

# About - AI Researcher / Engineer



이름 : 박창현

연락처 : 010-9270-3748

거주지 : 경기도 고양시 일산서구 대화동

주전공 : 응용통계학과

github : <https://github.com/sda96>

blog : <https://sda96.github.io/>

e-mail : [sda96@naver.com](mailto:sda96@naver.com)

안녕하세요. 항상 저 자신의 부족함을 느껴 부족한 것을 채우고 더 많은 것을 배우는 AI Research Engineer 직무 지원자 박창현입니다!

## 부족한 부분을 채우기 위해서 계속 공부하고 있습니다.

딥러닝을 공부하기 위해서 '데이터에듀'에서 기업현장실습을 하며 파이썬을 학습하였고, '모두의 연구소'에서 진행한 AIFEL 대전 2기 인공지능 개발자 교육과정을 통하여 딥러닝의 이론적 기초를 쌓고, 세부 분야인 NLP를 선택하여 Transformer, BERT, GPT등의 논문을 리뷰하고 기술 블로그에 기재했습니다.

NLP 분야를 선택하고 지속적인 공부를 이어가기 위해서 '집현전-자연어처리 스터디'에 참여하여 다양한 조가 선택한 논문을 함께 공부하고 제가 속한 조에서는 XLNet 논문을 리뷰하고 발표하였습니다.

## 프로젝트의 팀장으로써 성공적으로 프로젝트를 마친 경험이 있습니다.

(주)모두의 연구소 인공지능 개발자 과정을 신청하였고, 과정의 마지막으로 위협문장을 다중 분류 모델 성능 고도화를 하는 TUNiB 기업과제를 진행하였습니다.

해당 프로젝트에서 팀장으로써 팀원들의 능력에 맞는 역할분배와 서로 간의 일정을 공유할 수 있는 Github Project를 활용해 실험계획을 세워서 F1-score 성능을 0.673에서 0.882까지 끌어올렸습니다.

## 새로운 시도를 두려워하지 않습니다.

팀프로젝트 경험을 쌓기 위해서 한국데이터산업진흥원에서 주관한 빅데이터 인재양성 프로그램 경진대회에 자유주제로 참여하였고, '증상 입력을 통한 연관 질병 추천 웹 서비스 개발' 프로젝트를 수행하였습니다.

해당 프로젝트에서 모델 개발 업무를 맡아서 좀 더 입력과 연관성 높은 결과를 만들기 위해서 새로운 진료과 분류 모델을 추가하여 만족할 만한 결과를 만들었고 한국데이터산업진흥원장 상을 수상했습니다.

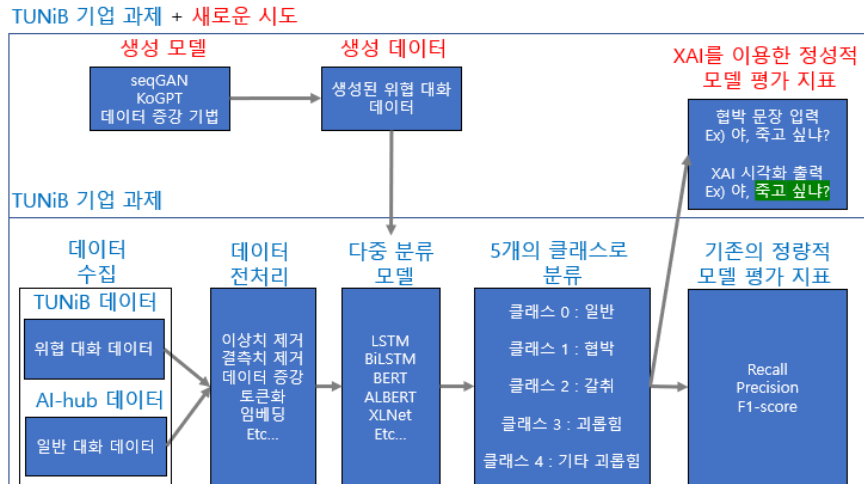
## 데이터 수집, 모델 구현, Azure를 통한 모델 배포 경험이 있습니다.

(주)퍼블릭에이아이에서 인공지능 개발자 직무를 6개월간 인턴하며 고등학생 대상 인공지능 교육을 위한 <AI-KIT> 상품 개발을 위하여 Image Classification 모델 파이프라인 개발 및 이슈 해결과 Azure Lab Service를 활용한 실습환경 구축에 참여하였습니다.

# Project

- 모두의 연구소 AIFFELTHON TUNiB 기업과제 (2022.01 ~ 2022.03)

- 사용 언어 및 패키지
  - Python, Tensorflow, HuggingFace
- 팀 프로젝트 소개 (팀장)



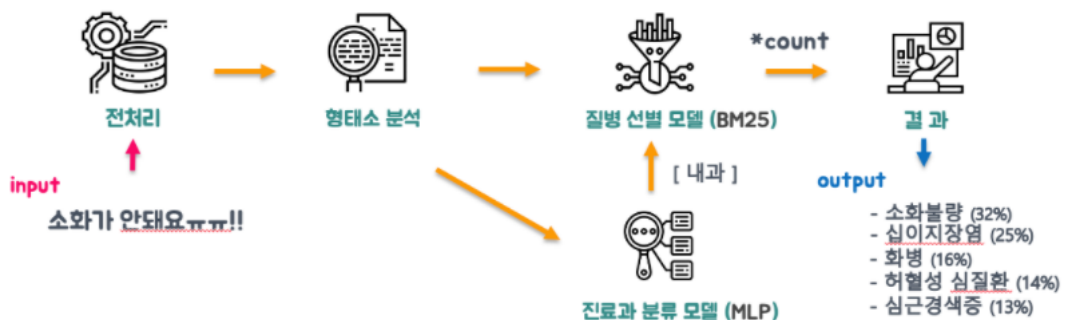
- TUNiB에서 제공한 위협 문장 데이터셋(DKTC)에 있는 '협박 대화', '갈취 대화', '직장내 괴롭힘 대화', '기타 괴롭힘 대화' 클래스가 존재하고 추가적으로 '일반 대화' 클래스를 AI-Hub에서 가져와서 **총 5가지의 클래스로 분류하는 다중 분류 모델을 구축하고 성능을 고도화** 시키는 것이 목적입니다.

- 역할

- 팀장으로써 각 팀원들 역할 분배 및 **Github proejct** 를 활용한 **일정 관리**를 하였습니다.
- 팀원과 동일한 작업환경을 맞추기 위해서 Google Cloud Platfrom 환경을 구축하고 직접 제작한 분류 모델 **프로토타입 Baseline 코드**를 공유하였습니다.
- Huggingface에서 사전학습된 모델을 불러와 프로토 타입 모델로써 Tensorflow 기반의 다중 분류 모델 구현하였으며, 처음에는 **0.673** 의 F1-score를 가진 프로토 타입 모델을 모델 앙상블 기법, 새로운 종류의 데이터, Task Adaptive Pre-Training(TAPT), Explainable attetnion visaulization을 통한 인사이트 도출등을 통하여 약 31% 상승한 **0.882** 의 F1- score 모델로 고도화 작업을 진행하였습니다.

- 빅데이터 인재양성 프로그램 프로젝트 경진대회 한국데이터진흥원장상 수상 (2020.10)

- 사용 언어 및 패키지
  - Python, tensorflow, cdqa(BM25)
- 팀 프로젝트 소개



- '문장단위 증상 입력을 통한 연관 질병 추천 웹 서비스 개발'을 목표로 하였으며 해당 서비스의 파이프라인은 문장을 입력받으면 입력된 문장과 관련된 진료과를 분류하고 진료

과에서 BM25 알고리즘을 사용하여 입력 문장과 유사정도 순으로 상위 5개 병명 문서를 출력하게 됩니다.

- 역할

- 형태소 분석기를 거친 입력 문장을 딥러닝 프레임워크인 Tensorflow2.x를 사용하여 해당 증상에 따른 진료과를 반환해주는 다중 분류 모델을 구현하여 기존 **모델에서 진료과 분류 모델(MLP)를 추가**시켰습니다.

진료과 분류 모델을 추가한 이유는 "소화가 안돼요." 라는 입력된 증상과 너무 연관성이 떨어지는 "이비인후과"등의 진료과가 출력되며 오히려 관련성이 없고 납득하기 어려운 질병문서를 출력하는 문제가 있었는데, 진료과 분류 모델을 추가함으로써 입력 증상과 보다 관련성 높은 질병을 도출하는 결과를 얻을 수 있었습니다.

## Career

- (주) 퍼블릭에이아이 AI개발자 인턴 6개월 (2020.09 ~ 2021.02)

- 사용 언어 및 패키지

- Python, Tensorflow

- 프로젝트 소개

- 고등학생 대상 인공지능 교육 패키지인 프로젝트 개발에 참여하였으며, 학생들의 작업 환경은 'Microsoft Azure Lab Services' 였고, 프로젝트의 목표는 Tensorflow 2.x를 활용하여 책의 주인공을 라즈베리 파이 카메라로 촬영하여 **주인공의 존재 여부를 판단하고 존재하는 경우 bounding box로 표시**를 하는 프로젝트입니다.
- 데이터 수집 목적의 라즈베리파이 카메라 환경 설정을 하였고 Azure Lab Service를 통하여 작업환경을 구축하였습니다.

- 수행 역할

- 모델 성능 고도화 작업

1. 동일 사진이 다른 환경으로 찍은 경우 성능이 떨어지는 문제 해결

- imgaug 패키지를 통해서 다양한 효과가 적용된 **훈련 데이터 생성** 적용된 imgaug 효과는 gaussian blur(이미지 노이즈 제거), brightness(밝기 조절), contrast(대비 조절)을 사용하여 모델의 robustness를 강화시켜주었습니다

2. 주인공과 비슷하게 생긴 인물을 분간해야하는 문제 해결

- 주인공과 비슷하게 생긴 인물의 **공통점을 제거하면 두 객체의 차이점에 주목하는 모델을 만들 수 있다는 가설**을 세웠고 모자를 제거한 뒤 학습을 시킨 결과 '월리'만을 분류하는 모델을 구현할 수 있었습니다.

- Azure Lab Services 선생용, 학생용 작업환경 구축

1. 프로젝트에서 진행하던 작업환경은 Azure Lab Services에서 진행하였으며 프로젝트에 필요한 라즈베리파이 카메라 설정과 패키지 설치등의 사전 작업환경 구축 작업을 진행
2. 선생님들이 Azure Lab Service를 사용할 수 있도록 가이드라인을 제작