

개요

- 한양대학교 기계공학부 전공(2019.02 졸업)
- Python 개발, 머신러닝, 딥러닝, 자연어처리
- SK하이닉스 양산기술 정규직 근무(2019.01~2020.03)

사용 가능 기술

- 프로그래밍 언어 : Python, SQL
- 분석 툴 : sklearn, Tensorflow, Pytorch, matlab
- 협업 툴 : github, slack, notion

관련 교육

- 혁신성장 청년인재 집중양성
인공지능 자연어처리(NLP) 기반 기업데이터 분석과정
- 멀티캠퍼스, 강남
2020.05~

관련 경험

- 텍스트 마이닝을 활용한 금융통화위원회 의사록 분석
- 20.07.10~20.08.03

한국은행에서 발간하는 ‘한국은행 금융통화위원회(이하 금통위) 의사록’에는 중앙은행이 시행하는 정책의 방향과 현 경제 상황에 대한 중앙은행의 판단이 포함되어 있다. 그러나 금통위 의사록의 문제는 매우 절제되어 있기 때문에 일반전인 독해로는 명확한 의미를 파악하는 것이 불가능하다. 따라서 텍스트 마이닝을 활용하여 금통위 의사록에 담겨있는 어조를 추출하여 수치화하고, 기준 금리의 변동과 얼마나 유사한지를 살펴보면서 지수의 설명력과 예측력을 검증한다.

맡은 부분

- Scrapy(crawling tool)을 이용해 20만개의 금리관련 뉴스기사 크롤링 후 전처리(tokenize, ngramize)
- 전체적인 pipeline 코딩
모든 수집된 데이터 통합/ 전월대비 콜금리 변동량을 통해 텍스트 문서들의 경향성(hawkish/dovish)라벨링/
라벨링된 문서에 나타난 토큰들의 빈도 수를 통해 token, ngram의 경향성 분류/
의사록의 문장&문서의 경향성 수치화/ 실제금리와 corr계산

- CNN 을 이용한 네이버 영화리뷰 데이터 감성분석
- 20.08.05~20.08.11

Convolution Neural Networks for Sentence Classification 논문을 바탕으로 자연어처리에 기존 RNN기반이 아닌 CNN구조를 활용 ,filter size로 ngram의 기능을 수행하며 좋은 성능낼 수 있음 보임.
위키피디아문서 기반인 Facebook의 Fasttext, 네이버영화리뷰 training set으로 word2vec, Contextualized Embedding 등의 Token embedding 방식 차이에 따라 성능을 비교해 보았다.

맡은 부분

네이버리뷰데이터 전처리(Konlpy이용)/fasttext embedding를 이용한 modeling(tensorflow keras)/
Github 코드, README정리

관련 경험

- Attention+seq2seq을 이용한 한영 번역 구현

20.08.13~20.08.19

Sequence to Sequence Learning with Neural Networks 논문 및 tensorflow 사이트를 기반으로 attention 기법이 포함된 seq2seq 한영변환을 구현해보고 BLEU로 성능을 테스트 해보고, 논문리뷰 및 코드 정리를 하였습니다. korpus는 aihub의 한국어-영어 번역 말뭉치를 사용하였습니다.

맡은 부분

-tensorflow 코드 기반으로 전체 flow modeling/

input 문장의 전처리 방법(글자단위, Okt, mecab) input문장배열순서 (정방향/역방향-논문 참고), 은닉층의 초기화방법을 달리하면서 번역기의 성능을 높여나감/ github 정리

- Big contest 공모전

20.08.23.~

분야 : 감염병으로 인한 소비/경제/행동 변화, 사회적 영향 분석을 통해 뉴노멀시대의 서비스 아이디어 제시
주제 : 인구 밀집도 기반 카페/음식점 추천 서비스

맡은 부분

- 신한카드 이용데이터 EDA - 코로나 전후 연령별 카테고리별 변화 분석
- 공공데이터 포털 서울시 상권데이터, 지하철 하차인원데이터를 이용하여 코로나 전후 동별 상권 수 변화량과 지하철역 하차 인원수 변화량과의 상관관계 분석
- google map api이용하여 상권정보 및 이용자 데이터 수집 및 데이터 전처리
- 지도위에 특정구역 위 격자화를 하여 격자 별 통계계산(상권 수, 혼잡도)
- Folium 이용하여 지도 visualizing, 추천 장소 list up 구현