

E-커머스 사용자의 평점과 리뷰 유용성이 상품 추천 시스템의 성능 향상에 미치는 영향 분석*

LIU FAN
경희대학교 대학원 경영학과
(liufan@khu.ac.kr)

이병현
경희대학교 대학원 빅데이터응용학과
(leebh1129@khu.ac.kr)

최일영
경희대학교 경영대학원
(choice102@khu.ac.kr)

정재호
경희대학교 대학원 경영학과
(jih8007@msn.com)

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

정보통신기술 발달로 스마트폰이 보급되면서, 온라인 쇼핑물 서비스는 컴퓨터가 아닌 모바일로도 사용이 가능해졌다. 그로 인해 온라인 쇼핑물 서비스를 이용하는 사용자는 급격히 증가하게 되고, 거래되는 제품의 종류 또한 방대해지고 있다. 따라서 기업은 이익을 최대화하기 위해서는 사용자가 관심을 가질만한 정보를 제공해주는 것이 중요하다. 이를 위해 사용자의 과거 행동 데이터나 행동 구매 기록을 기반으로 사용자에게 필요한 정보 또는 제품을 제시하는 것을 추천 시스템이라 한다. 현재 추천 서비스를 제공하는 대표적인 해외 기업으로는 Netflix, Amazon, YouTube 등이 있다. 최근 이러한 전자상거래 사이트에서는 사용자가 해당 제품에 대한 리뷰가 유용한지에 대해 투표할 수 있는 기능을 제공하고 있다. 이를 통해, 사용자는 유용하다고 판단되는 제품에 대한 리뷰와 평점을 참고하여 구매 의사결정을 내린다. 따라서 본 연구에서는 제품에 대한 평점과 리뷰의 유용성 정보 간의 상관관계를 파악하고, 리뷰의 유용성 정보를 추천 시스템에 반영하여 추천 성능을 확인하고자 한다. 또한 대부분의 사용자들은 만족한 제품에만 평점을 부여하는 경향이 있고 제품에 대한 평점이 높을수록 구매 의도가 높아지는 경향이 있다. 따라서 전통적인 협업 필터링 기법에 모든 평점을 반영한 결과와 4점과 5점 평점만을 반영한 추천 성능 결과를 비교하고자 한다. 이를 위해 본 연구에서는 Amazon에서 수집한 전자 제품 데이터를 사용하였으며, 실험 결과는 평점과 리뷰 유용성 정보 간 상관관계가 있는 것으로 확인되었다. 또한 모든 평점과 4점과 5점 평점만을 추천 시스템에 반영하여 추천 성능을 비교한 결과, 4점과 5점 평점만을 추천 시스템에 반영한 결과의 추천 성능이 더 높게 나타났다. 그리고 리뷰 유용성 정보를 추천 시스템에 반영한 결과는 리뷰가 유용할수록 추천 성능은 높게 나타나는 것으로 확인하였다. 따라서 이러한 실험 결과는 향후 개인화 추천 서비스의 성능 향상에 기여하고, 전자상거래 사이트에 시사점을 제공할 수 있을 것으로 본다.

주제어 : E-커머스, 추천시스템, 아마존 데이터, 리뷰 유용성, 협업 필터링

논문접수일 : 2022년 2월 27일 논문수정일 : 2022년 3월 21일 게재확정일 : 2022년 3월 23일

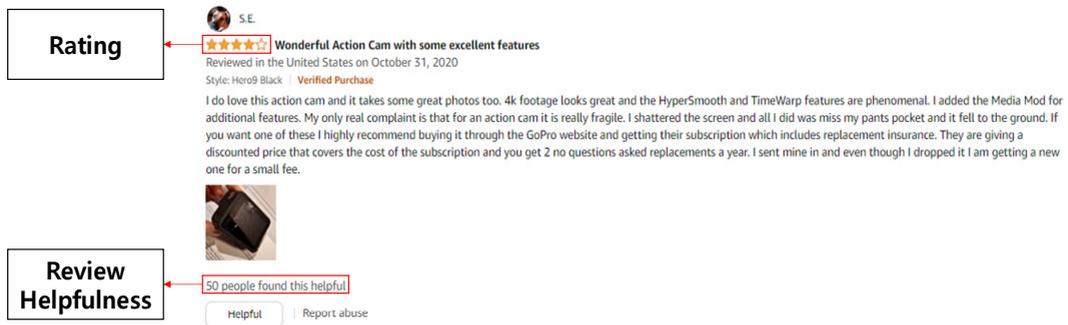
원고유형 : 급행논문 교신저자 : 정재호

1. 서론

최근 정보통신기술의 발달로 스마트폰이 보급되면서, 온라인 쇼핑물 서비스는 컴퓨터가 아닌

모바일로도 사용이 가능해졌다. 그로 인해 온라인 쇼핑물 서비스를 이용하는 사용자는 급격히 증가하게 되고, 거래되는 제품의 종류 또한 방대해지고 있다. 따라서 기업은 이익을 최대화하기

* 이 논문 또는 저서는 2020년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2020S1A5B8103855)



(Figure 1) Amazon Electronic Products of Ratings & Review Helpfulness

위해서는 사용자가 관심을 가질만한 정보를 제공해주는 것이 중요하다. 이를 위해 사용자의 과거 행동 데이터나 행동 구매 기록을 기반으로 사용자에게 필요한 정보 또는 제품을 제시하는 것을 추천 시스템이라 한다(Linden et al., 2003). 현재 이러한 추천 시스템을 활용하고 있는 대표적인 기업은 Netflix, Amazon, YouTube 등이 있다(Gomez-Uribe et al., 2015). 이들 기업은 사용자가 해당 제품에 부여한 평점, 구매 기록, 클릭 스트림(Click stream) 데이터 등을 활용하여 사용자에게 제품 및 서비스를 추천해주고 있다(Smith and Linden, 2017). 추천 시스템에서는 협업 필터링(Collaborative Filtering)이라는 기법이 대표적으로 활용된다(Su and Khoshgoftaar, 2009). 협업 필터링은 사용자가 제품에 부여한 평점 정보를 활용하여 선호할만한 상품을 예측하는 것이며, 기존의 협업 필터링 연구에서는 평점 정보를 사용하여 추천 성능을 검증하였다. 또한 대부분의 사용자들은 만족한 제품에만 평점을 부여하는 경향이 있고 제품에 대한 평점이 높을수록 구매 의도가 높아지는 경향이 있다(Chintagunta et al., 2010; Pradel et al., 2012).

최근에는 정성적 정보인 사용자가 작성한 리뷰를 반영하여 추천 시스템을 구축한 연구들이

진행되고 있다(Choe et al. 2013). 사용자가 작성한 리뷰는 평점보다 상세하고 신뢰할 수 있는 정보를 담고 있으며, 대부분의 전자상거래 사이트에서는 사용자가 구매한 제품에 대해 리뷰를 작성할 수 있는 서비스를 제공하고 있다. 또한 전자상거래 사이트의 사용자 중 88%는 제품 구매 시, 해당 제품에 대한 리뷰를 참고하여 구매를 결정한다(Moore, 2018). 이러한 리뷰에는 사용자가 제품을 구매한 이유와 제품에 대해 상세하고 신뢰할 수 있는 정보 등이 포함되어 있으므로 추천 시스템을 구축하는데 유용하게 사용된다(Qiu et al., 2016). 기존의 리뷰를 활용한 추천 시스템은 주로 리뷰의 감성 특성만을 추출하고, 이를 협업 필터링에 반영하는데 중점을 두었다(Srifi et al., 2020; Yun et al., 2018). 그러나 사용자가 제품에 대해 작성한 리뷰의 내용에는 거짓된 후기, 홍보성 내용, 전혀 제품과 관련이 없는 내용 등과 같이 사용자의 구매 의사결정을 저해하는 내용 또한 포함되어 있다(Ge et al., 2019). 따라서 최근 전자상거래 사이트에서는 사용자가 작성한 리뷰가 도움이 되었는지에 대한 여부를 투표할 수 있는 리뷰 유용성 투표 시스템을 도입하였다. 이러한 리뷰 유용성 정보는 제품을 구매하려는 사용자의 의사결정에 큰 영향을 끼치기 때

문에 추천 시스템을 구축하는데 있어 중요한 요인으로 작용된다(Castelli et al., 2017). <Figure 1>은 제품에 평점을 4점 부여한 리뷰에 50명의 사용자가 해당 리뷰가 유용하다고 투표하였다. 이는 제품에 높은 평점을 부여한 사용자의 리뷰가 유용하다는 것을 파악할 수 있다. 또한 제품에 부여된 평점은 리뷰의 유용성 평가에 있어 중요한 요인으로 작용한다는 것을 확인할 수 있다(Lee et al., 2017). 따라서 본 연구에서는 다음과 같은 3가지 연구를 수행하여, 기존 추천 시스템의 성능 향상에 기여하고자 한다.

- Q1 : 평점과 리뷰 유용성 정보 간의 상관관계가 있는가?
 Q2 : 추천 시스템에 모든 평점을 반영한 경우와 높은 평점만을 반영하였을 때의 추천 성능은 어떠한 차이가 있는가?
 Q3 : 리뷰가 유용할수록 추천 성능은 어떠한 변화가 있는가?

본 연구의 구성은 다음과 같다. 제1장에서는 본 연구의 배경과 목적에 대해 설명하였다. 제2장에서는 리뷰 유용성과 추천 시스템에 대한 이론적 배경을 설명한다. 제3장에서는 본 연구의 방법론에 대해 단계적으로 설명한다. 제4장에서는 실험 데이터와 평가지표 그리고 실험 결과에 대하여 설명한다. 마지막 5장에서는 결론과 연구의 한계점 및 향후 연구에 대해 설명한다.

2. 이론적 배경

2.1. 추천 시스템

추천 시스템은 사용자의 구매 기록, 평점, 리

뷰 등의 정보를 사용하여, 사용자가 선호할만한 제품을 추천하는 서비스를 의미한다(Shani et al., 2019). 또한 현재 추천 시스템에서 널리 활용되는 방식으로는 Goldberg et al.(1992)에 의해 처음으로 소개된 협업 필터링(Collaborative Filtering; CF) 방식이 있다. CF 방식은 메모리 기반 CF(Memory-based CF) 방식과 모델 기반 CF(Model-based CF) 방식으로 구분할 수 있으며, 메모리 기반 필터링 방식은 다시 User-based CF 방식과 Item-based CF 방식으로 구분할 수 있다(AI-Bashiri et al., 2018). User-based CF 방식은 사용자가 제품에 부여한 평점 등의 선호도 정보를 사용하여 유사한 성향을 지닌 이웃 사용자를 선별한 후, 선별된 이웃들이 공통적으로 선호하는 제품을 사용자에게 추천하는 방식이다(Koochi and Kiani, 2016). Item-based CF 방식은 사용자가 이전에 구매한 제품과 가장 유사한 제품을 추천하는 방식이다(Thakkar et al., 2019). 모델 기반 CF 방식은 사용자가 제품에 부여한 평점 정보에 분류, 군집화 등의 데이터 마이닝 기법을 적용하여 사용자가 구매하지 않은 제품에 부여할 평점을 예측하는 방식이다(Sun et al., 2013). 평점을 예측하는 방식으로는 주로 행렬 분해(Matrix Factorization)를 활용하며(Abdollahi and Nasraoui, 2017), 행렬 분해는 사용자와 사용자가 제품에 부여한 평점으로 구성된 행렬을 두 개의 행렬로 분해하는 방식이다. 대표적인 행렬 분해 기법으로는 특이값 분해(Singular Value Decomposition; SVD) 기법이 있다(Bang et al., 2015). SVD는 차원 축소의 개념이며, 고차원으로 이루어진 행렬을 저차원 행렬로 축소하는 방식이다. 따라서 SVD 기법을 활용하여 행렬을 차원 축소하고, 다시 원래의 행렬로 복원을 하면 기존의 결측치 값들이 채워지게 되며 이를 기반으로 사용자에게 제품을 추

천하는 방식이다. 따라서 기존 CF 방식은 주로 정량화된 구매 내역, 평점 정보 등의 정보를 활용하여 사용자의 선호도를 예측하여 제품을 추천하였다. 그러나 이러한 정량적 평점 정보만을 활용하면, 사용자의 선호도를 구체적으로 반영하지 못하므로 추천 성능이 저하되는 문제가 발생할 수 있다(Lu et al., 2015; Qiu et al., 2016). 본 연구에서는 평점과 리뷰 유용성 정보를 CF 방식에 반영하여 추천 성능을 확인하고자 하며, 이를 통해 기존 CF 방식의 추천 성능을 개선할 수 있을 것으로 본다.

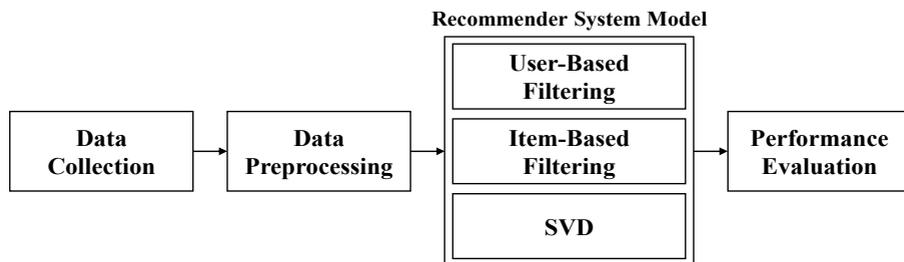
2.2. 리뷰 유용성

리뷰 유용성은 제품 및 서비스에 대한 유용한 정보를 제공함으로써 사용자의 구매 의사결정에 기여할 수 있다(Eslami et al., 2018; Shengli & Fan, 2019). 현재 전자상거래 사이트에서 제공하는 리뷰 유용성 시스템은 사용자가 제품 및 서비스에 대한 리뷰를 확인하고, 리뷰가 구매 의사결정에 도움이 되었다면 "Helpful", 도움이 되지 않았다면 "Unhelpful"을 클릭하여 투표를 하는 방식이다. 또한 리뷰 유용성에 대한 많은 연구들이 진행되었으며, 이에 대한 주요 선행 연구는 다음과 같다. Fang et al. (2016)은 제품 및 서비스에 대한 리뷰의 가독성이 높을수록 사용자에게 유

용하다는 것을 확인하였다. Wang and Yao (2019)는 리뷰의 가독성뿐만 아니라 제품 및 서비스 품질에 대한 내용이 리뷰에 충실히 담겨 있을수록 사용자에게 유용한 것을 확인하였다. 따라서 선행 연구에서는 주로 가독성, 길이, 내용 등 리뷰의 어떠한 요인이 사용자가 인지하는 리뷰 유용성에 영향을 끼치는지에 대해 연구가 진행되었다(Eslami et al., 2018; Al-Smadi et al., 2019; Nakayama and Wan, 2019; Ren and Hong, 2019). 그러나 이외에도 제품 및 서비스에 대한 평점 또한 리뷰의 유용성 평가에 있어 중요한 요인으로 작용한다(Lee et al., 2017). 따라서 본 연구에서는 평점과 리뷰의 유용성 정보 간의 상관관계를 파악하고, 이를 추천 시스템에 반영하여 추천 성능을 확인하고자 한다.

3. 연구 방법

본 연구에서는 <Figure 2>에 표현한 연구 절차를 따라 분석을 진행한다. 첫 번째 단계에서는 전자상거래 사이트 Amazon에서 제공하고 있는 데이터를 활용한다. 두 번째 단계에서는 수집한 데이터를 본 연구의 목적에 적합하도록 전처리한다. 세 번째 단계에서는 추천 성능을 평가하기



(Figure 2) Research Process

위해 User-based CF 방식과 Item-based CF 방식 그리고 SVD 모델을 사용한다. 네 번째 단계에서는 사용자가 제품에 부여할 평점을 예측하고, 마지막 다섯 번째 단계에서는 추천 성능을 평가한다.

1단계: 데이터 수집

첫 번째 단계에서는 전자상 거래 사이트 Amazon에서 제공하고 있는 1999년부터 2014년까지의 전자 제품 데이터를 사용한다. 데이터에는 사용자 정보, 제품 정보, 유용성 투표 수, 리뷰, 평점, 리뷰 작성 시간 등에 대한 정보가 포함되어 있고, 본 연구에서는 사용자 정보, 제품 정보, 평점 그리고 유용성 투표 수에 대한 정보를 활용한다.

2단계: 데이터 전처리

두 번째 단계에서는 수집한 데이터를 본 연구의 목적에 적합하도록 전처리하는 단계로, 리뷰의 유용성 투표 수를 활용하여 유용성 지수(Helpfulness Score)를 산출한다. 유용성 지수 산출 방식은 총 유용성 전체 투표 수를 유용한 리뷰 투표 수로 나눈다. 예를 들어, 총 30명의 사용자가 리뷰의 유용성에 대해 투표하였고, 그 중 15명이 유용하다고 투표를 하였으면 [15, 30]의 형태로 표현된다. 따라서 유용성 지수는 유용하다고 투표한 15명을 총 투표자 수 30명으로 나누면 0.5가 산출된다. 따라서 유용성 지수는 0에서 1사이의 값으로 도출되며, 1에 가까울수록 리뷰가 매우 유용하다는 것을 의미한다. 본 연구에서는 리뷰의 유용성 정보 반영에 따른 추천 성능을 평가하기 위해, 유용성 지수가 0인 데이터는 제외한다.

3단계: 추천 시스템 모델

세 번째 단계에서는 추천 성능을 평가하기 위해 추천 시스템에서 대표적으로 활용되고 있는 User-based CF 방식, Item-based CF 방식 그리고 SVD 모델을 적용하여 사용자가 제품에 부여할 평점을 예측한다(Postmus and Bhulai, 2018). User-based CF 방식은 사용자 간의 유사도를 기반으로 추천 대상 사용자에게 제품을 추천한다. 이를 위해 사용자 간의 유사도를 산출하며, 유사도는 식(1)과 같이 산출된다.

$$S_{(x,y)} = \frac{\sum_i (R_{x,i} - \bar{R}_x) \cdot \sum_i (R_{y,i} - \bar{R}_y)}{\sqrt{\sum_i (R_{x,i} - \bar{R}_x)^2} \cdot \sqrt{\sum_i (R_{y,i} - \bar{R}_y)^2}} \quad (1)$$

식(1)에서 $S_{(x,y)}$ 는 사용자 x 와 y 의 유사도를 나타내고, i 는 사용자 x 와 사용자 y 가 공통적으로 평가한 제품을 의미한다. $R_{x,i}$ 은 제품 i 에 대한 사용자 x 의 평점이며, $R_{y,i}$ 은 제품 i 에 대한 사용자 y 의 평점을 의미한다. \bar{R}_x 은 사용자 x 의 평점 평균값이고, \bar{R}_y 은 사용자 y 의 평점 평균값을 나타낸다. 따라서 사용자 간의 유사도가 산출되면, 이를 기반으로 추천 대상 사용자와 유사한 N 명의 이웃을 선택하게 된다. 그리고 선택된 이웃 사용자들이 제품에 부여한 평점을 기반으로 추천 대상 사용자가 제품에 부여할 평점을 예측하며, 식(2)와 같이 산출된다.

$$P_{x,i} = \bar{R}_x + \sum_{z \in N} (R_{z,i} - \bar{R}_z) \cdot \frac{s_{x,z}}{\sum_{z \in N} |s_{x,z}|} \quad (2)$$

식(2)에서 $P_{x,i}$ 는 제품 i 에 대한 사용자 x 의

의 평점을 나타낸다. \bar{R}_x 는 사용자 x 의 평점 평균값이며, $S_{x,z}$ 는 추천 대상 사용자 x 와 이웃 사용자 z 사이의 유사도를 의미한다. 그리고 N 은 선택된 가장 근접한 이웃들의 집합이며, z 는 각각의 이웃을 의미한다.

Item-based CF 방식은 사용자가 평점을 부여한 제품과 유사한 제품을 선정하고, 사용자가 제품에 부여할 평점을 예측한다. 따라서 $S(x, y)$ 는 제품 x 와 제품 y 간의 유사도를 의미하고 식(3)과 같이 산출된다.

$$S(x, y) = \frac{\sum_{u \in U} (R_{u,x} - \bar{R}_x)(R_{u,y} - \bar{R}_y)}{\sqrt{\sum_{u \in U} (R_{u,x} - \bar{R}_x)^2} \sqrt{\sum_{u \in U} (R_{u,y} - \bar{R}_y)^2}} \quad (3)$$

식(3)에서 U 는 제품 x 와 y 에 모두 평점을 부여한 전체 사용자 집합을 나타낸다. $R_{u,x}$ 와 $R_{u,y}$ 는 사용자 u 가 제품 x 와 제품 y 에 부여한 평점 평균값을 나타낸다. 그리고 가중 평균을 통해 추천 대상 사용자가 제품에 부여할 평점을 식(4)와 같이 산출하여 예측할 수 있다.

$$\hat{R}_{a,x} = \frac{\sum_{n \in N} R_{a,y} \cdot S(x, y)}{\sum_{n \in N} S(x, y)} \quad (4)$$

식(4)에서 선호도에 대한 평점은 제품 x 와 제품 y 간의 유사도 $S(x, y)$ 에 따라 계산되는 방식이다.

SVD 모델은 고차원 행렬을 저차원 행렬로 축소 시키는 원리이며, 하나의 행렬을 여러 행렬의 곱으로 분해하는 방식이다. 따라서 사용자와 제

품에 대한 $m \times n$ 크기의 행렬 $M = U \Sigma V^T$ 을 분해하면, 사용자 행렬 $U_{m \times n}$ 과 아이템 행렬 $V_{n \times n}^T$ 그리고 대각 행렬 $\Sigma_{m \times n}$ 로 구성된 총 3개의 행렬로 나타낼 수 있다. 따라서 행렬을 차원 축소한 다음 원래의 행렬로 복원하게 되면 결측값들이 채워지며, 이를 기반으로 추천을 한다. 따라서 본 연구에서는 User-based CF 방식, Item-based CF 방식 그리고 SVD 모델을 사용하여 추천 성능을 비교하여 평가하고자 한다.

4단계: 추천 성능 평가

본 연구에서는 추천 성능을 평가하기 위해 추천 대상 사용자의 실제 평점과 예측한 평점 간의 차이를 비교할 수 있는 MAE(Mean Absolute Error)를 활용하였다. MAE는 값이 작을수록 예측 정확도가 높다는 것을 의미하며, 식(5)와 같이 산출된다.

$$MAE = \frac{\sum_{i=1}^n |A_i - B_i|}{n} \quad (5)$$

식(5)에서 A_i 는 예측값, B_i 는 추천 대상 사용자의 평가 값, n 은 모든 사용자들이 평가한 아이템의 수를 나타낸다.

4. 연구 결과

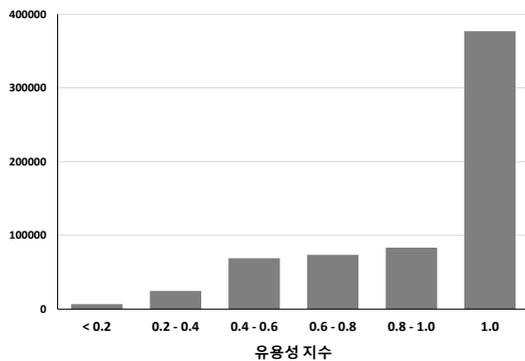
4.1. 실험 데이터

본 연구에서는 1999년부터 2014년까지의 아마존 전자 제품 데이터 중 사용자, 제품, 평점 그리고 리뷰에 대한 유용성 정보를 추출하여 실험을

진행하였다. 따라서 최종적으로 <Table 1>과 같이 총 192,403명의 사용자가 63,001개의 제품에 대해 부여한 1,689,188개의 평점과 리뷰 유용성 정보를 활용하였다. 또한 유용성 지수를 산출한 분포도 결과는 <Figure 3>과 같다.

<Table 1> Statistics of Amazon Electronic Products Dataset

사용자	아이템	평점 & 리뷰	리뷰 유용성
192,403	63,001	1,689,188	1,689,188



<Figure 3> Helpfulness Score Distribution

4.1.1. 평점과 유용성 지수 상관분석

평점과 유용성 지수 간의 상관관계수는 <Table 2>와 같이 0.314로 양(+)의 상관관계가 있는 것으로 확인되었으며, 0.001% 수준에서 통계적으로 유의미한 것으로 나타났다. 즉, 평점이 높을수록 유용성 지수는 증가한다고 볼 수 있다.

<Table 2> Correlation Analysis Results

	평점	유용성 지수
평점	1	
유용성 지수	0.314***	1

***p < 0.001

4.1.2. 평점과 유용성 지수 회귀분석

평점이 리뷰 유용성에 끼치는 영향을 파악하기 위해 회귀분석을 이용하였으며, 그 결과는 <Table 3>과 같다. 이 모형에 대한 설명력은 31.4%, 수정된 설명력은 9.8%, Durbin-Watson은 1.744으로 나타났고 0.001% 수준에서 통계적으로 유의미한 것으로 확인되었다. 분석 결과, 평점은 0.001% 수준에서 통계적으로 유의하며 리뷰 유용성에 정(+)의 영향을 끼치는 것으로 확인되었다. 따라서 평점이 높을수록 유용성 지수도 높아지므로 평점이 높은 리뷰일수록 사용자가 유용하다고 인지하는 것으로 볼 수 있다.

<Table 3> Regression Analysis Results

모형	비표준화 계수		표준화 계수	t	유의 확률
	B	SE	β		
(상수)	0.648	0.001		768.169	0.000
평점	0.052	0.000	0.314	262.83	0.000

a. 종속변수 : 유용성 지수

4.2. 분석 결과

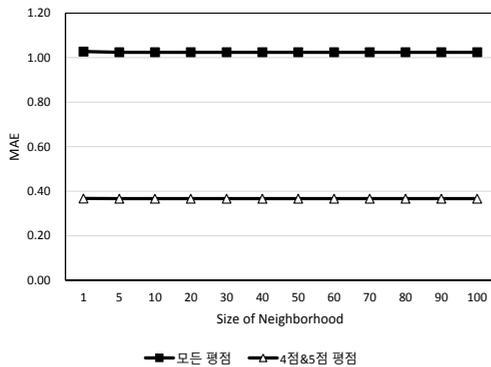
본 연구에서는 User-based CF, Item-based CF 그리고 SVD 모델의 추천 성능을 비교하여 평가하였다. User-based CF와 Item-based CF의 최근 접 이웃 수는 1, 5, 10, 20, 30, 40, 50, 60, 70, 80,

90, 100으로 설정하였고, SVD 모델의 특이값 (Singular Value) 수는 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100으로 설정하여 평점을 예측하였다. 그리고 실험 데이터는 <Table 4>와 같이 모든 평점과 4점과 5점 평점 데이터를 분리하여 실험을 진행하였다.

<Table 4> All Rating & 4, 5 Rating Dataset

데이터	User	Item	Review & Helpfulness
전체 평점	192,403	63,001	1,689,188
4 점, 5 점 평점	153,363	57,688	475,239

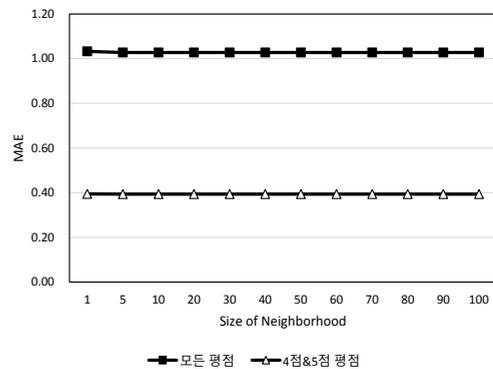
모든 평점과 4점과 5점 평점 데이터를 반영한 User-based CF의 MAE 결과는 <Figure 4>와 같다.



<Figure 4> User-based CF MAE 측정 결과

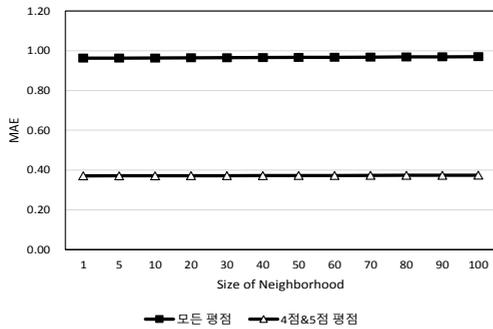
User-based CF의 MAE 측정 결과, 모든 평점을 사용하여 실험을 진행한 경우에는 최근접 이웃 수를 40으로 설정하였을 때 가장 높은 성능을 보였고, 4점과 5점 평점만 사용하여 실험을 진행한 경우에는 최근접 이웃 수가 100일 때 가장 높은

성능을 보였다. 또한 4점과 5점 평점만을 사용한 경우가 모든 평점을 사용한 것보다 전반적으로 높은 성능을 보였다. Item-based CF의 MAE 측정 결과는 <Figure 5>와 같으며 모든 평점을 사용하여 실험을 진행한 경우에는 최근접 이웃 수를 20으로 설정하였을 때 가장 높은 성능을 보였고, 4점과 5점 평점만 사용하여 실험을 진행한 경우에는 최근접 이웃 수가 5일 때 가장 높은 성능을 보였다.



<Figure 5> Item-based CF MAE 측정 결과

또한 User-based CF와 마찬가지로 4점과 5점 평점만을 사용한 경우가 전체 데이터를 사용한 것보다 전반적으로 높은 성능을 보였다. SVD 모델의 MAE 결과는 <Figure 6>과 같으며, 모든 평점을 사용하여 실험을 진행한 경우와 4점과 5점 평점만을 사용한 경우 모두 이웃 수를 1로 설정하였을 때 가장 높은 성능을 보였다. 또한 4점과 5점 평점만 사용하여 실험을 진행한 경우가 모든 평점을 사용한 것보다 전반적으로 높은 성능을 보였다. 따라서 User-based CF, Item-based CF 그리고 SVD 모델 모두 4점과 5점 평점만을 사용하여 실험을 진행한 결과가 모든 평점을 사용한 결과보다 높은 성능을 보였다.



(Figure 6) SVD 모델 MAE 측정 결과

그리고 모든 평점과 4점과 5점 평점만을 사용하여 실험을 진행하였을 때의 User-based CF, Item-based CF 그리고 SVD 모델 간의 추천 성능에 변별이 존재하는지에 대해 Kruskal-Wallis 검정을 통하여 <Table 5>와 같이 검정 통계량을 구하였다.

(Table 5) Kruskal-Wallis Test

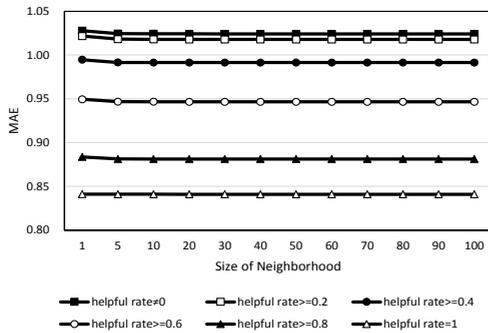
평정 항목	검정 통계량	수정된 유의확률
Group (4)-(6)	-12.500	1.000
Group (4)-(5)	-22.654	0.101
Group (4)-(3)	36.000	0.000
Group (4)-(1)	48.917	0.000
Group (4)-(2)	59.083	0.000
Group (6)-(5)	10.154	1.000
Group (6)-(3)	23.500	0.106
Group (6)-(1)	36.417	0.000
Group (6)-(2)	46.583	0.000
Group (5)-(3)	13.346	1.000
Group (5)-(1)	26.263	0.025
Group (5)-(2)	36.429	0.000
Group (3)-(1)	12.917	1.000
Group (3)-(2)	23.083	0.102
Group (1)-(2)	-10.167	1.000

Group (1) : Used-based CF 에 모든 평점을 반영한 MAE 측정 결과

- Group (2) : Item-based CF 에 모든 평점을 반영한 MAE 측정 결과
- Group (3) : SVD 모델에 모든 평점을 반영한 MAE 측정 결과
- Group (4) : Used-based CF 에 4 점과 5 점 평점만 반영한 MAE 측정 결과
- Group (5) : Item-based CF 에 4 점과 5 점 평점만 반영한 MAE 측정 결과
- Group (6) : SVD 모델에 4 점과 5 점 평점만 반영한 MAE 측정 결과

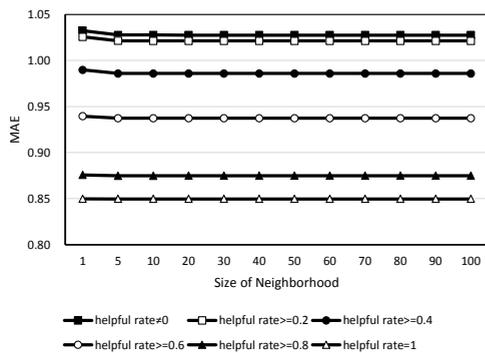
<Table 5>에서 분포를 검정한 결과 (4)와 (3), (1), (2)의 분포가 .000의 유의수준에서 유의미한 추천 성능의 차이를 보이는 것으로 나타났다. (6)과 (1), (2)의 분포 또한 .000의 유의수준에서 유의미한 추천 성능의 차이를 보이는 것으로 나타났다. (5)와 (1)은 .05의 유의수준에서 (5)와 (2)는 .000의 유의수준에서 유의미한 추천 성능의 차이를 보이는 것으로 확인되었다.

본 연구에서는 산출한 유용성 지수를 6개 범위로 구분하고, 이를 User-based CF, Item-based CF 그리고 SVD 모델에 반영하여 추천 성능을 비교하여 평가하였다. User-based CF와 Item-based CF의 최근접 이웃 수는 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100으로 설정하였고, SVD 모델의 특이값 수는 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100으로 설정하여 평점을 예측하였고, 이에 대한 MAE를 측정하였다. User-based CF는 <Figure 7>과 같이 유용성 지수가 0보다 큰 경우에는 최근접 이웃 수가 40, 유용성 지수가 0.2 이상인 경우에는 최근접 이웃 수가 30, 유용성 지수가 0.4 이상인 경우에는 최근접 이웃 수가 60, 유용성 지수가 0.6 이상인 경우에는 최근접 이웃 수가 30, 유용성 지수가 0.8 이상인 경우에는 최근접 이웃 수가 30일 때 그리고 유용성 지수가 1인 경우에는 최근접 이웃 수가 20일 때 가장 높은 성능을 보였으며, 유용성 지수가 증가할수록 MAE 값은 감소하였다.



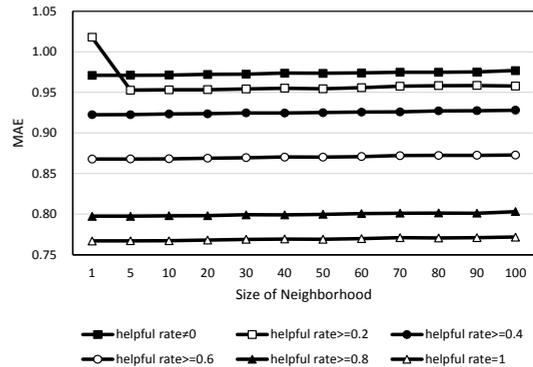
〈Figure 7〉 유용성 지수를 반영한 User-based CF MAE 측정 결과

Item-based CF는 <Figure 8>과 같이 유용성 지수가 0보다 큰 경우에는 최근접 이웃 수가 40일 때, 유용성 지수가 0.2 이상인 경우에는 최근접 이웃 수가 20일 때, 유용성 지수가 0.4 이상인 경우에는 최근접 이웃 수가 10일 때, 유용성 지수가 0.6 이상인 경우에는 최근접 이웃 수가 20일 때, 유용성 지수가 0.8 이상인 경우에는 최근접 이웃 수가 5일 때 그리고 유용성 지수가 1인 경우에는 최근접 이웃 수가 5일 때 가장 우수한 성능을 보였으며, 유용성 지수가 증가할수록 MAE 값은 감소하였다.



〈Figure 8〉 유용성 지수를 반영한 Item-based CF MAE 측정 결과

SVD 모델은 <Figure 9>와 같이 유용성 지수가 0보다 큰 경우에는 최근접 이웃 수가 1일 때, 유용성 지수가 0.2 이상인 경우에는 최근접 이웃 수가 5일 때, 유용성 지수가 0.4 이상인 경우에는 최근접 이웃 수가 1일 때, 유용성 지수가 0.6 이상인 경우에는 최근접 이웃 수가 1일 때, 유용성 지수가 0.8 이상인 경우에는 최근접 이웃 수가 5일 때 그리고 유용성 지수가 1인 경우에는 최근접 이웃 수가 1일 때 결과 가장 높은 성능을 보였으며, 유용성 지수가 증가할수록 MAE 값은 감소하였다. 따라서 User-based CF, Item-based CF, SVD 모델 모두 유용성 지수가 증가할수록 MAE 값이 감소하고 추천 성능은 향상되는 것으로 확인하였다.



〈Figure 9〉 유용성 지수를 반영한 SVD 모델 MAE 측정 결과

그리고 유용성 지수를 반영한 User-based CF, Item-based CF 그리고 SVD 모델 간의 추천 성능에 변별이 존재하는지에 대해 Kruskal-Wallis 검정을 통하여 <Table 6>과 같이 검정 통계량을 구하였다. <Table 6>에서 분포를 검정한 결과 (6)과 (4), (3), (2), (1)의 분포가 .000의 유의수준에서 유의미한 추천 성능의 차이를 보이는 것으로 나

타났다. (5)와 (3), (2), (1)의 분포 또한 .000의 유의수준에서 유의미한 추천 성능의 차이를 보이는 것으로 나타났고, (4)와 (3) 그리고 (3)과 (1)은 .05의 유의수준에서 (4)와 (2)는 .000의 유의수준에서 유의미한 추천 성능의 차이를 보이는 것으로 확인되었다.

〈Table 6〉 Kruskal-Wallis Test with Helpfulness Score

표본 1-표본 2	검정 통계량	수정된 유의확률
Group (6)-(5)	28.000	0.859
Group (6)-(4)	64.000	0.000
Group (6)-(3)	107.333	0.000
Group (6)-(2)	137.306	0.000
Group (6)-(1)	155.361	0.000
Group (5)-(4)	36.000	0.218
Group (5)-(3)	79.333	0.000
Group (5)-(2)	109.306	0.000
Group (5)-(1)	127.361	0.000
Group (4)-(3)	43.333	0.049
Group (4)-(2)	73.306	0.000
Group (4)-(1)	91.361	0.000
Group (3)-(2)	29.972	0.628
Group (3)-(1)	48.028	0.017
Group (2)-(1)	18.056	1.000

Group (1) : Used-based CF, Item-based CF, SVD 모델에 유용성 지수가 0 보다 큰 데이터를 반영한 MAE 측정 결과

Group (2) : Used-based CF, Item-based CF, SVD 모델에 유용성 지수가 0.2 이상인 데이터를 반영한 MAE 측정 결과

Group (3) : Used-based CF, Item-based CF, SVD 모델에 유용성 지수가 0.4 이상인 데이터를 반영한 MAE 측정 결과

Group (4) : Used-based CF, Item-based CF, SVD 모델에 유용성 지수가 0.6 이상인 데이터를 반영한 MAE 측정 결과

Group (5) : Used-based CF, Item-based CF, SVD 모델에 유용성 지수가 0.8 이상인 데이터를 반영한 MAE 측정 결과

Group (6) : Used-based CF, Item-based CF, SVD 모델에 유용성 지수가 1.0 인 데이터를 반영한 MAE 측정 결과

5. 결론

본 연구에서는 평점에 따른 추천 성능과 리뷰 유용성을 반영한 추천 성능을 비교하였다. 이를 위해, User-based CF, Item-based CF 그리고 SVD 모델과 Amazon에서 수집한 전자 제품 데이터를 사용하여 실험을 진행하고 추천 성능을 평가하였다. 실험은 모든 평점을 추천 시스템에 반영한 결과와 4점과 5점 평점만을 추천 시스템에 반영한 결과를 비교하였고, 마지막으로 리뷰 유용성 정보를 추천 시스템에 반영한 결과를 살펴보았다. 따라서 이에 대한 본 연구의 학술적 의의는 다음과 같다. 첫째, 모든 평점을 활용한 경우보다 4점과 5점 평점만을 사용하였을 때의 추천 성능이 전반적으로 향상되었다. 사용자가 제품에 부여한 평점이 4점 또는 5점의 경우에는 해당 제품에 대한 사용자의 만족도가 높은 것으로 볼 수 있다. 이는 추천 대상 사용자의 최근접 이웃을 선정할 때 유사하지 않은 사용자가 포함될 확률이 감소하기 때문에 추천 성능이 향상된 것으로 보이며, 이를 통해 특이 취향의 사용자에 대해서도 비교적 정확한 예측이 가능하다. 둘째, 리뷰 유용성 정보를 추천 시스템에 반영한 결과는 유용성 지수가 높을수록 추천 성능이 향상되는 것을 확인하였다. 따라서 리뷰의 유용성 정보가 추천 성능에 유의미한 영향을 끼칠 수 있으며, 이는 제품 및 서비스에 대한 리뷰가 유용하면 기존의 사용자 선호도 평점 정보보다 더 높은 신뢰성을 제공하는 것을 의미한다. 본 연구의 실무적 시사점으로는 첫째, 리뷰가 사용자에게 유용할수록 추천 성능이 향상됨을 확인하였다. 그러나 현재 전자상거래 사이트에서는 리뷰의 유용성 여부에 대해 투표할 수 있는 기능을 제공하고 있지만, 이를 추천 서비스에 적용한 경우는 현재

많지 않은 실정이다. 따라서 기업은 리뷰의 유용성 정보를 추천 서비스에 적용한다면 기존보다 더 사용자의 서비스 만족도를 향상시킬 수 있을 것으로 본다. 둘째, 기존에는 단순히 제품 및 서비스에 대한 리뷰를 작성할 수 있도록 유도하였는데, 사용자들이 리뷰를 상세하게 작성할 수 있도록 유도할 필요가 있다고 본다. 따라서 유용성 투표수가 많은 리뷰의 작성자에게는 할인 쿠폰, 마일리지 등의 보상을 제공한다면 리뷰의 품질을 높일 수 있고, 이를 통해 제품을 사용자의 구매 의사결정에 기여할 수 있을 것으로 본다. 셋째, 본 연구의 결과는 리뷰의 유용성 정보를 제공하는 다양한 분야의 전자상거래 사이트에 적용할 수 있으며, 이를 통해 사용자들의 서비스 만족도와 편의성을 개선할 수 있을 것으로 본다. 또한 기업의 고객 관리, 마케팅 등의 다양한 분야에서도 유용하게 활용할 수 있을 것으로 보며, 더 나아가 기업의 매출 상승에도 기여할 수 있을 것으로 본다. 그러나 본 연구는 이러한 시사점을 제공하고 있음에도 불구하고 다음과 같은 한계점이 존재한다. 첫째, 본 연구에서는 전자 제품 데이터를 사용하여 추천 성능을 평가하였다. 그러나 타 분야에 적용하기 위해서는 추가적으로 데이터를 수집하여 추천 성능을 검증할 필요가 있다. 또한 본 연구에서는 2014년까지의 데이터를 사용하여 결론을 도출하였다. 따라서 본 연구에서 도출한 결론이 현재에도 적용이 가능한지에 대한 검증이 필요하며, 이를 위해 최근 몇 년 이내의 데이터를 추가적으로 수집하여 분석을 진행할 필요가 있다. 둘째, 본 연구에서는 추천 시스템에서 대표적으로 사용되는 협업 필터링 방식을 사용하여 추천 성능을 검증하였는데, 최근에는 다양한 딥러닝 모델들이 추천 시스템 분야에 적용되고 우수한 성능을 보이고 있다. 따라

서 향후에는 딥러닝 모델을 활용하여 추천 성능을 검증하고 개선할 수 있다. 셋째, 본 연구에서는 리뷰의 유용성 여부에 대한 사용자의 투표수를 기반으로 유용성 지수를 산출하였는데, 향후에는 리뷰의 길이, 사진의 유무 여부 등의 정보를 포함한다면 더욱 정교한 리뷰 유용성 지수를 산출할 것으로 본다. 넷째, 일반적으로 전자상거래 사이트에서 기존에 출시된 제품에 대한 리뷰는 새롭게 출시된 제품의 리뷰보다 더 많은 유용성 평가를 받게 된다. 따라서 향후에는 제품의 출시 날짜 등의 정보를 고려하여 리뷰의 유용성 정보를 반영할 필요가 있다.

참고문헌(References)

- Abdollahi, B., and O. Nasraoui, "Using explainability for constrained matrix factorization," *In 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.
- Al-Smadi, M., B. Talafha, M. Al-Ayyoub, and Y. Jararweh, "Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews," *International Journal of Machine Learning and Cybernetics*, Vol.10, No.8(2019), 2163-2175.
- Bang, H., H. Lee, and J. H. Lee, "TV Program recommender system using viewing time patterns," *Journal of the Korean Institute of Intelligent Systems*, Vol.25, No.5(2015), 431-436.

- Castelli, M., L. Manzoni, L. Vanneschi, and A. Popovic, "An expert system for extracting knowledge from customers' reviews: The case of amazon. com, inc.," *Expert Systems with Applications*, Vol.84, 117-126.
- Chintagunta, P. K., S. Gopinath, and S. Venkataraman, "The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets," *Marketing science*, Vol.29, No.5(2010), 944-957.
- Choeh, J. Y., S. K. Lee, and Y. B. Cho, "Applying rating score's reliability of customers to enhance prediction accuracy in recommender system," *The Journal of the Korea Contents Association*, Vol.13, No.7(2013), 379-385.
- Eslami, S. P., M. Ghasemaghaei, and K. Hassanein, "Which online reviews do consumers find most helpful? A multi-method investigation," *Decision Support Systems*, Vol.113, (2018) 32-42.
- Fang, Q., C. Xu, M. S. Hossain, and G. Muhammad, "Stcaplrs: A spatial-temporal context-aware personalized location recommendation system," *ACM Transactions on Intelligent systems and technology (TIST)*, Vol.7, No.4(2016), 1-30.
- 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, L. R., "The development of markers for the Big-Five factor structure," *Psychological assessment*, Vol.4, No.1(1992), 26.
- Gomez-Uribe, C. A., and N. Hunt, "The netflix recommender system: Algorithms, business value, and innovation," *ACM Transactions on Management Information Systems (TMIS)*, Vol.6, No.4(2015), 1-19.
- Koohi, H., & K. Kiani, "User based collaborative filtering using fuzzy C-means," *Measurement*, 91, (2016), 134-139.
- Lee, S. H., A. R. Jo, and H. Y. Lee, "The Medical Service Customers Satisfaction Factors Extracted from Online Hospital Review Data Using Latent Dirichlet Allocation Method," *Journal of Korea Service Management Society*, Vol.18, No.5(2017), 23-44.
- 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.
- Lu, L., W. Xu, and S. Qiao, "A fast SVD for multilevel block Hankel matrices with minimal memory storage," *Numerical Algorithms*, Vol.69, No.4(2015), 875-891.
- 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.
- Nakayama, M., and Y. Wan, "The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews," *Information & Management*, Vol.56, No.2(2019), 271-279.
- Pradel, B., N. Usunier, and P. Gallinari, "Ranking with non-random missing ratings: influence of popularity and positivity on evaluation metrics," *In Proceedings of the sixth ACM conference on Recommender systems*, (2012), 147-154
- Qiu, J., J. Wang, S. Yao, K. Guo, B. Li, E. Zhou, and H. Yang, "Going deeper with embedded

- fpga platform for convolutional neural network,” *In Proceedings of the 2016 ACM/ SIGDA International Symposium on Field-Programmable Gate Arrays*, (2016), 26-35.
- Qiu, J., Q. Wu, G. Ding, Y. Xu, and S. Feng, “A survey of machine learning for big data processing,” *EURASIP Journal on Advances in Signal Processing*, Vol.2016, No.1(2016), 1-16.
- Ren, G., and T. Hong, “Examining the relationship between specific negative emotions and the perceived helpfulness of online reviews,” *Information Processing & Management*, Vol.56, No.4(2019), 1425-1438.
- Shani, G., and A. Gunawardana, “Evaluating recommendation systems,” *In Recommender systems handbook*, Springer, Boston, 2011.
- Shengli, L., and L. Fan, “The interaction effects of online reviews and free samples on consumers’ downloads: An empirical analysis,” *Information Processing & Management*, Vol.56, No.6(2019), 102071.
- Smith, B., and G. Linden, “Two decades of recommender systems at Amazon. com,” *Ieee internet computing*, Vol.21, No.3(2017), 12-18.
- Srifi, M., A. Oussous, A. Ait 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), 2009.
- Sun, Y., Z. Wang, P. Fu, Q. Jiang, T. Yang, J. Li, and X. Ge, “The impact of relative humidity on aerosol composition and evolution processes during wintertime in Beijing, China,” *Atmospheric Environment*, Vol.77, (2013), 927-934.
- Thakkar, P., K. Varma, V. Ukani, S. Mankad, and S. Tanwar, “Combining user-based and item-based collaborative filtering using machine learning,” *In Information and Communication Technology for Intelligent Systems*, Springer, Singapore, 2019.
- Wang, Y., J. Wang, and T. Yao, “What makes a helpful online review? A meta-analysis of review characteristics,” *Electronic Commerce Research*, Vol.19, No.2(2019), 257-284.
- 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.
- Zarzour, H., Z. Al-Sharif, M. Al-Ayyoub, and Y. Jararweh, “A new collaborative filtering recommendation algorithm based on dimensionality reduction and clustering techniques,” *In 2018 9th international conference on information and communication systems (ICICS)*, (2018), 102-106.

Abstract

Analysis of the Effects of E-commerce User Ratings and Review Helpfulness on Performance Improvement of Product Recommender System

LIU FAN* · Byunghyun Lee** · Ilyoung Choi*** · Jaeho Jeong**** · Jaekyeong Kim*****

Because of the spread of smartphones due to the development of information and communication technology, online shopping mall services can be used on computers and mobile devices. As a result, the number of users using the online shopping mall service increases rapidly, and the types of products traded are also growing. Therefore, to maximize profits, companies need to provide information that may interest users. To this end, the recommendation system presents necessary information or products to the user based on the user's past behavioral data or behavioral purchase records. Representative overseas companies that currently provide recommendation services include Netflix, Amazon, and YouTube. These companies support users' purchase decisions by recommending products to users using ratings, purchase records, and clickstream data that users give to the items. In addition, users refer to the ratings left by other users about the product before buying a product. Most users tend to provide ratings only to products they are satisfied with, and the higher the rating, the higher the purchase intention. And recently, e-commerce sites have provided users with the ability to vote on whether product reviews are helpful. Through this, the user makes a purchase decision by referring to reviews and ratings of products judged to be beneficial. Therefore, in this study, the correlation between the product rating and the helpful information of the review is identified. The valuable data of the evaluation is reflected in the recommendation system to check the recommendation performance. In addition, we want to compare the results of skipping all the ratings in the traditional collaborative filtering technique with the recommended performance results that reflect only the 4 and 5 ratings. For this purpose, electronic product data collected from Amazon was used in

* Department of Business Administration, Graduate School, Kyung Hee University

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

*** Graduate School of Business Administration, Kyung Hee University

**** Corresponding author: Jaeho Jeong

Department of Business Administration, Graduate School, Kyung Hee University

26, Kyungheedaero-ro, Dongdaemun-gu, Seoul 02447, Republic of Korea

Tel: +82-2-961-935, Fax: +82-2-961-935, E-mail: jjh8007@msn.com

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

this study, and the experimental results confirmed a correlation between ratings and review usefulness information. In addition, as a result of comparing the recommendation performance by reflecting all the ratings and only the 4 and 5 points in the recommendation system, the recommendation performance of remembering only the 4 and 5 points in the recommendation system was higher. In addition, as a result of reflecting review usefulness information in the recommendation system, it was confirmed that the more valuable the review, the higher the recommendation performance. Therefore, these experimental results are expected to improve the performance of personalized recommendation services in the future and provide implications for e-commerce sites.

Key Words : E-Commerce, Recommender System, Amazon Data, Review Helpfulness, Collaborative Filtering

Received : February 27, 2022 Revised : March 21, 2022 Accepted : March 23, 2022

Corresponding Author : Jaeho Jeong

저자 소개



LIU FAN

경희대학교 대학원 경영학과에서 경영학 석사과정을 수료하였다. 주요 관심 분야는 빅데이터 분석, 딥러닝, 개인화 추천 서비스, 자연어 처리 등이다.



이병현

경희대학교 대학원 빅데이터응용학과에서 경영학 석사학위를 취득하고, 현재 빅데이터 응용학과 박사과정에 재학 중이다. 주요 관심 분야는 개인화 추천 서비스, 딥러닝, 자연어 처리, 데이터 마이닝, 소셜 네트워크 분석 등이다.



최일영

경희대학교에서 경제학 학사, 동 대학원에서 경영정보시스템 전공으로 경영학 석사, 박사 학위를 취득하였다. 주요 관심분야는 빅데이터 분석, 딥러닝, 추천 시스템, 그린 비즈니스/IT, 비즈니스 인텔리전스, 사회네트워크분석 등이며 *Information Technology and Management*, *International Journal of Information Management*, 경영과학회지, 경영과학, 정보관리학회지, 지능정보연구 등 다수의 학술지에 논문을 게재하였다.



정재호

경희대학교에서 MBA 석사학위를 취득하였으며, 동 대학원에서 경영학 박사과정을 수료하였다. 현재 농업회사법인(주) 초록들 대표이사를 맡고 있다. 주요 관심분야는 빅데이터 분석, 딥러닝, 추천 시스템 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 (빅데이터분야) 및 AI 비즈니스 연구센터 센터장을 맡고 있다.