



최건우

📍 대한민국 경기도 김포시

☎ +82 10 6260-6320

✉ easter9211@gmail.com

🌐 <https://github.com/KeonwooChoi>

Born 1994년 3월 10일

EDUCATION

MAR 2015 – FEB 2021

컴퓨터과학과

상명대, 서울

학점 3.5/4.5

FEB 2021 – JUNE 2021(plan)

부스트캠프 AI Tech

NAVER Connect Foundation, Seoul

Languages & Certifications

TOEIC 895/990

정보처리기사

SQL Developer (SQLD)

SKILLS

Languages

Python/Javascript

Libraries & Tools

Pytorch/HuggingFace/Pandas/Scikit-learn/React/Nodejs/MySQL

PROJECT

NOV 2019 – OCT 2020

챗봇 프레임워크

- React, Nodejs, MySQL, AWS
- 사용자가 입력 한 키워드와 답변의 쌍으로 구성된 챗봇 데이터를 메신저 플랫폼 (e.g., LINE, Telegram)에 연결하는 API 개발
- React를 사용하여 메신저 플랫폼과 유사한 사용자 친화적인 UI구현
- <https://github.com/KeonwooChoi/EZ.ai>

Competitions

BoostCamp

마스크 착용 상태 분류

- 18,900개의 이미지를 18개의 클래스로 분류 (마스크 상태 3 개, 성별 2 개, 연령 3 개)
- What I've done:: Resnet, EfficientNet의 각각 5Fold CV를 한 후 앙상블
- What I've learned: 데이터 유출을 고려한 교차 검증 전략 설정 방법, Wandb를 이용한 하이퍼 파라미터 튜닝, cutmix를 포함한 정확도 향상을위한 다양한 이미지 증강.

BoostCamp

문장 내 개체간 관계 추출

- 관계 추출 후 42 개 클래스로 분류 (KLUE 데이터 9,000 문장)
- What I've done: 관계 추출 task을 Multi-turn QA task로 해결
- What I've learned: 교착어라는 한국어의 특성을 고려하여 음절 단위로 나눌 수 있는 커스텀 토큰라이저를 만들

BoostCamp

다중 도메인 대화 상태 추적

- 대화에서 각 턴의 상태 추적 (Wizard-of-Seoul의 7,000개의 대화 데이터)
- What I've done: 우리의 task에 맞게 여러 SOTA 모델 (e.g., CHAN-DST, Transformer-DST)을 구현하고 앙상블. 추적하려는 상태가 온톨로지에 없으면 정확도가 훨씬 낮아지기 때문에 생성 모델에 집중.
- What I've learned: 온톨로지 기반 및 생성 모델의 구조와 장단점

- BoostCamp** Deep Knowledge Tracing (In progress)
- 사용자의 공부 데이터를 기반으로 사용자가 마지막 질문을 맞출 것인지 예측 (i-screamedu의 사용자 7,442 명의 데이터)
 - What I've done: 유사한 캐글 대회 1등 솔루션을 참고하여 구현 후 윈도우 슬라이딩을 통해 데이터를 10배 증강 후 학습
 - What I've learned: 성능 향상을 위한 피쳐엔지니어링의 중요성

INTERESTS

성능을 향상시키기 위해 논문을 읽고 해결하고자 하는 task에 맞게 구현하는 것에 흥미를 느낍니다.

긴 대화에서 상태를 잃지 않고 추적하는 task, 대용량의 문서에서 질문과 답변을 추출하는 task에 관심이 있습니다.

성능 뿐 아니라 서버까지 고려하여 사용자를 만족시킬 있는 서비스 구현에 관심이 있습니다.

자연어 처리에서 더 나아가, 인간의 실제 경험을 표현할 수있는 Multi-modal task에 관심이 있습니다.

Military Service

MAY 2016 – FEB 2018

대한민국 육군 만기제대

- 제3기동단에서 의경으로 복무