

Neural Collaborative Filtering*

<https://arxiv.org/abs/1708.05031>

0. Introduction

- 추천 시스템에서 널리 쓰이는 행렬 분해(Matrix Factorization)는 내적(inner product) 기반이라 표현력이 제한적
- 딥러닝을 활용해 사용자의 선호 패턴을 더 복잡하게 학습할 수 있는 가능성 제시
- 기존 MF 한계를 극복하고 더 일반화된 프레임워크로 신경망 기반 협업 필터링(Neural Collaborative Filtering, NCF) 제안
- 핵심 기여
 - 일반화 가능한 협업 필터링 신경망 프레임워크 제시
 - GMF(Generalized Matrix Factorization)와 MLP 결합한 NeuMF 모델 제안
 - 다양한 벤치마크에서 기존 MF 대비 개선된 성능 검증

1. Overview

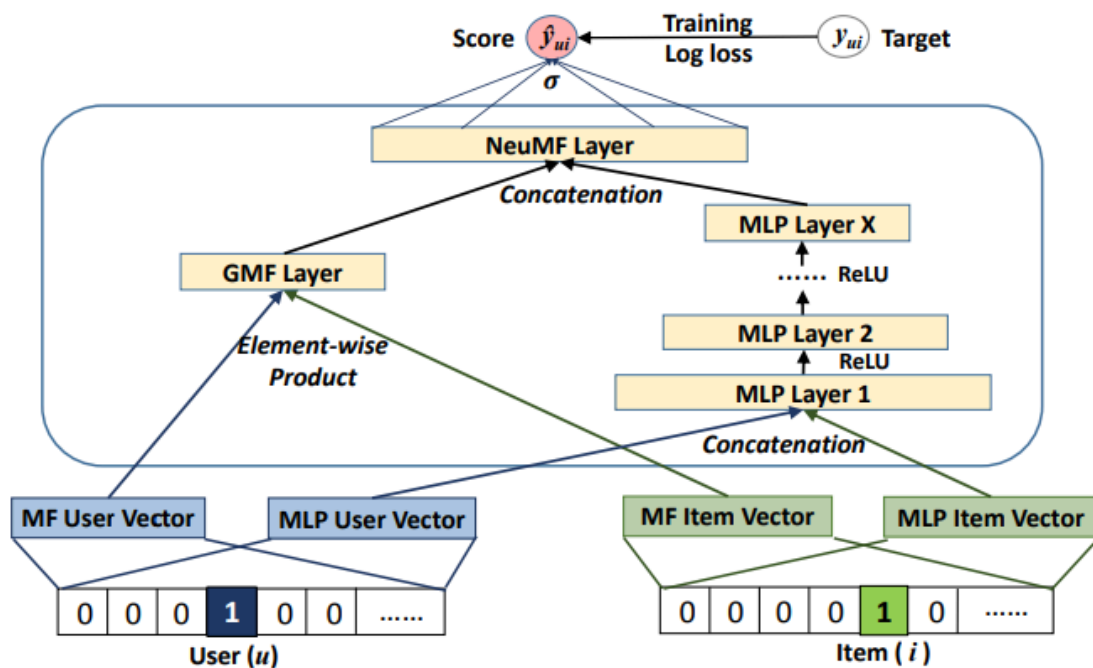
- 내적 기반 MF 대신 신경망을 활용해 user-item 상호작용 함수를 학습
- 기존 MF는 선형적이지만 NCF는 다층 신경망(MLP)을 통해 복잡한 비선형 함수를 학습 가능
- 구조적 변형 :
 - GMF (Matrix Factorization을 일반화한 선형 모델)
 - MLP (user/item latent vector를 concat 후 비선형 변환)
 - NeuMF (GMF + MLP 결합 모델, 두 representation을 합쳐 최종 예측)

2. Challenges

- 전통적 MF는 효율적이지만, 복잡한 사용자-아이템 상호작용을 충분히 포착하지 못함

- 신경망을 도입하면 성능 향상 가능하지만 계산 비용 증가 문제 발생
- 학습 데이터가 sparse한 추천 환경에서 overfitting 가능성 큼
- 다양한 모델(선형/비선형)을 어떻게 균형 있게 결합할지가 중요한 설계 문제

3. Method



- 입력 : 사용자 ID, 아이템 ID를 각각 embedding vector로 매핑
- GMF branch : user embedding과 item embedding의 element-wise product 계산
- MLP branch : user embedding과 item embedding을 concat하여 다층 퍼셉트론에 입력
- NeuMF : GMF와 MLP의 출력을 합쳐 최종 prediction layer로 전달 → sigmoid 활성화 함수로 확률값 출력
- 학습 : binary cross-entropy loss, Adam optimizer 사용
- Negative sampling 활용 : explicit feedback 대신 implicit feedback 데이터 기반 학습

4. Experiments

Table 1: Statistics of the evaluation datasets.

Dataset	Interaction#	Item#	User#	Sparsity
MovieLens	1,000,209	3,706	6,040	95.53%
Pinterest	1,500,809	9,916	55,187	99.73%

- 데이터셋 : MovieLens, Pinterest dataset 등 implicit feedback 기반 추천 데이터
- 평가 지표 : Hit Ratio (HR@10), Normalized Discounted Cumulative Gain (NDCG@10)
- 비교 모델 : MF, ItemPop, BPR, eALS 등 전통적 추천 알고리즘
- 실험 항목
 - GMF vs MLP vs NeuMF 성능 비교
 - embedding 차원, hidden layer 깊이 등 하이퍼파라미터 영향 분석
 - pre-training 전략(별도로 학습한 GMF/MLP 파라미터 활용)이 성능에 미치는 영향

5. Results

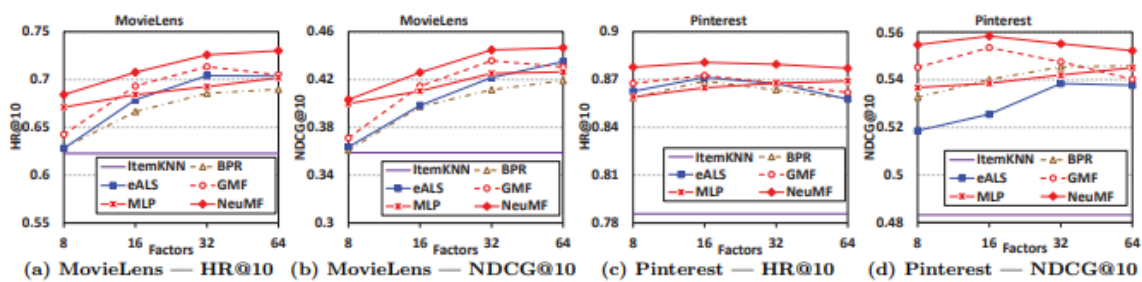


Figure 4: Performance of HR@10 and NDCG@10 w.r.t. the number of predictive factors on the two datasets.

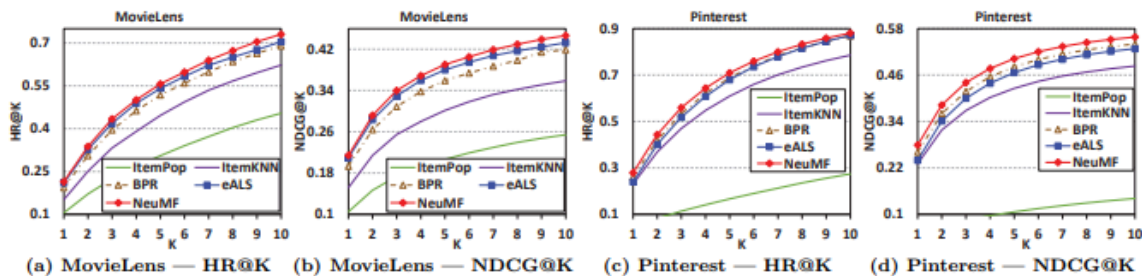


Figure 5: Evaluation of Top-K item recommendation where K ranges from 1 to 10 on the two datasets.

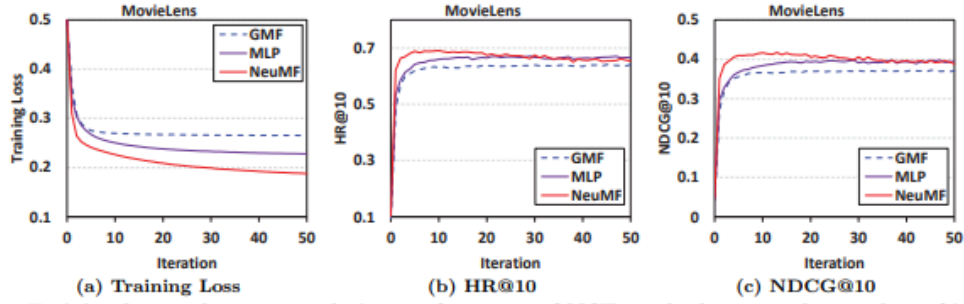


Figure 6: Training loss and recommendation performance of NCF methods *w.r.t.* the number of iterations on MovieLens (factors=8).

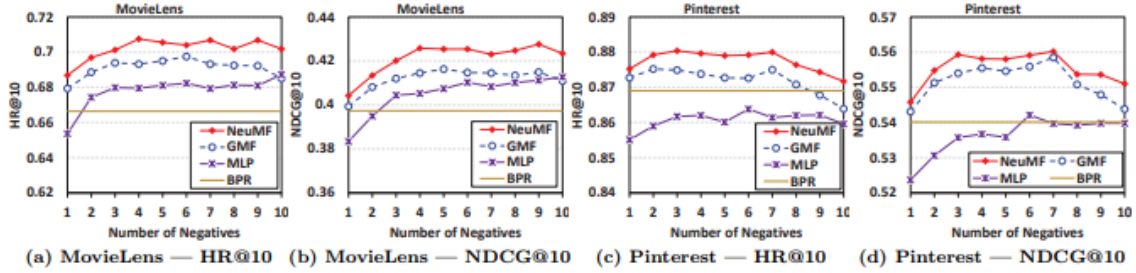


Figure 7: Performance of NCF methods *w.r.t.* the number of negative samples per positive instance (factors=16). The performance of BPR is also shown, which samples only one negative instance to pair with a positive instance for learning.

Table 3: HR@10 of MLP with different layers.

Factors	MLP-0	MLP-1	MLP-2	MLP-3	MLP-4
MovieLens					
8	0.452	0.628	0.655	0.671	0.678
16	0.454	0.663	0.674	0.684	0.690
32	0.453	0.682	0.687	0.692	0.699
64	0.453	0.687	0.696	0.702	0.707
Pinterest					
8	0.275	0.848	0.855	0.859	0.862
16	0.274	0.855	0.861	0.865	0.867
32	0.273	0.861	0.863	0.868	0.867
64	0.274	0.864	0.867	0.869	0.873

- NeuMF는 GMF와 MLP보다 consistently 높은 HR, NDCG 달성
- MLP branch만 사용해도 MF보다 성능 개선, 비선형 학습의 효과 확인
- Pre-training된 GMF와 MLP를 결합한 NeuMF가 가장 높은 성능
- embedding 차원을 늘릴수록 성능 향상되지만 계산 비용 증가
- MovieLens, Pinterest 모두에서 NeuMF가 SOTA 달성

6. Insight

- NCF는 단순한 내적 기반 MF를 신경망 프레임워크로 확장해 추천 시스템 연구에 큰 영향을 줌

- GMF와 MLP를 통합한 NeuMF 구조는 선형·비선형 상호작용을 모두 학습 가능
- 실무적으로는 implicit feedback 데이터에서 효과적인 추천 성능 제공
- 다만 대규모 데이터에서 계산량과 학습 비용이 크다는 단점 존재