

개인 소개



안녕하세요. 박창현입니다!

어떻게 하면 보다 자연스럽게 대화하고 사람과 같이 공감할 수 있는 오픈 도메인형의 챗봇을 만들 수 있을까 고민하고,
최종 목표로는 사람과 감정적으로 상호작용이 가능한 게임 NPC의 두뇌가 되는 챗봇을 만들고 싶습니다.
이러한 챗봇을 만들기 위해서 현재 자연어 처리(NLP)분야를 공부하고 있습니다.

- 전화번호 : 010-9270-3748
- 거주지 : 경기도 고양시 일산서구 대화동
- 전공 : 응용통계학과
- 관심분야 : 자연어 처리, 챗봇
- github : <https://github.com/sda96>
- blog : <https://sda96.github.io/>
- e-mail : sda96@naver.com

사용 기술

사용 기술 스택	기술 학습 계기 및 기간
	데이터 분석 준전문가 자격증 취득 2018.06 ~ 2018.08
	(주)데이터에듀 기업현장실습 2019.07.22 ~ 2019.08.16
	데이터 청년 캠퍼스 동국대학교 교육과정 이수 2020.06.22 ~ 2020.08.31
 	(주)퍼블릭에이아이 인공지능 개발자 인턴 2020.09.01 ~ 2021.02.28

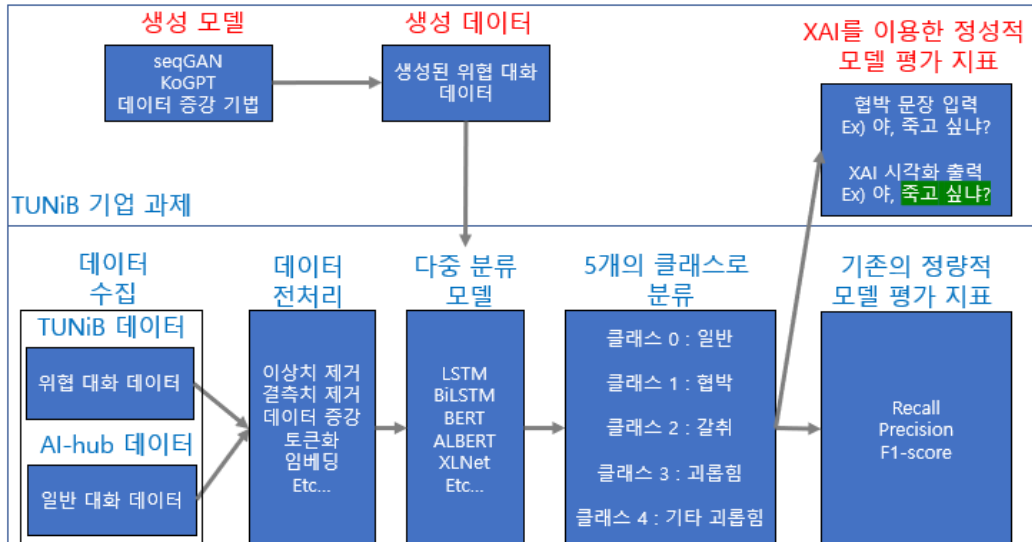
프로젝트

딥러닝

- 모두의 연구소 **AIFFELTHON TUNiB 기업과제** (2022.01.31 ~ 진행중)

- 사용 언어 및 패키지
 - Python3.x, Tensorflow2.x, HuggingFace
- 팀 프로젝트 소개(팀장)

TUNiB 기업 과제 + 새로운 시도

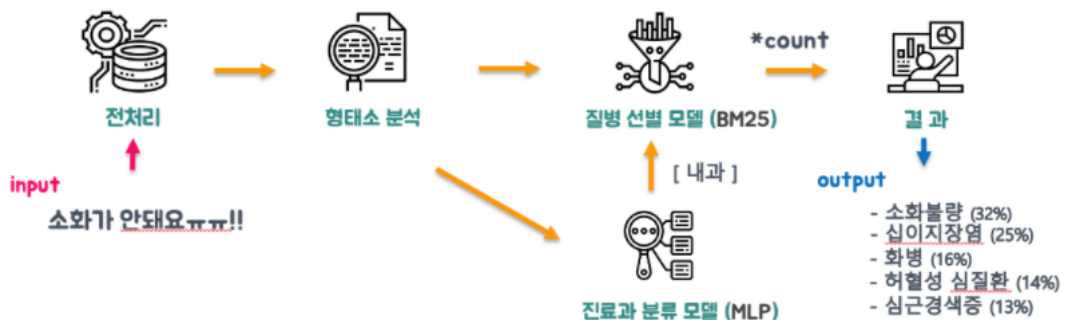


- TUNiB에서 제공한 위험 문장 데이터셋(DKTC)과 일반 대화 클래스를 ai-hub에서 가져와서 학습 5가지의 클래스로 분류시키는 텍스트 다중 분류 모델을 구축하여 가장 좋은 성능의 모델을 만들었습니다.

- 역할
 - 분류 모델 프로토타입 제작
 - 프로토타입을 기반으로한 모델 고도화 작업 진행
 - 팀장으로써 각 팀원들 역할 분배 및 일정 관리

- 데이터 청년 캠퍼스 동국대학교 프로젝트 경진대회 한국데이터진흥원장상 수상** (2020.10.28)

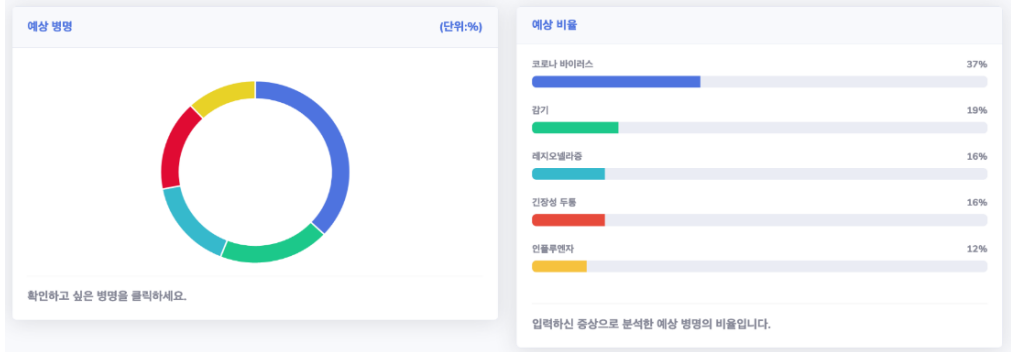
- 사용 언어 및 패키지
 - Python 3.x, tensorflow2.x, cdqa
- 팀 프로젝트 소개



- 환자 증상을 통한 연관 질병 추천 웹 서비스 개발을 목표로 하였으며 해당 서비스의 파이프라인은 문장을 입력받으면 입력된 문장과 관련된 진료과를 분류하고 진료과에서 BM25 알고리즘을 사용하여 입력 문장과 유사한 증상을 가진 병명을 반환해주며 구현한 웹서비스 결과는 아래의 그림과 같습니다.

증상을 통한 병명 분석

예상 병명 5개를 시각화한 모습입니다.



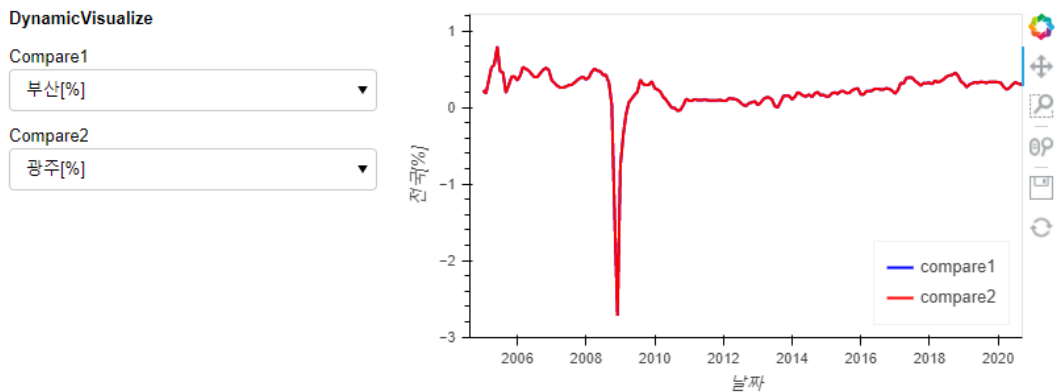
역할

- Tensorflow2.x 모델을 사용한 진료과 분류 모델 구현
- BM25 알고리즘이 구현되어 있는 cdqa패키지 적용

데이터 분석

데이콘 부동산 데이터 시각화 경진대회 (2021.05.01 ~ 2021.05.08)

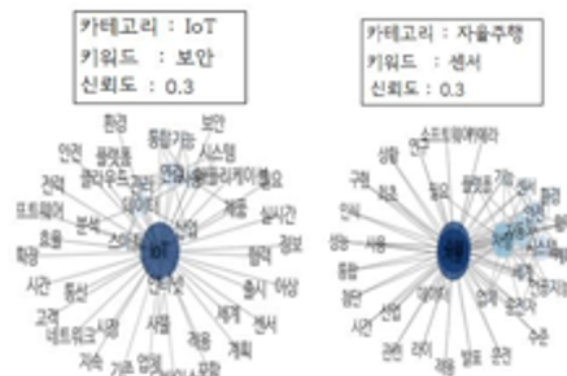
- 사용 언어 및 패키지
 - Python3.x, holoviews, matplotlib
- 토이 프로젝트 소개



- 데이콘에서 진행한 [한국 부동산 데이터 시각화 경진대회](#)에 참여하여 제주시 지가변동률 동적 시각화 및 분석내용을 작성하여 해당 [링크](#)로 공유하였습니다.

키워드 네트워크 분석을 활용한 뉴스 기사 카테고리별 키워드 분석 (2020.10.01 ~ 2020.12.17)

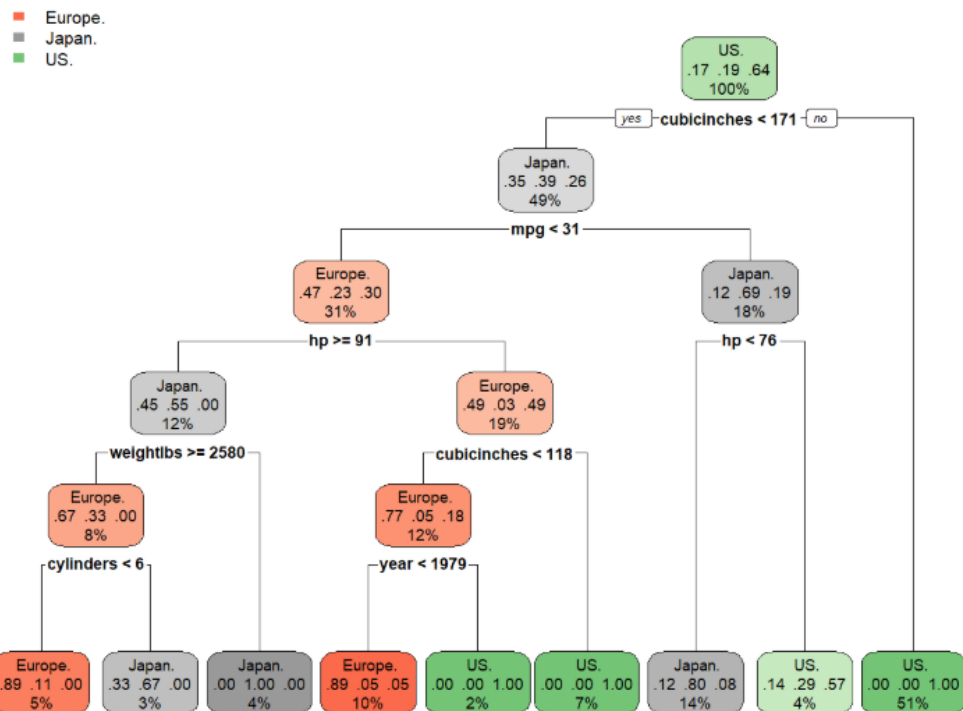
- 사용 언어 및 패키지
 - Python3.x, networkx
- 토이 프로젝트 소개



- 인공지능 관련 뉴스기사 데이터를 활용하여 각 카테고리별 핵심 키워드를 선정하여 선정된 키워드와 관련성 높은 연관 키워드를 추출한 프로젝트입니다.

- [Kaggle 자동차 데이터를 활용한 머신러닝 모델 성능 비교 분석 과제](#) (2020.04.12 ~ 2020.06.22)

- 사용 언어 및 패키지
 - R, randomForest, glmnet
- 토이 프로젝트 소개



- R 프로그래밍 언어를 활용하여 kaggle에 있는 cars 데이터셋을 분석하였으며 머신러닝 모델 의사결정나무, 다항 로지스틱 회귀 모델, 랜덤포레스트 3가지 모델의 성능을 비교하였으며 성능을 높이기 위한 주성분 분석을 진행하였습니다.

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)
- [밑바닥부터 시작하는 딥러닝 1권 내용 정리](#) (2021.01.01 ~ 2021.02.28)

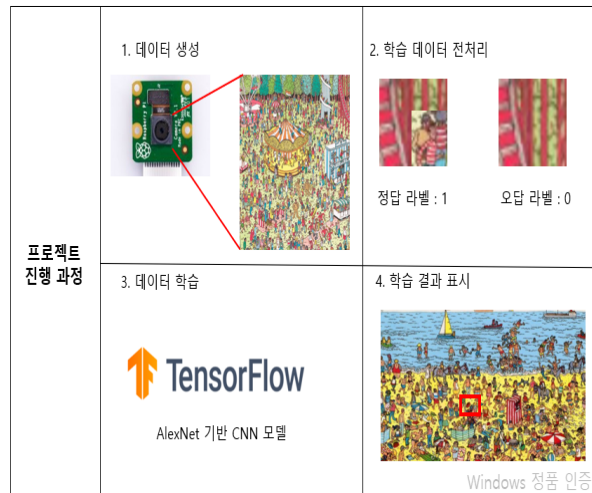
교육 및 자격증

- [모두의 연구소 AIFEL 대전 2기 딥러닝 자연어 처리 과정 이수증](#) (2021.12.07 ~ 진행중)
- [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)

경력

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

- 사용 기술
 - Python 3.x, Tensorflow 2.x
- 프로젝트 소개
 - 고등학생 대상 인공지능 교육용 모델 <AI-KIT로 배우는 인공지능 기초> 프로젝트 참여하여, Tensorflow 패키지를 활용하여 율리를 찾는 object detection 모델 구현
 - 데이터 수집 목적의 라즈베리파이 카메라 환경 설정 및 모델 배포
- 프로젝트 진행 과정



- 수행 역할
 - 모델 issue 관리
 - 같은 이미지라도 다른 환경으로 찍은 사진에서 성능이 떨어지는 issue
 - imgaug 패키지를 통해서 다양한 효과가 적용된 훈련 데이터 생성
 - 적용된 imgaug 효과는 gaussian blur(이미지 노이즈 제거), brightness(밝기 조절), contrast(대비 조절)
 - 서로 비슷하게 생긴 2개의 object를 구분하지 못하는 issue
 - 기존 학습 데이터에 공통점을 제거하여 차이점에 주목하는 모델을 만들어 원하는 object를 찾게 만듦

- (주) 데이터에듀 기업 현장실습 1개월 (2019.07.22 ~ 2019.08.16)

- 사용 기술
 - Python 3.x, Selenium
- 프로젝트 소개
 - 언어모델 구축을 위한 text data 수집
 - 파이썬 관련 내용 정리 : [점프 투 파이썬 독학](#)
 - 파이썬의 Selenium 패키지를 활용한 빅카인즈 사이트 데이터 스크래핑
- 수행 역할
 - Selenium 패키지를 활용한 raw text data 수집