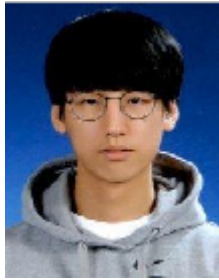


개인 소개



안녕하세요. 신입 NLP 엔지니어가 되기 위해서 공부하는 박창현입니다!

우선, 저 자신을 표현하는 키워드와 함께 키워드와 맞는 프로젝트와 경력을 간략하게 소개드리겠습니다.

🔗 주도적인 프로젝트 진행 능력 [프로젝트 이동]

- (주)모두의 연구소에서 진행된 인공지능 개발자 교육과정(AIFFEL)에서 TUNiB기업과제로 위협문장 클래스 분류 모델 고도화 프로젝트를 진행하였습니다.
- 팀장으로 활동하며 효율적인 업무를 위하여 프로토타입 모델과 베이스라인을 직접 제작하여 Github로 배포하였고, 이를 바탕으로 처음에는 모델의 F1-score가 **0.673** 이었지만 마지막 시간에는 **0.882** 까지 성능을 끌어올렸습니다.

🔗 집요한 문제 해결 능력 [프로젝트 이동]

- (주)퍼블릭에이아이에서 인공지능 개발자 인턴을 하는 동안 Tensorflow 2.x를 기반으로 고등학생 대상 인공지능 교육용 AI-KIT 개발에 참여하였습니다
- 이슈가 발생하면 이슈를 끊임없이 고민하고, 다양한 각도로 생각하여 어떻게든 해결하려는 성격 덕분에 모델 개발 이슈 관리 업무를 맡았습니다.

🔗 처음이라는 두려움을 이기는 추진력 [프로젝트 이동]

- 한국데이터산업진흥원에서 주관한 데이터 청년 캠퍼스 동국대 과정에서 자유 프로젝트로 '환자 증상을 통한 연관 질병 안내 웹서비스'를 구현하였습니다.
- tensorflow와 같은 딥러닝 프레임워크의 사용이 처음이었지만, 처음이라는 막연한 두려움을 이겨내고 tensorflow를 적극적으로 분석하여 결국에는 입력 증상에 맞는 진료과를 분류 모델을 직접 구현하여 성능 개선에 이바지하였으며, 결과적으로 한국데이터진흥원장상을 수상하게 되었습니다.

개인 정보

☎ 전화번호 : 010-9270-3748

🏠 거주지 : 경기도 고양시 일산서구 대화동

🎓 전공 : 응용통계학과

💬 관심분야 : 자연어 처리(NLP), 챗봇

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

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

e-mail : sda96@naver.com

사용 기술



프로젝트

- [모두의 연구소 AIFELTHON TUNiB 기업과제 \(2022.01.31 ~ 2022.03.10\)](#)
 - 사용 언어 및 패키지
 - Python3.x, Tensorflow2.x, 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 구축하여 가장 좋은 성능을 가진 모델을 구현하는 것**이 목적입니다.
 - 역할
 - 팀장으로써 각 팀원들 역할 분배 및 일정 관리를 위하여 Github Project를 관리하였습니다.
 - 팀원과 동일한 작업 환경을 제공하기 위해서 GCP 환경을 구축하고 직접 제작한 분류 모델 프로토타입 Baseline 코드를 팀원과 공유하였습니다.
 - Huggingface에서 사전학습된 모델을 불러와 프로토타입 모델로써 Tensorflow 기반의 다중 분류 모델 구현하였으며, 처음에는 **0.673**의 F1-score를 가진 프로토타입 모델을 모델 앙상블 기법, 새로운 종류의 데이터, Task Adaptive Pre-Training(TAPT), Explainable attention visualization을 통한 인사이트 도출등을 통하여 **0.882**의 F1-score 모델로 고도화 작업을 진행하였습니다.
- [데이터 청년 캠퍼스 동국대학교 프로젝트 경진대회 한국데이터진흥원장상 수상 \(2020.10.28\)](#)
 - 사용 언어 및 패키지
 - Python 3.x, tensorflow2.x, cdqa
 - 팀 프로젝트 소개
 - '환자 증상을 통한 연관 질병 추천 웹 서비스 개발'을 목표로 하였으며 해당 서비스의 파이프라인은 문장을 입력받으면 입력된 문장과 관련된 진료과를 분류하고 진료과에서 BM25 알고리즘을 사용하여 입력 문장과 유사한 증상을 가진 병명을 반환하는 웹서비스를 구현하였습니다.

- 역할

- 형태소 분석기를 거친 입력 문장을 딥러닝 프레임워크인 Tensorflow2.x를 사용하여 해당 증상에 따른 진료과를 반환해주는 다중 분류 모델을 구현하여 기존 one-stage 구조의 모델에서 진료과 분류 모델을 포함한 two-stage 모델로 변환하여 파이프라인의 결과 더욱 납득할 만한 결과를 얻을 수가 있었습니다.

경력

- (주) 퍼블릭에이아이 AI개발자 인턴 6개월 (2020.09.01 ~ 2021.02.28)

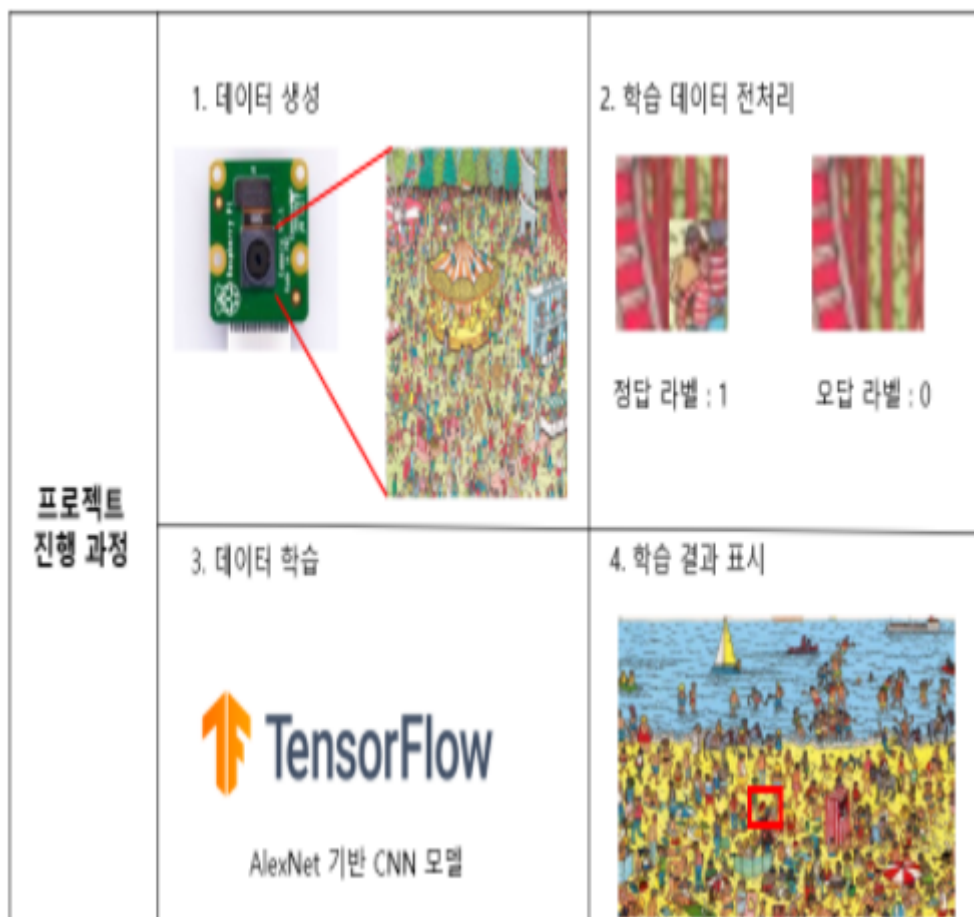
- 사용 언어 및 패키지

- Python 3.x, Tensorflow 2.x

- 프로젝트 소개

- 고등학생 대상 인공지능 교육 패키지인 <AI-KIT로 배우는 인공지능 기초> 프로젝트 개발에 참여하였으며, 학생들의 작업환경은 'Microsoft Azure Lab Services' 였고, 프로젝트의 목표는 Tensorflow 2.x를 활용하여 '월리를 찾아서'라는 책을 라즈베리 파이 카메라로 촬영하여 '월리'라는 object를 detection하는 프로젝트입니다.

- 프로젝트 진행 과정



- 역할

- 모델 이슈 관리

- 같은 이미지라도 새롭게 찍은 사진에서 성능이 떨어지는 이슈가 발생하였습니다.
- 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.

- '월리'라는 객체만을 분류해야하는데 '월리 여자친구'까지 함께 분류하는 이슈가 발생하였습니다.
 - 처음에는 '월리'와 '월리 여자친구' 두 객체의 차이점에 고려하여 안경의 형태를 모델이 학습할 수 있도록 edge detection의 'sovel filter'를 사용했지만 실패하였습니다.
 - 해당 이슈를 해결하기 위해서 일주일간 고민한 결과 기존 '월리'와 '월리 여자친구'의 공통점인 모자를 제거하면 두 객체의 차이점에 주목하는 모델을 만들 수 있다는 가설을 세웠고 모자를 제거한 뒤 학습을 시킨 결과 '월리'만을 분류하는 모델을 구현할 수 있었습니다.
- Azure Lab Services 선생용, 학생용 작업환경 구축
 - 프로젝트에서 진행하던 작업환경은 Azure Lab Services에서 진행하였으며 프로젝트에 필요한 라즈베리파이 카메라 설정과 패키지 설치등의 사전 작업환경 구축 작업을 진행하였습니다.
 - 선생님들이 Azure Lab Service를 사용할 수 있도록 사용 가이드라인을 제작하였습니다.

관련 공부

머신러닝

- [ISLR ch1~ ch6까지 파이썬으로 내용정리](#) (2020.10.01 ~ 2020.11.15)

딥러닝

- [핸즈온 머신러닝 2부 딥러닝 파트 내용 정리](#) (2021.03.01 ~ 중단됨)
- [밑바닥 부터 시작하는 딥러닝 1권 내용 정리](#) (2021.01.01 ~ 2021.02.28)

교육 및 자격증

- [모두의 연구소 AIFEL 대전 2기 딥러닝 자연어 처리 과정 이수](#) (2021.12.07 ~ 2022.03.10)
- [모두의 연구소 AIFEL 대전 2기 딥러닝 공통 과정 이수](#) (2021.09.06 ~ 2021.12.06)
- [네이버 커넥트재단 부스트 코스 AI- tech 2기 pre-course](#) (2021.06.01 ~ 2021.06.23)
- 한국데이터산업진흥원 주관 데이터 청년 캠퍼스 프로그램 동국대학교 과정 이수 (2020.06.22 ~ 2020.09.01)
- 데이터 분석 준전문가(Adsp) 자격증 취득 (2018.12.28)