

3.2 기억 기반 협력 필터링

- 기억 기반 협력 필터링은 User 기반 협력 필터링과 Item 기반 협력 필터링으로 나뉜다.
- (1) User 기반 협력 필터링 : User가 입력한 Rating을 기반으로 해당 User와 유사한 성향을 가지는 User들이 공통으로 선호하는 Item을 추천
- (2) Item 기반 협력 필터링 : 특정 Item을 기준으로 User들이 평가한 Rating이 유사한 Item을 선정하고 Rating을 예측

*** 유사도 측정

- User 혹은 Item 간의 유사도를 계산하는 것은 기억 기반 협력 필터링에서 가장 중요한 단계이다.
- 대표적인 유사도 측정 척도로는 **피어슨 상관계수**, **스피어만 순위상관계수**, **코사인 유사도**가 있다.

*** 선호도 예측

- 선호도의 예측 값을 얻기 위한 기법으로는 **가중합**과 **단순가중평균**이 있다.
- 가중합은 추천 대상 User a가 Item i에 대해 갖는 예측 선호도를 나타내고 아래 식과 같이 계산할 수 있다.

$$P_{a,i} = \bar{r}_a + \frac{\sum_{u \in U} (r_{u,i} - \bar{r}_u) \times w_{a,u}}{\sum_{u \in U} |w_{a,u}|}$$

\bar{r}_a 는 추천 대상 User가 모든 Item에 준 Rating의 평균이며 \bar{r}_u 는 다른 User가 모든 Item에 준 평균 Rating이다.

$w_{a,u}$ 는 추천 대상 User a와 User u 사이의 유사도로 유사할수록 큰 가중치 값을 가진다.

상위 N개 Item을 추천하기 위해서 추천 대상 user와 유사한 User를 K명 선택한 후 각 Item에 대해 구매 횟수를 계산한다.

최종적으로 User가 구매하지 않은 Item 중에서 구매 빈도가 높은 N개 Item을 추천한다.

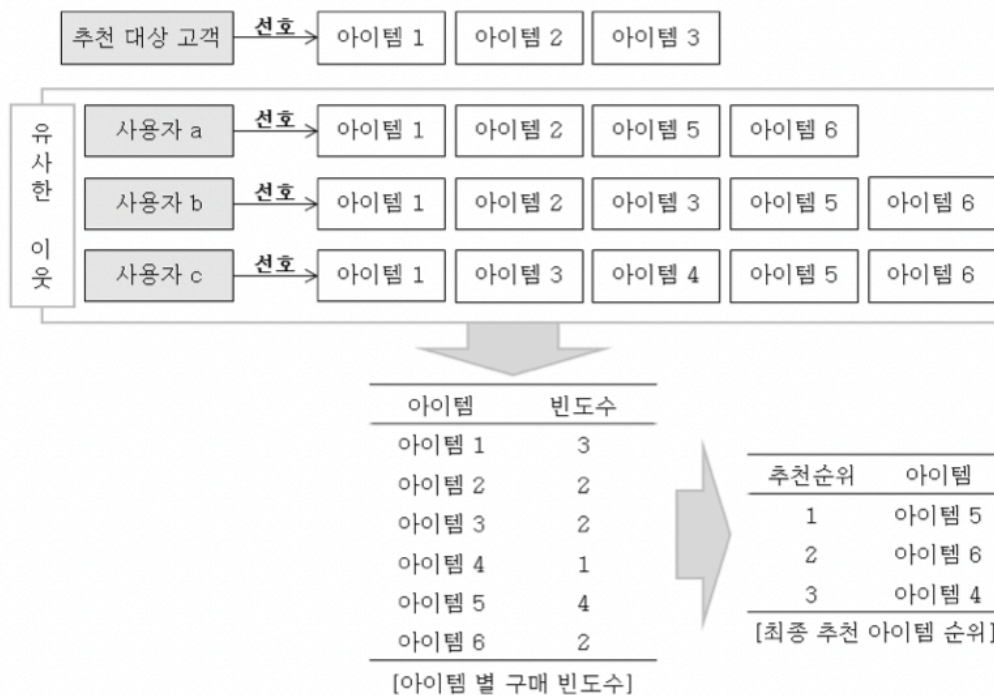
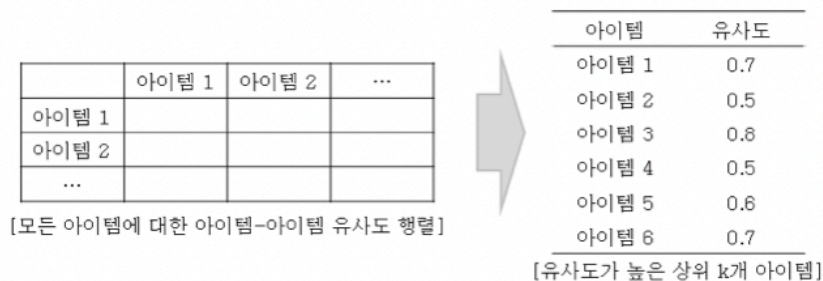


Figure 5. User-based top-n recommendation algorithm

아래 Figure6은 Item 기반 상위 N개 Item 추천 기법을 보여준다. Item - Item 행렬을 만든 후 유사도를 구하는데 이 때, 이미 구매한 Item은 유사도 계산에서 제외된다.



(a) Top-n lists of recommended items



(b) Ranking of recommendation

Figure 6. Item-based top-n recommendation algorithm

Item1, 2, 3은 이미 추천 대상 User가 구매를 했기 때문에 나머지 4,5,6에 대해서 유사도를 계산한다.

3.3 모델 기반 협력 필터링

기억 기반 협력필터링은 모델을 구축하지 않고 추천이 요구될 때 마다 **휴리스틱 기법**을 통해 결과를 도출하는 **Lazy Learning**이다.

실무자가 설정해야 할 파라미터 수가 적은 장점이 있는 반면 Sparsity, Grey Sheep과 같은 문제점이 발생한다.

기존의 기억 기반 협력필터링에 **기계학습과 데이터 마이닝 기법을 적용**하여 여러 문제점을 보완하고 있다.

3.4 차원 축소

협업 필터링에서는 User가 구매한 Item이 너무 적으면 상품 추천이 불가능해지는 문제점이 있다. 또한 User, Item이 많을수록 많은 계산 비용이 발생한다.

이런 문제를 개선하기 위해 정보검색 분야에서 사용되는 Latent Semantic Indexing 기법이 적용되었다.