

일반화 적응 심층 잠재요인 추천모형

김정하

국민대학교 시빅데이터융합경영학과
(emily577@naver.com)

장성현

국민대학교 시빅데이터융합경영학과
(jjang2142@naver.com)

이지평

국민대학교 시빅데이터융합경영학과
(easypyeong@naver.com)

조윤희

국민대학교 시빅데이터융합경영학과
(www4u@kookmin.ac.kr)

대표적인 추천 시스템 방법론인 협업 필터링(Collaborative Filtering)에는 이웃기반 방법(Neighbor Methods)과 잠재 요인 모델(Latent Factor model)이라는 두 가지 접근법이 있다. 이중 행렬 분해(Matrix Factorization)를 이용하는 잠재 요인 모델은 사용자-아이템 상호작용 행렬을 두 개의 보다 낮은 차원의 직사각형 행렬로 분해하고 이들의 행렬 곱으로 아이템의 평점(Rating)을 예측한다. 평점 패턴으로부터 추출된 요인 벡터들을 통해 사용자와 아이템 속성을 포착할 수 있기 때문에 확장성, 정확도, 유연성 측면에서 이웃기반 방법보다 우수하다고 알려져 있다. 하지만 평점이 지정되지 않은 아이템에 대해서는 선호도가 다른 개개인의 다양성을 반영하지 못하는 근본적인 한계가 있고 이는 반복적이고 부정확한 추천을 초래하게 된다.

이러한 잠재요인 모델의 한계를 개선하고자 각각의 아이템 별로 사용자의 선호도를 적응적으로 학습하는 적응 심층 잠재요인 모형(Adaptive Deep Latent Factor Model; ADLFM)이 등장하였다. ADLFM은 아이템의 특징을 설명하는 텍스트인 아이템 설명(Item Description)을 입력으로 받아 사용자와 아이템의 잠재 벡터를 구하고 어텐션 스코어(Attention Score)를 활용하여 개인의 다양성을 반영할 수 있는 방법을 제시한다. 하지만 아이템 설명을 포함하는 데이터 셋을 요구하기 때문에 이 방법을 적용할 수 있는 대상이 많지 않은 즉 일반화에 있어 한계가 있다. 본 연구에서는 아이템 설명 대신 추천시스템에서 보편적으로 사용하는 아이템 ID를 입력으로 하고 Self-Attention, Multi-head attention, Multi-Conv1d 등 보다 개선된 딥러닝 모델 구조를 적용함으로써 ADLFM의 한계를 개선할 수 있는 일반화된 적응 심층 잠재요인 추천모형 G-ADLFM을 제안한다.

다양한 도메인의 데이터셋을 가지고 입력과 모델 구조 변경에 대한 실험을 진행한 결과, 입력만 변경했을 경우 동반되는 정보손실로 인해 ADLFM 대비 MAE(Mean Absolute Error)가 소폭 높아지며 추천성능이 하락했지만, 처리할 정보량이 적어지면서 epoch 당 평균 학습속도는 대폭 향상되었다. 입력 뿐만 아니라 모델 구조까지 바꿨을 경우에는 가장 성능이 우수한 Multi-Conv1d 구조가 ADLFM과 유사한 성능을 나타내며 입력변경으로 인한 정보손실을 충분히 상쇄시킬 수 있음을 보여주었다. 결론적으로 본 논문에서 제시한 모형은 기존 ADLFM의 성능은 최대한 유지하면서 빠른 학습과 추론이 가능하고(경량화) 다양한 도메인에 적용할 수 있는(일반화) 새로운 모형임을 알 수 있다.

주제어 : 추천시스템, 협업 필터링, 행렬 분해, 딥러닝, 심층 잠재요인

논문접수일 : 2023년 2월 15일 논문수정일 : 2023년 2월 23일 게재확정일 : 2023년 2월 27일
원고유형 : Regular Track 교신저자 : 조윤희

* 본 논문은 산업통상자원부 지식서비스산업핵심기술개발사업으로 지원된 연구결과입니다. (20015152, 빅데이터 가공 및 공급 자동화를 기반한 통합 스포 데이터 분석 기술과 비대면 시장조사 시스템 융합 기술 개발)

1. 서론 및 관련 연구

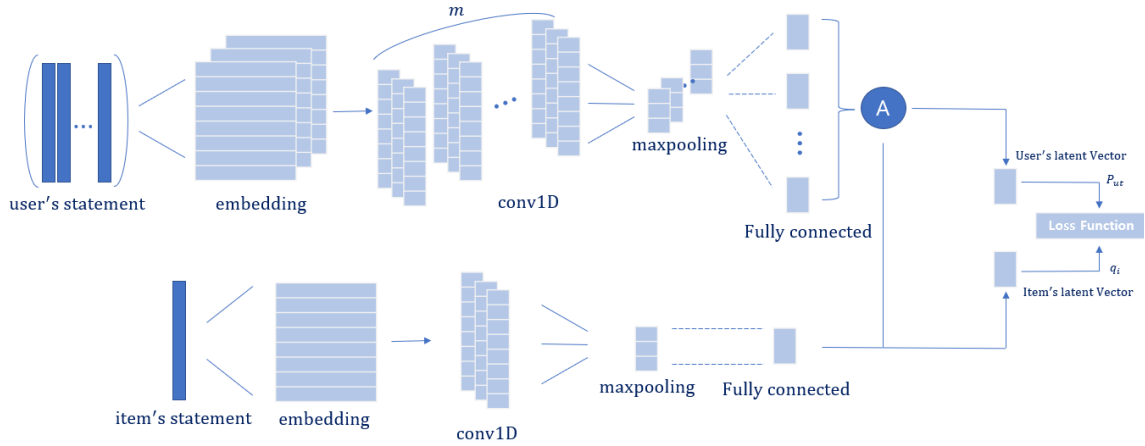
추천 시스템은 e-commerce 환경에서 고객에게 편리함을 제공함과 동시에 구매를 유도하여 기업의 이윤을 높인다. 이는 추천 시스템을 활용하는 온라인 쇼핑몰들의 성장을 통해 증명되고 있는데, Netflix에서는 미디어 선택의 60%가 아마존에서는 매출의 약 35%가 추천 시스템으로부터 기인한다고 알려져 있다(Lee et al., 2019). 이에 따라 추천 시스템은 매우 활용 가치가 높은 분야로 활발하게 연구가 진행되고 있다. 추천 시스템의 여러 방법론 중 가장 대표적인 것으로 협업 필터링(Collaborative Filtering)이 있다. 협업 필터링은 이전 구매 기록이나 제품 평가 기록 등 사용자의 과거 행동에 의존하는 방식이다(Goldberg et al., 1992; 박호연, 김경재, 2021; 권유진 등, 2022). 과거 행동으로부터 사용자-아이템 간의 상관관계를 찾아내는 것이 목적이며 특정 분야에 대한 지식이 필요하지 않다. 이러한 협업 필터링에는 이웃기반 방법(Neighbor Methods)과 잠재 요인 모델(Latent Factor model)이라는 두 가지 대표적인 접근법이 있다(홍태호 등, 2022).

먼저, 이웃기반 방법은 아이템 혹은 사용자 간 유사도를 계산하여 추천 대상과 가장 유사한 사용자/아이템으로부터 추천 아이템을 선정한다. 유사도 대상에 따라 사용자 기반과 아이템 기반으로 나눌 수 있으며 직관적이고 구현이 쉽다는 장점이 있다. 다음으로, 잠재 요인 모델은 사용자의 평점을 설명하는 잠재된 특징(요인)을 찾고 이를 추천에 활용하는 방법으로서 잠재 요인을 찾기 위해 행렬 분해(Matrix Factorization)를 주로 사용한다(Koren, 2008). 행렬 분해 기반 잠재 요인 모델은 사용자-아이템 상호작용 행렬을 두 개의 보다 낮은 차원의 직사각형 행렬로 분해하고

이들의 행렬 곱으로 아이템의 평점(Rating)을 예측한다. 평점 패턴으로부터 추출된 요인 벡터들을 통해 사용자와 아이템 속성을 포착할 수 있기 때문에 확장성, 정확도, 유연성 측면에서 이웃기반 방법보다 우수하다고 알려져 있다. 하지만 평점이 지정되지 않은 아이템에 대해서는 선호도가 다른 개인의 다양성(Individual Diversity)을 반영하지 못한다는 근본적인 단점을 지닌다. 이는 결국 반복적이고 부정확한 추천을 초래하게 된다. 이러한 단점을 보완하고자 최근에 각각의 아이템 별로 사용자의 선호도를 적응적으로 학습하는 적응 심층 잠재요인 모형(Adaptive Deep Latent Factor Model; ADLFM)이 등장하였다(Han et al., 2020).

ADLFM은 사용자-아이템 평점 대신에 평점이 있는 아이템 설명(Item Description)을 입력으로 받아 사용자와 아이템의 잠재 벡터를 구하고 어텐션 스코어(Attention Score)를 활용하여 개인의 다양성을 반영할 수 있는 방법을 제시한다. 여기서 아이템 설명이란 아이템의 특징을 설명하는 텍스트를 의미하는데 예를 들어, 옷의 경우에는 색상, 소재, 사이즈 등을 설명하는 텍스트를 말하며 책의 경우에는 제목, 저자, 출판사, 출판일, 책 소개, 책 속의 내용 등을 포함한다. 사용자의 평점 기록에서는 사용자가 관심을 가질 수 있는 상위 카테고리만을 알 수 있기 때문에 선호도 추론에 한계가 있지만, 아이템에 대한 더 자세한 정보를 담고 있는 아이템 설명을 사용하면 보다 다양하고 의미 있는 사용자의 선호도 벡터를 구성할 수 있게 된다(Han et al., 2020).

ADLFM의 모델 아키텍처를 개략적으로 도식화하면 <그림 1a>와 같다. 크게 사용자에게 대한 잠재 벡터를 구하는 과정과 타겟 아이템에 대한 잠재 벡터를 구하는 과정으로 나누어지며, 각 과정은 합성곱층(Convolution Layer)으로부터 어텐션



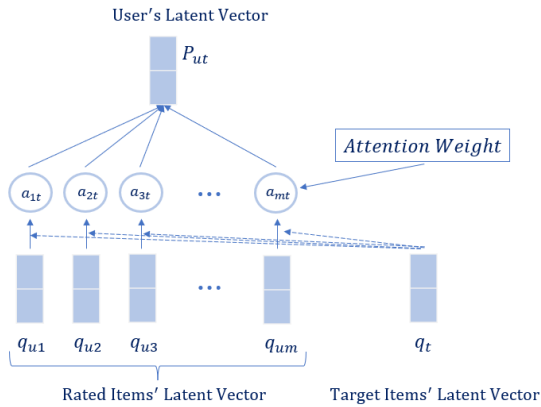
<그림 1a> ADLFM 모델 아키텍처(Han et al., 2020)

층(Attention Layer)(<그림 1a>의 A)까지 총 4개의 은닉층(Hidden Layer)으로 구성된다. 사용자의 잠재 벡터를 구하기 위한 입력으로는 사용자가 이용한 복수의 아이템 설명이, 타겟 아이템의 잠재 벡터를 구하기 위한 입력으로는 해당 아이템 설명이 사용된다. 모든 아이템 설명은 원-핫 인코딩(One-Hot Encoding)된 형태로 각각의 임베딩층(Embedding Layer)을 통과한다. 이후 합성곱층을 통과하며 서로 다른 필터(filter)를 사용하여 병렬적으로 특징을 추출한다. 이렇게 추출된 특징들은 최대풀링층(Max-Pooling Layer)을 거쳐 가장 중요한 문맥적(contextual) 특징만을 보존하게 되는데 이는 합성곱층을 통해 얻은 모든 특징이 성능 향상에 도움이 되지 않기 때문에 필요한 과정이다(Chollet, 2018). 마지막으로, 완전연결층(Fully Connected Layer)을 거치면 사용자와 관련된 아이템과 타겟 아이템에 대한 잠재 벡터를 얻게 된다. 지금까지의 처리과정은 텍스트 데이터를 처리하는 1D-CNN(Chollet, 2018)과 큰 차이가 없다.

ADLFM이 1D-CNN과 크게 다른 점은 이후 과정에서 어텐션을 사용한다는 것이다. ADLFM의

어텐션 모듈은 각 아이템 설명의 단어나 구절에 가중치를 부여하여 해당 아이템의 사용자별 취향을 보다 정확하게 반영하기 위하여 타겟 아이템의 잠재 벡터를 사용자와 관련된 아이템에 대한 어텐션 가중치로 적용하여 사용자에게 대한 잠재 벡터를 구한다. 이 과정은 <그림 1a>의 A 부분에 해당하며, A에 대한 세부 구조는 <그림 1b>와 같다. 이러한 방식으로, 어텐션 모듈은 사용자별 선호도의 다양성을 캡처하면서, 텍스트 속성을 활용하여 추천을 개인화한다. 이와 같이 사용자 잠재 벡터와 타겟 아이템 잠재 벡터를 얻은 후 기존의 잠재 요인 모델과 동일한 방식으로 두 벡터를 곱한 값이 특정 아이템에 대한 사용자의 평점과 같아지도록 학습이 진행된다.

Amazon 데이터 세트를 이용한 광범위한 실험을 통해 ADLFM은 기존 대표적인 추천 시스템 모델인 MF(Koren et al., 2009), PMF(Mnih et al., 2008), HFT(McAuley et al., 2013), RMR(Ling et al., 2014), DeepCoNN(Zheng et al., 2017)보다 높은 성능을 보였고, 모델 아키텍처에서 Attention의 기여도가 가장 높음을 알 수 있었다(Han et al., 2020).



〈그림 1b〉 ADLFM의 Attention Module
(Han et al., 2020)

ADLFM은 아이템 설명을 활용하여 추천을 제공하고 개인의 다양성을 고려하는 장점이 있지만, 이로 인해 두 가지 문제점이 발생한다. 첫째, 아이템 설명이 포함된 데이터가 드물어서 해당 모델이 일반화하기 어려워지는 단점이 있다. 대표적인 추천 시스템 데이터인 Movie Lens와 Netflix Prize는 아이템 설명 대신 사용자, 아이템, 평점, 평점 부여 시간 등을 포함하고 있다(Harper et al, 2015; Narayanan et al., 2006). 이로 인해 ADLFM을 적용할 수 있는 데이터가 제한된다. 둘째, 아이템 설명은 텍스트 데이터로 학습 속도가 현저히 저하되는 문제가 있다.

본 논문에서는 이러한 문제점을 해결할 수 있는 G-ADLFM(Generalized Adaptive Deep Latent Factor Recommendation Model)을 제안한다. G-ADLFM은 크게 두 가지 개선 사항을 갖는다. 첫째, 아이템 설명 대신 아이템 ID를 입력으로 사용하여 ADLFM보다 적용 가능한 데이터 범위를 넓혀 일반화 가능성을 제고하고자 한다. 아울러 아이템 설명의 데이터 크기보다 아이템 ID의 크기가 월등히 작기 때문에 모델의 경량화의 효과도 기대한다. 둘째,

아이템 설명 대신 아이템 ID를 사용하기 때문에 발생하는 정보의 손실로 인한 성능 하락을 막기 위해 최신의 딥러닝 레이어를 추가적으로 도입하여 ADLFM의 모델 아키텍처를 개선하고자 한다.

이 연구는 Self-Attention, Multi-Head Attention, Multi-Conv1D를 각각 적용한 3가지 버전의 G-ADLFM을 제안하고, 다양한 도메인의 데이터셋을 이용하여 입력과 모델 구조 변경에 대한 실험을 수행하였다. 실험 결과, 입력만 변경했을 경우 ADLFM 대비 MAE가 소폭 증가하고 추천 성능이 하락하지만, 처리할 정보량이 줄어들어 따라 epoch 당 평균 학습 속도가 대폭 향상되는 것을 발견하였다. 입력 및 모델 구조를 변경한 경우, Multi-Conv1d 구조가 ADLFM과 유사한 성능을 나타내며 입력 변경으로 인한 정보손실을 충분히 보상할 수 있음을 보여주었다. 따라서, 이 논문에서 제시한 모델은 기존 ADLFM의 성능을 최대한 유지하면서 빠른 학습 및 추론이 가능하며(경량화), 다양한 도메인에 일반화하여 적용할 수 있는 새로운 모델임을 확인할 수 있다.

2. 제안 추천 모델

2.1. G-ADLFM with Self-Attention

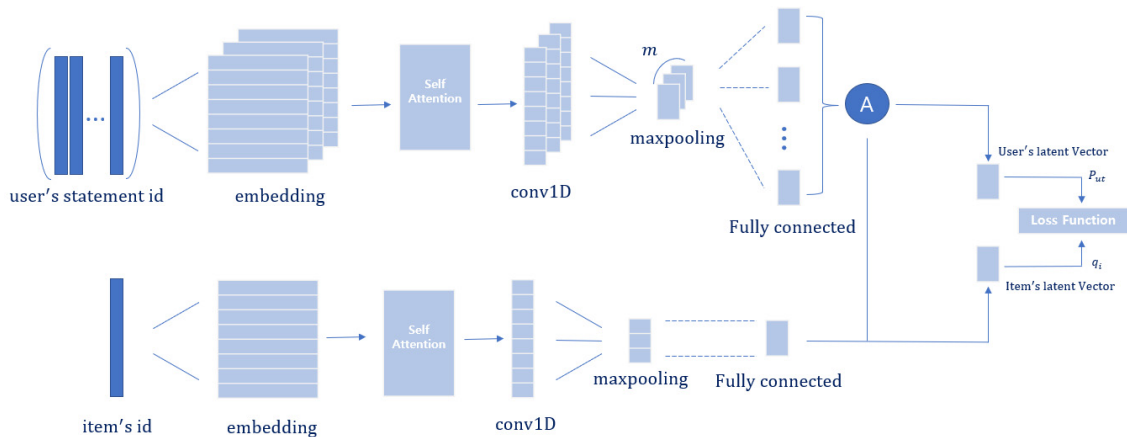
가장 먼저, Self-Attention을 활용한 G-ADLFM이다. 전반적인 모델 구조를 설명하기 전에 Self-Attention이 무엇인지 Vaswani et al(2017)의 논문에 기반하여 설명하고자 한다. Self-Attention이 아닌 일반적인 어텐션 함수는 쿼리(Query)와 키(Key)의 모든 조합에 대해 유사도를 계산한다. 이 유사도를 바탕으로 각 키에 대한 어텐션 가중치를 계산하고 이 가중치를 이용해 해당 키와 연

관된 값에 가중합을 취하여 반환한다. 즉, I love you라는 문장이 있을 때 I가 I love you 모든 단어에 대해 유사도를 구하고 그것을 값에 반영하고 최종적으로 가중합하여 반환하는 것이다(Vaswani et al., 2017).

이러한 일반적인 어텐션과는 달리, Self-Attention은 어텐션을 자기 자신에게 수행하는 것이다. 번역이나 자연어 처리(Natural Language Processing; NLP)를 진행할 때에 가장 중요한 점은 단어와 단어 간의 관계를 확실히 파악하는 것이다. 순환 신경망(Recurrent Neural Network; RNN)에서 단어 간의 간격이 먼 단어들의 관계를 잘 파악하지 못하는 단점이 있어, 이를 극복하고자 장단기 메모리(Long Short-Term Memory, LSTM) 모델이 등장했듯이, Self-Attention을 활용하면 본래 단어와도 잘 연관이 되며 문장 내 다른 단어들과의 유사도도 잘 파악할 수 있다. 한마디로 Self-Attention은 임베딩 벡터를 보완하여 보다 정확한 문장의 표현을 만드는 데 기여한다고 볼 수 있다.

앞서 설명한 기능을 하는 Self-Attention을 추가하여 새롭게 제안하는 Self-Attention G-ADLFRM의 모델 구조는 <그림 2>와 같다. 기존 ADLFRM

에서는 총 4개의 은닉층을 사용했다면, Self-Attention G-ADLFRM은 총 5개의 은닉층을 사용한다. 사용자 잠재 벡터와 고려 중인 아이템(Target Item)의 잠재 벡터를 얻기 위해 아이템 ID가 입력으로 사용된다. 아이템 ID는 Embedding 이후 Self-Attention Layer를 지나게 된다. 이 과정은 아이템 설명에서 아이템 ID로의 입력 변경이 주는 정보의 손실과 그에 따른 성능 하락을 극복하기 위해서 Embedding을 강화할 수 있도록 한다. 또한 합성곱 신경망이나 순환 신경망과 비교했을 때, 레이어 당 총 계산 복잡도를 낮춰 경량화 및 빨라진 학습 속도를 유지할 수 있다. 입력 시퀀스의 길이 n 이 임베딩 차원 d 보다 작은 경우가 대부분인데, 이는 Self-Attention이 RNN이나 CNN과 같은 다른 방법보다 층당 계산 복잡도에서 유리하다는 것을 의미한다. 그리고 네트워크 내부에서 장거리 의존성(Long-Range Dependency) 사이의 경로 길이, 즉 최대 경로 길이(Maximum Path)에 따라 의존성을 가질 수도 있는데, Self-Attention은 문장 내의 모든 단어에 대해 $O(1)$ 의 복잡도로 한 번에 접근할 수 있기 때문에 합성곱 신경망이나 순환 신경망 보다 더 우수한 성능을 보인다(Vaswani



<그림 2> Self-Attention G-ADLFRM

et al., 2017). 이러한 이유로 Self-Attention Layer를 추가한 Self-Attention G-ADLFRM을 제안한다.

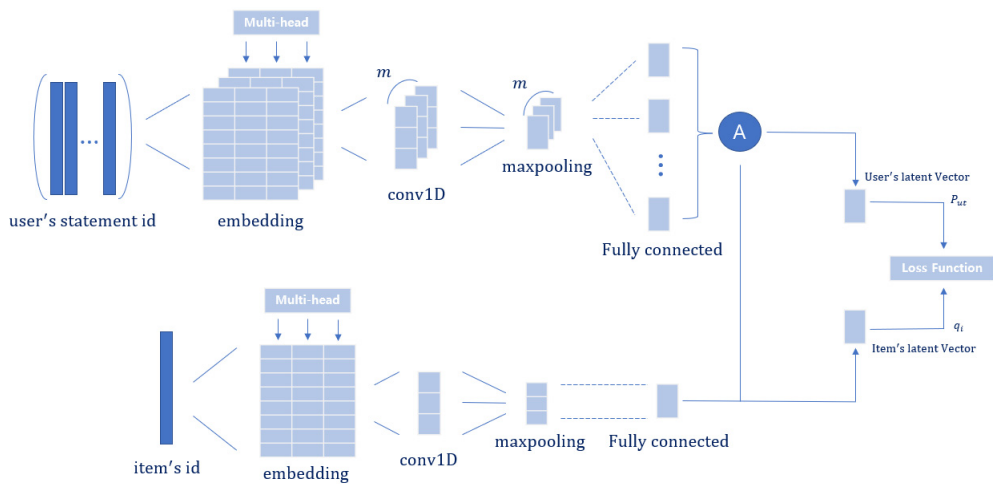
2.2. G-ADLFRM with Multi-Head Attention

두 번째로 Multi-Head 모듈을 적용한 G-ADLFRM이다. 먼저 제안한 Self-Attention G-ADLFRM처럼 이 모델도 아이템 설명 대신 아이템 ID를 넣어 사용자와 아이템의 잠재 벡터를 구하게 된다. 정보의 손실과 성능 하락을 개선하기 위해서는 다수의 레이어를 추가하거나 더욱 복잡한 모델 구조를 설계할 수도 있지만, 입력의 변경으로 얻을 수 있는 장점인 경량화와 빠른 학습 속도를 유지하기 위해서는 약간의 구조 수정을 통해 더욱 강화된 임베딩을 얻어 잠재 벡터를 효과적으로 만들어야 한다. 이를 위해 임베딩 과정에 Multi-Head를 추가했다.

트랜스포머 모델에서는 어텐션 함수를 한 번 적용하는 것보다 어텐션 함수를 병렬로 여러 번 적용하는 것이 더 효과적일 거라는 가정 하에 d_model 의 차원을 Num_Heads 개로 나누어서

d_Model/Num_Heads 의 차원을 갖는 쿼리, 키, 값에 대해 Num_Heads 개의 병렬 어텐션을 수행했다 (Vaswani et al., 2017). 해당 논문에서는 하이퍼 파라미터인 Num_Heads를 8로 정했고, 8개의 병렬 어텐션이 이뤄진다. 다시 말해 앞서 설명한 어텐션이 8개의 병렬로 이루어지는데, 각각의 어텐션 값 행렬은 “어텐션 헤드”로 표현되며 가중치 행렬의 값은 8개의 어텐션 헤드마다 모두 다르게 설정된다. 이는 병렬로 수행되는 여러 어텐션 헤드를 통해 서로 다른 시각에서 정보를 수집하겠다는 의미다. 병렬 어텐션 연산을 완료하면 각 어텐션 헤드의 결과를 연결(concatenate)하여 하나의 행렬로 만든다.

임베딩 과정에서 위와 같은 기능을 통해 여러 시각을 반영한 결과를 얻도록 Multi-Head Attention G-ADLFRM을 제안한다. Multi-Head Attention을 적용한 G-ADLFRM으로 전반적인 파이프라인은 <그림 3>과 같다. 입력은 앞서 설명한 Self Attention을 적용한 G-ADLFRM과 동일하게 아이템 설명이 아닌 아이템 ID를 넣었다. 그리고 임베딩 과정에



<그림 3> Multi-Head Attention G-ADLFRM

서 Multi-head Attention을 사용했다. Multi-Head Attention은 여러 가지 관점에서 어텐션을 적용해 다양한 도메인에 대해 학습을 할 수 있게 해주는 모듈이므로 다양한 관점에서 학습한다는 아이디어를 반영하고자 하고자, Multi-Head의 기능을 구현하여 기존 임베딩층에 추가했다. 이때 구현을 위하여 Keras의 Multi-Head 모듈을 사용했다(Zheng et al., 2017). 이후의 과정은 동일하다.

2.3. G-ADLFRM with Multi-Conv1D

세 번째로 <그림 4>와 같이 Multi-Conv1D를 적용한 G-ADLFRM을 제안한다. 앞에서 제안한 두 가지 경우는, 임베딩을 강화하여 성능을 향상시키고자 한 반면, Multi-Conv1D G-ADLFRM은 임베딩 이후 모델 구조 변경을 통해 성능 향상을 도모하고자 한다. 이에 기존의 합성곱층을 수정하여 임베딩 값에서 더 다양한 시각의 정보를 반영한다. 기존 ADLFRM 모델에서는 합성곱층에서 커널 크기가 3인 모델 구조를 사용해 지역(local)적인 정보만을 반영하고 있으나, 지역적인 정보뿐만 아니라 전역(global)적인 정보도 함께 반영하면 잠재 벡터의 다양성을 더욱 높일 수 있다.

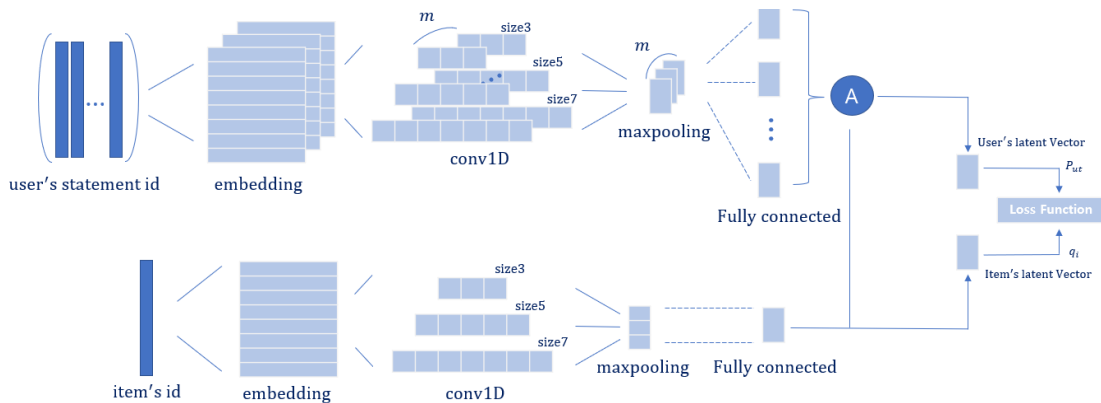
이에 본 연구에서는 이러한 고려를 바탕으로 모델 구조를 다음과 같이 개선하였다.

입력은 앞선 경우와 동일하게, 아이템의 설명 대신 아이템 ID를 사용한다. 아이템 ID는 임베딩을 거쳐 Multi-Conv1D Layer를 통과하게 되는데, 이 과정에서 다중 레이어를 사용하여 필터 크기를 각각 3, 5, 7로 바꾸어 지역 및 전역 정보를 모두 고려한다. 3 사이즈의 커널에서는 지역적인 정보를, 5, 7 사이즈의 커널에서는 상대적으로 전역적인 정보를 추출할 수 있게 되고, 이렇게 추출된 3개의 특성맵(Feature Map)을 모두 반영하여(Multiply) 잠재 벡터를 생성한다. 따라서 Multi-Conv1D Layer를 사용하면 전역 및 지역 정보가 모두 잘 반영될 수 있는 추천모델을 기대할 수 있다.

3. 실험 및 결과

3.1 실험 방법

본 연구에서는 다음과 같은 가설을 확인하기 위해 총 3가지 실험을 진행했다.



<그림 4> Multi-Conv1D G-ADLFRM

〈표 1〉 실험 데이터 개요

Data	Row	User	Items	Reference num		
				3	4	5
Anime Recommendation Database(2016)	57,633,278	325,772	17,562	90408 /3 = 30,136	119700/4 = 29,925	151070/5 = 30,214
Amazon Fine Food Reviews(2017)	568,454	256,059	74,258	130524/3 = 43,508	129796/4 = 32,449	114450/5 = 22,890
Amazon Beauty Products Ratings(2018)	2,023,270	1,210,271	249,274	193008/3 = 64,336	225196/4 = 56,299	210270/5 = 42,054

- ① 기존 ADLFM에서 입력의 변화를 주었을 때, 즉 입력을 아이템 설명에서 아이템 ID 로 변경했을 때 추천성능은 얼마나 하락하고 학습속도는 얼마나 향상되는가?
- ② 본 논문에서 제안하는 3가지 유형의 G-ADLFM 이 경량화 및 속도를 유지하며 성능을 향상시키는가?
- ③ 특정 데이터에 한정되지 않고 여러 도메인의 데이터에서도 유사한 결과를 얻을 수 있는가?

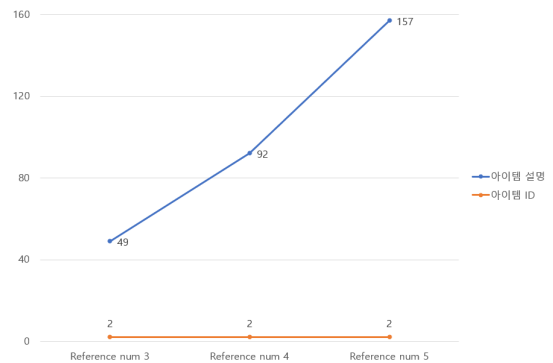
실험 데이터셋으로 <표 1>에 나와있는 Anime Recommendation Database(2016), Amazon Fine Food Reviews(2017), Amazon Beauty Product Ratings(2018) 등 3가지 데이터를 이용하였다. 각 데이터는 ADLFM의 초매개변수(hyper-parameter)인 reference num에 따라 3개의 데이터로 다시 구성하였는데, reference num이란 사용자의 잠재 벡터를 구하기 위해 사용되는 해당 사용자가 이용한 아이템의 개수를 의미한다. 학습데이터, 검증데이터, 평가 데이터를 각각 75:10:15의 비율로 나눠서 홀드아웃(hold-out) 교차검증을 Google Colab Pro 환경에서 실시했다. 심층신경망 초매개변수로서 Optimizer(최적화 알고리즘)는 Adam, Conv1D Layer의 필터 수는 64, 배치크기는 32로 설정하였고 30 epoch으로 학습을 진행했다. 추천 성능지표로서 이상치

에 덜 민감하여(Lemire et al., 2005) 추천 시스템 연구에서 많이 사용하고 있는 MAE를 사용했으며, 모델 별로 학습속도를 측정하기 위해 epoch 당 소요시간(sec)도 함께 사용하였다.

3.2 실험 결과

〈표 2〉 입력만 변경한 실험 결과

Item Description	Reference num		
	3	4	5
Item Description	1.1067	1.1166	1.1330
Item Id	1.1972	1.1918	1.2093
decrease rate	-8.2%	-6.7%	-6.7%



〈그림 5〉 입력 변경 전 후의 학습속도

〈표 3〉 G-ADLFRM 성능실험 결과

Model	Input	Modified Part	Reference num		
			3	4	5
ADLFM	Description	-	1.1067 / 157s	1.1166 / 92s	1.1330 / 49s
	ID	-	1.1972 / 3.1s	1.1918 / 2.8s	1.2093 / 2.7s
G-ADLFRM	ID	Self-Attention	1.2093 / 3.1s	1.1918 / 3.1s	1.1972 / 2.8s
		Multi-Head Conv1D	1.1822 / 3.1s	1.1847 / 2.9s	1.2147 / 2.9s
		Multi Conv1D	*1.1102 / 3.1s	*1.1045 / 3.1s	*1.1390 / 3.1s
Restruc. ADLFM Description/ Multi Conv1D			90.35%	101.08%	99.47%
Improv. ADLFM ID / Multi Conv1D			+7.26%	+7.32%	+5.81%

먼저, ADLFM의 입력으로 아이템 설명을 사용했을 때와 아이템 ID로 대체했을 때의 성능을 비교하는 실험에서는 Anime Recommendation Database(2016)를 사용했다. <표 2>의 결과를 보면, 정보손실로 인하여 MAE가 평균 7.2%(6.7% ~ 8.2%) 상승하며 입력의 변화로 인해 성능이 하락했다. 하지만 투입되는 정보량이 대폭 감소했음에도 불구하고 성능이 나빠지는 정도는 그리 크지 않으며, reference num이 증가할수록 그 차이는 오히려 적어지고 있음을 알 수 있다. 또한 <그림 5>에서 보는 바와 같이 처리할 정보량이 적어 짐에 따라 학습 속도가 평균 97.4% 감소하는 즉, 모델 경량화 효과가 크게 나타남을 확인할 수 있었고 reference num이 이러한 속도변화에 지대한 영향을 미치는 초매개변수임을 알 수 있다.

두 번째로, 본 논문에서 제안한 3가지 유형의 G-ADLFRM이 정보손실로 인한 성능 감소의 정도를 어느 정도 개선할 수 있는지 알아보는 성능비교 실험에서도 Anime Recommendation Database(2016)를 사용했다. <표 3>에 나타나는 것처럼 기존 ADLFM에서 아이템 설명을 입력으로 사용한 경우와 비교하였을 때, 투입되는 정보가 적은 아이템 ID를

사용하였음에도 모델 구조의 개선으로 전반적으로 성능이 향상된 것을 확인할 수 있었다. 평균적으로 기존 ADLFM 대비 96.9%의 성능에 도달했으며, 이중 Multi-Conv1D를 적용한 G-ADLFRM이 가장 높은 성능을 보였다. Multi-Conv1D의 경우 reference num이 3일 때 성능비가 90.35%로 가장 작았고, 5일 때는 99.47%에 도달하고 4일 때는 101.08%의 성능을 보이며 ADLFM 보다 우수한 성능을 보였다. 또한 학습속도가 92초에서 3.1초로 획기적으로 개선되었다. 이러한 결과로 미루어 볼 때, 아이템 ID로의 입력 변화로 인해 발생했던 성능 저하 문제를 본 연구에서 제시한 모델의 구조의 변경을 통해 충분히 해결할 수 있음을 알 수 있다.

마지막으로, G-ADLFRM을 Amazon Fine Food Reviews(2017)와 Amazon Beauty Product Ratings (2018) 등 다른 데이터에 적용해도 성능에 차이가 없는지 실험했다. <표 4>에 이와 관련된 실험 결과가 제시되어 있는데, Amazon Beauty Product Ratings(2018) 데이터의 평균 MAE가 1.0274, Amazon Fine Food Reviews(2017) 데이터의 평균 MAE가 0.91873로서 기존 ADLFM 보다 더 나은

〈표 4〉 다양한 데이터를 적용한 실험 결과

DATA	Model	Modified Part	Reference num		
			3	4	5
Beauty	ADLFM	-	1.0456 / 3.1s	1.0514 / 3.1s	1.0546 / 3.1s
		Self-Attention	1.0639 / 3.1s	1.0535 / 3.1s	1.0407 / 3.1s
	G-ADLFM	Multi-Head Conv1D	1.0227 / 3.1s	1.0083 / 3.1s	1.0069 / 3.1s
		Multi Conv1D	1.0086 / 4.1s	0.9745 / 3.5s	1.0072 / 3.1s
Food	ADLFM	-	0.9176 / 2.9s	0.8946 / 2.7s	0.9393 / 2.6s
		Self-Attention	0.9195 / 3.1s	0.8951 / 2.8s	0.9401 / 2.7s
	G-ADLFM	Multi-Head Conv1D	0.9132 / 2.9s	0.8911 / 2.9s	0.9391 / 2.7s
		Multi Conv1D	0.9161 / 3.1s	0.8994 / 2.9s	0.9563 / 3.1s

성능을 보여주고 있다. 이와 같이 G-ADLFM이 다양한 도메인의 데이터에서도 성능차이가 거의 없음을 보여줌으로써 도메인의 일반화 가능성이 충분하다는 것을 확인할 수 있다. 다만, 앞서 실험했던 Anime Recommendation Database(2016)의 경우 Multi-Conv1D 구조가 reference num과 관계 없이 다른 유형보다 높은 성능을 보였던 것과는 달리, 두 데이터의 경우에는 가장 높은 성능을 보이는 모델 구조가 달라지는 점을 보아 한 가지 모델 구조가 모든 도메인에 대해 항상 제일 높은 성능을 도출한다는 것은 아니라는 점을 확인할 수 있다.

4. 결론

본 논문에서는 추천 시스템의 성능을 더욱 개선하기 위해 사용자의 선호도 요인에 따른 개인의 다양성을 고려한 G-ADLFM 모델을 제안하였다. 이 모델은 기존의 대표적인 추천 시스템 모델보다 우수한 성능을 보이며, 구조가 복잡하

지 않아 추천 시스템을 처음 공부하는 사람들도 쉽게 접근할 수 있는 장점이 있다. G-ADLFM 모델은 이전에 제안된 ADLFM 모델과 비교하여 입력으로 아이템 ID를 사용함으로써 해당 모델이 적용 가능한 데이터의 범위를 확장시켰다는 점에서 큰 차이를 보인다. 또한, 정보의 손실로 인한 성능 하락을 보완하고자 모델 구조를 수정하여 성능을 더욱 향상시켰다. G-ADLFM 모델은 경량화된 구조를 가지고 있으며, 일반화가 가능하다는 점에서 매우 유용하다.

모델 구조 변화를 총 세 가지 방향으로 시도하였다. 먼저, Self-Attention과 Multi-Head Attention을 적용하여 임베딩을 강화하여 더욱 효과적으로 잠재 벡터를 만들 수 있도록 하였다. 다음으로, Conv1D Layer를 수정하여 지역적인 정보와 전역적인 정보를 모두 반영할 수 있는 Multi-Conv1D G-ADLFM을 제안하였다. 이러한 시도들로 인해 G-ADLFM 모델은 기존의 모델들보다 더욱 높은 성능을 보여주며, 사용자의 다양한 선호도 요인을 고려하여 추천을 할 수 있는 더욱 효과적인 방법을 제공하게 된다.

본 논문에서는 G-ADLFRM 모델이 기존의 ADLFRM 모델과 비교하여 입력으로 아이템 ID를 사용함으로써 해당 모델이 적용 가능한 데이터의 범위를 확장시키는 것을 목표로 하였다. 이를 위해 다양한 실험을 통해 입력을 아이템 ID로 변경하고 제안된 3개의 G-ADLFRM 모델을 적용하여 입력의 단순화로 인한 정보 손실에도 불구하고 ADLFRM와 유사한 성능을 보였다. 이는 G-ADLFRM 모델이 다양한 데이터에 적용될 수 있으며 유의미한 성능과 빠른 학습 속도를 가지고 있다는 것을 보여준다. 또한 모델 구조가 복잡하지 않아 추천 시스템을 처음 공부하는 사람들도 쉽게 접근할 수 있어 추천 시스템의 새로운 베이스 라인 모델로서 활용도가 높다.

하지만, 본 연구에서는 여러 데이터를 활용하여 실험을 수행한 결과, 데이터의 크기에 따라 최고 성능을 보이는 모델 구조가 상이하게 나타났다. 모든 경우에서 기존 ADLFRM 모델을 능가하는 모델 구조는 발견하지 못했다. 이러한 한계는 모델 아키텍처를 단순화하기 위해, 즉 경량화를 위해 기존 모델에 하나의 레이어만을 추가한 것에서 기인한 것으로 추정된다. 따라서, 추후 하나 이상의 레이어 추가를 통해 보다 나은 성능향상의 효과를 얻을 수 있는 모델에 대해 연구하고자 한다. 이러한 연구를 통해 G-ADLFRM 모델의 성능을 보다 향상시켜 다양한 분야에서 더욱 유용하게 활용될 수 있으리라 기대한다.

참고문헌(References)

[국내 문헌]

권유진, 최민석, 조운호. (2022). 부가 정보를 활용한 비전 트랜스포머 기반의 추천시스템. *지능정보연구*, 28(3), 119-137.

박호연, & 김경재. (2021). BERT 기반 감성분석을 이용한 추천시스템. *지능정보연구*, 27(2), 1-15.

홍태호, 홍준우, 김은미, 김민수. (2022). 영화 리뷰의 상품 속성과 고객 속성을 통합한 지능형 추천시스템. *지능정보연구*, 28(2), 1-18.

[국외 문헌]

Amazon - Ratings (Beauty Products). <https://www.kaggle.com/datasets/skillsmuggler/amazon-ratings>.

Amazon Fine Food Reviews. <https://www.kaggle.com/datasets/snap/amazon-fine-food-reviews>.

Anime Recommendation Database. <https://www.kaggle.com/datasets/CooperUnion/anime-recommendations-database>.

Andriy, M., & Salakhutdinov, R. R. (2007). Probabilistic matrix factorization. *Advances in Neural Information Processing Systems 20 (NIPS 2007)*.

Chollet, F. (2018). *Deep Learning with Python*. Manning Publications.

Goldberg, D., Nichols, D., Oki, B. M., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12), 61-70, 426-434.

Han, J., Zheng, L., Xu, Y., Zhang, B., Zhuang, F., Philip, S. Y., & Zuo, W. (2020). Adaptive deep modeling of users and items using side information for recommendation. *IEEE Transactions on Neural Networks and Learning Systems*, 31(3), 737-748.

Harper, F. M., & Konstan, J. A. (2015). The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, 5(4), 1-19.

- Koren, Y. (2008). Factorization meets the neighborhood: a multifaceted collaborative filtering model. *Proceedings of the 14th ACM SIGKDD*.
- Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42(8), 30-37.
- Lee, D., & Kartik H. (2019). How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment. *Information Systems Research*, 30(1), 239-259.
- Lemire, D., & Maclachlan, A. (2005). Slope one predictors for online rating-based collaborative filtering. *Proceedings of the 2005 SIAM International Conference on Data Mining*.
- Ling, G., Lyu, M. R., & King, I. (2014). Ratings meet reviews, a combined approach to recommend. *Proceedings of the 8th ACM Conference on Recommender systems*, 105-112.
- McAuley, J., & Leskovec, J. (2013). Hidden factors and hidden topics: understanding rating dimensions with review text. *Proceedings of the 7th ACM conference on Recommender systems*, 165-172.
- Narayanan, A., & Shmatikov, V. (2006). How to break anonymity of the Netflix prize dataset. arXiv preprint cs/0610105.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems 30 (NIPS 2017)*.
- Voita, E., Talbot, D., Moiseev, F., Sennrich, R., & Titov, I. (2019). Analyzing multi-head self-attention: Specialized heads do the heavy lifting, the rest can be pruned. arXiv preprint arXiv:1905.09418.
- Zheng, L., Noroozi, V., & Yu, P. S. (2017). Joint deep modeling of users and items using reviews for recommendation. *Proceedings of the tenth ACM International Conference on Web Search and Data Mining*, 425-434.

Abstract

A Generalized Adaptive Deep Latent Factor Recommendation Model

Kim, Jeongha* · Lee, Jipyong* · Jang, Seonghyun* · Cho, Yoonho**

Collaborative Filtering, a representative recommendation system methodology, consists of two approaches: neighbor methods and latent factor models. Among these, the latent factor model using matrix factorization decomposes the user-item interaction matrix into two lower-dimensional rectangular matrices, predicting the item's rating through the product of these matrices. Due to the factor vectors inferred from rating patterns capturing user and item characteristics, this method is superior in scalability, accuracy, and flexibility compared to neighbor-based methods. However, it has a fundamental drawback: the need to reflect the diversity of preferences of different individuals for items with no ratings. This limitation leads to repetitive and inaccurate recommendations. The Adaptive Deep Latent Factor Model (ADLFM) was developed to address this issue. This model adaptively learns the preferences for each item by using the item description, which provides a detailed summary and explanation of the item. ADLFM takes in item description as input, calculates latent vectors of the user and item, and presents a method that can reflect personal diversity using an attention score. However, due to the requirement of a dataset that includes item descriptions, the domain that can apply ADLFM is limited, resulting in generalization limitations. This study proposes a Generalized Adaptive Deep Latent Factor Recommendation Model, G-ADLFRM, to improve the limitations of ADLFM. Firstly, we use item ID, commonly used in recommendation systems, as input instead of the item description. Additionally, we apply improved deep learning model structures such as Self-Attention, Multi-head Attention, and Multi-Conv1D. We conducted experiments on various datasets with input and model structure changes. The results showed that when only the input was changed, MAE increased slightly compared to ADLFM due to accompanying information loss, resulting in decreased recommendation performance. However, the average learning speed per epoch significantly improved as the amount of information to be processed decreased. When both the input and the model structure were

* Department of AI, Big Data & Management, Kookmin University

** Corresponding Author: Cho, Yoonho

Department of AI, Big Data & Management, Kookmin University

77 Jeongneung-ro, Seongbuk-gu, Seoul 02707, Korea

Tel: +82-2-910-4950, Fax: +82-2-910-5209, E-mail: www4u@kookmin.ac.kr

changed, the best-performing Multi-Conv1d structure showed similar performance to ADLFM, sufficiently counteracting the information loss caused by the input change. We conclude that G-ADLFRM is a new, lightweight, and generalizable model that maintains the performance of the existing ADLFM while enabling fast learning and inference.

Key Words : Recommender System, Deep Learning, Collaborative Filtering, Deep Latent Factor, Attention

Received : February 15, 2023 Revised : February 23, 2023 Accepted : February 27, 2023

Corresponding Author : Cho, Yoonho

저자 소개



김정하

현재 국민대학교 경영대학 AI빅데이터융합경영학과 학사과정에 재학 중이다. 주 연구 관심 분야는 빅데이터 분석, 딥러닝, 추천 시스템 등이다.



이지평

현재 국민대학교 경영대학 AI빅데이터융합경영학과 학사과정에 재학 중이다. 주 연구 관심 분야는 빅데이터 분석, 머신러닝, 추천시스템 등이다.



장성현

현재 국민대학교 경영대학 AI빅데이터융합경영학과 학사과정에 재학 중이다. 주 연구 관심 분야는 빅데이터 분석, 머신러닝, 추천시스템 등이다.



조윤희

현재 국민대학교 AI빅데이터융합경영학과 교수로 재직 중이다. 서울대학교 계산통계학과를 졸업하고, KAIST 경영정보공학과에서 석사, KAIST 경영공학과에서 박사학위를 취득하였으며, LG전자(주)에서 6년간 주임연구원으로 재직하였다. 주 연구분야는 추천 시스템, 머신러닝, 딥러닝, 빅데이터분석, 디지털마케팅 등이다.