

조동빈

이메일: ken.dbc.career@gmail.com

연락처: +821045207849

추천 시스템 ML Engineer로 일하고 있으며, 다음과 같은 생각들에 중점을 두고 업무를 수행합니다.

- 방향성: 주어진 task에 적합한 데이터 전략을 수립하며, 모델 및 배포 환경을 개발
- 모델: 데이터 유형(공업, 텍스트, 음성, 이미지, 그래프, 유저 정보) 별 특성을 반영한 모델 설계

경력

아프리카TV

데이터기술팀/팀원

2023.02 - 현재 재직중

● [추천 파이프라인] 신규 글로벌 서비스의 전체 추천 파이프라인 개발 (진행중) 2025.02 -

완료:

- 서비스 추천 파이프라인 개발을 위한 장기 계획 수립 (4단계, 포트폴리오 참고)
- ' 선호 ' 스트리머 탐색 로직 개발

진행중:

- 선호 스트리머 데이터 Feature Store 활용한 파이프라인 개발

예정:

- 선호 스트리머 기반, LIVE 방송 및 VOD 추천 모듈 개발 (쿼리 조회)
- MAB를 이용한 '연관' 스트리머 탐색 모듈 개발
- ML/DL 도입 및 유저 별 스트리머/VOD 관계 추론

● [MLOps 도입] MLflow 도입 2025.01 - 2025.02

내용:

- MLOps 툴 중 하나인 MLflow 도입
- 쉽게 다양한 프로젝트에서 쉽게 적용할 수 있도록 wrapping class 개발
- HDFS와 RDB 연동 및 MLflow 서버 이중화 (장애 대비)

성과: Library화 하여 팀 및 외부 ML 유관 부서에 공유

● [숏폼 플레이리스트] 개인화 추천에 스트리머 탐색 로직 추가 2024.11 - 2024.12

내용:

- 숏폼 동영상의 경우 평소 선호하지 않는 스트리머들도 시청하는 특성이 존재
- 숏폼 개인화 추천 데이터에 유저 별로 현재 선호 중이지 않은 다른 인기 스트리머를 추천 데이터에 삽입하는 로직을 추가

성과: 개발한 네 가지 탐색 로직들 중 세 가지 로직들에서 긍정적인 반응 (노출 대비 재생 완료 확률 40% 이상)

● [즐거찾기 페이지] 클립&숏폼 동영상 개인화 추천 2024.08 - 2024.10

내용:

- 기존 즐겨찾기 페이지의 클립 구좌와 숏폼 구좌는 즐겨찾기 된 모든 스트리머들 중 최신/인기 순서로 노출 됨
- 따라서 대체로 선호 여부와 관계 없이 인기 스트리머들의 동영상이 노출됨 (오래된 유저의 경우 즐겨찾기는 100명을 넘기도 함)
- 선호 스트리머를 short-term, long-term 별로 선정 후 후보군을 제한하여 ML 모델로 추천

성과: 2024.12.26 서비스 배포 후 현재 MY에서의 VOD 재생 횟수 약 20% 증가

● **조희수 가치 평가 파이프라인 개발** 2024.08 - 2024.10

내용: 유저들이 동영상을 시청했을 때, 유저 별 특성을 고려하여 유효 시청 여부를 판단

1. spark 작업

- 현재 batch에 생성된 시청 이력들 중 abusing을 제외
- 현재 batch에 생성된 시청 이력들을 추천 모델이 input으로 사용할 수 있도록 빠르게 점수화 및 모델에게 전달

2. hive 작업

- abusing table update 등

성과: 정확한 서비스 현황 파악 및 추천 데이터 input 품질 개선

● **[스트리머 개인 채널] 스트리머 대표 동영상 추천** 2024.05 - 2024.07

내용:

- 스트리머를 표현하는 대표 동영상들을 선정하여, 스트리머 페이지 최상단에 노출하기 위한 목적
- 스트리머의 모든 VOD(다시보기, 클립, 숏폼 등)들을 점수화
- 8 개의 통계적 아이템 평가 지표 설계 (인기성, 최신성, 희소성, 양질성 등)
- 6 개의 스트리머 유형 분류 및 유형 별 아이템 평가 지표에 대한 가중치 설계

성과: 개발 작업 완료, 2025년 배포 예정

● **오프라인 시뮬레이션 파이프라인 개발 및 자동화** 2023.11 - 2024.01

내용: 여러 ML 모델들을 online 환경을 최대한 모방하여 오프라인에서 평가할 수 있는 시뮬레이션 시스템 개발

- 지표 예시: Hit@k, Hit@1, nDCG, MAP, user coverage, streamer coverage, freshness, ...
- 각 모델의 성능 및 특성 파악
- 모델 간 겹치는 타겟 유저들에 대한 지표들을 비교 평가

성과:

- online test는 cost와 risk가 매우 크므로, 배포 후보가 되는 모델의 성능을 대략적으로 체크할 수 있게 됨
- 다양한 측면에서 추천 데이터를 정량/정성 평가 가능

● **[메인 페이지] 숏폼 동영상 개인화 추천** 2023.08 - 2023.10

내용: 메인 페이지 숏폼 동영상 개인화 추천 모델 개발 및 배포

- 기존 모델의 경우 다양한 스트리머를 추천하는 특성이 있으며, 많은 수의 유저들에게 이러한 추천이 부적절 하다 판단
- '항상 보던 스트리머' 위주로 시청하는 유저들에게 개발한 모델을 이용한 추천 데이터를 배포

성과: 기존 모델 대비 CTR 7% 성능 향상

- **[메인 페이지] 라이브 방송 추천** 2023.04 - 2023.06

내용: 메인 페이지 라이브 방송 개인화 추천 알고리즘(통계 기반) 개발 및 배포

성과: 기존 추천 알고리즘 대비 CTR 3% 향상 -> 기존 로직을 대체하여 현재 배포 중

언디파인드

2022.06 - 2022.12

개발팀/팀원 (회사 내부 사정으로 퇴사)

- **매치 결과 자동 기록 시스템 개발** 2022.11 - 2022.11

eSports 대회 참가자가 매치 결과를 촬영하여 서버로 전송했을 때,
OCR 모듈을 이용하여 순위/킬 횟수 등을 자동으로 기록해주는 시스템 개발

- **대회 규정 추천 시스템 개발** 2022.08 - 2022.11

eSports 대회를 대회 운영자가 생성할 때,
운영자가 기재 한 몇가지 대회 규정을 기반으로, 나머지 규정들을 추천 및 자동으로 기입하는 추천 시스템 개발

- **FAQ 챗봇 개발** 2022.06 - 2022.11

FAQ 챗봇을 개발 및 배포

AI Server를 만들고, 영어/한글 모델을 각각 다른 ECS(AWS)를 이용하여 운영

카카오

2019.12 - 2020.02

추천팀/인턴

- **만화책 추천 알고리즘 개발** 2022.02 - 2022.02

Text Analysis, Collaborative Filtering, Computer Vision, Multi-armed Bandit 등을
활용하여 추천 모델을 배포 및 테스트

- **자동차 동영상 추천 알고리즘 개발** 2022.01 - 2022.02

Text Analysis, Collaborative Filtering, Multi-armed Bandit 등을 활용하여 추천 모델을
배포 및 테스트

한빛소프트

2018.11 - 2019.04

AR팀/팀원 (대학원 진학을 목표로 퇴사)

- **한글 음성 합성 모듈 개발** 2019.01 - 2019.04

다중화자 단일 모델의 음성 합성 모듈 개발

- **영어 챗봇 개발** 2018.11 - 2019.01

데이터 탐색 및 전처리

학력	한양대학교 대학원 컴퓨터소프트웨어학과 석사 2020.03 - 2022.02
	1. Stanford와 Almond라는 인공지능 비서 시스템의 한글 Seq2SQL 모델 개발 2. [논문, 2022 BIB] RAMP: Response-Aware Multi-task Learning with Contrastive Regularization for Cancer Drug Response Prediction 3. [논문, 2022 ICEIC] Quantization training with two-level bit width
	강원대학교 산업공학과 학사 2012.03 - 2018.08
	프로그래밍(C++), 데이터베이스, 정보시스템 응용통계, 최적화 알고리즘, 경영과학 HCI 게임공학(Gamification), 기술경영 생산/품질/물류 관리
스킬	PyTorch, Tensorflow, Python, C, AWS, Docker, Ubuntu, CUDA, SQL, Git, Scikit-Learn, Kubernetes, Redis, ElasticSearch, Linux, 딥 러닝, Machine Learning, Apache Airflow, Apache Spark
수상 및 기타	삼성청년SW아카데미(SSAFY) 교육 수료 2019.12 Back-end, Front-end 기초 공부
	ADsP, 데이터분석준전문가 자격증 2017.12 취득
외국어	영어 일상회화 ● OPIc 2017.05.12 AL