

Introduction to Recommender Systems

<https://product.kyobobook.co.kr/detail/S000201794203>

추천 시스템

입력 > 프로세스 > 출력

입력 : 콘텐츠와 인터랙션, 내용 기반 필터링 혹은 협조 필터링

프로세스 : 개요 추천, 연관 아이템 추천, 개인화 추천

출력

pull 타입 / push 타입

추천 시스템 개발에 필요한 3가지 스킬

: 비즈니스 문제 해결 능력

: 데이터 사이언스 능력

: 데이터 엔지니어링 능력

추천 시스템 프로젝트 진행 방법

과제 설립 > 가설 수립 > 데이터 설계/수집/가공 > 알고리즘 선정 > 학습/파라미터 튜닝 > 시스템 구현 > 평가 및 개선

클라우드 소싱

데이터 편향

배치 추천, 실시간 추천

적합 아이템 발견

개요 추천 - 신규 혹은 사용 빈도 낮은 사용자 정착

알림서비스 - 사용 빈도 향상 및 이탈자 복귀

연관 아이템 추천 - 교차 판매

개인화 - 장기적 사용자 충성도 향상

신선함, 흥미로움, 다양성

협조 필터링 : 메모리 기반 방법, 모델 기반 방법

콜드 스타트 문제

간접 지정형

사용자-사용자 메모리 기반 방법, 아이템-아이템 메모리 기반 방법

기호 데이터

명시적 피드백과 암묵적 피드백

평갓값 ~ 평갓값 행렬

토픽 모델

행렬 분해

LDA

명시적 획득과 암묵적 획득

데이터의 희박함

평갓값의 변동과 편향

배치 추천

실시간 추천

개요 추천

다단계 추천

근사 최근법 탐지

nDCG

클라이언트 사이드 로그

서버 사이드 로그

사용자 커버리지, 카달로그 커버리지

ERR

오프라인 평가 - 모델 정밀도 평가, 모델 밸리데이션, 모델 튜닝

온라인 평가 - A/B테스트, 인터리빙

사용자 스터디를 통한 평가 - 조사 설계, 설문

편향 - 선택 편향, 동조 편향, 노출 편향, 위치 편향, 인지도 편향, 불공정성

상호추천시스템

OEC 지표

가드레일 지표

감도

신뢰성

효율성

디버깅 간으성과 액션 가능성

해석 가능성과 방향성

인터리빙 ~ 멀티리빙

TDM - Team Draft Multileaving

PM - Probabilistic Multileaving

OM - Optimized Multileaving

PPM - Pairwise Preference Multileaving

F1-measure

Recall

Precision

집합 평가 지표

PR 곡선

AUC 곡선

MRR : 사용자 순위에 대해 최초의 적합 아이템이 순위에서 얼마나 상위에 위치하는지 평가하는 지표

AP@K : 평균 정밀도

MAP@K : 중앙 평균 정밀도

MAE : 평균 절대값 오차

MSE : 평균 제곱 오차

RMSE : 평균 제곱근 오차