데이터 전처리를 통한 Bert4Rec 성능 개선 : 학습 용량 증가의 해결책

서강대학교 컴퓨터공학과
120240287 장래영

서강대학교 컴퓨터공학과
120240288 홍문기

목차

서론

SASRec vs BERT4Rec

순차적 추천 시스템	사용자 로그로부터 관심사를 추출하고 이를 바탕으로 사용자가 다음에 **선호할만한 항목을 추천**
대표 모델	SASRec 과 BERT4Rec *트랜스포머 기반의 대표적 순차적 추천 모델

SASRec vs BERT4Rec

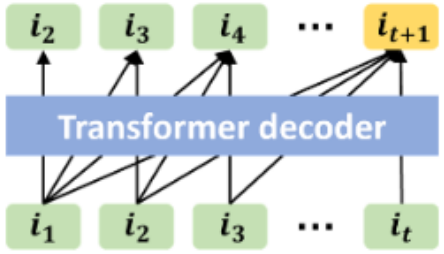
SASRec

- 트랜스포머의 디코더 구조를 활용한 단방향 셀프 어텐션

특정 시점에 소비한 항목과 그 이전에 소비한 항목들과의 상호 의존성 고려

항목 간의 관계를 고려할 때,
과거에서 현재로의
한 가지 방향만을 고려

모델 구조



(a) SASRec

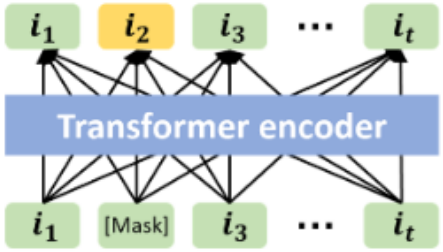
BERT4Rec

- 트랜스포머의 인코더 구조를 활용한 양방향 셀프 언텐션

특정 시점의 항목 & 그 이전과 이후에 소비한 모든 항목들과의 상호 의존성 고려

항목 간의 관계를 고려할 때,
과거에서 현재, 현재에서 과거의
두 방향 모두 고려

모델 구조



(b) BERT4Rec

BERT4Rec은 SASRec보다 더 많은 학습 모수를 가지며, 이에 따라 모델 학습 용량도 증가

본론

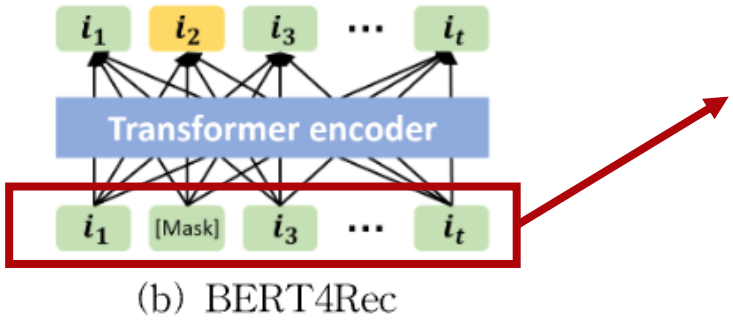
문제 정의

구간별 평균화

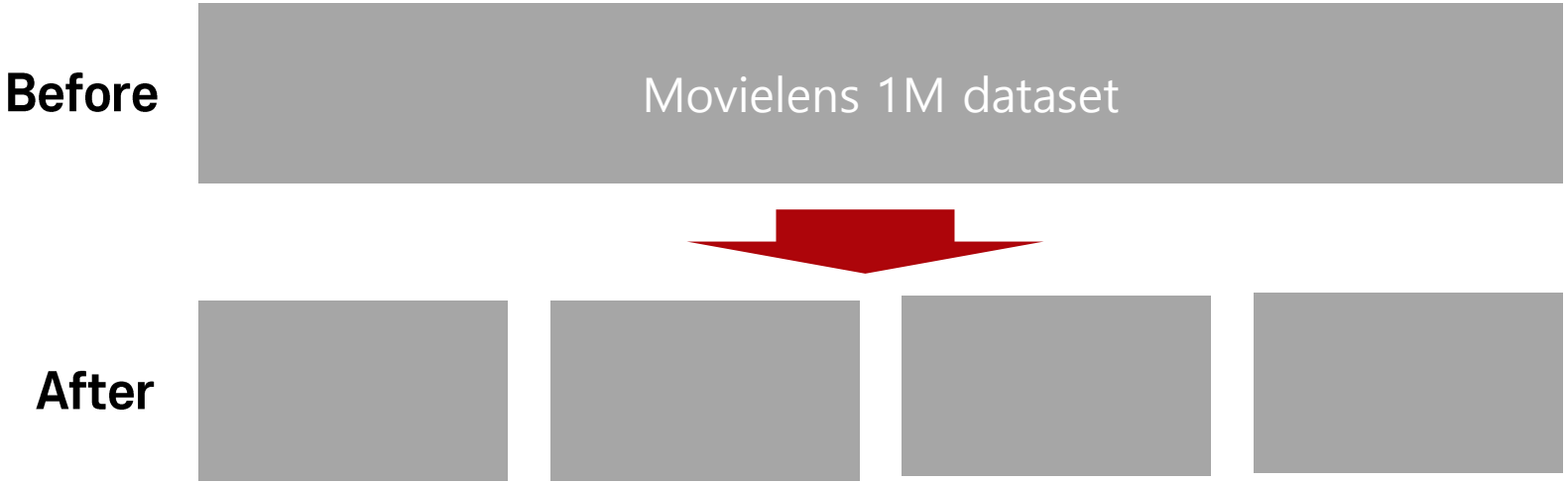
구간을 나누어 평균을 내는 경우

- 입력 시퀀스의 일부를 평균화하여 입력 크기를 줄이고, 모델 처리 속도 향상
- 양방향 모델의 경우 각 구간 내에서의 전체적인 시간적 패턴을 더 정확히 학습

모델 구조



(b) BERT4Rec



Patch Size를 줄이며 모델 처리 속도 향상

구현방식

시퀀스 구성

- 시퀀스 구성

- 각 사용자는 여러 영화를 시청, 이 시청 이력을 시간 순서대로 나열하여 시퀀스 형성

구간 설정

- Dataset 구간 설정
  - 큰 Dataset

큰 구간 고려하여 구간 설정 (ex, 5000,10000)

  - 작은 Dataset

작은 구간 고려하여 구간 설정 (ex, 1000)

평균 계산

- 구간에 대한 평균 계산
  - 각 구간에 대해 timestamp 값의 평균 계산

최종적으로, 구간별 평균을 사용하여 사용자의 시간적 흐름 포착하고, 이를 기반으로 다음 영화 예측

예측

- 구간별 평균 값을 이용하여 영화 예측

- 시퀀스 기반 추천 모델인 Bert4Rec or SASRec 사용

모델 적용

- MovileLens 1M dataset
  - 상대적으로 구간을 크게 나누며 적용

실험

실험 환경

학습 설정

```
config = {
    #'data_path' : "/content/drive/MyDrive/RecsysTutorial/Data/MovieLens" , # 데이터 경로
    'data_path' : pd.read_csv('/content/drive/MyDrive/Data/ratings.csv') , # 데이터 경로

    'max_len' : 50,
    'hidden_units' : 50, # Embedding size
    'num_heads' : 1, # Multi-head layer 의 수 (병렬 처리)
    'num_layers': 2, # block의 개수 (encoder layer의 개수)
    'dropout_rate' : 0.5, # dropout 비율
    'lr' : 0.001,
    'batch_size' : 128,
    'num_epochs' : 50,
    'num_workers' : 2,
    'mask_prob' : 0.15, # for cloze task
}
```

평가 방식

Hit Ratio@K

*추천된 상위 N개의 아이템 중 실제 사용자가 선호하는 아이템이 포함된 비율

NDCG@K

*추천된 상위 N개의 아이템의 관련성을 고려하여 순위매기기를 평가

K = 10 설정

BERT4Rec 성능 분석

MEAN

Datasets	Metric	Bert4Rec(Base)	Ours@1000	Ours@5000	Ours@10000	Improv.
ML-1m	HIT@10	54.3	57.7	59.7	58.1	3.4
	NDCG@10	30.9	32.7	34.4	33.3	1.8

MEDIAN

Datasets	Metric	Bert4Rec(Base)	Ours@1000	Ours@5000	Ours@10000	Improv.
ML-1m	HIT@10	54.3	56.2	57.7	57.5	1.9
	NDCG@10	30.9	32.0	32.7	32.5	1.1

SASRec 성능 분석

MEAN

Datasets	Metric	SASRec(Base)	Ours@1000	Ours@5000	Ours@10000	Improv.
ML-1m	HIT@10	47.8	47.7	47.9	48.0	0.1
	NDCG@10	26.4	26.4	26.4	26.6	0.2

단방향 모델로, 시퀀스의 순서를 엄격하게 따르므로
평균 타임스탬프를 사용하는 것이 성능에 큰 영향을 미치지 않을 수 있음

결론

결론 및 후속 연구 방향

다양한 데이터셋 추가 분석

- Amazon Beauty data
  - 2M dataset
- Amazon BookStore data
  - 3M dataset

논문제출

- 국내학회
  - 인공지능학회 (6월 30일 마감)
- 국외저널
  - IEEE Access

후속 연구

- 모델링을 통한 Bert4Rec 성능 개선

참고논문

(1) Zhankui He, Handong Zhao, Zhe Lin, ZhaowenWang, Ajinkya Kale, and Julian McAuley. Locallyconstrained self-attentive sequential recommendation.CIKM, 2021.

(2) Wang-Cheng Kang and Julian J. McAuley.Self-Attentive Sequential Recommendation. ICDM,2018.

(3) Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu,Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, andJi-Rong Wen. S3-Rec: Self-Supervised Learningfor Sequential Recommendation with MutualInformation Maximization. CIKM, 2020.

(4) Chen Ma, Peng Kang, and Xue Liu. Hierarchicalgating networks for sequential recommendation.KDD, 2019.

(5) Jiaxi Tang and Ke Wang. Personalized top-nsequential recommendation via convolutional sequenceembedding. WSDM, 2018.

(6) Fajie Yuan, Alexandros Karatzoglou, Ioannis Arapakis,Joemon M Jose, and Xiangnan He. A simpleconvolutional generative network for next itemrecommendation. WSDM, 2019.

(7) Tim Donkers, Benedikt Loepp, and Jürgen Ziegler,Sequential User-based Recurrent Neural NetworkRecommendations. RecSys, 2017.

(8) Balázs Hidasi and Alexandros Karatzoglou.Recurrent Neural Networks with Top-k Gains forSession-based Recommendations., CIKM, 2018.

(9) Chao-Yuan Wu, Amr Ahmed, Alex Beutel,Alexander J. Smola, and How Jing. RecurrentRecommender Networks. WSDM, 2017.

(10) Shuqing Bian, Wayne Xin Zhao, Kun Zhou, JingCai, Yancheng He, Cunxiang Yin, and Ji-Rong Wen.Contrastive Curriculum Learning for SequentialUser Behavior Modeling via Data Augmentation.CIKM, 2021.

감사합니다