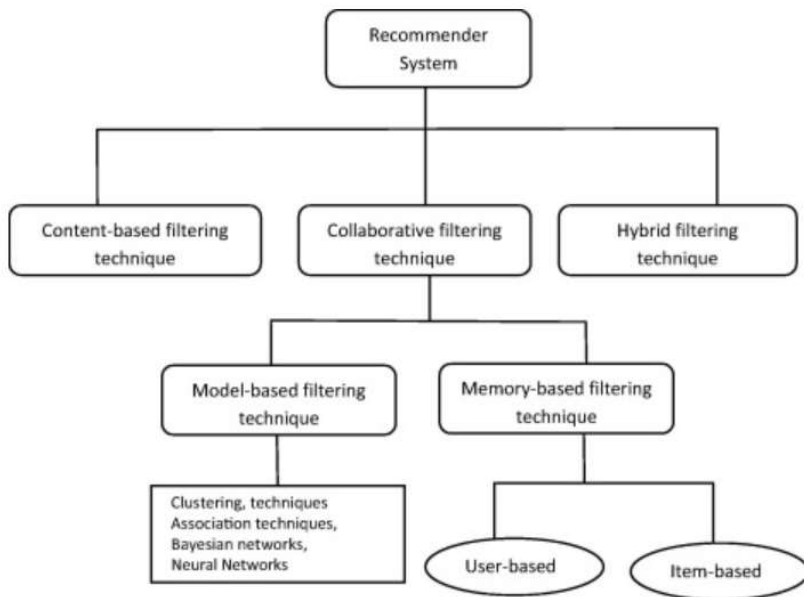


추천시스템

추천 시스템



1. Content-Based Filtering (콘텐츠 기반 필터링)

- INPUT: 사용자 프로필, 제품 속성
- OUTPUT: 특정 아이템과 비교하여 비슷한 best-matching 아이템을 추천
- 예: 영화 아이언맨을 좋아하면 동일 주인공, 장르의 어벤저스도 좋아할 확률이 높다

장점:

추천 가능한 아이템 범위가 넓다 (=새로운 아이템이나 인기 없는 아이템도 추천 가능)

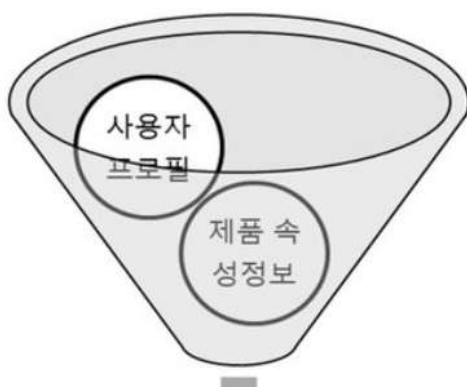
다른 유저 데이터가 필요하지 않다

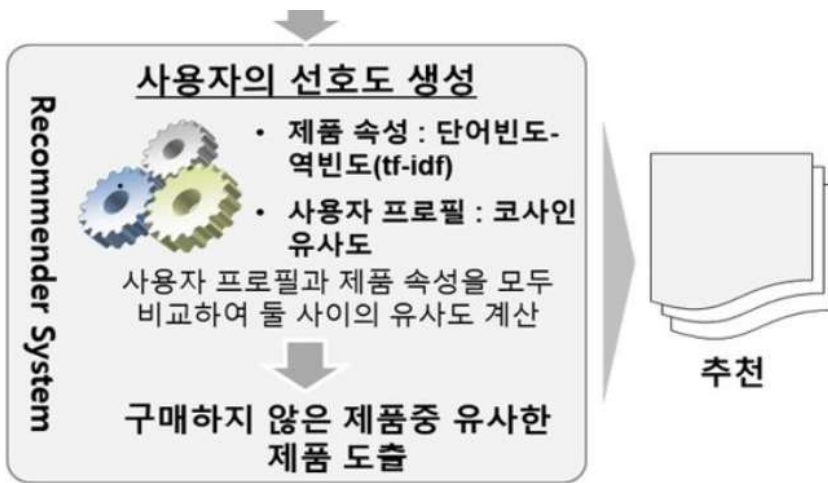
객관적인 제품 속성 기반으로 협업 필터링 대비 정확도 높다

단점:

새로운 유저를 위한 추천이 어렵다

유저의 다양한 취향을 반영하기 어렵다





- 아이템 특징을 끌어내기 위해 tf-idf (tf 단어 빈도 * 1 / df 문서 빈도)나 Word2Vec 사용

2. Collaborative Filtering (협업 필터링)

- INPUT: 사용자의 취향 정보 (구매 이력 및 평가 결과)
- OUTPUT: 유사 선호도를 통해 평가 결과를 기반으로 추천

장점:

제품, 사용자 정보 불필요해 구현 용이 (= 정확하게 아이템 그 자체를 이해하지 않아도 추천 가능)

단점:

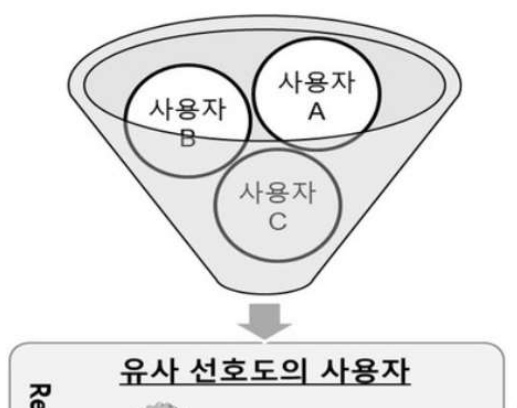
평가 정보 만으로는 정확한 추천 불가

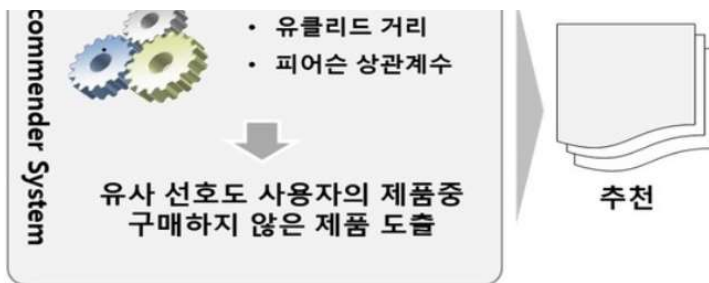
평가되지 않은 아이템은 추천되지 않는다

인기 있는 아이템만 추천하는 경향

1) Memory-based Filtering (이웃 기반 협업 필터링) : user-item matrix 통해 유사도 기반으로 아이템 추천

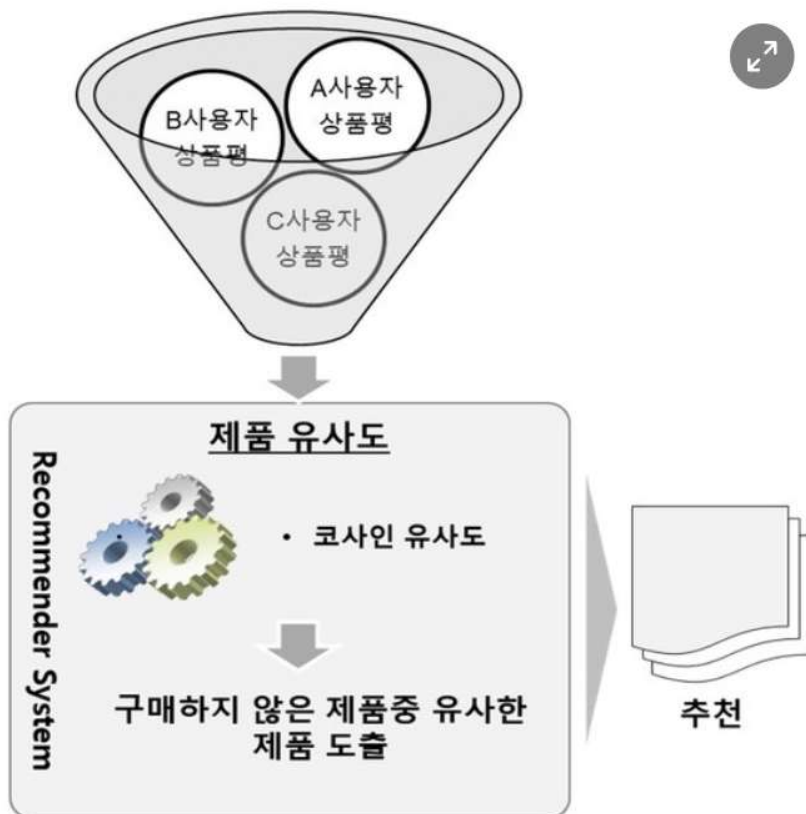
(1) 사용자 기반 협업 필터링





- 유사도 : 두 아이템을 얼마나 유사한 사용자가 선호했는지
- 예 : 맥주를 구매한 당신과 비슷한 사람들은 기저귀도 구입

(2) 아이템 기반 협업 필터링



- 유사도 : 두 사용자가 얼마나 유사한 아이템을 선호했는지
- 예 : 공조와 라라랜드의 유사도가 높다면, 공조를 좋아하는 사람은 라라랜드를 좋아할 확률이 높다

2) Latent based collaborative (잠재 요인 기반의 협업 필터링)

- 사용자-아이템 평점 행렬에 잠재 되어 있는 어떤 요인이 있다고 가정
- 행렬 분해를 기반
- 대규모 다차원 행렬을 SVD와 같은 차원 감소 기법으로 분해하는 과정에서 잠재 요인(latent factor) 찾아내는 방법
- 예 : 영화 관람한 사용자와 영화 정보의 다차원 행렬에서 차원 감소로 영화 장르 라는 잠재 요인 찾아내기

3. Hybrid Filtering (하이브리드 필터링)

- 1) Combining Filtering : 두 가지 알고리즘을 모두 적용하고, 이의 가중 평균을 구하는 기법
- 2) Collaboration via Content : 평점 데이터와 아이템 프로필을 조합하여 사용자 프로필을 만들어 추천하는 기법