

# 조동빈

## Data Scientist

Birthday	1993.1.5
Email	<a href="mailto:ken.dbc.career@gmail.com">ken.dbc.career@gmail.com</a>
Mobile	010-4520-7849
Address	경기도 남양주시 오남읍 양지로 240 번길 37

### 소개 / About Me

- 공업/이미지/텍스트/음성/그래프/유저 데이터들을 다루며 DS 역량을 키워왔습니다.
- 데이터 특성에 기반한 응용 SW 를 설계/최적화 하는 것에 주된 노력들을 해 왔습니다.
- 데이터 분석, ML modeling, 실험 설계 및 배포까지 전 과정을 즐깁니다.
- 팀원 뿐만 아니라, 팀의 리더 로서도 업무를 수행해 왔습니다.

### 기술스택 / Skill Set

구분	Skill
Programing Languages	Python, C, C++, CUDA, JavaScript, Shell Scripting
Framework / Library	Pytorch, TensorFlow, Scikit-learn, Pandas, cuDNN, cuBlas(LT), TensorRT
Server	Django, Flask, FastAPI
MLOps / DevOps	Docker, Docker-compose, Kafka, Kubernetes, Nginx
OS	Ubuntu, OS X
DB	MySQL, MongoDB, DynamoDB, PostgreSQL
Etc	AWS, Make

### 경력 사항 / Work Experience 총 3년 2개월

재직 기간	회사명	부서 및 R&R
2022.06.22 ~ 2022.12.30	엔디파인드	개발팀, AI 개발자
2020.03.01 ~ 2022.02.25	(한양대) ML system lab	컴퓨터소프트웨어학과, 석사과정
2019.12.26 ~ 2020.02.29	카카오	추천팀, 인턴
2018.11.05 ~ 2019.04.22	한빛소프트	AR 팀 인공지능 파트, 전임연구원

1. 2022.06 ~ 2022.12 언디파인드 / 개발팀 / AI 개발자 / 사원

\*회사 소개 : Undefined 는 iScrim 이라는 플랫폼을 개발하는 회사이며, iScrim 은 eSports 대회를 쉽게 개최할 수 있도록 돕는 플랫폼입니다. 앱을 이용한다면 복잡한 대회 개최 및 운영 절차들을 간편하게 수행할 수 있습니다.

A. 2022.11 ~ 2022.11 / 경기 결과 자동 기록 시스템

개발 환경	Python 개발 및 AWS 배포
인력구성 및 기여도	- 회사 내 유일 AI 개발자 - 사업/기획/프론트엔드/백엔드 직원들과 협의하여 개발/배포/운영을 진행
주요업무 및 상세역할	전략 수립, 모델 개발, Pipelining, 배포
성과/결과	개발 및 배포 후 유지/보수

B. 2022.08 ~ 2022.11 / 대회 규정 추천 시스템

개발 환경	Python 개발 및 AWS 배포
인력구성 및 기여도	회사 내 유일 AI 개발자 사업/기획/프론트엔드/백엔드 직원들과 협의하여 개발/배포/운영을 진행
주요업무 및 상세역할	전략 수립, 모델 개발, 모델 최적화, Pipelining, 배포
성과/결과	개발 및 배포 후 유지/보수

C. 2022.06 ~ 2022.11 / FAQ Chatbot 시스템

개발 환경	Python 개발 및 AWS 배포
인력구성 및 기여도	회사 내 유일 AI 개발자 사업/기획/프론트엔드/백엔드 직원들과 협의하여 개발/배포/운영을 진행
주요업무 및 상세역할	전략 수립, 모델 개발, 모델 최적화, Pipelining, 배포
성과/결과	개발 및 배포 후 유지/보수

2. 2020.03 ~ 2022.02 한양대학교 대학원 / ML system lab / 석사과정

\*회사 소개 : MLsys 연구실은 딥러닝 모델들의 학습/추론 과정을 system 적으로 최적화하는 기법들을 연구합니다. 그 예로 분산학습의 효율성 향상과 GPU 연산 최적화가 있습니다. 저는 그 중에서도 Quantization 과 Graph Embedding 학습과 같은 연구를 주로 진행했습니다.

A. 2020.03 ~ 2022.12 / Graph Embedding Generation

개발 환경	- OS: Ubuntu - Language: Python, Cython - Framework: PyTorch, TensorFlow2 - Library: Scikit-learn, Node2Vec
인력구성 및 기여도	- 항암제와 세포의 반응성을 연구하는 프로젝트를 세 연구실이 협업 - 그 중 본인의 연구실은 node 들의 embedding 을 학습하는 역할

	- 연구실 내의 참여 연구원은 본인 1명이므로, 리드 및 개발 작업 모두 수행
주요업무 및 상세역할	<ul style="list-style-type: none"> <li>- Negative Sampling 을 개선한 RA-NS 를 개발</li> <li>- Embedding vector 의 quality 를 측정하는 metric 인 Embedding Score 개발</li> <li>- Cython 으로 모델 학습 속도 최적화</li> <li>- Node2Vec, GCN, GTN 등에 개발하여 실험</li> <li>- 논문 작성</li> </ul>
성과/결과	2022 년 12 월, Briefings in Bioinformatics 저널에 논문 publish

B. 2020.03 ~ 2022.02 / DNN Model Quantization - 1

개발 환경	<ul style="list-style-type: none"> <li>- OS: Ubuntu</li> <li>- Language: Python</li> <li>- Framework: PyTorch, Darknet</li> <li>- Library: Scikit-learn</li> </ul>
인력구성 및 기여도	<ul style="list-style-type: none"> <li>- 2021년에는 단독 진행</li> <li>- 2022년부터 후배 석사과정 학생 2명을 리드</li> <li>- 개발을 포함한, 모든 작업을 주로 진행</li> </ul>
주요업무 및 상세역할	<ul style="list-style-type: none"> <li>- Google's Quantization 기술 follow up</li> <li>- Statistics 관점에서의 개선점 파악</li> <li>- Framework를 이용한 코드 및 개선점 적용</li> </ul>
성과/결과	현재 후배 학생들에게 조언을 주고 있으며, 2023 ICML 논문 제출 예상

C. 2021.10 ~ 2021.12 / DNN Model Quantization - 2

개발 환경	<ul style="list-style-type: none"> <li>- OS: Ubuntu</li> <li>- Language: Python</li> <li>- Framework: PyTorch</li> </ul>
인력구성 및 기여도	<ul style="list-style-type: none"> <li>- 후배 석사과정 학생 2명을 리드</li> <li>- 본인이 직접 코드를 개발</li> <li>- 후배 학생이 실험 및 논문 작성</li> </ul>
주요업무 및 상세역할	<ul style="list-style-type: none"> <li>- 논문 분석 및 한계점 도출 (Quantization Aware Training, QuantNoise)</li> <li>- 개선점 파악 및 코드 구현</li> </ul>
성과/결과	2022 년 ICEIC 학회에 논문 publish

D. 2020.07 ~ 2020.09 / Korean AI Assistant

개발 환경	<ul style="list-style-type: none"> <li>- OS: Ubuntu</li> <li>- Language: Python</li> <li>- Framework: TensorFlow2</li> </ul>
인력구성 및 기여도	<ul style="list-style-type: none"> <li>- Stansford OVAL Lab.과 협업</li> <li>- 동료 석사과정생과 함께 작업을 진행</li> <li>- AI Assistant 인 Almond 의 기술들을 분석한 후 연구 과정을 설계</li> </ul>
주요업무 및 상세역할	<ul style="list-style-type: none"> <li>- Almond 기술의 논문 research</li> <li>- AI Assistant pipeline 분석</li> <li>- KR Seq2SQL model 데이터 생성</li> </ul>
성과/결과	당시 2 개월 진행 후 중단됐으며, 최근 연구실 내부적으로 재개됨

### 3. 2019.12 ~ 2020.02

#### 카카오 / 추천팀 / 추천 알고리즘 개발자 / 인턴

\*회사 소개 : 카카오 추천팀은 카카오 내부적으로 추천 알고리즘이 필요한 서비스들의 추천 시스템을 개발해 주는 부서입니다. 저는 인턴끼리 3인 1팀을 이루었고, 총 두 가지 서비스에 대한 추천 알고리즘을 개발했습니다.

#### A. 2020.02 ~ 2020.02 / 만화 추천 알고리즘 개발

개발 환경	Ubuntu, Python, Docker, Kubernetes, Kafka
인력구성 및 기여도	- 인턴 팀원 2명과 함께 진행 - 데이터 분석과 Project Manager 역할로 기여
주요업무 및 상세역할	- 데이터 수집/분석 - 실험 설계/분석 - 실험 log 기록 - 중간/최종 점검에 대한 PT 제작 및 발표
성과/결과	Click Through Rate(CTR) 상승

#### B. 2020.01 ~ 2020.02 / 자동차 동영상 추천 알고리즘 개발

개발 환경	Ubuntu, Python, Docker, Kubernetes, Kafka
인력구성 및 기여도	- 인턴 팀원 2명과 함께 진행 - 데이터 분석과 Project Manager 역할로 기여
주요업무 및 상세역할	- 데이터 수집/분석 - 실험 설계/분석 - 실험 log 기록 - 중간/최종 점검에 대한 PT 제작 및 발표
성과/결과	Click Through Rate(CTR) 상승

### 4. 2018.11 ~ 2019.04

#### 한빛소프트 / AR 팀 / 인공지능파트 / 전임연구원

\*회사 소개 : 당시 속해 있었던 팀은 한빛소프트에서 처음으로 개설한 인공지능 연구파트로, 원어민 강사를 대체할 수 있는 음성 챗봇을 만드는 것을 목표로 연구를 진행했습니다.

#### A. 2019.01 ~ 2019.04 / Speech Synthesis 모듈 개발

개발 환경	Python, Window, Ubuntu, TensorFlow1
인력구성 및 기여도	단독 진행
주요업무 및 상세역할	- Speech dataset 확보 및 전처리 - 모델 탐색 - 모델 학습
성과/결과	사내 초기 데이터셋 및 모델 확보

#### B. 2018.11 ~ 2018.12 / 음성 Chatbot 개발

개발 환경	Python, Window, Ubuntu, TensorFlow1
인력구성 및 기여도	- 인공지능 파트장 1명의 리드 하에 작업 - 기존 prototype 모델의 데이터셋과 모델의 퀄리티 개선 model 개서

주요업무 및 상세역할	<ul style="list-style-type: none"> <li>- 기본적인 NLU, NLG 모델을 사용하는 기존 모듈들의 최신화</li> <li>- 기존 dialogue 데이터셋의 전처리 개선(label 재배치 포함)</li> <li>- Metric 변경으로 모델 평가 모듈 개선</li> </ul>
성과/결과	NLU/NLG 모델 성능 향상, 사내/사외 시연 진행

## 기타 사항

### 학력

- 2020.03 ~ 2022.02 한양대학교 컴퓨터소프트웨어학과 졸업
- 2012.03 ~ 2018.08 강원대학교 졸업
- 2008.03 ~ 2011.02 보성고등학교 졸업

### 자격증

- ADsP(2017.12 취득)

### 교육 및 대외활동

- 2019.07 ~ 2019.12 삼성청년 SW 아카데미