

딥러닝을 공부하는 수학자입니다

윤도연

Content



AI Project

Senticle : 뉴스 데이터를 활용한 기업별 주가 예측 프로젝트

심장 CT영상 이미지 분할 Contest

Samsung Semiconductor Smart Interlock

ETC Projects and Studies

1. Senticle : 뉴스 데이터 기반 주가 예측

LINK : <https://github.com/ydy8989/senticle-proj>



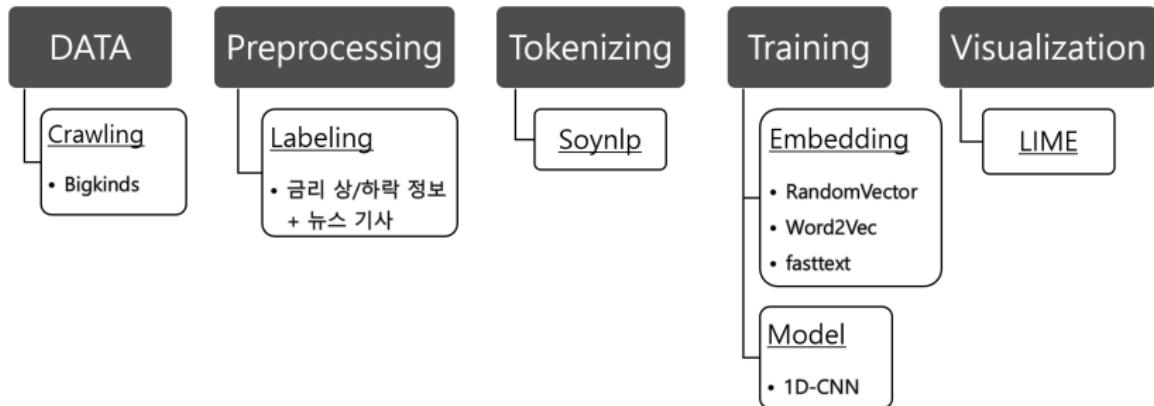
문제 정의 및 데이터셋

- ▶ 특정 기업의 당일 뉴스 데이터를 통해 익일 주가를 예측하는 모델 구현 & 앱 서비스
- ▶ 기업별 뉴스 데이터 5년치 크롤링 후 정제를 통한 학습 데이터 구축

사용 기술

- ▶ NLP, RNN(LSTM), Tensorflow, Crawling

전체 프로세스



1. Senticle : 뉴스 데이터 기반 주가 예측

LINK : <https://github.com/ydy8989/senticle-proj>



3

수행 역할(참여 인원: 안드로이드, 백엔드, ML 교육생 등 4명)

- ▶ 전처리(뉴스 데이터 레이블링, 토크나이징 일부)
- ▶ 임베딩 및 모델링

결과

- ▶ 당일 뉴스 기사 학습을 통한 익일 주가 예측 (ACC 72%)
- ▶ LIME 시각화로부터 확률값 제시를 통한 모델 신뢰성 객관화

"中 감산폭 완화로 큰 가격 약화 없을 것"-포스코 컨콜

"당초 중국에서 최초 발표하기에는 동절기 감산 기간을

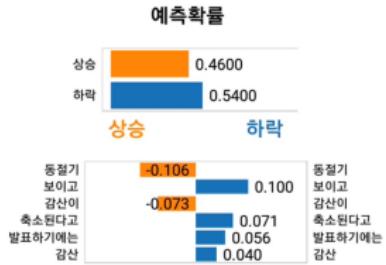
작년보다 늘어난 6개월로 확대한다고 발표했으나 최근 미·중

무역분쟁 영향으로 약화되는 현상을 보이고 있다. 감산이

완화된다는 이유로 인해 실망감이 표출이되며 가격이 약세를

보이고 있다. 다만 4분기는 전통적인 성수기로 수요가

증가하는 특성상 가격은 상승할 것으로 예상된다.



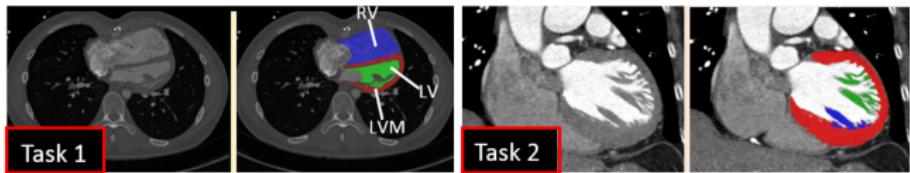
2. 심장 CT이미지 분할 Contest

github.com/ydy8989/Cardiac_Segmentation



문제 정의 및 데이터셋

- ▶ 3차원 CT image에서의 병변 검출을 위한 범용 장기 분할(Organ Segmentation) 모델의 성능을 겨루는 컨테스트
- ▶ 모델을 통한 이익 ⇒ 진단까지 걸리는 비용과 시간 절감 + 진단 주관성 배제
- ▶ 병변을 보유한 환자 100여명의 원본 CT이미지와 해당 병변에 label된 CT이미지



사용 기술

- ▶ Keras, 3D-UNet, Docker, Segmentation, SimpleITK

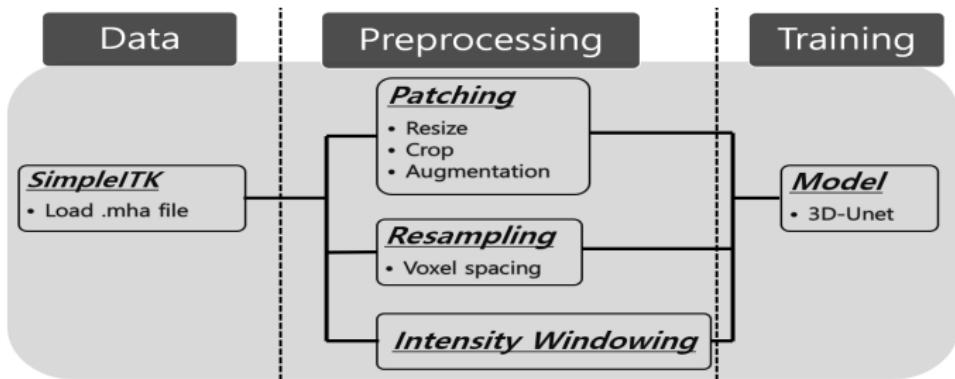
2. 심장 CT이미지 분할 Contest

github.com/ydy8989/Cardiac_Segmentation



5

프로세스

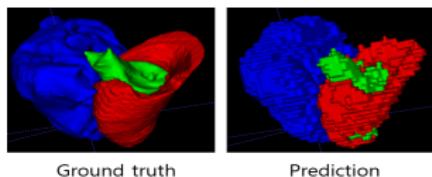


수행 역할(참여 인원: 인턴 교육생 3명)

- ▶ 전처리(resampling, intensity windowing)
- ▶ 후처리(Inference시 test data set 전처리 fine tuning)

결과

- ▶ acc 71%, 의료 이미지 영상 처리 학습 기반 마련



3. Samsung Semiconductor Smart Interlock

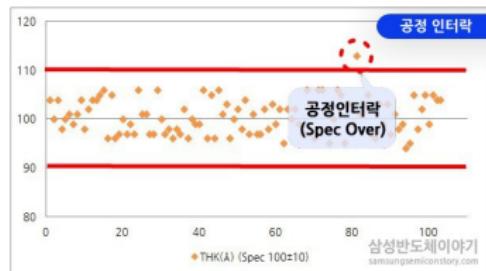
Hbee Co. AI team, Machine learning engineer



Process

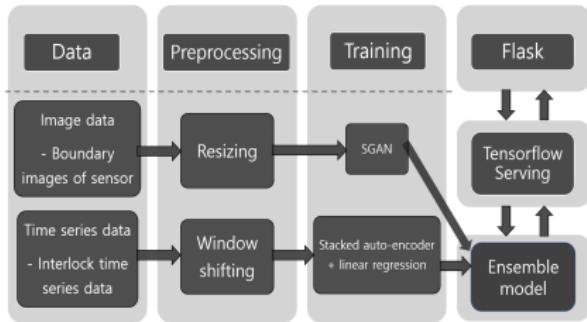
Motivation

센서값이 boundary 밖으로 벗어날 때
자동으로 공정이 멈추는 "인터락" 상황 발생



진성 실제 공정의 결함으로 인한 중지 ⇒ 빠른 수리 및 수율 보전을 위한 재가동 시급

가성 센서 오작동으로 인한 공정 중지 ⇒ 수율 손실 최소화를 위한 공정 재가동 시급



수행역할

▶ 센서 시계열 모델 구현 전 과정 + 전처리를 제외한 센서 이미지 모델 구현

성과

▶ 현재 삼성전자 반도체 공정라인 80% 가량에 적용되어 사용중

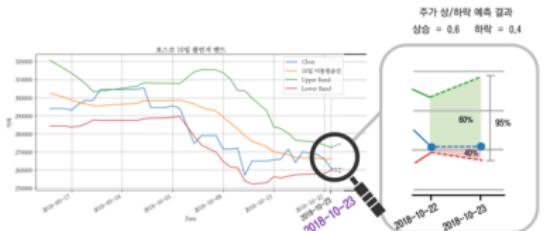
ETC Projects and Studies



7

▶ 주가 예측 백테스팅 프로그램

1. RNN, LSTM을 활용한 주가 시계열 예측
2. 기업별 이격도 선정을 통한 백테스팅 시스템 구현
3. 이동평균선과 볼린저 밴드를 예측 신뢰도 제공



▶ 블로그 포스팅

1. 관심분야(GNN, 추천시스템) 관련 글 포스팅
(링크 :<https://ydy8989.github.io>)



Semi-Supervised Classification with
Graph Convolutional Networks(GCN) 논
문 리뷰

온지 6년 약 3000 헤이트스피어 글을 처리하는 Graph neural networks 분야에서 가장 잘 알려진 Semi-Supervised Classification with Graph Neural Networks(SGC)에 대해 살펴보았다. 그 외에 다른 학습 방법인 GAT, GraphSAGE, GraphConv, GraphSAGE++ 등에 대해서도 살펴보았다. 특히 SGC와 GAT는 각각 유망한 분야로 보고자 한다.

Thank You!