

댓글 특성 기반 고객 프로파일 구축을 통한 개인화 추천 서비스에 관한 연구

엄금철^{1*} · 이가은² · 이청용³

¹가천대학교 경영대학 글로벌경영학과 교수

²경희대학교 무역학과 박사수료

³경희대학교 소셜네트워크과학과 석사과정

A Study on the Personalization Recommendation Service Incorporating Review-based Customer Profile

JinZhe Yan^{1*} · JiaEn Li² · QingLong Li³

¹Professor, College of Business, Gachon University, Seongnam 13120, Korea

²Ph.D. Candidate, Department of International Business&Trade, Kyunghee University, Seoul 02447, Korea

³Master's Course, Department of Social Network Science, Kyunghee University, Seoul 02447, Korea

[요 약]

전 세계적으로 전자상거래 시장의 규모가 급속하게 커지면서 개인화 추천 서비스 관련 연구가 꾸준히 이루어지고 있다. 기존 개인화 추천 서비스 연구에서는 주로 고객의 명시적 평점 (Explicit Rating)과 암묵적 데이터 (Implicit Feedback)를 활용하여 고객의 선호도를 예측했다. 하지만 정량적인 데이터만 사용하면 추천의 정확도가 떨어진다. 본 연구에서는 고객 평점만을 고려하는 기존 추천 방법론의 한계를 극복하기 위해 고객의 정성적 선호도를 나타내는 댓글 데이터를 사용하여 고객에게 맞춤형 상품을 추천하는 방법을 제안하고자 한다. 본 연구에서는 CRSE 모델을 구축하고, 이를 통해 상품 구매할 때 고객에게 중요한 정보를 제공하는 댓글 데이터를 정교하게 분석하여 정량적인 선호도 점수를 산출하여 새로운 고객 프로파일을 구축하였다. 실험 결과, 본 연구에서 제안한 방법은 기존 추천 기법들보다 예측 정확도가 더 우수한 것으로 나타났다.

[Abstract]

Global e-commerce companies offer personalized recommendation services to gain sustainable competitive advantages. Previous personalized recommendation service studies mainly used customer's explicit rating and implicit data to predict customer preferences. However, the recommendation accuracy is low when only using quantitative data. To overcome the limitations of the existing recommender system that only considers customer ratings, the authors proposed a novel method to recommend products to customers using review data that represent their qualitative preferences. This study proposed a CRSE model and created a new customer profile. The experiment results show that the predictive accuracy of the proposed methods is superior to existing recommended techniques.

색인어 : 개인화 추천 서비스, 협업 필터링, 심층 신경망, 감성 분석

Key word : Personalization Recommender service, Collaboration Filtering, Deep Neural Network, Sensitivity Analysis

<http://dx.doi.org/10.9728/dcs.2020.21.9.1575>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 06 August 2020; **Revised** 02 September 2020

Accepted 02 September 2020

***Corresponding Author; JinZhe Yan**

Tel: +82-31-750-4460

E-mail: yanjz@gachon.ac.kr

I. 서론

정보 기술의 발달과 모바일 기기의 대중화로 인해 전 세계적으로 전자상거래 시장의 규모가 급속하게 커지고 있다. 이러한 배경하에 매년 수많은 새로운 상품이 계속 출시되고 있지만, 고객들은 원하는 상품을 선택할 때 오히려 더 많은 시간이 걸리고 어려움을 겪고 있다. 이에 따라 고객의 적성에 적합한 상품이나 서비스를 제공하는 개인화 추천 서비스의 중요성이 대두되고 있다. Amazon[1], Netflix[2], Youtube[3] 등 세계적인 기업들은 오래전부터 축적된 고객의 구매 기록을 기반으로 고객에게 개인화 추천 서비스를 제공하여 기업 경쟁력을 강화하고 있다.

고객에게 맞춤형 추천을 제공하는 기존 개인화 추천 서비스 연구에서는 주로 고객의 명시적 평점 (Explicit Rating)과 구매 여부 혹은 방문 여부를 나타내는 암묵적 데이터 (Implicit Data)를 활용하여 고객의 선호도를 예측했다[4]. 최근 연구에서 이러한 정량적인 데이터를 기반으로 개인화 추천 서비스를 제공하면 추천의 정확도가 떨어질 수 있다는 문제가 꾸준히 제기되고 있다[5]. 예를 들어, 고객 1과 고객 2는 같은 상품을 구매 후 모두 평점 5점을 부여했을 때 해당 상품에 대한 정량적인 선호도는 같다고 볼 수 있다. 그러나 고객 1은 배송 속도와 사은품 서비스 등 전반적으로 만족하여 평점 5점을 부여했고, 고객 2는 배송 속도에 만족하지 못하지만, 상품의 품질과 디자인에 만족하여 평점 5점을 부여했다. 이처럼 두 고객은 같은 상품에 같은 평점을 부여했지만, 해당 상품에 대한 선호도는 같다고 보기는 어렵다. 고객이 특정 상품을 구매하거나 선호하는 이유 등을 파악할 수 있는 정성적인 선호도를 고려하지 않고, 단순히 정량적인 평점이나 구매 여부 등 데이터만을 고려하여 상품을 추천하는 것은 추천의 정확도를 떨어뜨리는 요소로 될 수 있다.

최근에는 이러한 기존 연구의 한계를 극복하기 위해 고객의 선호도 예측함에 있어서 다양한 유형의 데이터들이 사용되고 있다. 그중에서 가장 대표적으로 사용되는 것이 고객이 상품에 대하여 정성적인 선호도를 나타낼 수 있는 댓글 데이터이다[6]. 정보 기술이 발달하면서 대부분의 전자상거래 플랫폼에서는 고객이 구매한 제품에 대한 선호를 나타낼 수 있는 댓글 서비스를 제공하기 시작하면서, 고객들은 직접 댓글을 통해 상품에 자신의 선호도를 나타내고 있다[7]. 이와 같은 댓글 데이터는 상품에 대한 고객이 선호도가 구체적이고 신뢰할 수 있는 정보가 포함하기에 개인화 추천 서비스 제공에 유용하게 사용될 수 있다[8]. 이에 따라, 본 연구는 감성 분석을 통해 정성적인 댓글 데이터를 정량적인 데이터로 변환하여 새로운 고객 프로파일을 구축하여 고객에 맞춤형 상품을 추천하는 방법을 제안하고자 한다.

본 연구에서는 새롭게 제안하는 개인화 추천 서비스 방법론은 평점, 구매 여부 등 정량적 데이터만을 사용하여 고객에게 추천을 제공하는 한계를 개선하기 위해 고객이 남긴 댓글 데이터를 사용하여 새로운 선호도 프로파일을 구축하여 고객에 상품을 추천하도록 설계되었다. 이를 위해, 최근 댓글의 감성 특

성 추출에 많이 사용되고 있는 CNN(convolutional neural network) 알고리즘을 적용하여 고객이 남긴 댓글에서 감성 특성을 추출하여 상품에 대한 선호도를 정교하게 예측하고자 한다. 본 연구에서는 기존 연구에서 시도하지 않은 고객이 작성한 댓글 활용하여 상품에 대한 선호도를 예측하여 새로운 고객 프로파일을 구축하여 상품을 추천하는 방법을 새롭게 제안한다.

본 연구의 구성은 다음과 같다. 2장에서는 개인화 추천 서비스와 감성 분석 기반 개인화 추천 서비스에 대한 이론적 배경을 서술하고, 관련 선행연구에 대하여 살펴본다. 3장에서는 본 연구에서 제안하는 추천 방법에 대해 자세하게 설명한다. 4장에서는 제안하는 추천 방법의 성과를 평가하는 방법과 실험 결과를 기술한다. 마지막으로 5장에서는 본 연구의 결론을 기술하고, 본 연구의 한계점 및 향후 연구계획에 대하여 기술한다.

II. 관련 연구

2-1 개인화 추천 서비스

개인화 추천 서비스는 고객의 과거 구매 기록이나 평점 등을 활용하여 고객에게 가장 적합한 상품 또는 서비스를 제공하는 것을 목적으로 하고 있다. 기본적으로 많이 사용되고 있는 CF(Collaborative filtering) 기법은 1990년대에 처음으로 소개된 이후 개인화 추천 서비스 중에서 현재까지 가장 우수한 성능을 나타내고 있는 기법이다[9]. CF 기법은 기본적으로 고객 혹은 상품 간 유사도를 기반으로 고객의 선호를 예측한다. 추천 대상 고객이 정해지면 해당 고객의 구매 기록을 바탕으로 다른 고객들 간의 유사도를 측정하고 상품을 추천한다. 하지만 새로운 고객의 경우 유사도를 측정할 수 있는 충분한 구매 기록이나 평점 정보가 존재하지 않을 수 있다. 이처럼 CF 기법은 새로운 고객의 선호도를 예측할 수 없는 Cold Star 문제와 고객이 선호하는 상품이라도 아직 구매가 이루어지지 않아서 추천 리스트를 생성할 수 없는 First Star 문제도 존재하고 있다[10]. 또한, 고객의 구매 기록이 기하급수적으로 증가하며 따라 추천 알고리즘의 계산량이 급증하게 되어 연산에 많은 시간과 비용이 발생하는 확장성 (Scalability) 문제도 존재한다[11].

최근 많은 연구에서는 CF 기법의 장점을 극대화하고 단점을 보완하기 위해 다양한 연구를 시도하고 있으며, 그중에서도 대표적인 예는 CF 기법에 DNN(Deep Neural Network) 알고리즘을 적용한 것이다. 예를 들어 RBM(Restricted Boltzmann Machine), Autoencoder와 VAE(Variational Autoencoder)와 같은 DNN 알고리즘은 고객의 선호도 예측에 적용하였다[12, 13, 14]. NCF[15]는 사용자와 상품의 잠재적 요인 사이의 비선형 관계를 학습하기 위해 심층 신경망을 적용한 NeuMF를 제안했다. LRML[16]는 메모리 모델과 Attention Mechanism을 사용하여 잠재적 요인 대신 잠재적 관계를 학습하여 모델의 유연성을 제고하였다. CMN[17]은 CF 기법에 메모리 신경망을 적용하여

상품 간의 상위 이웃 기반 관계를 학습시켰다. 이처럼 CF 기법은 놀라운 성과를 거두고 있지만, 여전히 데이터가 희박하고 상호운용성을 개선해야 하는 문제가 존재한다. 최근에는 이미지, 텍스트, 영상 등 정보들을 활용하여 CF 기법의 한계를 극복하기 위해 다양한 연구가 진행되고 있다[18, 19].

2-1 감성 분석 기반 개인화 추천 서비스

최근 전자상거래 플랫폼에서는 추천 알고리즘의 정확도를 향상하기 위해 수집되는 댓글 데이터는 감성 분석을 통해 정량적 데이터로 변환하여 연구에 사용하고 있다. 감성 분석은 기본적으로 댓글에서 고객의 주관적인 의견과 감정을 예측하고 분류하는 기법이다. 기존 감성 분석 연구는 단순히 고객의 감정을 긍정과 부정으로 분류하는 데이터 마이닝 기법에 국한되지만, 최근 기계학습을 활용한 감성 분석 연구는 고객의 선호와 감정을 정교하게 예측할 수 있다[20, 21].

기계학습 기반 감성 분석에서 대표적으로 많이 사용되고 기법은 Naive Bayes, Decision Tree, KNN(k-Nearest Neighbors), SVM(Support Vector Machine), Maximum Entropy 등이 있다. Pang et al.[22]은 영화 데이터를 사용하여 n-gram 기법으로 감성 분석을 수행하였으며, Naive Bayes, SVM, Maximum Entropy 기법으로 분석한 결과로 SVM 기법이 가장 우수한 정확도를 보였다. 이후 다양한 분야의 연구에서 SVM 기법을 활용한 감성 분석을 진행했다[23, 24, 25]. 그러나 Ferguson et al.[26]의 연구는 감성 분석에 기계학습 기법을 적용한 결과, Multi Naive Bayes 기법이 SVM 기법보다 우수한 성능을 나타냈으며, Pak and Paroubek[27]의 관련 연구에서도 유사한 연구 결과를 도출하였다.

기계학습 기법을 활용한 감성 분석에서 최근에는 심층 신경망을 활용하는 연구가 증가하는 추세이다. 심층 신경망을 활용한 감성 분석은 기존의 기법에서 입력 데이터 특성이 증가할 때 모델의 정확도가 하락하는 문제점을 보완할 수 있다. 또한, 기존 기계학습 기법에서는 값을 분류할 때 특징 집합을 따로 추출하여야 하지만, 심층 신경망 기법은 복잡한 특징을 자동으로 정교하게 추출할 수 있다는 장점이 있다. 예를 들어, JRL[28]은 댓글, 상품 이미지, 평점 등 여러 정보를 사용하여 사용자와 상품의 표현(Representations)을 학습하고, Top-N 추천 리스트에 순위를 부여했다. DeepCoNN[29]는 사용자의 댓글 집합과 상품의 댓글 집합에 각각 CNN 모델을 적용하여 사용자와 상품의 표현을 학습시키고, 다음으로 각각의 표현을 연결하여 최종적으로 평점을 예측했다. NARRE[30]는 DeepCoNN에서 소개된 유사한 구조를 사용하지만 어텐션 메커니즘을 사용하여 사용자와 상품 표현 학습에서 과거 댓글의 유용성(helpfulness)을 재측정한다. MPCN[31]은 포인터 네트워크(Pointer Network)를 사용하여 다양한 댓글의 역할을 구분하고 사용자와 상품의 표현을 학습할 때 대표적인 것을 선택한다.

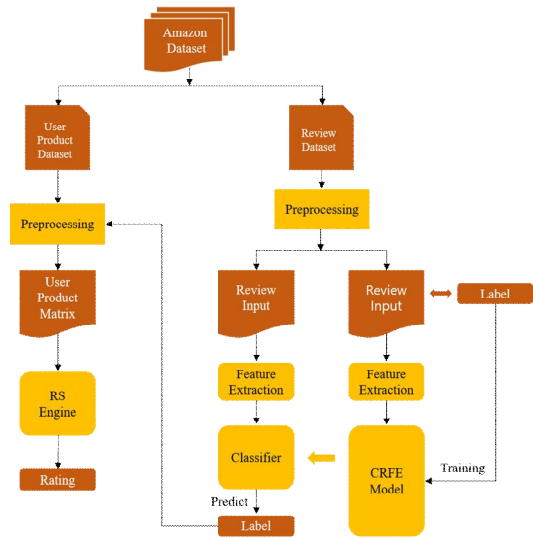


그림 1. 제안 방법 프레임워크
Fig. 1. Proposed Methodology Framework

III. 제안 방법

기존 개인화 추천 서비스 기법은 주로 고객이 직접 부여한 평점 데이터를 사용하여 모델을 설계하였다. 그러나 고객의 구매 동기 즉, 정성적인 선호도를 고려하지 않고, 오직 구매 여부나 평점을 바탕으로 상품을 추천하면 추천시스템의 정확도가 떨어지는 문제가 발생한다. 최근에는 고객의 정성적 선호도를 나타내는 댓글 데이터를 활용한 개인화 추천 서비스의 필요성이 증대하지만 이와 관련된 연구는 부족한 실정이다. 따라서 본 연구에서는 고객이 작성한 댓글 데이터에서 감성 특성을 정교하게 추출하여 새로운 고객 프로파일을 구축하고 이를 통해 고객에게 상품을 추천하는 방법을 그림 1과 같이 새롭게 제안한다.

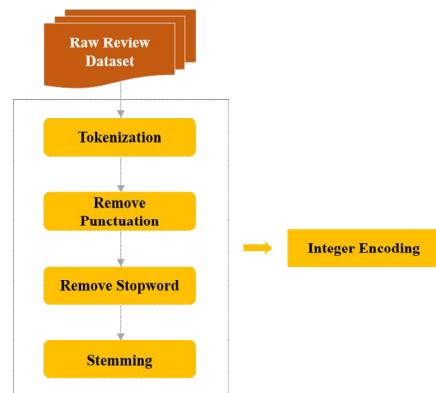


그림 2. 댓글 전처리 프로세스
Fig. 2. Review preprocessing process

3-1 Review Preprocessing

고객의 남긴 댓글은 여러 단어로 구성되지만 모든 단어가 고객의 감성 특성을 나타내는 것은 아니다. 예를 들어, ‘design’, ‘very’, ‘beautiful’ 등은 고객의 감성 특성을 나타내는 중요한 단어지만 ‘I’, ‘my’, ‘is’ 등은 크게 중요하지 않다. 댓글에서 감성의 특성을 반영할 수 있는 중요한 특성만 추출하여 분석하면 데이터의 크기를 줄이고 모델의 연산 속도 및 효율성 개선에도 도움을 줄 수 있다. 따라서 본 연구에서는 그림 2과 같은 과정을 거쳐 전처리를 수행했다.

온라인에서 수집된 댓글의 경우 HTML 태그나 다양한 부호 혹은 기호를 포함하고 있다. 이러한 노이즈 데이터(Noise Data)는 실제 분석에서 아무 의미를 나타내지 않기 때문에 사전에 제거하는 작업을 수행해야 한다. 본 연구에서는 노이즈 데이터의 등장 패턴을 파악하고 정규 표현식(Regular Expression)을 정의하여 값을 제거했다. 그리고 ‘I’, ‘my’, ‘me’, ‘over’, 조사, 접미사 등과 같은 단어들은 댓글 데이터에 나타나는 빈도수는 많으나 실제 분석을 수행할 때 의미가 없으며, 이러한 단어들을 불용어(Stopword)라고 부른다. 본 연구에서는 100개 이상의 불용어를 정의하고 있는 NLTK 라이브러리를 통해 댓글에 포함된 불용어를 제거했다. 또한, 댓글에서 영어 대문자와 소문자를 통합하거나 비슷한 의미의 단어를 하나의 단어로 정규화를 하면 단어의 개수를 줄일 수 있는 또 다른 전처리 방법이다. 따라서 댓글에 포함된 비슷한 의미를 나타내는 단어의 개수를 줄이기 위해 어간 추출(Stemming) 작업을 진행하였다.

위와 같은 전처리 단계를 마무리하면 알고리즘이 댓글을 숫자로 처리할 수 있도록 정수 인코딩 작업을 수행해야 한다. 즉, 댓글에 포함된 단어를 숫자 텐서로 변환하는 벡터화 작업을 수행해야 한다. 이를 위해 Keras에서 제공하는 Tokenizer 함수를 사용하여 댓글에 포함된 각 단어에 대하여 고유한 벡터값을 부여했다.

3-2 CNN-Based Review Semantics Extractor Model

최근에는 CNN, RNN 등과 같이 텍스트 데이터에서 문장 특성(Feature)을 추출하는 알고리즘이 여러 가지가 있다. 본 연구에서는 기존 RNN 기법보다 성능이 우수한 CNN 기법을 기반으로 CRSE(CNN-Based Review Semantics Extractor Model) 모델을 구축하고 댓글의 문장 특성을 추출하여 선호도 점수를 예측하여 새로운 고객 프로파일을 구축하고자 한다.

1) Word Embedding Layer

댓글과 같은 텍스트 데이터를 모델에 학습시키기 위해서는 단어의 의미를 벡터화시키는 작업을 거쳐야 한다. 기존의 벡터화 작업은 대부분 One-hot Encoding 기법을 사용했으나, 이 기법은 대부분 값이 0으로 나타내는 희소 표현(Sparse Representation) 문제가 발생한다. 본 연구에서는 텍스트를 밀집 벡터(Dense Vector)의 형태로 나타내는 워드 임베딩(Word Embedding) 기법을 사용하여 벡터화 작업을 진행했다. 워드 임

베딩 기법에는 Word2Vec, FastText 등 다양하게 있으며, 본 연구에서는 Keras에서 제공하는 함수를 사용한다.

예를 들어, $X_{u,i} = \{w_1, w_2, \dots, w_i, \dots, w_n\}$ 와 같은 댓글 데이터가 있다고 가정한다. $X_{u,i}$ 는 특정 고객 u 가 특정 상품 i 에 남긴 댓글을 의미하고, w_i 는 댓글에서 i 번째 단어를 의미하고 n 는 댓글에 포함된 단어의 개수를 의미한다. 각 구매 후기의 모든 단어에 임베딩 기법을 적용하여 단어를 밀집 벡터(e_1, e_2, \dots, e_n) 형태로 표현한다. 그리고 각 구매 후기는 식 1과 같이 임베딩 행렬로 표현된다.

$$E \in R^{n \times D} \tag{1}$$

식 1 중 D 는 임베딩 차원 크기를 의미하고 n 는 각 구매 후기에 포함된 단어의 개수를 의미하고 있다. 그리고 각 단어는 아래 식 2와 같이 정의할 수 있다.

$$e_k \in R^{D \times 1} \tag{2}$$

2) Convolution Layer

컨볼루션 레이어(Convolution layer)의 주요 기능은 입력되는 행렬에서 지역적 특성(Local Feature)을 추출하는 것이다. 이때 특성 추출에 사용되는 필터 커널(Filter Kernel)의 크기는 일반적으로 임베딩 차원의 크기가 같다. 입력 데이터의 행렬을 $E = [e_1, e_2, \dots, e_i, \dots, e_n]$ 라고 한다면, 컨볼루션 연산은 수식 3과 같이 정의된다.

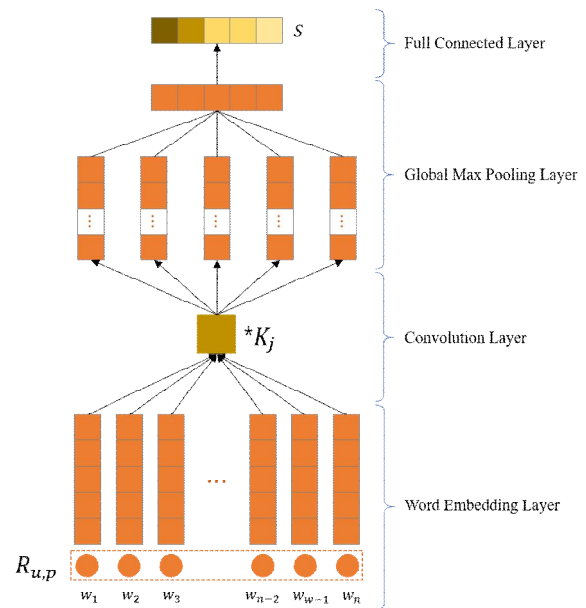


그림 3. CRSE 모델
Fig. 3. CRSE Model

$$c_j = f(W \cdot V + b_j) \quad (3)$$

식 3 중 $W \in R^{k \times m}$ 은 필터 커널을 나타내며, k 와 m 은 각각 컨볼루션 커널의 높이와 너비를 의미한다. b_j 는 편향(Bias)을 의미하고 f 는 활성화 함수 Relu(식 4)를 나타내고 있다.

$$f(x) = \max(0, x) \quad (4)$$

3) Pooling Layer

Pooling Layer의 주요 기능은 입력되는 컨볼루션에서 추출한 댓글 특성을 압축한 다음에 중요한 특성만 추출하는 것이다. 구매 후기를 분석할 때 전체 문장에서 일부분의 단어 혹은 문장만이 분석에 영향을 미친다. 따라서 본 연구에서는 식 5, 6와 같이 최대 풀링(Max Pooling) 적용하여 연산을 수행한다.

$$x = [x_1, x_2, \dots, x_i, \dots, x_{m-l+1}] \quad (5)$$

$$x_i = \max(c_i, c_{i+1}, \dots, c_{i+k-1}) \quad (6)$$

m 은 c_j 벡터의 차원을 의미하고, \max 는 최대값 함수를 나타낸다.

4) Fully Connected Layer

Fully Connected Layer의 주요 기능은 구매 후기에서 추출한 특성 행렬을 기반으로 결과를 분류하는 것이다. 분류 결과는 다음 식 7과 같이 정의할 수 있다.

$$Y = f(W \cdot X) + b \quad (7)$$

식 7 중, f 는 활성화 함수 Softmax를 나타내고, W 와 b 는 각각 행렬의 가중치와 편향을 의미한다. 이 계층은 입력 특성을 평점(1점부터 5점)으로 분류하도록 설정한다. 분류된 값이 0점과 가까우면 해당 상품에 대한 고객의 선호도는 낮고, 분류된 값이 5점과 가까우면 선호도는 높다고 판단할 수 있다.

3-3 Review Semantics Extractor

새롭게 구성된 프로파일의 성능을 측정하기 위해 최근 CF 연구에서 우수한 성능을 보여준 NCF(Neural Collaborative Filtering) 모델과 NeuMF(Neural Matrix Factorization) 모델을 사용하여 실험을 진행했다.

1) NCF(Neural Collaborative Filtering)

NCF 모델은 기존의 MF(Matrix Factorization) 기법의 한계를 극복하기 위해 제안되었다. MF 모델은 간단한 벡터의 내적을 통해 고객의 선호도를 추정하므로 성능 측면에서 한계가 있다.

이를 극복하기 위해 심층 신경망을 적용하여 고객과 상품의 잠재적 요인 사이의 비선형 관계를 파악하는 연구들이 제안되었으며, NCF는 사용자와 고객 사이의 잠재적 요인을 다층 신경망을 통해 학습시킨다. 이를 위해 고객의 잠재적 벡터(Latent Vector)와 상품의 잠재적 벡터를 아래 식 8, 9과 같이 각각 p_u 와 q_i 로 임베딩을 한다.

$$p_u = p^T v_u^U \quad (8)$$

$$q_i = Q^T v_i^I \quad (9)$$

식 8 중, $p \in R^{|U| \times d_u}$ 는 사용자 임베딩 행렬을 나타내고 $Q \in R^{|I| \times d_i}$ 는 상품의 임베딩 행렬을 나타내고 있다. $|U|$ 는 사용자의 개수를 의미하고 $|I|$ 는 상품의 개수를 의미하고 있다. 또한, v_u^U 는 사용자 u 를 식별할 수 있는 고유의 값을 나타내고, v_i^I 는 상품 i 를 식별하는 고유의 값을 나타낸다. 다음으로 사용자와 상품의 잠재적 표현을 식 10과 같이 연결하여 다층 퍼셉트론(MLP; Multi Layer Perceptron)으로 입력으로 전송한다.

$$z_1 = [p_u; q_u] \quad (10)$$

식 10중, z 는 사용자와 상품의 잠재적 표현을 연결한 값을 의미하고, $[\cdot; \cdot]$ 는 벡터를 연결하는 연산자를 나타내고 있다.

입력된 잠재적 표현은 MLP에서 고객과 상품의 잠재적 특성의 비선형 관계를 학습하게 된다. MLP 학습 과정은 아래 식 11과 같이 정의된다.

$$\varnothing_2(z_1) = a_2(W_2^T z_1 + b_2) \quad (11)$$

.....

$$\varnothing_L(z_{L-1}) = a_L(W_L^T z_{L-1} + b_L)$$

$$\hat{y}_{u,i} = \sigma(h^T \varnothing_L(z_{L-1}))$$

여기서 W_x , b_x 와 a_L 는 각각 가중치 행렬 (Weight Matrix), 편향과 활성화 함수 Relu를 의미하고 있다. $\hat{y}_{u,i}$ 는 사용자 u 가 상품 i 에 예측 평점을 나타내며, σ 는 최종 출력층의 활성화 함수를 의미하며 본 연구에서는 Relu로 설정했다.

2) NeuMF(Neural Collaborative Filtering)

NCF 모델에서는 고객과 상품 사이의 잠재적 요인을 다층 신경망을 통해 비선형 관계를 학습시켰다. NeuMF 기법은 기존의 비선형 모델(GMF)에 다층 선형 모델(MLP)을 결합하여 잠재적 요인의 선형과 비선형 관계를 모두 고려하여 학습을 진행한다.

$$\varnothing^{GMF} = p_u^G \odot q_i^G \quad (12)$$

$$\varnothing^{MLP} = a_L(W_L^T(a_{L-1}(\dots a_2(W_2^T \begin{bmatrix} p_u^M \\ q_i^M \end{bmatrix} + b_2)\dots)) + b_L)$$

$$\hat{y}_{ui} = \sigma(h^T \begin{bmatrix} \varnothing^{GMF} \\ \varnothing^{MLP} \end{bmatrix} a)$$

p_u^G 와 p_u^M 는 GMF와 MLP에 대한 사용자 임베딩을 나타내
고, 마찬가지로 p_i^G 와 p_i^M 는 상품 임베딩을 나타내고 있다. 앞
서 방법과 같이 MLP의 활성화 함수는 Relu를 사용한다.

IV. 실험 설계 및 결과

4-1 데이터 및 평가지표

본 연구에서 제안한 방법론을 검증하기 위해 상거래 사이트 Amazon에서 1996년 5월부터 2014년 7월까지 수집된 Kindle Store Category 데이터 세트를 사용했다. 데이터 세트에는 기본적인 고객 정보, 상품 정보, 평점 및 상품 후기 정보와 함께 다양한 추가 속성을 포함하고 있다. 데이터 세트 속성에 대한 자세한 설명은 표 1의 예시와 같다. 본 데이터 세트는 최소 5개 이상의 댓글을 작성한 고객과 최소 5개 댓글을 받은 상품만 포함하고 있다. 최종 데이터 세트에는 68,223명의 고객과 61,934개의 상품으로 구성되어 있고 댓글의 수는 982,619개이다. 그중에서 ‘helpful’ 값이 1 이상인 댓글을 297,199개를 선택하여 CRSE 모델의 학습에 사용하고 남은 669,019개 댓글 데이터는 모델 검증에 사용하였다.

평점 예측 알고리즘의 성능평가는 일반적으로 예측 평점과 실제 평점의 차이를 비교한다[32]. 본 연구에서는 개인화 추천 서비스 연구에서 기본적으로 많이 사용되고 있는 MAE, MSE를 평가지표로 사용한다. MAE는 다음 식 13과 같이 정의된다[33].

$$MAE = \frac{\sum |\hat{r}_{u,i} - r_{u,i}|}{|\Omega|} \tag{13}$$

여기서 Ω 는 실험용 데이터의 개수이며, $\hat{r}_{u,i}$ 는 예측 평점, $r_{u,i}$ 는 실제 평점이다. MAE는 오차의 크기에 대하여 영향이 없으며 모두 같은 가중치를 가지게 된다.

MSE는 예측 평점과 실제 평점의 차이를 제곱하여 평균으로 구한값이며 식2와 같다[34].

$$MSE = \frac{1}{\Omega} \sum (\hat{r}_{u,i} - r_{u,i})^2 \tag{14}$$

여기서 $\hat{r}_{u,i}$ 는 예측 평점이고 $r_{u,i}$ 는 실제 평점이며 Ω 는 실험용 데이터의 개수이다. MSE는 오차가 큰 평점에 대해 가중치를 높게 부여하는 평가지표이다.

표 1. 데이터 세트 속성 설명

Table 1. Description of Dataset Attributes

Column Name	Description
asin	ID of the product, like B000FA64PK
helpful	helpfulness rating of the review – example: 2/3
overall	rating of the product
reviewText	text of the review (heading)
reviewTime	time of the review (raw)
reviewerID	ID of the reviewer, like A3SPTOKDG7WBLN
reviewerName	name of the reviewer
summary	summary of the review (description)
unixReviewTime	unix timestamp

4-2 실험 설계

본 연구에서 제안한 추천 방법론의 우수성을 검증하기 위해 CRSE 모델을 만들고 이를 통해 새로운 고객 프로파일을 구축하였다. 또한, 기존의 고객 평점만을 고려하는 방법론과 본 연구에서 제안하는 방법론의 성능을 비교를 위해 최근 개인화 추천 서비스에 연구에서 주로 사용되고 있는 NCF, NeuMF 기법을 적용하였다[15].

댓글 데이터를 정량화하기 위한 CRSE 모델의 파라미터를 표 2와 같이 설정하고 실험을 진행했다. 댓글의 길이는 100, 200, 300, 500, 1000]로 설정하고, 패딩 작업을 통해 모든 댓글의 길이를 특정 길이로 동일하게 설정했다. NCF 알고리즘에서 고객과 상품의 임베딩 차원을 [8, 16, 32, 64]로 설정하고, 학습률은 0.01, 최적화 함수는 Adam를 사용하였으며, 배치 크기는 64로 설정했다. NeuMF 알고리즘에서 GMF의 임베딩 차원은 8, MLP의 임베딩 차원은 [8, 16, 32, 64]로 설정하였으며, MLP의 각 층의 뉴런을 [64, 32, 16, 8]로 설정하고 배치 크기는 64, 학습률은 0.01, 최적화 함수는 Adam를 적용했다.

표 2. CRSE 모델 파라미터

Table 2. CRSE model parameters

Parameters	Value
Input Sentence Length	[100, 200, 300, 500, 1000]
World Vector Dimension	120
Vocabulary Size	[10000, 20000, 30000, 40000, 50000]
Convolution Kernel Size	3 x 3
Hidden Layer Size	256
Dropout	0.2

4-3 실험 결과

댓글 데이터를 정량화하기 위한 CRSE 모델에서 단어 집합의 크기(Vocabulary Size)가 10,000과 20,000일 때 댓글의 길이는 200에서 예측 정확도가 가장 높고 500에서 가장 낮게 나타나고 있으며, 단어 집합의 크기가 30,000일 때 댓글의 길이는 300에서 예측 정확도가 가장 높고 200에서 가장 낮게 나타나고 있다. 또한, 단어 집합의 크기가 40,000일 때 댓글의 길이는 100에서 예측 정확도가 가장 높고 300에서 가장 낮게 나타나고 있으며, 50,000을 설정하면 댓글의 길이가 200일 때 정확도가 가장 높고, 500일 때 가장 낮게 나타나고 있다. 결론적으로 그림 4와 같이 단어 집합의 크기가 10,000을 설정하고 댓글의 길이를 200으로 설정했을 때 가장 높은 정확도를 보이므로, 이를 CRSE 모델의 최종 파라미터로 설정하고 고객의 새로운 프로파일을 구축했다.

본 연구에서 제안한 추천 방법론의 우수성을 검증하기 위해 최근 개인화 추천 서비스에 연구에서 주로 사용되고 있는 NCF, NeuMF 알고리즘을 적용하여 성능을 비교했다. 성능 비교를 위해 앞서 제시한 MAE와 MSE 평가지표를 사용하여 실제 평점과 예측 평점 간의 차이를 확인했다. Base Profile은 기존 고객의 프로파일에서 고객의 평점을 활용한 알고리즘의 성능을 나타내며, Proposal Profile은 댓글 데이터를 분석하여 새롭게 구축한 프로파일의 평점을 활용한 알고리즘의 성능을 나타낸다. NCF에서 MAE 평가지표를 사용할 때 잠재적 요인(Latent Factor)이 8에서는 0.079, 16은 0.089, 32는 0.087, 64는 0.075만큼 성능이 개선되었다. MSE 평가지표를 사용할 때 잠재적 요인(Latent Factor)이 8에서는 0.004만큼 성능이 감소하고 16은 0.006, 32는 0.049, 64는 0.052만큼 성능이 개선되었다.

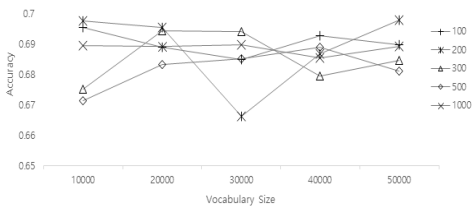


그림 4. 단어 빈도수 및 댓글 길이가 모델에 미치는 영향
Fig. 4. The impact of the word frequency and review length on the model.

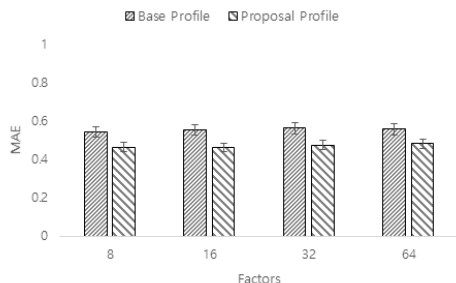


그림 5. NCF에서 MAE 평가지표 성능 비교
Fig. 5. Comparison of performance of MAE in NCF

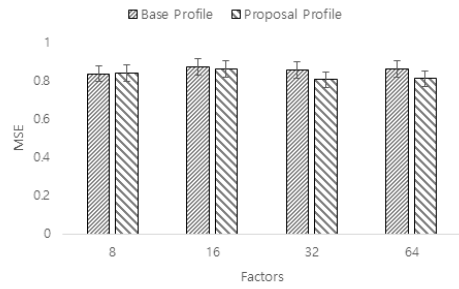


그림 6. NCF 모델의 MSE 성능 비교
Fig. 6. Comparison of performance of MSE in NCF

NeuMF에서 MAE 평가지표를 사용할 때 잠재적 요인(Latent Factor)이 8에서는 0.099, 16은 0.102, 32는 0.133, 64는 0.118만큼 성능이 개선되었다. MSE 평가지표를 사용할 때 잠재적 요인(Latent Factor)이 8에서는 성능에 큰 변화가 없으며, 16은 0.097, 32는 0.118, 64는 0.039만큼 성능이 개선되었다. 연구 결과 본 연구에서 제안하는 추천 방법론이 고객이 프로파일을 사용하는 방법보다 전체적으로 우수한 성능을 나타내고 있다.

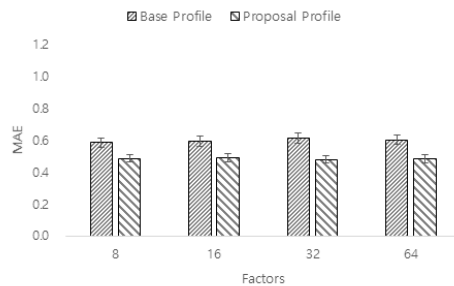


그림 7. NeuMF에서 MAE 평가지표 성능 비교
Fig. 7. Comparison of performance of MAE in NeuMF

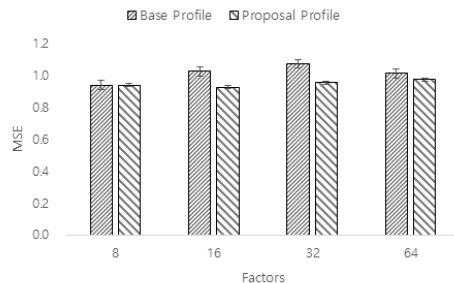


그림 8. NeuMF에서 MSE 평가지표 성능 비교
Fig. 8. Comparison of performance of MSE in NeuMF

또한, 본 연구에서 제안하는 방법론이 기존의 고객 평점만을 고려한 전통적인 방법론에 비교했을 때 통계적으로 얼마나 유의한 성능의 차이를 보이는지 검증하기 위해 대응표본 t-검정을 수행하였으며 결과는 표2와 같다.

표 3. 대응표본 t-검정 결과
Table 3. Paired t-test result

Paired t-test	Mean	SD	Mean difference	t	P
NCF Base Profile Proposal Profile	.675	.536	.002	442.921	.000
NeuMF Base Profile Proposal Profile	.712	.490	.001	510.133	.000

NCF의 경우 95% 신뢰수준 하에서 유의한 것($t=442.921$, $p\text{-value}=0.00$)으로 나타났으며, NeuMF도 95% 신뢰수준 하에서 유의한 것($t=510.133$, $p\text{-value}=0.00$)으로 나타났다. 이는 고객의 평점만을 고려하는 기존 추천 방법론과 본 연구에서 제안한 방법론 간의 성능에 유의한 차이가 있다고 볼 수 있다. 결과적으로 고객 평점만을 고려하는 전통적인 추천 방법론보다 고객의 남긴 댓글 데이터를 정교하게 분석하여 새로운 고객 프로파일을 구축하여 이를 추천시스템에 적용하는 방법론이 더 우수한 성능을 나타내고 있음을 알 수 있다.

V. 결론

본 연구에서는 고객의 정량적 선호도를 나타내는 명시적 평점 (Explicit Rating)과 구매 여부 혹은 방문 여부를 나타내는 암묵적 데이터 (Implicit Data)를 활용하여 상품을 추천하는 기존 개인화 추천 서비스의 한계를 극복하기 위해 새로운 추천 방법론을 제안했다. 실험을 위해 전자상거래에서 고객이 특정 상품을 구매할 때 중요하게 생각하는 댓글 데이터를 사용하였다. 또한, 단순히 통계 혹은 추론 기법을 기반으로 선호도를 분류하는 기존 연구의 한계를 극복하기 위해 최근 자연어 처리 분야에서 많이 사용되는 CNN 알고리즘을 활용하였다. 이를 통해 고객이 특정 상품에 작성한 댓글에서 문장 특성을 추출하여 이를 정교하게 수치화하여 새로운 고객 프로파일을 구축하였다. 실험 결과, 고객이 직접 부여한 정량적인 데이터만을 활용한 기존 방법론보다 본 연구에서 제안한 추천 방법론이 더 우수한 예측 정확도를 보여주었다.

본 연구는 다음과 같은 학술적 시사점을 갖는다. 전자상거래 플랫폼에서 댓글 데이터는 고객이 해당 상품의 구매를 결정할 때 중요한 역할을 하고 있다. 따라서 고객에게 유용한 댓글 데이터를 정량화하여 고객에게 맞춤형 추천 서비스 제공이 필요하다. 하지만, 기존 개인화 추천 서비스 연구에서는 주로 고객이 직접 부여한 평점 데이터를 사용하여 상품을 추천했다. 본 연구에서는 고객이 직접 작성한 댓글 데이터에서 감성 특성을 추출하고 이를 정량화하여 새로운 고객 프로파일을 구축했다. 특히, 기존 상품 후기 관련 연구는 단순히 긍정과 부정으로 분류하는 데이터 마이닝 기법에 국한되었다. 본 연구에서는 CNN 알고리즘을 활용하여 댓글 데이터를 정교하게 분석하여 새로운 고객 프로파일을 구축하고 고객에게 맞춤형 상품을 추천하는 새로운 추천 방법론을 제안했다. 본 연구는 개인화 추천 서

비스 연구에 댓글 데이터 분석에 심층 신경망을 적용함으로써 개인화 추천 서비스 연구의 확장에 기여하고 있다.

또한, 본 연구에서 제안한 방법론은 평점과 댓글 데이터를 같이 사용하는 다양한 분야에 적용할 수 있다. 최근 다양한 분야의 전자상거래에서 고객이 직접 상품에 대한 구매 후기를 작성할 수 있는 시스템을 구축하기 시작하면서 다양한 분석 기법을 활용한 자연어 처리 연구도 꾸준히 진행되고 있다. 본 연구에서는 전자상거래 분야에 대하여 새로운 고객 프로파일을 구축하고 상품을 추천하였지만, 평점과 댓글 데이터를 같이 사용하는 다양한 분야에 적용할 수 있다. 이를 통해 댓글 데이터의 감성 특성을 정교하게 예측하여 정량적인 값으로 변환하여 새로운 고객의 프로파일 구축이 가능하다.

마지막으로, 본 연구 결과는 전자상거래 관련 실무자들에게 개인화 추천 서비스 전략 수립에 대한 시사점을 제공하고 있다. 기존의 전자상거래 개인화 추천 서비스는 단순히 고객, 상품 정보와 평점 정보를 바탕으로 고객의 선호도를 예측하여 상품을 추천해주는 전략을 일관되게 사용했다. 이는 고객이 부여한 명시적인 평점이 상품에 대한 실제 선호도라는 믿음 때문이었다. 그러나 적지 않은 평점들이 고객이 실제 선호도를 나타내지 못하고 있다. 본 연구는 댓글 데이터를 사용하는 새로운 추천 방법론을 제안하여 통계적으로 검증함으로써 기존 전자상거래 개인화 추천 서비스 전략에 재고의 여지가 있음을 시사하고 있다.

본 연구의 한계점과 향후 연구 계획은 다음과 같다. 첫째, 본 연구에서 제안한 추천 방법론은 전자상거래의 식품 분야에만 적용된다. 따라서 본 연구에서 제안한 방법론을 일반화하기 위해서는 다양한 분야의 데이터 세트에 적용하는 연구를 추가로 수행해야 한다. 둘째, 본 연구의 데이터 전처리 단계에서는 이미 구축된 사전을 기반으로 불용어 제거, 토큰화 등 작업을 진행하였다. 하지만 도메인에 따라 적용되는 단어 사전이 다르므로 추후 연구에서는 해당 도메인의 단어 사전 구축하여 전처리 시도하는 방안을 고려해 볼 수 있다. 셋째, 본 연구에서 사용한 CNN 알고리즘의 성능을 개선할 수 있는 추가적인 연구가 필요하다. 최근 자연어 처리 분야에서 Attention Mechanism, Bert 등 알고리즘을 활용한 연구가 꾸준히 발표되고 있으며, 추후 연구에서는 최신 기법을 활용한 새로운 방법론을 제안하면 본 연구에서 제안한 방법론보다 정확도가 더욱 향상될 것으로 기대하고 있다.

참고문헌

[1] G. Linden, B. Smith, and J. York, "Amazon. com recommendations: Item-to-item collaborative filtering," IEEE Internet computing, no. 1, pp. 76-80, 2003.
 [2] J. Bennett, and S. Lanning, "The netflix prize." p. 35.
 [3] A. S. Das, M. Datar, A. Garg, and S. Rajaram, "Google news personalization: scalable online collaborative

- filtering." pp. 271-280.
- [4] F. Ricci, L. Rokach, and B. Shapira, "Introduction to recommender systems handbook," *Recommender systems handbook*, pp. 1-35: Springer, 2011.
- [5] Z. Zhang, D. Zhang, and J. Lai, "urCF: user review enhanced collaborative filtering," 2014.
- [6] Kim, K. N., Choi, D. S., Lee, K. M., & Lee, M. J., "Analysis on interesting element of mobile application using user reviews." *Journal of Digital Contents Society*, vol. 13, no. 3, pp. 431-438, 2012.
- [7] S. Dhanasobhon, P.-Y. Chen, M. Smith, and P.-y. Chen, "An analysis of the differential impact of reviews and reviewers at Amazon. com," *ICIS 2007 Proceedings*, pp. 94, 2007.
- [8] J. Y. Choeh, 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, pp. 379-385, 2013.
- [9] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry," *Communications of the ACM*, vol. 35, no. 12, pp. 61-71, 1992.
- [10] X. Su, and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Advances in artificial intelligence*, vol. 2009, 2009.
- [11] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Transactions on Information Systems (TOIS)*, vol. 22, no. 1, pp. 5-53, 2004.
- [12] R. Salakhutdinov, A. Mnih, and G. Hinton, "Restricted Boltzmann machines for collaborative filtering." pp. 791-798.
- [13] S. Sedhain, A. K. Menon, S. Sanner, and L. Xie, "Autorec: Autoencoders meet collaborative filtering." pp. 111-112.
- [14] X. Li, and J. She, "Collaborative variational autoencoder for recommender systems." pp. 305-314.
- [15] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering." pp. 173-182.
- [16] Y. Tay, L. Anh Tuan, and S. C. Hui, "Latent relational metric learning via memory-based attention for collaborative ranking." pp. 729-739.
- [17] T. Ebesu, B. Shen, and Y. Fang, "Collaborative memory network for recommendation systems." pp. 515-524.
- [18] Z. Ji, H. Pi, W. Wei, B. Xiong, M. Woźniak, and R. Damasevicius, "Recommendation based on review texts and social communities: A hybrid model," *IEEE Access*, vol. 7, pp. 40416-40427, 2019.
- [19] X. Wang, X. Yang, L. Guo, Y. Han, F. Liu, and B. Gao, "Exploiting social review-enhanced convolutional matrix factorization for social recommendation," *IEEE Access*, vol. 7, pp. 82826-82837, 2019.
- [20] Q. Diao, M. Qiu, C.-Y. Wu, A. J. Smola, J. Jiang, and C. Wang, "Jointly modeling aspects, ratings and sentiments for movie recommendation (JMARS)." pp. 193-202.
- [21] X. Guan, Z. Cheng, X. He, Y. Zhang, Z. Zhu, Q. Peng, and T.-S. Chua, "Attentive aspect modeling for review-aware recommendation," *ACM Transactions on Information Systems (TOIS)*, vol. 37, no. 3, pp. 1-27, 2019.
- [22] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," *arXiv preprint cs/0205070*, 2002.
- [23] A. C. König, and E. Brill, "Reducing the human overhead in text categorization." pp. 598-603.
- [24] A. Kennedy, and D. Inkpen, "Sentiment classification of movie reviews using contextual valence shifters," *Computational intelligence*, vol. 22, no. 2, pp. 110-125, 2006.
- [25] M. Annett, and G. Kondrak, "A comparison of sentiment analysis techniques: Polarizing movie blogs." pp. 25-35.
- [26] P. Ferguson, N. O'Hare, M. Davy, A. Bermingham, P. Sheridan, C. Gurrin, and A. F. Smeaton, "Exploring the use of paragraph-level annotations for sentiment analysis of financial blogs," 2009.
- [27] A. Pak, and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining." pp. 1320-1326.
- [28] Y. Zhang, Q. Ai, X. Chen, and W. B. Croft, "Joint representation learning for top-n recommendation with heterogeneous information sources." pp. 1449-1458.
- [29] L. Zheng, V. Noroozi, and P. S. Yu, "Joint deep modeling of users and items using reviews for recommendation." pp. 425-434.
- [30] C. Chen, M. Zhang, Y. Liu, and S. Ma, "Neural attentional rating regression with review-level explanations." pp. 1583-1592.
- [31] Y. Tay, A. T. Luu, and S. C. Hui, "Multi-pointer co-attention networks for recommendation." pp. 2309-2318.
- [32] G. Shani, and A. Gunawardana, "Evaluating recommendation systems," *Recommender systems handbook*, pp. 257-297: Springer, 2011.
- [33] K. Goldberg, T. Roeder, D. Gupta, and C. Perkins, "Eigentaste: A constant time collaborative filtering algorithm," *information retrieval*, vol. 4, no. 2, pp. 133-151, 2001.
- [34] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egyptian Informatics Journal*, vol. 16, no. 3, pp. 261-273, 2015.



엄금철(Jinzhe Yan)

2020년 : 성균관대학교 일반대학원 (경영학 박사)

2020년~현 재: 가천대학교 경영대학 글로벌경영학과 교수

※관심분야 : 마케팅 모델링, 플랫폼 비즈니스, 빅데이터 마케팅, 머신러닝, 중국경영



이가은(Jiaen Li)

2017년 : 경희대학교 일반대학원 (경영학 석사)

2020년 : 경희대학교 일반대학원 (박사과정 수료)

2020년~현 재: 경희대학교 무역학과 객원교수

※관심분야 : 글로벌 가치사슬, 비관세장벽, 수출부가가치, 무역빅데이터, 텍스트 분석 등



이청용(Qinglong Li)

2019년 : 경희대학교 경영학과 (경영학 학사)

2020년~현 재: 경희대학교 일반대학원 소셜네트워크과학과 석사과정

※관심분야 : 추천시스템, 데이터마이닝, 텍스트마이닝, 딥러닝 등