

[Review]

추천 시스템을 만들기 위해서는 유저의 선호도를 표현할 수 있는 데이터가 필요하다. 대부분의 추천 시스템에서는 이러한 데이터를 유저의 아이템에 대한 평점, CTR 등으로 나타냈다. 그러나 본 논문에서는 이러한 유저의 선호도를 표현하는 데이터를 dwell time이라는 유저가 아이템에 머문 시간으로 나타냈다.

dwell time의 경우 사용기기에 따라서 분포의 차이가 발생하기 때문에 이러한 차이를 제거해주기 위해서 각 사용기기의 log-dwell time의 평균과 표준편차를 구해서 정규화를 해줄 필요가 있다.

그리고 dwell time의 정보가 없을 때는 dwell time을 예측할 수 있는 모델을 만들 수 있다. 기사의 길이, 주제, 카테고리 등을 feature로 활용하여 Support Vector Regression 모델을 구축하면 된다.

dwell time을 이용하여 Machine Learning to Rank 추천 시스템 모델을 구축하기 위한 알고리즘으로 본 논문에서는 크게 Gradient Boosted Decision Tree와 Matrix Factorization을 제시했다.

Gradient Boosted Decision Tree을 이용할 경우에는 목표변수를 각 아이템의 평균 dwell time으로 둘 수 있으며, 각 아이템에 대한 weight를 dwell time으로 활용할 수 있다.

Matrix Factorization을 이용할 경우에는 목표 변수를 user-item dwell time으로 두어 물음표 부분을 추론하는 방식으로 활용할 수 있다.

[본문 URL]

<https://velog.io/@tobigs-recsys/%EC%9E%84%EC%8B%9C-Beyond-Clicks-Dwell-Time-for-Personalization>