뷰를 정제할 때 NLTK 패키지에서 기존에 구축한 사전을 바탕으로 불용어 제거, 어간 추출 등 다양한 작업을 수행했다. 하지만 적용하는 분야에 따라 사용되는 단어, 표현 등이 다르므로 추가 연구를 통해 특정 분야에 최적화된 사전을 구축하여 리뷰를 정제하는 작업을 수행하는 방법을 고려할 수 있다.

셋째, 본 연구에서는 감성분석에서 널리 사용되고 있는 딥러닝 기법을 사용하여 감성분석 모델을 구축했다. 최근에는 Attention Mechanism, Bert 등 기법들이 자연어 처리 분야에서 우수한 성능을 보이면서 많은 주목을 받고 있다. 따라서 추후 연구에서는 이와 같은 기법을 적용한다면 리뷰의 감성 특성을 정교하게 추출하여 개인화 추천 서비스의 예측 정확도가 더욱 개선될 수 있을 것으로 기대한다. 마지막으로 사용자가 작성한 리뷰의 유용성 여부에 관한 추가적인 연구가 필요하다. Amazon, Coupang 등 대표적인 전자상거래 플랫폼에서는 사용자가 작성한 리뷰가 실제 도움이 되었는지를 투표하는 서비스를 제공하고 있다. 이와 같은 리뷰 유용성 정보를 사용하면 광고성 리뷰, 악성 리뷰 등 실제 분석에 영향을 줄 수 있는 리뷰를 제거할 수 있다. 추후 연구를 통해 유용성 점수를 개인화 추천 서비스에 반영하여 추천 성능을 높이는 방법론을 고려할 수 있다.

5.3 학술적 의의와 실무적 시사점

본 연구는 리뷰에 포함된 정성적인 선호도 정

보와 정량적인 선호도 정보를 비교하여 선호도 정보가 일치하는 사용자를 바탕으로 프로파일을 구축하고 이를 통해 사용자의 선호도를 예측하였다. 이와 같은 본 연구의 학술적 의의는 다음과 같다. 첫째, 대표적인 비정형 데이터 형태를 나타내는 리뷰 데이터를 개인화 추천 서비스 연구에서 효과적으로 사용할 수 있는 방법론을 제안했다. 기존 개인화 추천 서비스의 연구는 주로 평점, 구매 여부 등 수리적으로 분석하기 쉬운 정량적 정보를 사용하여 추천 시스템을 구축하고 사용자 선호도를 예측하였다(Choi *et al.*, 2016; Li *et al.*, 2014; Ricci *et al.*, 2011). 최근 온라인 쇼핑 문화의 확산으로 온라인을 통해 제품을 구매하는 사용자가 증가하고 있다. 특히 제품을 구매할 때 리뷰 데이터는 사용자가 제품 구매를 결정할 때 중요한 역할을 하고 있다. 따라서 사용자 맞춤형 추천 서비스를 제공할 때 기존 평점 데이터와 함께 리뷰 데이터를 동시에 고려하는 것이 중요하다. 본 연구에서는 사용자가 작성한 리뷰를 분석하여 정량적 선호도와 비교하여 선호도 정보가 일치하는 사용자 프로파일을 구축하여 최종적으로 사용자의 선호도를 예측하였다. 따라서 기존 정량적 선호도만을 고려한 연구에 국한되지 않고 리뷰 데이터를 동시에 고려하여 추천 성능을 제고시키는 방법론은 기존 연구와 차별화되는 부분이다. 특히, 기존 개인화 추천 서비스 연구에서 시도하지 않은 정량적 선호도 정보와 정성적 선호도 정보 간의 비교를 통해 선호도 정보가 일치한 사용자를 선별하는 추천 방법론을 제안하여 추천 성능을 높임으로써 개인화 추천 서비스 연구의 확장에 기여할 수 있다.

둘째, 기존 감성분석 연구는 주로 긍정과 부정으로 분류하는 통계 및 데이터 마이닝 기법에 국한되어 있다. 최근 딥러닝 기술이 다양한 분야에서 우수한 성능을 보여주면서 감성분석 관련 연구에서도 이를 활용하는 연구들이 활발히 진행되고 있다. 본 연구에서는 CNN, LSTM, CNN + LSTM 등 다양한 유형의 딥러닝 기법을 활용하여 사용자

가 작성한 리뷰의 의미론적 감성 특성을 정교하게 추출하는 감성분석 모델을 구축하였다. 이를 통해 사용자가 작성한 리뷰의 정성적 선호도 정보를 예측하여 정량적 선호도 정보와 일치하는지를 비교하였다. 또한, 딥러닝 기법에 적합한 파라미터 조합을 적용하기 위해 기존 감성분석 관련 연구를 참고하여 다양하게 모델을 학습하여 높은 분류 성능을 나타내는 파라미터 조합을 설정하였다. 마지막으로, 딥러닝 기법을 활용한 감성분석 모델에서 CNN 및 LSTM 단일 모델을 적용하는 경우보다 CNN+LSTM 결합 모델을 적용했을 때 더 우수한 분류 정확도를 나타내고 있음을 다시 한번 확인했다(Ombabi *et al.*, 2020). 또한, CNN+LSTM을 적용하여 리뷰의 정성적 선호도 정보와 정량적 선호도 정보가 일치하는 사용자 프로파일을 구축하고 이를 통해 사용자의 선호도를 예측했을 때 예측 정확도 가장 많이 개선되었음을 확인할 수 있다.

본 연구의 결과는 다음과 같은 실무적 시사점을 제공한다. 첫째, 본 연구에서 제안한 추천 방법론은 평점과 리뷰를 작성할 수 있는 다양한 분야의 전자상거래 플랫폼에 적용할 수 있다. 기존 전자상거래 플랫폼에서 개인화 추천 서비스 전략을 구축할 때 주로 평점, 구매 여부 등 정량적인 선호도 정보를 사용하여 제품이나 서비스를 추천했다. 하지만 사용자가 작성한 선호도 등급이 실제 선호도를 나타내지 못하고 있으며 오히려 추천 성능을 저하하는 문제가 발생한다. 본 연구에서 제안하는 방법론이 기존 방법론보다 우수한 예측 성능을 확인함으로써 기존 개인화 추천 서비스 전략에 재고의 여지가 있음을 확인할 수 있다.

둘째, 대부분의 전자상거래 플랫폼은 사용자가 부여한 평점을 바탕으로 제품이나 서비스에 대한 종합적인 평점 정보를 제공한다. 이와 같은 평점 정보는 사용자가 특정 제품이나 서비스를 구매할 때 고려하는 중요한 요소이므로 매우 중요하다. 기존 전자상거래 사이트에는 대부분 사용자가 부여한 정량적인 평점만을 고려하여 종합적인 평점을 산출하였다. 그러나 이와 같은 정량적인 평점

선호도가 실제 사용자의 선호도를 제대로 반영하지 못하기에 평점 정보에 대한 신뢰성이 떨어질 수 있다. 이에 따라 정량적 평점을 평균으로 계산하는 기존 방법의 한계를 개선하여, 정량적 선호도 정보와 정성적 선호도 정보가 일치하는 사용자의 정보를 활용한다면 더 정교하게 종합적인 평점 정보를 제공할 수 있다. 마지막으로 전자상거래에서 제공하는 제품이나 서비스 항목은 다양하지만, 인기 있는 제품이나 서비스를 반복적으로 제공하는 롱테일(Long tail) 문제점을 개선할 수 있다. 사용자의 구매 내역은 서로 다르지만, 사용자가 작성한 리뷰는 다양한 감성 특성을 포함하고 있다. 이처럼 리뷰의 감성 특성을 정교하게 분석하여 감성 특성의 유사성을 추천 서비스에 반영한다면 사용자들의 관심이 다소 부족한 다양한 제품이나 서비스를 제공할 수 있다.

참고 문헌

- [1] 박승택, 성인재, 서상원, 황지수, 노지성, 김대원, “기계학습 기반의 뉴스 추천 서비스 구조와 그 효과에 대한 고찰: 카카오의 루빅스를 중심으로”, *사이버커뮤니케이션학보*, 제34권, 제1호, 2017, pp. 5-48.
- [2] 배상현, 최병구, “변이형 오토인코더와 어텐션 메커니즘을 결합한 차트기반 주가 예측”, *Information Systems Review*, 제23권, 제1호, 2021, pp. 23-43.
- [3] 손지은, 김성범, 김현중, 조성준, “추천 시스템 기법 연구동향 분석”, *대한산업공학회지*, 제41권, 제2호, 2015, pp. 185-208.
- [4] 전병국, 안현철, “사용자 리뷰 마이닝을 결합한 협업 필터링 시스템: 스마트폰 앱 추천에의 응용”, *지능정보연구*, 제21권, 제2호, 2015, pp. 1-18.
- [5] 현지연, 유상이, 이상용, “평점과 리뷰 텍스트 감성분석을 결합한 추천시스템 향상 방안 연구”, *지능정보연구*, 제25권, 제1권, 2019, pp.

- 219-239.
- [6] 조승연, 최지은, 이규현, 김희웅, “고객 온라인 구매후기를 활용한 추천시스템 개발 및 적용”, *Information Systems Review*, 제17권, 제3호, 2015, pp. 95-111.
- [7] Adomavicius, G. and A. Tuzhilin, “Personalization technologies: A process-oriented perspective”, *Communications of the ACM*, Vol.48, No.10, 2005, pp. 83-90.
- [8] Alhuzali, H., M. Elaraby, and M. Abdul-Mageed, “Ubc-nlp at iest 2018: Learning implicit emotion with an ensemble of language models”, *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 2018, pp. 342-347.
- [9] Annett, M. and G. Kondrak, “A comparison of sentiment analysis techniques: Polarizing movie blogs”, *Conference of the Canadian Society for Computational Studies of Intelligence*, 2008, pp. 25-35.
- [10] Bennett, J. and S. Lanning, “The netflix prize”, *Proceedings of KDD Cup and Workshop*, 2007, p. 35.
- [11] Borgonovo, E. and E. Plischke, “Sensitivity analysis: A review of recent advances”, *European Journal of Operational Research*, Vol.2248, No.3, 2016, pp. 869-887.
- [12] Cakir, B. and E. Dogdu, “Malware classification using deep learning methods”, *Proceedings of the ACMSE 2018 Conference*, 2018, pp. 1-5.
- [13] Cheng, Z., Y. Ding, L. Zhu, and M. Kankanhalli, “Aspect-aware latent factor model: Rating prediction with ratings and reviews”, *Proceedings of the 2018 World Wide Web Conference*, 2018, pp. 639-648.
- [14] 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.36, No.3, 2016, pp. 397-402.
- [15] Christopher Frey, H. and S. R. Patil, “Identification and review of sensitivity analysis methods”, *Risk Analysis*, Vol.22, No.3, 2002, pp. 553-578.
- [16] 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, pp. 271-280.
- [17] Faker, O. and E. Dogdu, “Intrusion detection using big data and deep learning techniques”, *Proceedings of the 2019 ACM Southeast Conference*, 2019, pp. 86-93.
- [18] Ferguson, P., N. O’Hare, M. Davy, A. Bermingham, P. Sheridan, C. Gurrin, and F. meaton, “Exploring the use of paragraph-level annotations for sentiment analysis of financial blogs”, *Workshop on Opinion Mining and Sentiment Analysis*, 2009.
- [19] García-Cumbreras, M. Á., A. Montejó-Ráez, and M. C. Díaz-Galiano, “Pessimists and optimists: Improving collaborative filtering through sentiment analysis”, *Expert Systems With Applications*, Vol.40, No.17, 2013, pp. 6758-6765.
- [20] Hassan, A. and A. Mahmood, “Deep learning approach for sentiment analysis of short texts”, *2017 3rd International Conference on Control, Automation and Robotics (ICCAR)*, 2017, pp. 705-710.
- [21] Herlocker, J. L., J. A. Konstan, and J. Riedl, “Explaining collaborative filtering recommendations”, *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, 2000, pp. 241-250.
- [22] Hofmann, T., “Latent semantic models for collaborative filtering”, *ACM Transactions on Information Systems (TOIS)*, Vol.22, No.1, 2004, pp. 89-115.

- [23] Jangid, H., S. Singhal, R. R. Shah, and R. Zimmermann, "Aspect-based financial sentiment analysis using deep learning", *Companion Proceedings of the The Web Conference 2018*, 2018, pp. 1961-1966.
- [24] Jakob, N., S. H. Weber, M. C. Müller, and I. Gurevych, "Beyond the stars: exploiting free-text user reviews to improve the accuracy of movie recommendations", *Proceedings of the 1st International CIKM Workshop on Topic-sentiment Analysis for Mass Opinion*, 2009, pp. 57-64
- [25] Kazmaier, J. and J. H. van Vuuren, "A generic framework for sentiment analysis: Leveraging opinion-bearing data to inform decision making", *Decision Support Systems*, Vol.135, 2020, p. 113304.
- [26] Kennedy, A. and D. Inkpen, "Sentiment classification of movie reviews using contextual valence shifters", *Computational Intelligence*, Vol.22, No.2, 2006, pp. 110-125.
- [27] 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, pp. 140-151.
- [28] 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, pp. 212-219.
- [29] Kim, Y., "Convolutional Neural Networks for Sentence Classification", *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1746-1751.
- [30] Kingma, D. P. and J. Ba, "Adam: A method for stochastic optimization", *Proceedings of the International Conference on Learning Representations (ICLR)*, 2015.
- [31] König, A. C. and E. Brill, "Reducing the human overhead in text categorization", *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2006, pp. 598-603.
- [32] Konstan, J. A., B. N. Miller, D. Maltz, J. L. Herlocker, L. R. Gordon, and J. Riedl, "GroupLens: Applying collaborative filtering to usenet news", *Communications of the ACM*, Vol.40, No.3, 1997, pp. 77-87.
- [33] Lai, S., L. Xu, K. Liu, and J. Zhao, "Recurrent convolutional neural networks for text classification", *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015, pp. 2267-2273
- [34] 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, pp. 1910-1921.
- [35] Leung, C. W., S. C. Chan, and F. L. Chung, "Integrating collaborative filtering and sentiment analysis: A rating inference approach", *Proceedings of the ECAI 2006 Workshop on Recommender Systems*, 2006, pp. 62-66
- [36] Levi, A., O. Mokryn, C. Diot, and N. Taft, "Finding a needle in a haystack of reviews: cold start context-based hotel recommender system", *Proceedings of the Sixth ACM Conference on Recommender Systems*, 2012, pp. 115-122
- [37] 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, pp. 95-104
- [38] Linden, G., B. Smith, and J. York, "Amazon.com recommendations: Item-to-item collaborative filtering", *IEEE Internet Computing*, Vol.7, No.1, 2003, pp. 76-80.
- [39] 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, pp. 126-138.
- [40] Moshfeghi, Y., B. Piwowarski, and J. M. Jose, "Handling data sparsity in collaborative filtering using emotion and semantic based features", *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2011, pp. 625-634.
- [41] Nguyen, T. T., P. M. Hui, F. M. Harper, L. Terveen, and J. A. Konstan, "Exploring the filter bubble: the effect of using recommender systems on content diversity", *Proceedings of the 23rd International Conference on World Wide Web*, 2014, pp. 677-686.
- [42] Ombabi, A. H., W. Ouarda, and A. M. Alimi, "Deep learning CNN-LSTM framework for Arabic sentiment analysis using textual information shared in social networks", *Social Network Analysis and Mining*, Vol.10, No.1, 2020, pp. 1-13.
- [43] Pak, A. and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining", *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC)*, 2010, pp. 1320-1326.
- [44] Pang, B., L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques", *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL-2002)*, 2002.
- [45] 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, pp. 300-313.
- [46] 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, pp. 10059-10072.
- [47] Park, K., J. Lee, and J. Choi, "Deep neural networks for news recommendations", *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, pp. 2255-2258
- [48] Ricci, F., L. Rokach and B. Shapira, *Introduction to Recommender Systems Handbook*, Springer, Boston, MA, 2011.
- [49] Sarwar, B., G. Karypis, J. Konstan, and J. Riedl, "Application of dimensionality reduction in recommender system-a case study", *ACM WebKDD 2000 Workshop*, 2000.
- [50] Seo, S., J. Huang, H. Yang, and Y. Liu, "Interpretable Convolutional Neural Networks with Dual Local and Global Attention for Review Rating Prediction", *Proceedings of the Eleventh ACM Conference on Recommender Systems*, 2017, pp. 297-305
- [51] Severyn, A. and A. Moschitti, "Twitter sentiment analysis with deep convolutional neural networks", *Proceedings of the 38th international ACM SIGIR Conference on Research and Development in Information Retrieval*, 2015, pp. 959-962.
- [52] Shen, T., T. Zhou, G. Long, J. Jiang, S. Pan, and C. Zhang, "Disan: Directional self-attention network for rnn/cnn-free language understanding", *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018.
- [53] Su, X. and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques", *Advances in Artificial Intelligence*, Vol.2009, 2009, pp. 1-19
- [54] Tang, J., H. Gao, X. Hu, and H. Liu, "Context-aware review helpfulness rating prediction",

- Proceedings of the 7th ACM Conference on Recommender Systems*, 2013, pp. 1-8.
- [55] Tai, K. S., R. Socher, and C. D. Manning, “Improved semantic representations from tree-structured long short-term memory networks”, *Proceeding of IJCNLP*, 2015, pp. 1556-1566.
- [56] Wen, Y., W. Zhang, R. Luo, and J. Wang, “Learning text representation using recurrent convolutional neural network with highway layers”, *Proceedings of the 39th ACM SIGIR Workshop on Neural Information Retrieval*, 2016.
- [57] Wielgosz, M., A. Skoczeń, and M. Mertik, “Using LSTM recurrent neural networks for monitoring the LHC superconducting magnets”, *Nucl Instrum Methods A*, Vol.867, 2017, pp. 40-50.
- [58] Yenter, A. and A. Verma, “Deep CNN-LSTM with combined kernels from multiple branches for IMDb review sentiment analysis”, *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, 2017, pp. 540-546.
- [59] Zhang, L., S. Wang, and B. Liu, “Deep learning for sentiment analysis: A survey”, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol.8, No.4, 2018, pp. e1253.
- [60] Zhang, S., L. Yao, A. Sun, and Y. Tay, “Deep learning based recommender system: A survey and new perspectives”, *ACM Computing Surveys (CSUR)*, Vol.52, No.1, 2019, pp. 1-38.
- [61] 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.
- [62] Zhang, Z., D. Zhang, and J. Lai, “urCF: User review enhanced collaborative filtering”, *20th Americas Conference on Information Systems*, 2014.
- [63] Zheng, L., V. Noroozi, and S. Yu, “Joint deep modeling of users and items using reviews for recommendation”, *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, 2017, pp. 425-434.

A Study of Deep Learning-based Personalized Recommendation Service for Solving Online Hotel Review and Rating Mismatch Problem

Qinglong Li* · Shibo Cui** · Byunggyu Shin*** · Jaekyeong Kim****

Abstract

Global e-commerce websites offer personalized recommendation services to gain sustainable competitiveness. Existing studies have offered personalized recommendation services using quantitative preferences such as ratings. However, offering personalized recommendation services using only quantitative data has raised the problem of decreasing recommendation performance. For example, a user gave a five-star rating but wrote a review that the user was unsatisfied with hotel service and cleanliness. In such cases, has problems where quantitative and qualitative preferences are inconsistent. Recently, a growing number of studies have considered review data simultaneously to improve the limitations of existing personalized recommendation service studies. Therefore, in this study, we identify review and rating mismatches and build a new user profile to offer personalized recommendation services. To this end, we use deep learning algorithms such as CNN, LSTM, CNN + LSTM, which have been widely used in sentiment analysis studies. And extract sentiment features from reviews and compare with quantitative preferences. To evaluate the performance of the proposed methodology in this study, we collect user preference information using real-world hotel data from the world's largest travel platform TripAdvisor. Experiments show that the proposed methodology in this study outperforms the existing other methodologies, using only existing quantitative preferences.

Keywords: *Personalized recommendation services, collaborative filtering, sentiment analysis, deep learning*

* Ph.D. Student, Department of Big Data Analytics, Graduate School, Kyung Hee University

** Master, Department of Big Data Analytics, Graduate School, Kyung Hee University

*** Corresponding Author, Ph. D. Candidate, Department of Business Administration, Graduate School, Kyung Hee University

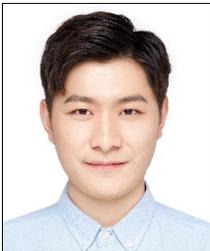
**** Professor, School of Management & Department of Big Data Analytics, Kyung Hee University

○ 저 자 소 개 ○



이 청 용 (leecy@khu.ac.kr)

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.



최 사 박 (cui@khu.ac.kr)

중국 Hebei Normal University Of Science & Technology에서 공학 학사학위를 취득하고, 경희대학교 대학원 빅데이터응용학과에서 경영학 전공으로 석사학위를 취득하였다. 주요 연구 관심 분야는 추천 시스템, 딥러닝, 자연어 처리 및 빅데이터에 분석 등이다.



신 병 규 (kbscandoit@naver.com)

경희대학교 경영대학원에서 석사학위를 취득하였으며, 동 대학원에서 경영정보시스템 전공으로 경영학 박사과정을 수료하였다. 주요 관심 분야는 비즈니스 애널리틱스(Business Analytics), 데이터 마이닝, 인공지능 및 자연어 처리, 빅데이터 분석, 소셜 네트워크 분석 등이다.



김 재 경 (jaek@khu.ac.kr)

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, 추천 시스템, 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.

논문접수일 : 2021년 03월 11일

게재확정일 : 2021년 05월 31일

1차 수정일 : 2021년 04월 30일