

# GAN기반의 하이브리드 협업필터링 추천기 연구

송 희 석\*

## A Study for GAN-based Hybrid Collaborative Filtering Recommender

Hee Seok Song\*

### Abstract

As deep learning technology in natural language and visual processing has rapidly developed, collaborative filtering-based recommendation systems using deep learning technology are being actively introduced in the recommendation field. In this study, OCF-GAN, a hybrid collaborative filtering model using GAN, was proposed to solve the one-class and cold-start problems, and its usefulness was verified through performance evaluation. OCF-GAN based on conditional GAN consists of a generator that generates a pattern similar to the actual user preference pattern and a discriminator that tries to distinguish the actual preference pattern from the generated preference pattern. When the training is completed, user preference vectors are generated based on the actual distribution of preferred items. In addition, the cold-start problem was solved by using a hybrid collaborative filtering recommendation method that additionally utilizes user and item profiles. As a result of the performance evaluation, it was found that the performance of the OCF-GAN with additional information was superior in all indicators of the Top 5 and Top 20 recommendations compared to the existing GAN-based recommender. This phenomenon was more clearly revealed in experiments with cold-start users and items.

Keywords : Deep Learning, GAN, Recommendation, Hybrid Collaborative Filtering, One-Class, Cold-Start

Received : 2022. 11. 23.    Revised : 2022. 12. 20.    Final Acceptance : 2022. 12. 23.

※ This research was supported by the National Research Foundation of Korea(NRF) grant (No.2021R1F1A1045815).

\* Professor, Hannam University, Department of MIS, 70 Hannamro Daeduckgu Daejeon 34430, Korea, Tel: +82-42-629-8344, e-mail : hssong@hnu.kr

## 1. 서론

4차 산업혁명 기술의 보급과 더불어 개인화와 맞춤화에 대한 요구 증대로 추천시스템의 산업체 활용이 급격히 확대되고 있다. 추천시스템은 활용되는 데이터와 추천방식에 따라 내용기반 추천, 협업필터링기반 추천 및 이들을 결합한 하이브리드 추천으로 구분되며 이 중 협업필터링기반 추천은 도메인에 관계없이 적용이 가능한 범용성과 우수한 추천성능으로 널리 산업체에서 활용되고 있다. 최근 자연어와 시각처리 분야의 딥러닝 기술이 급속히 발전하면서 추천 분야에서도 딥러닝 기술을 활용한 협업필터링기반 추천시스템이 활발히 소개되고 있다. Bokde et al.[2015]와 Koren et al.[2009]은 행렬분해(Matrix Factorization)를 심층신경망으로 표현한 행렬분해기반 신경망모델(MFNN)을 제안하였고, He et al.[2017]과 He and Chua[2017]은 뉴럴협업필터링모델(NCF)과 일반화행렬분해와 결합된 뉴럴협업필터링모델(NCF+GMF)을 제안하여 개선된 성능을 보여준 바 있다.

딥러닝기반 협업필터링 추천시스템에 대한 다양한 연구에도 불구하고 협업필터링 추천모델은 원클래스와 콜드스타트라는 두 가지 문제가 존재한다. 원클래스 문제는 OCCF (One-Class Collaborative Filtering) 문제라고도 불리며(Pan et al., 2008) 추천시스템 학습에 필요한 부정사례 데이터 수집의 어려움으로 인해 발생하는 문제이다. 최근 협업필터링은 사용자가 아이템에 대해 부여하는 평점과 같은 명시적인 피드백보다는 좋아요, 북마크, 클릭 등 사용자의 선호도를 추론하는 데 활용할 수 있는 묵시적 피드백 정보를 활용하는데 이는 묵시적 피드백이 명시적 피드백보다 더 풍부하고 편향이 적기 때문에 연구 커뮤니티에서 많은 관심을 받고 있기 때문이다. 딥러닝기반 협업필터링 추천모델을 구축하기 위해서는 사용자가 선호하는 아이템에 대한 정보인 긍정사례와 사용자가 선호하지 않는 아이템에 대한 정보인 부정사례가 함께 수집되어 훈련 모델에 입력되어야 하는데 묵시적 피드백을 이용하는 경우 긍정사례로만 훈련이 이루어지기 때문에 모든 결측값에 대한 예측이 긍정값으로 주어지는(Trivial solution) 문제가 발생한다. 원클래스 문제를 해결하기 위한 방법으로 그간 소개된 방법은 가중손실법과 데이터밸런싱 방법 등이 주류이다(Pan et al., 2008).

그러나 이들 방법 또한 사용자 선호가 확인되지 않은 경우를 미지 또는 싫어함 중 한쪽으로 처리하기 때문에 높은 추천성능을 보장하기 어렵다.

협업필터링 추천모델의 두 번째 문제는 콜드스타트 문제이다. 이는 사용자 아이디와 아이템 아이디로 구성된 상호작용 행렬만을 이용하여 학습이 이루어지기 때문에 상호작용이 없는 신규사용자나 신규아이템의 경우 추천이 이루어질 수 없는데서 비롯된 문제이다. Strub et al.[2016]은 협업필터링의 콜드스타트 문제를 극복하기 위한 아이디어로 하이브리드 협업필터링 모델(HCF: Hybrid Collaborative Filtering)을 제안한 바 있다. 일반적으로 HCF는 사용자와 아이템에 대한 부가정보로부터 신규사용자 또는 신규아이템에 대한 잠재요인을 추출하여 상호작용행렬에 반영함으로써 신규 아이템이나 신규 사용자의 상호작용 기록을 증강하기 때문에 콜드스타트 문제 해결의 유력한 대안으로 부상하고 있다. Strub et al.[2016]에 의하면 HCF는 순수 협업필터링 모델에 비해 높은 추천성능을 보여주고 있을 뿐 아니라 내용기반 추천시스템과 달리 부가정보를 이용하여 상호작용행렬의 희박도를 개선하는 방식을 사용하기 때문에 부가정보의 이용이 가능한 다양한 응용분야에 적용이 가능한 프레임워크가 될 것으로 전망하고 있다.

현재 딥러닝기반 협업필터링 추천모델의 두 가지 근원적 문제인 원클래스와 콜드스타트 문제는 별개로 연구가 진행되고 있으며 이 두 문제를 동시에 해결할 수 있는 방법은 아직 발표되지 않고 있다. 원클래스와 콜드스타트 문제 해결을 위해 본 연구에서는 GAN을 이용한 하이브리드 협업필터링 모델인 OCF-GAN (One-class and Cold-Start problem Free GAN)을 제안한다. Conditional GAN에 기반한 OCF-GAN은 실제 사용자 선호패턴과 유사한 패턴을 생성하려는 생성기와 실제 선호패턴과 생성된 선호패턴을 구분하려는 판별기로 구성되며, 생성기는 판별기와 경쟁을 통해 실제 사용자가 선호하는 아이템 분포를 학습하고 훈련이 종료되면 실제 선호아이템 분포에 기반하여 사용자가 좋아할 만한 아이템을 생성하게 되는 원리로 긍정사례만 있어도 학습이 가능하기 때문에 원클래스 문제의 해결이 가능해진다. 또한 사용자와 아이템에 대한 부가정보를 추가적으로 활용하는 하이브리드 협업필터링 추천방식을 사용함으로써 콜드스타트 문제를 해결

할 수 있다.

## 2. 기존연구

### 2.1 원클래스 문제를 다룬 기존연구

원클래스 문제를 OCCF(One-Class Collaborative Filtering)로 처음 명명한 연구는 Pan et al.[2008]이다. 이 연구에서는 OCCF를 해결하기 위한 가중치기반 저순위근사법과 부정사례 샘플링 등 두 가지 프레임워크를 제안하였고 실험결과 제안한 프레임워크가 순수 협업필터링에 비해 우수한 성능을 나타냄을 실증한 바 있다. 가중치기반 저순위근사법은 모든 결측값을 부정사례로 간주한 후 목적함수(손실함수)에서 긍정사례의 오류항과 부정사례의 오류항에 서로 다른 가중치를 부여하여 학습하는 방법이다. 또한 부정사례 샘플링법은 상호작용행렬에서 결측값을 샘플링하여 부정사례를 추출하고 모든 긍정사례를 포함하는 부분 상호작용행렬을 여러 번 걸쳐 만들고 여러 개의 부분 상호작용행렬을 통합하여 전체 상호작용행렬을 구성하는 일종의 앙상블 방법이다. 이들 방법은 사용자 선호가 확인되지 않은 경우를 미지 또는 싫어함 중 한쪽으로 처리하기 때문에 높은 추천성능을 보장하기 어렵다. Li et al.[2010]은 검색이력이나 구매이력 및 브라우징 이력 등의 커뮤니티 기반 상호작용 정보와 같은 풍부한 사용자 정보를 활용하여 OCCF문제의 정확도를 개선하고 있다. 이 연구에서는 OCCF문제를 위해 다양한 원천정보로부터 계산된 점수를 선형결합하는 방법과 사용자정보를 협업필터링 내 임베딩하여 사용하는 두 가지 방법을 제안하고 있으며 대규모 전자상거래에 적용한 실험에서 정확도가 개선됨을 보여주었다.

Pan et al.[2008]과 Li et al.[2010]의 연구 이후 원클래스 문제는 상호작용행렬의 희박도 문제를 개선하기 위한 연구로 확대되어 연구자의 관심을 받아왔다. Hwang et al.[2016]은 방대한 양의 평가되지 않은 항목에 대한 사용자의 사전 사용 선호도 개념을 협업필터링에 적용하여 Top N 추천의 정확성과 실행 시간을 개선하는 방법을 제안하였다. 이 연구에서는 아직 평가되지 않았지만 사용자로부터 매우 낮은 평가를 받을 가능성이 있는 항목을 효과적으로 식별하여 부정사례로 간주하여 추천모델을 구성한다. 단순하면서도

새로운 제로 인젝션(Zero injection)방법은 평점 행렬의 희박성 문제를 해결할 뿐만 아니라 관심 없는 항목이 상위 N 개 항목으로 추천되는 것을 방지하여 정확도를 향상시킨다. 제안된 아이디어는 MovieLens 데이터셋과 MyMediaLite에 적용되어 기존의 협업필터링 모델 대비 정확도를 2~5배 향상시킨 것을 확인하였다. OCCF 관련 기존연구들이 평점예측 또는 순위예측에 중점을 둔 반면 Li et al.[2017]은 추천아이템의 평점과 순위를 동시에 최적화하는 추천 접근 방식을 시도하였다. 이 연구에서는 최신 CLiMF(Collaborative Less-is-More Filtering) 접근 방식과 PMF(Probabilistic Matrix Factorization) 접근 방식을 기반으로 한 새로운 통합 OCCF 접근 방식(UOCCF)을 제안하고 있다. UOCCF에서는 사용자 및 항목에 대한 공통의 잠재특징들을 공유함으로써 순위중심 관점과 평점중심 관점 모두의 장점을 취할 수 있다.

원클래스 문제와 상호작용행렬의 희박도 문제해결을 위해 GAN의 적용 가능성을 최초로 보여준 연구는 Wang et al.[2017]의 IRGAN 모델이다. 이후 2018년 SIGIR의 GAN기반 정보검색 모델(GAN-based information retrieval models)에 관한 세미나(Zhang, 2018)에서 학자들은 GAN이 추천시스템의 고질적 문제인 데이터잡음과 희박성문제를 해소하는데 효과적인 수단이 될 것임에 동의하면서 GAN기반 추천시스템 연구가 더욱 주목받게 되었다. IRGAN은 더 많은 정보를 포함하는 부정사례를 샘플링하여 추천성능을 개선하고자 한 연구로 요약된다. IRGAN의 생성기는 주어진 사용자에 대해 적합한 아이템을 예측하고, IRGAN의 판별기는 각각의 사용자-아이템 쌍에 대해 관련도(relevancy)를 예측한 후 예측결과를 생성기에게 주어 보다 더 적합한 아이템을 생성하도록 유도한다. IRGAN의 생성기는 바닐라 GAN에서 경사하강법을 사용하는 것과 달리 아이템에 대한 샘플링이 이산적으로 이루어지기 때문에 정책경사(Policy gradient)에 의해 계산된 확률에 기반하여 아이템 아이디어를 생성한다는 점이 특징이다. Chae et al.[2018]은 IRGAN이 적대적 훈련의 장점을 온전히 활용하지 못하고 있다고 주장하며 CFGAN모델을 제안하였다. 즉 IRGAN의 생성기는 연속형 값이 아닌 이산적인 개별 아이템 인덱스를 생성하게 되는데 훈련이

반복되면서 생성기는 실제 사용자가 선호하는 아이템과 동일한 아이템을 생성하게 되고 이후 훈련데이터에는 동일한 아이템 인덱스에 대해 모순되는 레이블이 추가되어 판별기를 혼란에 빠뜨리게 된다. 결과적으로 생성기와 판별기는 서로의 능력을 향상시키지 못하며 경쟁이 서로 도움이 되지 않는 상황으로 치닫게 된다고 분석하였다. CFGAN에서는 생성기가 단일 아이템 인덱스를 생성하는 대신 실제값을 가지는 원소로 구성된 구매 벡터를 생성함으로써 훈련데이터에 모순된 레이블이 혼재하는 것을 방지하고 있다. Zhao et al. [2018]은 시퀀스 추천을 목표로 행렬분해와 순환신경망(Recurrent Neural Network) 및 GAN을 결합한 PLASTIC 모델을 제안하였다. PLASTIC의 적대적 훈련과정에서 생성기는 사용자와 시간을 입력으로 받아 사용자의 추천 목록을 직접 예측하고 판별기는 삼네트워크(Siamese network)를 기반으로 장기 및 단기 순위모델을 통합하여 실제 샘플과 생성된 샘플을 정확하게 구분하게 된다. Liu et al. [2020]은 IRGAN을 확장한 협업필터링 GAN인 CoFiGAN을 제안하였다. CoFiGAN의 생성기는 Pair-wise 손실 함수를 사용하여 실제와 유사한 항목을 생성하고 판별기는 생성된 항목을 실제 항목과 구별하는 역할을 수행한다.

기존 GAN기반 추천 연구를 요약하면 먼저 생성기의 출력 측면에서 CFGAN과 같이 구매벡터를 생성하는 시도가 있는가 하면 IRGAN 처럼 아이템을 생성하는 연구도 있으며 AugCF와 같이 사용자-아이템 튜플을 생성하는 연구도 있다. 추천시스템에 GAN을 적용하는 목적 측면에서 기존 연구들을 살펴보면 학습데이터의 노이즈 개선, 불균형클래스의 샘플링, 부정사례 생성, 희박도 개선 등 다양한 목적이 존재함을 알 수 있다. 기존문헌 연구에서 언급한 것처럼 추천시스템에 GAN을 적용하는 시도가 최근 늘어나고 있지만 대부분 협업필터링 시스템에 GAN을 적용하는 시도이며 원클래스와 콜드스타트 문제를 동시에 해결하기 위한 하이브리드 협업필터링에 GAN을 적용하는 시도는 찾아보기 어렵다.

## 2.2 하이브리드 협업필터링 연구

협업필터링 추천의 한계점은 사용자나 아이템에 대

한 평가가 없는 경우 사용자에게 아이템을 추천하기 어려운 콜드스타트 문제이다. 이 문제를 극복하기 위한 한 가지 아이디어는 협업필터링 추천에 사용자 또는 아이템에 대한 정보를 추가적으로 활용하여 상호작용 행렬(평점 행렬)의 희박도를 개선하는 것이다. Adams and Murray [2010]과 Porteous and Asuncion [2010]은 상호작용 행렬에 사용자와 아이템에 대한 부가정보를 통합하기 위한 베이지안 행렬분해(Matrix factorization) 방법을 제안하였으며 이들 알고리즘이 순수 협업필터링 추천에 비해 높은 추천성능을 나타내는 것을 확인한 바 있다 [Lee et al., 2012]. 이처럼 콜드스타트라는 협업필터링 추천의 한계를 극복하기 위한 초기 연구에서는 행렬분해 기법을 주로 사용하고 있는데 행렬분해를 이용한 추천은 희박도가 높은 상호작용 행렬에서 잠재요인을 추출하여 저 차원의 잠재요인 행렬을 생성하고 이들의 곱으로 관측되지 않은 상호작용 행렬의 평점을 예측하게 된다. 그러나 이 방법은 선형모델을 통한 예측이어서 미묘한 잠재요인을 포착하는데 한계가 있으며 저 차원의 행렬의 곱으로 원본 상호작용 행렬을 완벽히 재현하기는 어렵다. 따라서 최근에는 심층신경망과 같은 비선형모델을 활용하여 협업필터링에 사용자 또는 아이템에 대한 부가정보를 추가하는 모델이 제안되고 있는데 Strub et al. [2016]의 연구가 대표적이다.

Strub et al. [2016]은 오토인코더(Autoencoder)를 기반으로 한 하이브리드 협업필터링 접근 방식을 제안하였다. 이 연구에서는 희박한 상호작용 데이터와 부가정보를 통합하여 비선형적 행렬분해를 구현하는 협업필터링 신경망아키텍처를 제안하였고 MovieLens와 Douban 데이터 세트에 적용하여 추천성능이 개선되는 것을 보여주었다. 제안된 시스템은 사용자와 아이템의 비선형 표현을 학습하고 사용자와 아이템에 대한 부가정보를 통합하여 콜드스타트 문제를 완화하고 있다. Won et al. [2019]는 온라인 쇼핑물 고객의 검색 키워드와 구매 이력 데이터를 활용한 Doc2Vec기반의 새로운 하이브리드 협업필터링 모델을 제안하였다. 제안된 모델을 실제 온라인 쇼핑물 데이터를 사용하여 실험한 결과 검색 키워드 데이터가 고객의 선호도를 효과적으로 표현하고 기존 협업필터링의 개선에 기여할 수 있음을 주장하였다. 그러나 이 방법은 부가적인 정보를 활용하여 상호작용행렬의 희박도를 개선하기 보다는

협업필터링과 내용기반 필터링을 결합하는 방식이어서 순수 하이브리드 협업필터링 방식으로 보기는 어렵고 도메인마다 다른 방식의 구현이 필요하기 때문에 일반화하기 어려운 솔루션이라 할 수 있다.

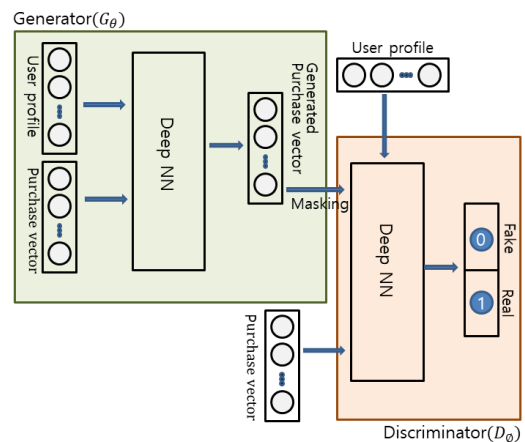
### 3. OCF-GAN(One-class and Cold-start problem Free GAN) 추천기 설계

본 연구에서는 원클래스와 콜드스타트 문제를 동시에 해결할 수 있는 추천기인 OCF-GAN을 제안한다. 심층신경망에 기반한 협업필터링이 높은 추천성능을 보이면서 추천시스템 연구 커뮤니티에서 활발히 연구되고 있지만 상호작용 데이터가 없는 신규사용자 또는 신규아이템에 대해서는 추천이 이루어질 수 없다. 또한 묵시적 피드백정보를 이용하여 추천기를 구현할 경우 필연적으로 원클래스 문제에 직면하게 된다. 두 가지 문제를 동시에 극복하기 위한 OCF-GAN의 아이디어는 사용자와 아이템 프로파일과 같은 부가정보를 활용하여 신규사용자 및 아이템에 대한 상호작용 행렬을 보강할 뿐 아니라 적대적생성 훈련을 기반으로 부정사례를 포함한 상호작용 행렬을 보강하는 방법으로 추천성능을 향상하는 것이다. 그러나 이 아이디어를 구현하는 데는 다음과 같은 몇 가지 주요 이슈가 해결되어야 한다.

먼저 GAN모델의 생성기가 무엇을 생성할 것인가에 관한 이슈이다. 딥러닝 기반의 협업필터링 모델은 특정사용자를 위한 아이템을 추천하는 것이 일반적이다. 따라서 GAN모델의 생성기 또한 입력된 사용자를 위한 단일아이템 인덱스를 생성하도록 설계할 수 있다. 실제 IRGAN[Wang et al., 2017]이나 CoFiGAN[Liu et al., 2020], DASO[Fan et al., 2019] 등 많은 연구에서 생성기가 아이템을 샘플링하는 식으로 설계하고 있다. 그러나 생성기가 단일아이템 인덱스를 생성하게 되면 이산값으로 인해 경사계산을 위한 미분이 어렵다는 단점이 존재하며, GAN 훈련시 훈련데이터에 모순된 레이블이 혼재하여 추천성능을 저하시킬 수 있다. 이에 본 연구에서는 CFGAN[Chae et al., 2018]에서 시도한 것처럼 생성기가 특정 사용자에게 추천할 아이템 인덱스 대신 아이템별 구매벡터를 생성함으로써 미분불능 문제를 해결하고 성능향상을 도모하고자 한다. 두 번째 이슈는 하이브리드 협업필터링과 GAN모델을 통합모델로 구성하는 방법에 관한 것이다. 즉 하이브리드 협업필

터링과 GAN모델이 단계별로 학습 또는 수행되는 것이 아니라 하나의 모델로 통합될 때 최적화가 가능하다. 따라서 두 가지 모델을 결합한 통합모델에 대한 구조설계를 어떻게 할 수 있는지 또한 주요 이슈이다. OCF-GAN에서는 하이브리드 협업필터링 구현을 위해 사용되는 부가정보를 임베딩 층을 통해 벡터화하여 GAN모델의 생성기에 입력하는 방식을 사용하여 두 가지 모델을 통합한 모델을 제시하기로 한다.

세 번째 이슈는 일반화 이슈이다. 추천시스템이 도메인마다 다양하고 이질적인 사용자와 아이템에 대한 부가정보를 사용하게 되면 상호작용정보만을 이용하는 순수 협업필터링 추천시스템과 달리 도메인마다 새로운 모델 구조가 필요하며 부가정보의 가공에 추가적인 노력과 비용이 소요되기 마련이다. 따라서 모든 도메인에 공통 적용이 가능한 일반화된 솔루션의 제시가 요구된다. 프레임워크의 일반화를 위한 아이디어는 사용자와 아이템에 대한 부가정보를 추천을 위해 직접 사용하기보다는 상호작용 행렬을 보강하는 용도로 활용하는 것이다. 상호작용 행렬이 보강되어 희박도(Sparsity)가 낮아지면 도메인에 독립적인 협업필터링모델처럼 별도의 모델구조 변경 없이 모든 도메인에 동일한 방법으로 추천시스템을 적용할 수 있게 된다. 상기 이슈를 해결할 수 있는 OCF-GAN의 구조는 <Figure 1>과 같다. OCF-GAN은 사용자별 개인화된 파라미터(parameters)에 대한 학습을 고려하는 conditional GAN(Mirza and Osindero, 2014)에 기반하여 구성된다.



<Figure 1> User-based OCF-GAN Architecture(Changing User to Item in Item-based OCF-GAN)

(Figure 1)에서 OCF-GAN에는 생성기( $G_\theta$ )와 판별기( $D_\phi$ ) 두 가지 주요 구성요소가 있다. 생성기는 사용자프로파일이 주어지면 해당사용자의 상호작용 데이터를 모방하여 가짜의 구매벡터를 생성하는 반면 판별기는 상호작용 행렬에서 샘플링 한 실제 구매벡터와 생성기에 의해 생성된 가짜 구매벡터를 구분하는 역할을 수행한다. 훈련과정을 보다 세부적으로 설계하면 다음과 같다. 여기서는 사용자기반 OCF-GAN을 기준으로 설명하며 아이템기반 OCF-GAN에서는 사용자를 아이템으로 대체하면 된다. 먼저 판별기의 목적은 실제 구매벡터와 생성된 구매벡터를 정확하게 구별할 수 있는 로그우도(log-likelihood)를 최대화하는 판별기의 파라미터  $\phi$ 를 발견하는 것이며 판별기 목적함수를 제시하면 다음과 같다.

$$\phi^* = \arg \max_{\phi} \sum_{u=1}^N (E_{r_u \sim p_{true}} [\log D_\phi(r_u, s_u | u)] + E_{\hat{r}_u \sim p_\theta} [\log(1 - D_\phi(\hat{r}_u \otimes e_u, s_u | u))])$$

$e_u$ :  $u$ 가 아이템  $i$ 를 구매했는지 여부를 나타내는 indicator vector

식에서 사용자  $u$ 에 대해 실제 구매벡터 분포인  $p_{true}$ 에서 가져온 구매벡터( $r_u$ )와 함께 사용자프로파일( $s_u$ )이 입력되면 판별기의 예측값인  $D_\phi(r_u, s_u | u)$ 은 1을 출력해야하고, 사용자  $u$ 에 대해 생성기에 의해 생성된 구매벡터 분포인  $p_\theta$ 에서 가져온 구매벡터( $\hat{r}_u$ )와 사용자프로파일

( $s_u$ )이 입력되면 판별기의 예측값인  $D_\phi(\hat{r}_u \otimes e_u, s_u | u)$ 가 0을 출력해야 목적함수가 최대가 된다. 즉 위의 목적함수를 최대로 하는 판별모델의 최적의 파라미터  $\phi^*$ 를 찾는 것이 판별기의 최종 목표이다. 이때 생성기의 출력  $\hat{r}_u$ 에  $u$ 가 아이템  $i$ 를 구매했는지 여부를 나타내는 indicator vector인  $e_u$ 를 곱하여 마스킹한다는 점이다. 즉 사용자  $u$ 에 의해 실제 구매하지 않은 아이템에 해당하는 생성기의 출력 노드를 제거하여 판별기로 하여금 제거된 노드를 무시하게 하고, 생성기로 하여금 제거된 노드에 대해 경사를 갱신하지 않게 함으로써 계산의 복잡도를 줄이는 효과가 있다.

한편 생성기의 목적함수는 다음과 같다.

$$\begin{aligned} \theta^* &= \arg \min_{\theta} \sum_{u=1}^N (E_{r_u \sim p_{true}} [\log D_\phi(r_u, s_u | u)] \\ &\quad + E_{\hat{r}_u \sim p_\theta} [\log(1 - D_\phi(\hat{r}_u \otimes e_u, s_u | u))]) \\ &= \arg \min_{\theta} \sum_{u=1}^N E_{\hat{r}_u \sim p_\theta} [\log(1 - D_\phi(\hat{r}_u \otimes e_u, s_u | u))] \end{aligned}$$

생성기는 판별기를 속일 수 있을 만큼 생성된 구매벡터 분포  $p_\theta$ 를 실제 구매벡터 분포인  $p_{true}$ 에 적합 시킴으로써 목적함수를 최소화한다. 생성기의 목적함수는  $D_\phi(\hat{r}_u \otimes e_u, s_u | u)$ 가 1일 때 최소화가 되며 이는 생성된 구매벡터가 사용자프로파일과 함께 판별기에 입력되었을 때 판별기가 1을 출력할 수 있도록 해야 함을 의미한다.

이상의 논의를 토대로 OCF-GAN의 훈련 알고리즘(Figure 2)을 구성하면 다음과 같다.

<p>Input: rating matrix <math>\{r_u\}</math>, user profile <math>\{s_u\}</math>  1: Initialize generator and discriminator with random weights <math>\theta, \phi</math>  2: repeat  3:   for g-steps do  4:     Sample minibatch of K users  5:     Generates K fake purchase vectors <math>\{\hat{r}_1, \hat{r}_2, \dots, \hat{r}_K\}</math> for each user <math>u</math> using <math>r_u, s_u</math>  6:     Update generator parameters by <math>\theta^{new} = \theta^{old} - \lambda \frac{1}{K} \nabla_{\theta} J^G</math>  7:   end for  8:   for d-steps do  9:     Sample minibatch of L users  10:     Get their real purchase vectors: <math>\{r_1, r_2, \dots, r_L\}</math>  11:     Generates L fake purchase vectors <math>\{\hat{r}_1, \hat{r}_2, \dots, \hat{r}_L\}</math> using <math>r_u, s_u</math> and masking with <math>e_u</math>  12:     Train discriminator to find optimal <math>\phi^*</math>    i.e. <math>\phi^* = \arg \max_{\phi} \sum_{u=1}^N (E_{r_u \sim p_{true}} [\log D_\phi(r_u, s_u   u)] + E_{\hat{r}_u \sim p_\theta} [\log(1 - D_\phi(\hat{r}_u \otimes e_u, s_u   u))])</math>  13:   end for  14: until <math>L^{G,D}</math> converges</p>
--

(Figure 2) OCF-GAN Algorithm

OCF-GAN의 학습 알고리즘은 요약하면 다음과 같다. 먼저 생성기와 판별기의 배치사이즈(여기서는 K와 L)를 결정한다. 생성기는 K명의 사용자를 임의로 샘플링하여 해당 사용자의 구매벡터( $r_u$ )와 사용자프로파일( $s_u$ )을 입력하여 K개의 구매벡터  $\{\hat{r}_1, \hat{r}_2, \dots, \hat{r}_K\}$ 를 생성한다. 이후 생성된 구매벡터( $\hat{r}_u$ )와 사용자프로파일( $s_u$ )을 입력하고 타겟값은 실제(1)로 부여하여 생성기의 파라미터를 갱신한다. 이때 판별기의 파라미터는 고정하고 훈련을 진행한다. 이후 판별기의 훈련에서는 생성된 L개의 구매벡터와 실제 해당사용자의 L개의 구매벡터를 합친 후 레이블을 각각 생성(0)과 실제(1)로 부여하여 판별기를 훈련한다. 이상의 학습과정을 다음 배치에 다시 적용하고 모든 배치에 대해 훈련이 끝나면 1 에포크가 완료된다.

## 4. 실험

### 4.1 데이터셋

본 연구에서는 성능평가를 위해 공개된 데이터셋인 movielens 데이터 중 상호작용 내역이 10만개로 구성된 ml-100K 데이터셋을 사용하기로 한다(Song, 2019; Song and Song, 2022). 이 데이터셋은 943명의 사용자와 1,682개의 영화에 대한 총 100,000개 평가 데이터로 구성되어있다. 평점의 범위는 1점부터 5점까지이며 각 사용자는 최소 20개 이상의 영화에 평점을 부여하고 있다. 본 연구에서는 평점이 1 이상인 모든 사례를 긍정사례로 간주하여 좋아함('1')으로 정의하고 상호작용이 없는 사례를 싫어함('0')으로 정의하여 상호작용 행렬을 구성하였다. 또한 사용자 정보 중 성별, 직업, 나이를 추출하여 사용자프로파일로 준비하고 영화에 대한 정보 중 장르와 출시연도를 선정하여 아이템프로파일 정보로 준비하여 OCF-GAN 훈련에 사용하였다. 전체 상호작용데이터는 10만 건이며 이 중 80%를 훈련데이터로, 나머지 20%를 테스트데이터로 구분하여 학습 및 성능평가에 사용하였다.

### 4.2 성능평가 방법

OCF-GAN의 성능평가 절차를 사용자 기반(User-based) OCF-GAN을 예로 들어 설명하면 다음과 같다.

먼저 훈련데이터를 이용하여 OCF-GAN을 학습한다. 학습이 완료되면 테스트데이터로부터 한 명의 사용자를 선택하여 구매벡터와 사용자프로파일을 모델에 입력한 후 해당사용자를 위한 아이템별 추천점수를 출력한다. 출력된 아이템별 추천점수 중 훈련데이터에서 이미 좋아요('1')한 아이템에 대해서는 추천점수를 마이너스 값으로 재설정 후 해당 사용자를 위한 Top N아이템을 선정한다. 이후 최종 선정된 Top N아이템과 테스트셋의 해당사용자의 실제 상호작용 내역과 비교하여 최종 성능평가지표를 계산한다.

한편 성능평가를 위한 지표로 본 연구에서는 정밀도(Precision), 재현율(Recall) 뿐 아니라 추천순위기반의 성능평가 지표인 NDCG(Normalized discounted cumulative gain)와 MRR(Mean reciprocal rank)을 사용하기로 한다. MRR은 사용자가 좋아하는 아이템이 추천목록의 몇 번째에 있는지를 고려하지 않고 오직 가장 상위의 추천 아이템만을 고려하여 성능평가 점수를 계산하는 방식으로 사용자가 실제 좋아한 아이템이 추천목록의 상위에 있을수록 높은 점수를 받게 된다. 이에 비해 NDCG는 사용자가 더 좋아하는 아이템을 덜 좋아하는 아이템보다 추천 목록상 상위에 노출되었는지를 평가하는 방식으로 로그함수를 이용하여 하위 아이템에 대한 영향을 줄이기 때문에 이전 추천문제의 성능평가에도 많이 사용되고 있다.

### 4.3 실험설계 및 실험환경

여기에서는 원클래스와 콜드스타트 문제해결을 목표로 제시한 OCF-GAN이 당초 목표대로 두 가지 문제를 해결하는지를 실험을 통해 확인하기로 한다. 이를 위해 크게 두 가지로 구분하여 실험을 진행하였다. 첫 번째는 기존의 GAN기반 추천기인 CFGAN과 비교하여 성능을 평가하는 것이다. CFGAN은 오직 상호작용 내역만을 이용하여 추천을 하기 때문에 CFGAN과 OCF-GAN의 성능비교는 사용자-아이템 상호작용 내역 외에 사용자 또는 아이템에 대한 부가정보의 추가가 추천 성능에 미치는 영향을 파악하는데 도움이 된다. 사용자-아이템 상호작용 내역만을 이용하여 추천한 성능보다 사용자 또는 아이템에 대한 부가정보를 추가하여 훈련했을 때의 추천성능이 더 높다면 하이브리드 협업필터링모델과 GAN모델의 통합이 성공적이

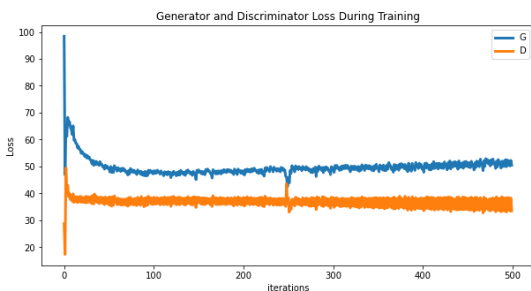
있음을 확인할 수 있다. 두 번째는 부가정보의 추가가 신규사용자 또는 신규아이템에 대한 추천성능에 얼마나 기여하는지를 실험하는 것이다. 알려진 바와 같이 단순 협업필터링 모델에서는 신규사용자와 신규아이템에 대한 추천이 불가하다. 따라서 본 연구에서는 신규 사용자 또는 신규아이템을 상호작용기록이 매우 적은 사용자 또는 아이템으로 정의하고 이들을 대상으로 OCF-GAN 모델을 학습한 후 추천을 수행하고 그 성능을 평가해 보기로 한다.

실험은 사용자기반 또는 아이템기반 여부, 사용자 또는 아이템프로파일 정보 포함여부, 전체 훈련데이터 또는 콜드스타트 대상으로 구분하여 총 8회에 걸쳐 성능평가 실험이 진행되었다. 각 실험에서 생성기와 판별기는 각각 3개의 완전연결층으로 구성하였고 은닉층의 노드는 각각 400개, 125개로 구성하였다. 또한 생성기의 배치사이즈는 32, 판별기의 배치사이즈는 64로 설정하였고 각 각의 최적화 알고리즘은 adam을 선정하였다. GAN 모델에서는 생성기와 판별기 중 한 모델을 개선하면 다른 모델이 나빠질 수 있으므로 두 경쟁 모델의 균형점을 찾는 것이 중요하다. 통상적으로 GAN이 안정적으로 학습되는 경우, 에포크가 진행되면서 판별기와 생성기의 손실값이 특정값으로 수렴해야 하며 생성기의 손실값이 판별기의 손실값보다 높게 나타나야 한다. 학습과정에서 GAN 모델의 수렴여부는 손실값의 변동(Loss Oscillator)을 보면 발견할 수 있다. 만일 GAN 모델이 수렴하지 않는 경우, 배치크기, 에포크 수, 학습률, 판별기와 생성기의 활성화 함수 등 하이퍼파라미터를 튜닝하여 해결할 수 있다. OCF-GAN이 안정적으로 학습이 되는지를 확인하기 위하여 <Figure 3>과 같이 에포크 진행에 따른 손실값의 변

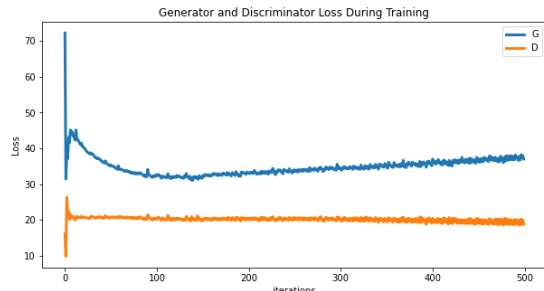
화를 관찰하였다. 그림에서 알 수 있듯이 생성기와 판별기의 손실값이 안정적으로 수렴하는 것을 확인할 수 있다.

#### 4.4 실험결과

<Table 1>은 사용자-아이템 상호작용 내역 외에 사용자 또는 아이템에 대한 부가정보의 추가가 OCF-GAN의 성능에 미치는 영향을 나타내고 있다. 즉 <Table 1>은 아이템기반 및 사용자기반 OCF-GAN을 훈련하기 위해 사용되는 정보유형에 따라 추천성능이 어떻게 다른지를 나타내고 있다. <Table 1>의 첫 번째 행은 CFGAN에서 시도한 것처럼 아이템기반으로 OCF-GAN을 구현한 후 사용자-아이템 상호작용 내역(Rating Matrix)만을 입력하여 Top 5 및 Top 20 아이템에 대한 추천을 수행하고 정밀도, 재현율, NDCG, MRR를 각각 측정된 결과이며, 두 번째 행은 아이템기반 OCF-GAN의 입력을 사용자-아이템 상호작용내역(Rating Matrix) 외에도 아이템에 대한 부가 정보까지 확장하여 훈련한 학습모델의 추천성능을 나타내고 있다. 아이템에 대한 부가정보는 무비렌즈 데이터셋에 포함된 영화에 대한 정보 중 장르와 출시년도로 이들 정보를 원-핫벡터로 인코딩하고 이후 임베딩과정을 거쳐 레이팅 정보와 함께 추천모델에 입력하였다. 성능평가 결과 정밀도 측면에서는 Top 5 추천과 Top 20 추천에서 공히 상호작용 내역 정보와 부가정보를 함께 입력한 훈련모델의 성능이 근소하게 높게 나타났다. 그러나 재현율, NDCG, MRR의 경우 상호작용 내역 정보만 입력한 훈련모델이 부가정보를 함께 학습한 훈련모델보다 근소하게 좋은 성능을 나타내기도 하였기 때문에 부가정보를 함께 활용한 추천



(a) Loss oscillator of item-based OCF-GAN



(b) Loss oscillator of user-based OCF-GAN

<Figure 3> Loss Oscillator of OCF-GAN



모델의 성능이 더 좋다고 평가할 수는 없었다. 한편 사용자 기반 OCF-GAN의 실험결과에서는 아이템 기반 OCF-GAN과는 다른 성능평가 결과가 나타났다. <Table 1>의 세 번째 행은 사용자기반으로 OCF-GAN을 구현한 후 사용자-아이템 상호작용 내역만을 입력하여 훈련모델을 학습한 후 Top 5 및 Top 20 추천 성능을 나타내고 있으며, 네 번째 행은 사용자기반 OCF-GAN의 입력을 사용자-아이템 상호작용내역 외에도 사용자에게 대한 부가정보까지 포함하여 학습한 모델의 추천성능을 나타내고 있다. 이때 사용자에게 대한 부가정보로는 성별, 직업, 나이를 활용하였다. 셋째 행과 넷째 행의 추천성능을 비교하면 Top 5 및 Top 20 추천의 모든 지표에서 부가정보를 포함하여 훈련한 추천모델의 성능이 우수한 것으로 나타났다. 아이템 기반 추천모델의 경우 부가정보가 추천성능향상에 기여하지 못한 것에 비해 사용자 기반 추천모델에서는 부가정보의 추가가 추천성능 향상에 도움이 된 것은 사용된 부가정보의 종류가 다르고 각 부가정보별로 추천성능에 미치는 기여도가 달랐기 때문일 것으로 판단된다.

<Table 2>는 부가정보의 추가가 신규사용자 또는 신규아이템에 대한 추천성능에 얼마나 기여하는지를

실험한 결과이다. 알려진 바와 같이 단순 협업필터링 모델에서는 신규사용자와 신규아이템에 대한 추천이 불가능하다. 따라서 본 연구에서는 신규사용자 또는 신규아이템을 상호작용 기록이 매우 적은 사용자 또는 아이템으로 정의하여 실험을 진행하였다. 신규사용자 또는 신규아이템에 대한 추천성능을 측정하기 위해 훈련데이터에서 특정 사용자(또는 아이템)를 선택하여 사용자-아이템 상호작용 내역을 몇 개만 남기고 다른 모든 상호작용내역을 삭제한 후 추천을 수행하고 추천결과가 삭제된 기록과 얼마나 일치하는 지로 신규사용자 또는 신규아이템에 대한 추천성능을 평가하였다. 보다 구체적으로 훈련데이터에 있는 1,683개의 아이템 중 1번에서 400번까지의 아이템에 대한 상호작용 내역은 그대로 남겨두고, 401번에서 1,683번 아이템에 대한 상호작용 내역은 제거한 후 총 46,093건으로 훈련데이터를 재 구성한 후 추천모델을 학습하였다. <Table 2>의 첫 번째 행은 아이템기반으로 OCF-GAN을 구현한 후 사용자-아이템 상호작용 내역만을 입력하여 Top N 추천을 수행한 후의 성능평가 결과이며, 두 번째 행은 사용자-아이템 상호작용 내역 뿐 아니라 아이템에 대한 부가정보(장르, 제작년도)를 이용하여 추천모델을 구

<Table 1> Experiment Results of Item-based and User-based OCF-GAN

Method	Input	Top5				Top20			
		Precision	Recall	NDCG	MRR	Precision	Recall	NDCG	MRR
Item-based	Rating (CFGAN)	0.4358	0.1518	0.4728	0.6841	0.2914	0.3596	0.4358	0.6963
	Rating+Side (OCF-GAN)	0.4426	0.1544	0.4725	0.6694	0.2920	0.3591	0.4334	0.6812
User-based	Rating (CFGAN)	0.3765	0.1247	0.3776	0.5210	0.2672	0.3315	0.3700	0.5383
	Rating+Side (OCF-GAN)	0.3962	0.1355	0.4268	0.6384	0.2708	0.3398	0.4017	0.6530

<Table 2> Experiment Results for Cold-start Cases

Train data	Method	Input	Top5				Top20			
			Precision	Recall	NDCG	MRR	Precision	Recall	NDCG	MRR
Cold-start Cases	Item-based OCF-GAN	Rating	0.3928	0.1350	0.4264	0.6440	0.2540	0.3082	0.3814	0.6572
		Rating+Side	0.3998	0.1357	0.4328	0.6465	0.2555	0.3089	0.3847	0.6585
	User-based OCF-GAN	Rating	0.2106	0.0676	0.2245	0.3426	0.1489	0.1752	0.2121	0.3573
		Rating+Side	0.2310	0.0759	0.2512	0.3940	0.1597	0.1864	0.2326	0.4131

성한 후의 성능평가 결과를 나타내고 있다. 또한 <Table 2>의 세 번째 행은 사용자기반으로 OCF-GAN을 구현한 후 사용자-아이템 상호작용 내역만을 입력하여 모델을 학습한 후 Top 5 및 Top 20 추천에 대한 추천성능을 나타내고 있으며, 네 번째 행은 사용자기반 OCF-GAN의 입력을 사용자-아이템 상호작용 내역 외에도 사용자에게 추가 정보까지 포함한 학습모델의 추천성능을 나타내고 있다. 셋째 행과 넷째 행의 추천성능을 비교하면 Top 5 및 Top 20의 모든 지표에서 추가정보를 포함한 추천모델의 성능이 우수한 것으로 나타났다. 결국 추가정보를 포함한 GAN기반 하이브리드 협업필터링모델인 OCF-GAN은 신규사용자나 신규아이템에 대한 추천에 있어서도 효과적인 것으로 나타났다.

## 5. 결론 및 향후 연구

본 연구에서는 원클래스와 콜드스타트 문제 해결을 위해 GAN을 이용한 하이브리드 협업필터링 모델인 OCF-GAN을 제안하고 성능평가를 통해 유용성을 확인하였다. Conditional GAN에 기반한 OCF-GAN은 실제 사용자 선호패턴과 유사한 패턴을 생성하려는 생성기와 실제 선호패턴과 생성된 선호패턴을 구분하려는 판별기로 구성되며, 생성기는 판별기와의 경쟁을 통해 실제 사용자가 선호하는 아이템 분포를 학습하고 훈련이 종료되면 실제 선호아이템 분포에 기반하여 사용자가 좋아할 만한 아이템을 생성하게 된다. 또한 협업필터링 추천기 대신 사용자와 아이템에 대한 추가정보를 추가적으로 활용하는 하이브리드 협업필터링 추천방식을 사용함으로써 콜드스타트 문제를 해결하였다.

성능평가 결과 Top 5 및 Top 20 추천의 모든 지표에서 기존 GAN기반 추천기 대비 추가정보를 포함하여 훈련한 OCF-GAN의 성능이 우수한 것으로 나타났다. 이러한 현상은 콜드스타트 사용자와 아이템에 대상으로 한 실험에서 보다 명확하게 드러났다. 제안된 OCF-GAN의 생성기는 부정사례 증강을 통해 원클래스 문제를 해소할 수 있을 뿐 아니라 신규사용자 및 아이템에 대해서도 재학습 없이 추천이 가능하다는 장점이 있다. 또한 본 연구에서 제안한 생성기는 개별 아이템과 같은 단일 아이템 인덱스를 생성하는 대신 실제값을 가지는 원소로 구성된 구매벡터를 직접 생성하기 때

문에 경사하강법(Gradient descent) 적용이 가능하며, 훈련데이터에 모순된 레이블이 존재하는 단일 아이템 인덱스 생성 방식의 문제를 해결할 수 있기 때문에 성능 향상을 기할 수 있다는 장점이 있다. 그러나 OCF-GAN은 구매벡터의 크기가 아이템의 개수가 많아지면 지나치게 커질 수 있고, 상호작용 내역이 전혀 없는 신규사용자나 신규아이템에 대한 추천은 어렵다는 실무적 한계도 여전히 존재한다.

향후 연구로써 추천모델에 대한 설명기능을 제공하기 위한 후속연구가 필요할 것이다. GAN 기반 추천모델의 생성기와 판별기는 주로 블랙박스 모델로 분류되는 DNN을 사용하여 구성되기 때문에 추천이유에 대한 설명이 어렵다. 따라서 GAN의 생성기와 판별기가 모델 설명 가능성을 향상시키기 위해 서로 경쟁하는 형태의 연구가 필요할 것으로 보인다. 또한 지속적으로 증가하는 데이터를 실시간으로 학습하여 추천에 반영하기 위해서는 GAN 기반 추천모델의 확장성(Scalability)도 중요한 연구 이슈가 될 것이다.

## References

- [1] Adams, R. P., and Murray, G. E. D. I., "Incorporating side information in probabilistic matrix factorization with gaussian processes", arXiv preprint arXiv: 1003.4944, 2010.
- [2] Bokde, D., Girase, S., and Mukhopadhyay, D., "Matrix factorization model in collaborative filtering algorithms: A survey", *Procedia Computer Science*, Vol. 49, 2015, pp. 136-146.
- [3] Chae, D., Kang, J., Kim, S., and Lee, J., "CFGAN: A generic collaborative filtering framework based on generative adversarial networks", In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM 2018)*, 2018, pp. 137-146.
- [4] Fan, W., Derr, T., Ma, Y., Wang, J., Tang, J., and Li, Q., "Deep adversarial social recommendation", In *Proceedings*

- of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI 2019), 2019, pp. 1351-1357.
- [5] He, X. and Chua, T., "Neural factorization machines for sparse predictive analytics", Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval. ACM, 2017.
- [6] He, X., Liao, L., Zhang, H., Nie, L., Hu, X., and Chua, T., "Neural collaborative filtering", Proceedings of the 26th International Conference on World Wide Web, 2017, pp. 173-182.
- [7] Hwang, W.-S., Parc, J., Kim, S.-W., Lee, J., and Lee, D., "Told You I Didn't Like It: Exploiting Uninteresting Items for Effective Collaborative Filtering", In Proceedings of the 32nd IEEE Int'l Conf. on Data Engineering (IEEE ICDE), 2016.
- [8] Koren, Y., Bell, R., and Volinsky, C., "Matrix factorization techniques for recommender systems", Computer, Vol. 42, No. 8, 2009, pp. 30-37.
- [9] Lee, J., Sun, M., and Lebanon, G., "A comparative study of collaborative filtering algorithms", arXiv preprint arXiv:1205.3193, 2012.
- [10] Li, Y., Hu, J., Zhai, C., and Chen, Y., "Improving one-class collaborative filtering by incorporating rich user information", In Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM '10), 2010, pp. 959-968.
- [11] Li, G., Zhang, Z., Wang, L., Chen, Q., and Pan, J., "One-class collaborative filtering based on rating prediction and ranking prediction", Knowledge-Based Systems, Vol. 124, 2017, pp. 46-54.
- [12] Liu, J., Pan, W., and Ming, Z., "Cofigan: Collaborative filtering by generative and discriminative training for one-class recommendation", Knowledge-based Systems, Vol. 191, 2020.
- [13] Mirza, M. and Osindero, S., "Conditional generative adversarial nets", arXiv preprint arXiv:1411.1784, 2014.
- [14] Pan, R., Zhou, Y., Cao, B., Liu, N. N., Lukose, R., Scholz, M., and Yang, Q., "One-class collaborative filtering", Eighth IEEE International Conference on Data Mining, 2008.
- [15] Porteous, I., and Asuncion, M. W. A. U., "Bayesian matrix factorization with side information and dirichlet process mixtures", In Proceedings of AAAI, 2010.
- [16] Song, H. S., "Deep neural network-based beauty product recommender", Journal of Information Technology Application & Management, Vol. 26, No. 6, 2019, pp. 89-101.
- [17] Song, G. and Song, H., "Algorithm for generating negative cases for collaborative filtering recommender", Expert Systems, Vol. 39, No. 12, 2022.
- [18] Strub, F., Mary, J., and Gaudel, R., "Hybrid Collaborative Filtering with Autoencoders", arXiv preprint arXiv:1603.00806, 2016.
- [19] Wang, J., Yu, L., Zhang, W., Gong, Y., Xu, Y., Wang, B., Zhang, P., and Zhang, D., "IRGAN: A minimax game for unifying generative and discriminative information retrieval models", In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2017, pp. 515-524.
- [20] Won, H., Lee, Y., Shim, J., and Ahn, H., "A Hybrid Collaborative Filtering Model Using Customer Search Keyword Data

- for Product Recommendation”, 18th IEEE International Conference On Machine Learning And Applications (ICMLA), 2019, pp. 1523-1526.
- [21] Zhang, W., “Generative adversarial nets for information retrieval: Fundamentals and advances”, The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, 2018, pp. 1375-1378.
- [22] Zhao, W., Wang, B., Ye, J., Gao, Y., Yang, M., and Chen, X., “PLASTIC: prioritize long and short-term information in top-n recommendation using adversarial training”, In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI 2018), 2018, pp. 3676-3682.

## ■ 저자소개



송희석

고려대학교 경영학과 학사, 한국과학기술원 경영과학과 석사 및 경영공학과 박사학위를 취득하였으며 대우전자, 대우정보시스템 기술연구소 근무를 거쳐 현재는 한남대학교 경영정보학과 교수로 재직 중이다.

관심분야는 CRM과 DataMining, 소셜네트워크, 빅데이터, 딥러닝, 비즈니스모델 등이며 주요연구 결과는 Knowledge-based systems, Telecommunications Policy, Expert systems, Expert systems with applications, Artificial Intelligence Review 등의 학술지에 발표하였다.