

[Review]

본 논문에서는 기존의 추천시스템에서 쓰이는 NLP 모델 중 Transformer 등의 단방향 모델은 시퀀스를 왼쪽에서 오른쪽의 순서로 읽어가며 빈칸을 예측하기 때문에 사용자의 행동에 대한 representation이 제한된다고 한다. 따라서 본 논문은 사용자의 행동에 대한 representation을 더 잘 캡처할 수 있도록 Deep한 양방향 Self-Attention 모델로서 BERT를 추천 시스템에 활용하는 것을 제안했다.

본 논문에서 제안한 BERT4Rec은 단어가 들어갈 위치에 아이템을 넣는다. 양방향의 아이템을 통해 맥락을 파악하고, Random 하게 마스킹 처리를 하여 주변 맥락으로 아이템을 예측한다. 이러한 방식을 통해서 Data leakage를 막고 Data augmentation의 효과도 얻을 수 있었다. 모델은 유저가 선택할 수 있는 모든 가능한 아이템들의 확률을 예측한다. 또한 Self-Attention을 통해서 거리가 먼 아이템의 의존성도 반영할 수 있다. 그리고 Mult-head Self-Attention을 활용하여 조금 더 다양한 아이템의 속성을 반영할 수 있게 된다.

그러나 본 논문에서 제시한 BERT4Rec의 경우 추천 시스템에 BERT를 적용하여 좋은 성능을 보였지만 많은 한계를 가진다. 아이템에 순서가 항상 정해져 있는 것은 아니며, 기술적으로 아이템의 시퀀스를 유저별로 저장하기는 어렵다. 또한 도메인에 따라 아이템의 Cycle이 존재한다. 따라서 BERT를 적용할 수 있는 Task가 매우 한정적이라는 단점을 가진다. 또한 BERT의 경우 다른 모델에 비하여 상대적으로 학습 속도가 느리기 때문에 실용적이지 않다.

[본문 URL]

<https://velog.io/@tobigs-recsys/%E3%85%A0BERT4Rec-Sequential-Recommendation-with-Bidirectional-Encoder-Representation-from-Transformer>