# 정확성-다양성 딜레마 완화를 위한 손실 함수 가중치 재조정 전략

한광석, 권원빈, 유환조*

포항공과대학교

{gshan, kwb4453, hwanjoyu}@postech.ac.kr

# Loss function Re-weight Strategy for alleviating the accuracy-diversity dilemma

GwangSeok Han, WonBin Kweon, HwanJo Yu*

POSTECH

요 약

본 논문은 추천 시스템이 가지는 편향 증폭(bias amplification) 문제를 완화하기 위해, 추천 결과를 다양화하는 전략으로 손실 함수에 페널티를 부여하는 방안을 제안한다. 전통적인 추천 시스템은 정확도를 위해 사용자가 많이 소비한 카테고리에 속하는 아이템을 과도하게 추천하는 경향이 있다. 편향 증폭이라고 불리는 해당 문제는 사용자가 특정 정보만 접하게 되는 필터 버블 현상을 발생시킨다. 확증편향을 유발하는 필터 버블을 해결하는 효과적인 방법 중 하나는 추천의 결과를 다양하게 만드는 것이다. 일반적으로 다양성을 증가시키면 정확도가 하락하는 상충 관계(trade-off)가 존재한다. 본 연구는 이를 완화하기 위해 손실 함수에 가중치 재조정 전략을 적용한다. 소비된 아이템 카테고리 간 데이터 불균형 문제를 해결함으로써, 정확도의 희생 없이 추천 결과의 다양성을 증가시킨다. 2개의 벤치마크 데이터 세트에 대한 실험을 통해 가중치 재조정 전략을 사용하면 정확도 및 다양성이 모두 향상됨을 확인하였다.

## Ⅰ. 서 론

정보 과부하 시대에 추천 시스템은 온라인 서비스에서 필수적인 요소이다. 추천 시스템은 과거 사용자 행동(예: 클릭, 평점)과 아이템의 특성(예: 카테고리)에 기반해 사용자의 선호도를 추론한다. 이때, 사용자의 과거 데이터에는 소비된 아이템 카테고리 간 불균형 문제가 존재한다. 따라서 정확도만 고려하는 추천 시스템은 편향 증폭 현상이 발생하기 쉽다. 게다가 피드백 루프로 인해 시간이 흐름에 따라 편향 증폭이 더욱 심화되어, 사용자 관심사가 좁아지고 필터 버블과 같은 문제가 발생한다 [1]. 장기적으로 필터 버블은 사용자의 만족도를 낮추고 아이템의 독창성을 줄여 전체적인 생태계에 악영향을 끼치기 때문에, 이를 해결하는 것은 중요하다. 사용자의 만족도를 향상하는 데 중요한 역할을 하는 추천의 다양화는 필터 버블을 완화하는 방법으로 주목받고 있다 [2].

추천의 다양화 기법은 2가지 유형으로 발전되어 왔다. 초기에는 후처리(post processing) 방식으로 후보군(candidates)을 생성한 후 다양성을 부여했다. 다양성을 부여하는 모듈은 후보군을 생성하는 과정과 독립적이기 때문에, 다양성 신호가 정확도 향상 프로세스에 반영되지 않아 최종 추천 결과가 차선책(sub optimal)이 될 위험이 커진다 [3]. 그리고 추론할 때 다양성을 고려하기 위한 추가적인 계산 비용이 필요하다. 이에 맞서 다양성과 정확성을 동시에 고려하는 엔드투엔드(end-to-end) 방식이 등장하였다. 하지만, 사용자의 행동 특성을 고려하지 않고 동일한 아이템에 동일한 페널티를 부여하여 추천의 결과가 비슷해지는 문제가 발생한다. 예를 들어, 평균적으로 인기가 많은 액션(action) 게임을 모든 사용자에게 추천이 적게 되도록, 인기가 적은 레이싱(racing) 게임을 모든 사용자에게 추천이 많이 되도록 만든다. 이는 개인화된 추천을 제공한다는 목표에 어긋나며 사용자의 만족도와 추천의 정확도를 해치는 것으로 널리 알려져 있다 [4].

앞선 문제들을 해결하기 위해 사용자의 카테고리 선호도를 기반으로 각 사용자에게 다양성을 부여하여 정확성-다양성 딜레마를 해결하는 방법이 등장했다 [2, 5]. 하지만 모델 구조가 복잡해지거나 사전 학습 과정이 추가되는 등 학습 과정이 복잡해지고 학습 시 더 많은 시간이 필요해졌다. 본 논문에서는 간단하지만, 효과적인 카테고리 선호도에 기반한 손실 함수 가중치 재조정 전략을 제시해 정확성-다양성 딜레마를 해결한다. 2개의 벤치마크 데이터 세트에 대한 실험 결과는 제안된 방법의 효과를 입증해 준다.

## Ⅱ. 본론

사용자의 불균형한 카테고리 소비 분포로 인해 학습 데이터에서 카테고리 간 편향이 발생한다. 이는 학습 과정에서 그라디언트(gradient) 편향을 초래하고 그 결과 인기 있는 카테고리가 많이 추천되는 편향 증폭 현상이 발생하게 된다. 본 논문에서는 클래스 불균형 문제를 효과적으로 해결하는 Focal 손실 함수[5]에 영감을 받아, 예측하기 어려운 아이템에 더 큰 가중치를 부여하고 예측하기 쉬운 아이템에 더 작은 가중치를 부여하는 재조정 전략을 제안한다. 예를 들어, 많이 소비한 액션에 속하는 게임을 사용자가 소비할 것으로 예측하기는 쉽다. 반면에 적게 소비한 레이싱에 속하는 게임을 소비할 것으로 예측하기는 어렵다. 유사하게 많이 소비한 액션에 속하는 게임을 사용자가 선택하지 않을 것으로 예측하기는 어렵고, 적게 소비한 레이싱 게임을 선택하지 않을 것으로 예측하기는 쉽다. 이 점을 고려해 각 경우에 대해 서로 다른 가중치 재조정 전략을 채택한다.

소비한 적이 있는 아이템(positive samples)의 경우, 적게 소비한 카테고리에 속한 아이템의 손실값에 높은 가중치를 부여하고, 많이 소비한 카테고리에 속한 아이템의 손실값에 낮은 가중치를 부여한다. 구체적으로 $C_u = [c_{u1}, c_{u2}, \cdots, c_{uK}]$를 훈련 세트에서 사용자 $u$가 소비한 카테고리 비율을 라고 하자. 카테고리 소비 비율에 반비례한 가중치를 부여하기 위해, $C_u' = [1-c_{u1}, \cdots, 1-c_{uK}] = [c_{u1}', \cdots, c_{uK}']$와 같이 1에서 각 카테고리 소비 비율을 뺀 가중치를 부여한다. 이때, 단순히 $C_u'$를 손실 함수 가

중치로 사용하면 몇 가지 문제점이 발생한다. 지나치게 큰 또는 작은 $c'_{uk}$를 부여하면 추천의 정확도가 떨어지고 모델 수렴에 도움이 되지 않는다. 따라서 선형 정규화(linear normalization)로 가중치를 적절한 구간 $[w_{\min}, w_{\max}]$에 속하도록 하여 문제를 해결한다.

$$c_{uk}^{pos} = w_{\min} + (w_{\max} - w_{\min}) \times \frac{c'_{uk} - \min(C'_u)}{\max(C'_u) - \min(C'_u)} \quad \cdots (1)$$

반대로, 소비한 적이 없는 아이템(negative samples)의 경우 많이 소비한 카테고리에 속한 아이템의 손실값에 높은 가중치를 부여하고 적게 소비한 카테고리에 속한 아이템의 손실값에 낮은 가중치를 부여한다.

$$c_{uk}^{neg} = w_{\min} + (w_{\max} - w_{\min}) \times \frac{c_{uk} - \min(C_u)}{\max(C_u) - \min(C_u)} \quad \cdots (2)$$

사용자의 집합을 $U$, 사용자가 소비한 아이템 집합을 $O_u^+$, 소비하지 않은 아이템 집합을 $O_u^-$ 라고 하자. 훈련과정에서 최종 손실 함수는 다음과 같이 정의된다.

$$loss_{re} = -\frac{1}{N}\left(\sum_{u \in U}\left(\sum_{(i,k) \in O_u^+} c_{uk}^{pos} \times loss(u,i) + \sum_{(j,k) \in O_u^-} c_{uk}^{neg} \times loss(u,j)\right)\right) \cdots (3)$$

본 논문에서는 벤치마크 데이터 세트인 ML-1M과 Steam Game을 이용하여 실험을 진행하였다. 각 데이터 세트의 통계량은 표 1과 같고, 6:2:2 비율로 학습, 검증, 테스트 데이터로 사용하였다.

표 1 데이터 세트 통계

| | 사용자 수 | 아이템 수 | 카테고리 수 | 상호작용 수 |
|---|---|---|---|---|
| ML-1M | 5,148 | 2,380 | 18 | 553,277 |
| Game | 31,572 | 6,070 | 12 | 1,351,659 |

기본 모델로 MultVAE [7], DiffRec [8]을 사용하였으며, 손실 함수로는 각 논문에서 제안된 것을 사용하였다. 다양성을 고려하는 DCRS [2]와 MultVAE., DiffRec에 가중치 재조정 전략을 적용한 MultVAE_Re, DiffRec_Re을 비교하여 그 효과를 검증하였다. 정확도 측정 지표로는 NDCG(NDCG@K)를 사용했고 다양화 측정 지표로는 Entropy(Etp@K)를 사용했다. NDCG는 추천된 결과 내 정답 아이템이 얼마나 높은 순위에 있는지를 고려한 지표이고, Entropy는 추천된 결과 내 얼마나 다양하게 카테고리가 분포되어 있는지를 나타내는 지표이다. 표 2는 MultVAE, DiffRec, DCRS 그리고 각 모델에 가중치 재조정 전략을 적용한 MultVAE_Re, DiffRec_Re의 실험 결과이다. 재조정 전략에서는 $w_{\min}, w_{\max}$ 외에 모든 하이퍼파라미터를 동일한 값으로 고정하였다. 실험 결과, 카테고리 선호도에 기반해 각 사용자에게 적합하도록 손실 함수를 재조정하여 NCDG와 Etp 값이 모두 향상되었다. 즉, 각 사용자에게 적합한 페널티를 적용하여 정확성-다양성 딜레마가 완화되었다. 더불어 그림 1을 통해 평균적으로 사용자들이 많이 소비한 액션 게임을 과도하게 추천하는 편향 증폭 현상이 줄어들었고, 캐주얼(casual), 레이싱 등 적게 소비한 카테고리에 속하는 게임이 더 많이 추천되고 있는 것을 살펴볼 수 있다. 이와 같은 실험 결과를 통해 가중치 재조정 전략이 간단하지만, 효과적으로 편향 증폭과 정확성-다양성 딜레마를 완화한 것을 확인할 수 있다.

표 2 기본 모델과 성능 비교

| | ML-1M | | Steam Game | |
|---|---|---|---|---|
| | NDCG@10 | Etp@10 | NDCG@10 | Etp@10 |
| MultVAE | 0.0884 | 2.5611 | 0.0519 | 2.2931 |
| DiffRec | 0.0963 | 2.6819 | 0.0501 | 2.3871 |
| DCRS | 0.0931 | 2.6800 | 0.0520 | 2.2943 |
| MultVAE_Re | 0.0977 | **3.0908** | **0.0535** | 2.5340 |
| DiffRec_Re | **0.0982** | 2.9008 | 0.0515 | **2.5955** |

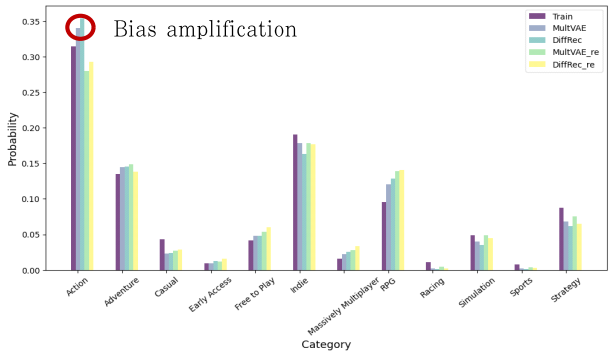


그림 2 Game 학습 데이터 세트와 top-20 추천의 카테고리 분포

## Ⅲ. 결론

본 논문은 추천 시스템의 편향 증폭 문제를 해결하는 것을 목표로 한다. 이를 위해 소비된 카테고리 간 불균형 문제를 해결하는 손실 함수 재조정 전략을 적용하였고, 다양한 실험을 통해 제시된 방법의 효과를 검증하였다. 그 결과 편향 증폭 문제뿐만 아니라 정확성-다양성 딜레마를 완화하여 정확성과 다양성 모두 성능이 향상되었다.

## 참 고 문 헌


[1] Wenjie Wang, Fuli Feng, Xiangnan He, Xiang Wang, and Tat-Seng Chua. 2021. Deconfounded recommendation for alleviating bias amplification. In KDD.

[2] Xiaoying Zhang, Hongning Wang, and Hang Li. 2023. Disentangled Representation for Diversified Recommendations. In WSDM.

[3] Yu Zheng, Chen Gao, Liang Chen, Depeng Jin, and Yong Li. 2021. DGCN: Diversified recommendation with graph convolutional networks. In WWW.

[4] Wentao Ning, Reynold Cheng, Xiao Yan, Ben Kao, Nan Huo, Nur Al Hasan Haldar, Bo Tang. 2024. Diffusion Recommender Model. In WWW.

[5] Yuanyi Ren, Hang Ni, Yingxue Zhang, Xi Wang, Guojie Song, Dong Li, and Jianye Hao. 2023. Dual-Process Graph Neural Network for Diversified Recommendation. In CIKM

[6] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, Piotr Dollar. 2017. Focal Loss for Dense Object Detection. In CVRP.

[7] Dawen Liang, Rahul G. Krishnan, Matthew D. Hoffman, Tony Jebara. 2018. Variational Autoencoders for Collaborative Filtering. In WWW.

[8] Wenjie Wang, Yiyan Xu, Fuli Feng, Xinyu Lin, Xiangnan He, Tat-Seng Chua. 2023. Diffusion Recommender Model. In SIGIR..