

Recommendation System with Implicit Feedback

<http://sanghyukchun.github.io/95/>

Explicit Feedback and Implicit Feedback

Explicit feedback - item에 호감이 있는지를 수치로 나타내는 것 (Netflix평점)

Implicit feedback - item에 대한 간접적인 정보 (상품 조회, 재생목록, 시청한 횟수 등등)

Item에 대한 구체적인 preference를 제공하지 않기 때문에 잠재적으로 흥미가 있는 item 모두 값이 0일 수도 있다.

Implicit feedback 문제

- Item을 굉장히 조금만 click하거나 사용하더라도 구매가 이루어지기 때문에 negative observation이 positive observation의 수를 압도한다. 이 점을 고려하지 않으면 model이 overfitting이 된다.
- 물건을 구매하여도 물건에 대해서 반드시 긍정적으로 생각할 수는 없다.

Collaborative filtering for Implicit Feedback

p_{ui} (preference vector)

- user u 가 item i 를 선호하는지 하지 않는지 여부
- $\text{sign}(r_{ui})$ 으로 정의된다. Input이 negative value이면 -1, positive value면 1
- 항상 신뢰할 수 있는 값이 아니기 때문에 confidence level c_{ui} 을 정의

c_{ui} (confidence level)

-사용자가 한 항목을 엄청 많이 재구매했다면 u 는 i 를 아주 높은 확률로 prefer한다는 사실을 가 정할 수 있다. 따라서 confidence level은 r 에 대한 increasing function으로 정의할 수 있다.

$$c_{ui} = 1 + \alpha r_{ui}$$

-Parameter α 가 positive observation과 negative observation 중요도를 조정하는 이점

(α 의 크기가 작다면 positive와 negative observation의 confidence차이가 큰 α 를 가질 때 보다 상대적으로 작을 것이라는 것을 기대할 수 있게 된다.)

Objective function

$$\min_{X,Y} \sum_{u,i \in \kappa} c_{ui} (p_{ui} - x_u^\top y_i)^2 + \lambda (\|x_u\|_2^2 + \|y_i\|_2^2).$$

Rating vector r 를 Preference vector p 로 변환

Confidence level c 를 정의 p 와 c 를 사용해 RMSE objective function을 optimize

Confidence level c 를 정의하면서 Implicit feedback문제 해결