

[Review]

추천 시스템에서 모델의 성능을 최대화 하기 위해서는 유저 행동에 숨겨진 복잡한 feature interaction을 학습하는 것이 매우 중요하다. 본 논문에서 제시한 DeepFM 이전의 모델들은 대부분 고차원 또는 저차원의 feature interaction에 치우쳐 학습하거나 전문가의 feature engineering이 필요했다. 그러나 본 논문에서 제시한 DeepFM 모델은 feature engineering 없이 저차원과 고차원의 feature interaction을 End-to-End 모델로 학습할 수 있다. 따라서 DeepFM 모델은 Factorization machines이 가지는 장점과 Deep Learning이 가지는 장점을 결합했다고 볼 수 있다.

DeepFM 모델은 Wide & Deep 모델을 발전시킨 모델로 볼 수 있다. 기존의 Wide & Deep 모델은 Wide 부분을 학습시키기 위해서 feature engineering이 필요했지만 DeepFM의 FM Component와 Deep Component가 같은 Embedding Layer를 공유하기 때문에 feature engineering이 필요 없으며, 가중치 초기화 없이 End-to-End 방식으로 학습이 가능하다.

DeepFM 모델은 FM Component와 Deep Component로 이루어져 있다. FM Component는 저차원의 feature interaction을 학습시키는 부분으로 input에 대한 각각의 feature latent vector의 내적으로 feature interaction을 학습한다. Deep Component는 고차원의 feature interaction을 학습시키는 부분으로 매우 sparse한 input data를 Embedding Layer를 통해서 저차원의 dense한 Vector로 압축하여 feature interaction을 학습한다.

DeepFM 모델은 기존의 CTR예측 SOTA 모델의 단점을 극복하고 더 좋은 성능을 보인 모델이며, pre-training이 필요하지 않고, 고차원과 저차원의 feature interaction을 같이 학습하고, feature embedding을 공유하여 feature engineering이 필요 없도록 했다는 장점을 가진다.

[본문 URL]

<https://velog.io/@tobigs-recsys/Paper-Review-2017-Huifeng-Guo-DeepFM-A-Factorization-Machine-based-Neural-Network-for-CTR>