* **추천 시스템의 중요성**

콘텐츠 포털, 전자상거래 등 추천 시스템을 통해 **사용자의 취향을 이해하고 맞춤 상품과** 콘텐츠를 제공해 고객을 머무르게 하기 위한 전력을 기울임

* **추천엔진의 필요성**

너무 많은 상품으로 가득 찬 온라인 스토어 / 한정된 시간, 어떤 상품을 골라야 할지 선택의 압박

**VS** 추천엔진은 사용자가 **무엇을 원하는지 빠르게 찾아내** 사용자의 온라인 쇼핑 이용 즐거움을 배가 시킨다.

* **추천 시스템의 묘미**

사용자 **자신도 좋아하는지 모르고 있었던 취향** 발견

추천 시스템에 **신뢰가 높아지면서** 사용자는 추천 아이템을 더 많이 선택하게 되고, 이로 인해 더 **많은 데이터가 추천 시스템에 축적**되며 **추천이 정확해지고 다양**해짐.

* **추천이 지배하는 사회**

정교한 추천 시스템은 사용자에게 **높은 신뢰**도를 얻게 되며, 맹목적으로 **사용자가 의존**하게 만듦.

서비스 제공자는 **고객 충성도를 크게 향상**시킬 수 있음.

* **추천 시스템 방식** (아래의 둘 중 하나를 선택하거나 두가지를 결합하여 hybrid방식으로 사용)

**콘텐츠 기반 필터링 Content Based Filtering**

감독, 배우, 영화 설명, 장르 등 영화를 구성하는 다양한 콘텐츠들을 텍스트 기반 문서 유사도로 비교하여 추천

영화구성 🡪 피처 벡터화 🡪 코사인 유사도 🡪 유사도 및 평점에 따른

콘텐츠 텍스트 (Count, TF-IDF) 영화 추천

**콘텐츠 기반 필터링 구현 프로세스**

1. 콘텐츠에 대한 여러 텍스트 정보들을 **피처 벡터화**
2. **코사인 유사도**로 콘텐츠별 유사도 계산
3. 콘텐츠 별로 **가중 평점을 계산**
4. 유사도가 높은 콘텐츠 중에 **평점이 좋은 콘텐츠 순으로 추천**

**협업 필터링 Collaborative Filtering**

개봉하는 영화 볼까 말까? “친구에게 물어보자” 🡪 **단, 취향이 비슷한 친구에게 물어봐야 함**

* **협업 필터링 유형**

**최근접 이웃기반(Nearest Neighbor) 🡪** 사용자 기반 (User-user CF) / 아이템 기반(Item-item CF)

**잠재 요인 기반(Latent Factor) 🡪** 행렬 분해 기반(Matrix Factorization)

* **협업 필터링의 특징**

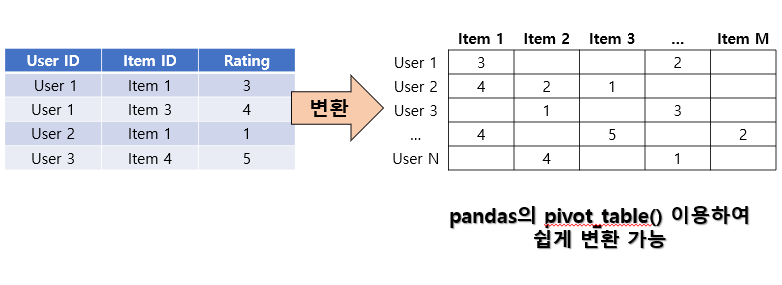
**User behavior**(item구매 이력, 영화 평점 이력)에만 기반하여 추천 알고리즘들을 전반적으로 지칭

상품, 영화 등사용자가 **아직 평가하지 않은 item에 대한 평가(rating)을 예측**하는 것이 주요 역할

테이블이(가) 표시된 사진

자동 생성된 설명

* **협업 필터링을 위한 데이터 세트 – 사용자 로우 아이템 컬럼**



* **사용자 기반과 아이템 기반 협업 필터링 이해**

**사용자 기반(User-User) :**

특정 사용자와 비슷한 고객들을 기반으로 이 비슷한 고객들이 선호하는 다른 상품을 추천

특정 사용자와 비슷한 상품을 구매해온 고객들은 비슷한 고객으로 간주

당신과 비슷한 고객들이 다음 상품도 구매함(Customers like you also bought these items)

테이블이(가) 표시된 사진

자동 생성된 설명

**아이템 기반(Item-item) :**

특정 상품과 유사한 좋은 평가를 받은 다른 비슷한 상품을 추천

사용자들로부터 특정상품과 비슷한 평가를 받은 상품들은 비슷한 상품으로 간주

이 상품을 선택한 다른 고객들은 다음 상품도 구매함(Customers who bought this item also bought these items)

테이블이(가) 표시된 사진

자동 생성된 설명

* **사용자 기반 VS 아이템 기반**

일반적으로 사용자 기반 보다는 **아이템 기반 방식이 더 선호됨.**

**사람 간의 특성**은 상대적으로 **다양한 요소들에 기반**함 🡪 단순히 동일한 상품을 구입하였다고,

유사한 사람이라고 판단하기 어려움.

* **아이템 기반 협업 필터링의 개인화된 영화 추천**

**아이템 기반의 협업 필터링에서 개인화된 평점 예측**

**Weighted Rating Sum**

사용자 u의 아이템 i에 대한 평점 예측을 사용자 u가 아이템 i와 유사한 다른 아이템들(N개의 다른 아이템)의 합으로 계산하되 **아이템 i와 다른 아이템들 간의 유사도를 반영한 합**으로 계산

: 사용자 u 아이템 i의 개인화된 예측 평점 값

: 아이템 i와 가장 유사도가 높은 TOP-N개 아이템의 유사도 벡터

: 사용자 u와 아이템 i와 가장 유사도가 높은 TOP-N개 아이템에 대한 실제 평점 벡터

테이블이(가) 표시된 사진

자동 생성된 설명

**아이템 기반 협업 필터링 구현 프로세스**

1. 사용자 – 아이템 행렬 데이터를 **아이템-사용자 행렬 데이터로 전환**
2. 아이템간의 **코사인 유사도로 아이템 유사도 산출**
3. 사용자가 관람(구매)하지 않은 아이템들 중 아이템간 유사도를 반영한 **예측 점수 계산**
4. 예측 점수가 가장 높은 순으로 아이템 추천

* **잠재 요인 협업 필터링의 개요**

**사용자-아이템 평점 행렬** 속에 숨어 있는 **잠재 요인을 추출해 추천 예측**을 할 수 있게 하는 기법

대규모 다차원 행렬을 **SVD와 같은 행렬 분해(Matrix Factorization) 기법으로 분해하는 과정에서 잠재 요인을 추출** 후, 사용자-아이템 평점 행렬을 재구성하며 추천을 구현.

텍스트이(가) 표시된 사진

자동 생성된 설명

잠재 요인 협업 필터링의 행렬 분해 목표는 **희소 행렬 형태의 사용자-아이템 평점 행렬**을

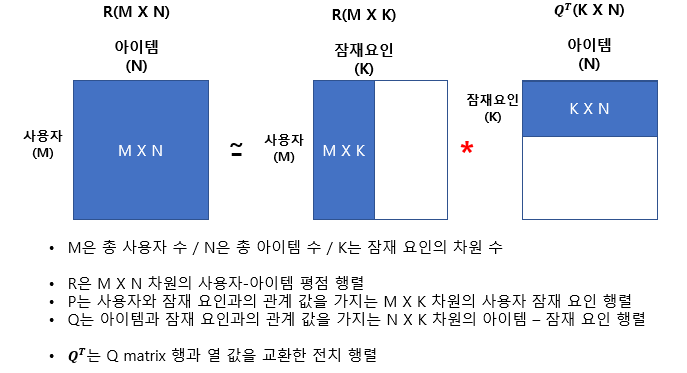
**밀집(Dense) 행렬형태의 사용자-잠재 요인 행렬과 잠재요인-아이템 행렬**로 분해한 뒤 이를 재결합하여 밀집 행렬 형태의 사용자-아이템 평점행렬을 생성하여 사용자에게 새로운 아이템을 추천하는 것

* **행렬 분해를 통한 잠재 요인 협업 필터링**

테이블이(가) 표시된 사진

자동 생성된 설명

* **잠재 요인 기반의 행렬 분해 이해**



* **사용자-아이템 평점 행렬 분해 이슈**

테이블이(가) 표시된 사진

자동 생성된 설명

* **경사 하강 법 기반의 행렬 분해**

1. P와 Q를 임의의 값을 가진 행렬로 설정
2. P와 Q값을 곱해 예측 R행렬 계산 후, 예측한 R행렬 값과 실제 R의 오류 값을 계산
3. **이 오류 값을 최소화**할 수 있도록 **P와 Q행렬을 적절한 값으로 각각 업데이트**
4. 만족할 만한 오류 값을 가질 때가지 2,3번 작업 반복 🡪 P, Q 값 업데이트하여 근사화

* **경사 하강 법 기반의 행렬 분해 비용 함수**

텍스트이(가) 표시된 사진

자동 생성된 설명

* **경사 하강 법 기반의 행렬 분해 업데이트 식**

실제 R행렬 값과 예측 R 행렬 값의 차이를 최소화하는 방향성을 가지고 P행렬과 Q행렬에 업데이트 값을 반복적으로 수행하며 예측 R 행렬을 구하는 방식이 경사 하강법 기반의 행렬 분해.

텍스트이(가) 표시된 사진

자동 생성된 설명