

LightGCN : Simplifying and Powering Graph Convolution Network for Recommendation

<https://arxiv.org/pdf/2002.02126>

0. Introduction

- 그래프 기반 협업 필터링은 사용자–아이템 관계 잘 모델링 가능함
- 기존 GCN 기반 추천 모델은 추천 문제에 불필요한 연산 많아 구조 복잡하고 계산 비용 큼
- 본 논문에서는 이러한 요소 제거하고 추천에 맞게 단순화한 모델 LightGCN 제안함

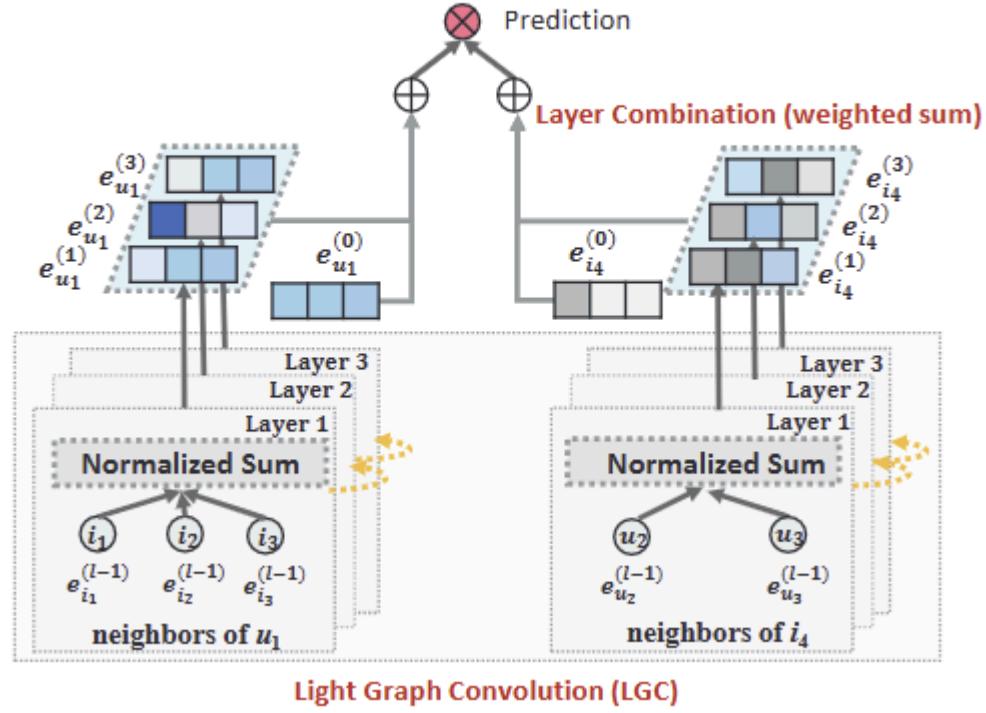
1. Overview

- 기존 GCN 추천 모델은 feature transformation, activation function 포함함
- 추천 문제에서는 이러한 요소가 핵심 역할 아님을 분석함
- LightGCN은 해당 요소 제거하고 이웃 노드 임베딩 전달만 수행하는 단순 구조 사용함
- 이를 통해 성능과 효율 동시에 개선함

2. Challenges

- 불필요한 변환 연산으로 계산 비용 증가 문제 존재함
- 모델 복잡성으로 학습 불안정 발생함
- 레이어 깊어질수록 성능 저하 발생함
- 대규모 추천 환경에서 확장성 떨어지는 문제 있음
- 추천 문제에 맞는 단순 구조 필요함

3. Method



- 사용자-아이템 관계 bipartite graph로 구성함
- 각 노드는 초기 임베딩만 유지함
- 레이어마다 이웃 임베딩 평균 방식으로 전달함
- transformation 및 activation 함수 제거함
- 각 레이어 임베딩을 가중 합하여 최종 표현 생성함
- BPR loss 기반으로 추천 모델 학습함

4. Experiments

Table 2: Statistics of the experimented data.

Dataset	User #	Item #	Interaction #	Density
Gowalla	29,858	40,981	1,027,370	0.00084
Yelp2018	31,668	38,048	1,561,406	0.00130
Amazon-Book	52,643	91,599	2,984,108	0.00062

- Gowalla, Yelp, Amazon 등 공개 데이터셋 사용해 실험 진행함
- 기존 Matrix Factorization 및 GCN 기반 모델과 성능 비교 수행함
- Recall, NDCG 지표 기준 성능 평가함
- 레이어 수 및 임베딩 결합 방식에 따른 성능 변화 분석함

5. Results

Table 3: Performance comparison between NGCF and LightGCN at different layers.

Dataset		Gowalla		Yelp2018		Amazon-Book	
Layer #	Method	recall	ndeg	recall	ndeg	recall	ndeg
1 Layer	NGCF	0.1556	0.1315	0.0543	0.0442	0.0313	0.0241
	LightGCN	0.1755(+12.79%)	0.1492(+13.46%)	0.0631(+16.20%)	0.0515(+16.51%)	0.0384(+22.68%)	0.0298(+23.65%)
2 Layers	NGCF	0.1547	0.1307	0.0566	0.0465	0.0330	0.0254
	LightGCN	0.1777(+14.84%)	0.1524(+16.60%)	0.0622(+9.89%)	0.0504(+8.38%)	0.0411(+24.54%)	0.0315(+24.02%)
3 Layers	NGCF	0.1569	0.1327	0.0579	0.0477	0.0337	0.0261
	LightGCN	0.1823(+16.19%)	0.1555(+17.18%)	0.0639(+10.38%)	0.0525(+10.06%)	0.0410(+21.66%)	0.0318(+21.84%)
4 Layers	NGCF	0.1570	0.1327	0.0566	0.0461	0.0344	0.0263
	LightGCN	0.1830(+16.56%)	0.1550(+16.80%)	0.0649(+14.58%)	0.0530(+15.02%)	0.0406(+17.92%)	0.0313(+18.92%)

*The scores of NGCF on Gowalla and Amazon-Book are directly copied from Table 3 of the NGCF paper (<https://arxiv.org/abs/1905.08108>)

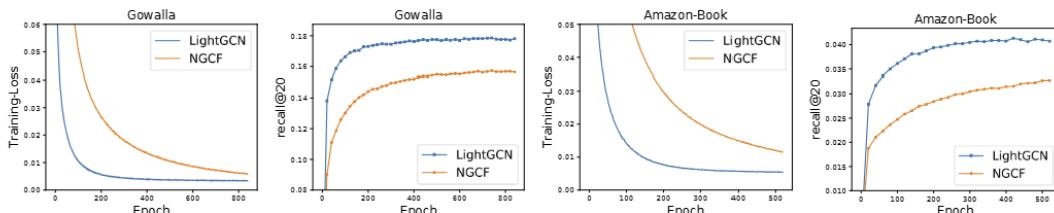


Figure 3: Training curves of LightGCN and NGCF, which are evaluated by training loss and testing recall per 20 epochs on Gowalla and Amazon-Book (results on Yelp2018 show exactly the same trend which are omitted for space).

Table 4: The comparison of overall performance among LightGCN and competing methods.

Dataset	Gowalla		Yelp2018		Amazon-Book	
	recall	ndcg	recall	ndcg	recall	ndcg
NGCF	0.1570	0.1327	0.0579	0.0477	0.0344	0.0263
Mult-VAE	0.1641	0.1335	0.0584	0.0450	0.0407	0.0315
GRMF	0.1477	0.1205	0.0571	0.0462	0.0354	0.0270
GRMF-norm	0.1557	0.1261	0.0561	0.0454	0.0352	0.0269
LightGCN	0.1830	0.1554	0.0649	0.0530	0.0411	0.0315

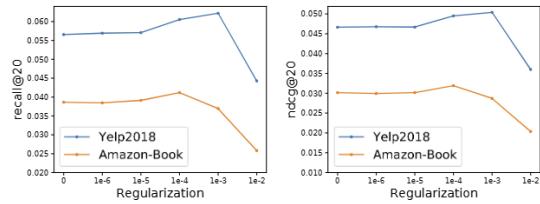


Figure 5: Performance of 2-layer LightGCN w.r.t. different regularization coefficient λ on Yelp and Amazon-Book.

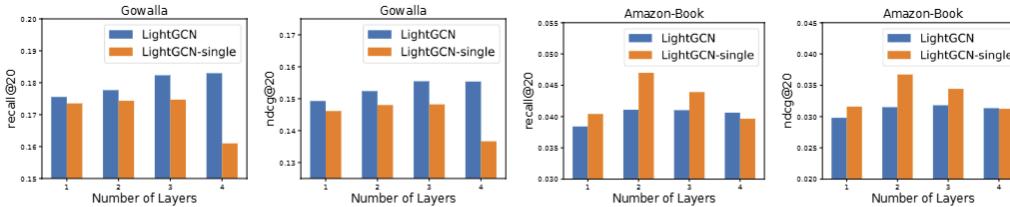


Figure 4: Results of LightGCN and the variant that does not use layer combination (i.e., LightGCN-single) at different layers on Gowalla and Amazon-Book (results on Yelp2018 shows the same trend with Amazon-Book which are omitted for space).

Table 5: Performance of the 3-layer LightGCN with different choices of normalization schemes in graph convolution.

Dataset	Gowalla		Yelp2018		Amazon-Book	
	recall	ndcg	recall	ndcg	recall	ndcg
LightGCN-L ₁ -L	0.1724	0.1414	0.0630	0.0511	0.0419	0.0320
LightGCN-L ₁ -R	0.1578	0.1348	0.0587	0.0477	0.0334	0.0259
LightGCN-L ₁	0.159	0.1319	0.0573	0.0465	0.0361	0.0275
LightGCN-L	0.1589	0.1317	0.0619	0.0509	0.0383	0.0299
LightGCN-R	0.1420	0.1156	0.0521	0.0401	0.0252	0.0196
LightGCN	0.1830	0.1554	0.0649	0.0530	0.0411	0.0315

Table 6: Smoothness loss of the embeddings learned by LightGCN and MF (the lower the smoother).

Dataset	Gowalla	Yelp2018	Amazon-book
	Smoothness of User Embeddings		
MF	15449.3	16258.2	38034.2
LightGCN-single	12872.7	10091.7	32191.1
Smoothness of Item Embeddings			
MF	12106.7	16632.1	28307.9
LightGCN-single	5829.0	6459.8	16866.0

- 기존 GCN 기반 추천 모델 대비 성능 향상 확인됨
- 계산량 감소하고 학습 속도 개선됨
- 깊은 레이어에서도 안정적인 성능 유지함
- 모델 단순화가 오히려 추천 성능 향상으로 이어짐 확인함

6. Insight

- 추천 시스템에서는 복잡한 변환보다 이웃 정보 전달이 핵심 요소임 확인함
- 단순한 구조가 성능과 효율 동시에 개선 가능함 보여줌
- LightGCN은 이후 그래프 추천 모델의 기본 베이스라인으로 자리 잡음