

Domain에 강한 데이터 분석가 임태규



Projects

- AI를 활용한 조미니 경도 예측 프로젝트(2021.07~2021.08)
 - 모델개발 : pandas, numpy, catboost, xgboost, lightgbm, tensorflow, hyperopt
 - 모델배포 : flask
 - 프로젝트 주소 : https://github.com/TaeKyulm/seah_project

Projects

- WeaTo(2022.07.01~2022.07.23)
 - 해당 지역의 실제 날씨를 알 수 있는 웹페이지 개발
 - 사용stack : selenium, django, django-restframework
 - 프로젝트 관리 : notion, git
 - selenium을 사용하여 네이버날씨의 정보를 크롤링하고 이를 react 프로젝트에서 rendering 할 수 있도록 json 형식으로 결과를 변환해주는 코드를 작성하였음.
 - 사용자가 작성한 메모에 대한 crud 모델을 만들었음
 - 프로젝트 주소 : <https://github.com/TaeKyulm/WeaTo>

Skill

- Major
 - 4학년 1학기 기준 전체평점 4.33/4.5 전공 평점 4.42/4.5로 제조업에서 사용되는 도메인 분야의 내용에 대해 흥미가 높습니다.
 - 기계공학과 물리학 복수전공을 통해 철강, 공작법, 반도체, 디스플레이 등 다양한 산업군에서의 변수에 대해 이해를 할 수 있습니다.
 - 이를 기반으로 새로운 파생변수 생성, 또는 통계학적인 eda로만 알 수 없는 가설에 대해 검증을 해보는 등 다양하게 활용할 수 있습니다.

Skill

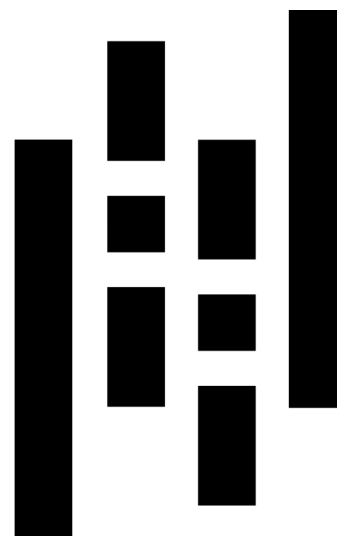
- Communication

- 전공 멘토 동아리 활동을 통해 다양한 사람 앞에서의 발표를 좋아합니다.
- 전공 TA 조교 활동 및 코딩 동아리 활동을 통해 어려운 내용을 남들에게 쉽게 설명하는 것에 자신이 있습니다.
 - 미적분학, 통계학을 이수하지 않는 팀원에 대해 Generative Model의 원리에 대해 이해를 시킨 경험이 있고, ta활동을 통해 기계공학 전공을 베이스가 없는 학생들을 이해시킨 적이 있습니다.

Stack

- pandas

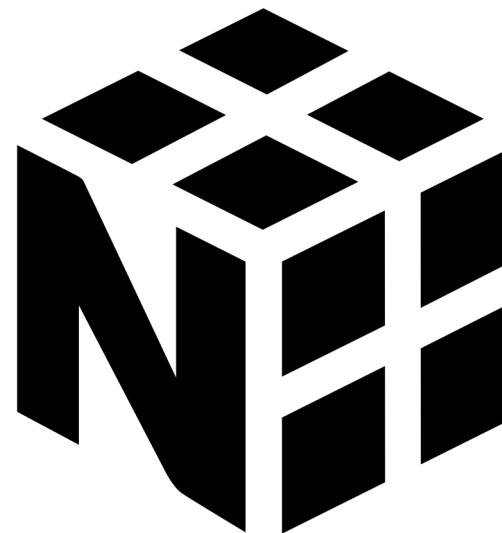
Tabular 데이터를 다루는데 익숙합니다. 날짜 형태의 데이터, 위치 정보 형태의 데이터 등을 처리해보았고, 여러 데이터들을 foreign key를 활용하여 join하여 새로운 데이터로 변환 가능합니다. 여러 열들을 적용하여 파생변수를 만든 데 익숙하고, 데이터의 전처리에 있어 도메인의 이해에 기반하여 진행합니다.



Stack

- numpy

numpy를 활용한 수학 연산에 익숙합니다.
딥러닝에서 Numpy를 활용한 평가지표 작성이나
차원 조절을 할 수 있습니다. 또한 학부수준의
선형대수학의 경우 다양한 수학적인 상황을
코드로 구현할 수 있습니다.



Stack

- scikit-learn

Scikit-learn을 활용하여 데이터의 전처리 및 비지도학습, 통계해석을 진행합니다. 가령 변수가 많은 경우 Lasso 회귀, Ridge 회귀 및 Random Forest를 활용하여 변수 선택을 하거나 다른 모델들을 활용하여 하나의 ensemble 모델을 만들 수 있습니다. 모델의 검증에 있어 k-fold 검증 코드를 작성하여 모델의 과적합 유무를 판단할 수 있습니다. 모델링의 경우 보통 gradient boosting 모델을 선호합니다.



Stack

- Hyperopt

기존 scikit-learn을 활용한 grid-search 방식의 파라미터 튜닝은 시간이 오래걸리고, global optimum 값이라는 보장이 없습니다.

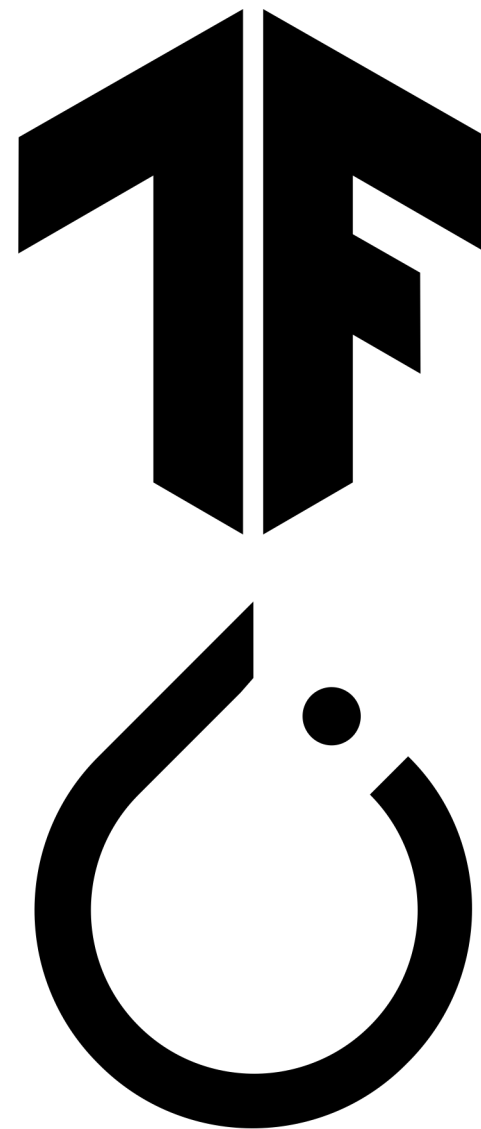
Hyperopt라이브러리를 활용하여 Bayesian 통계학에 활용한 파라미터 튜닝을 진행하여 빠르고 global한 최적의 파라미터 튜닝을 찾을 수 있습니다.



Stack

- tensorflow / pytorch

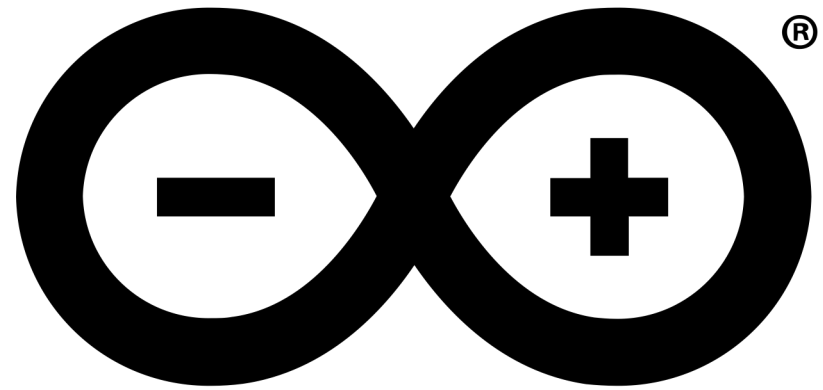
CNN, RNN, LSTM 등 다양한 layer에 대해 이해하고 있습니다. 또한 callback 함수들을 활용하여 robust한 모델을 만들 줄 알고, pre-trained 모델을 활용하여 fine tuning을 할 수 있습니다. 최근에는 논문 구현을 통해 실제 문제를 적용해보는 연습을 해보고 있습니다.



Stack

- arduino

학부시절 관련 교과목을 들으며 다양한 센서에 맞는 제어 코드를 작성할 수 있습니다. 적외선 센서를 활용한 자율주행, python을 활용한 6자유도 로봇암 설계 등 다양한 프로젝트를 진행하였습니다.



Github

- <https://github.com/TaeKyulm>

Education

- 강북고등학교 : 자연계열
-2013.03~2016.02
- 영남대학교 기계공학부(물리학 복수전공)
-2017.03~2023.02
-현재 졸업예정
-평점 : 4.32 / 4.5

About Me

- 개발태도
 - 다양한 기술들을 실제 산업의 문제에 적용을 하려 합니다.
 - 내가 아닌 실 사용자가 만족하는 결과물을 만들려고 합니다.
 - 부족함을 인정하고 늘 새로 배울려고 합니다.
 - 학부시절 동아리 활동 등을 통해 남의 말을 경청하는 법을 배웠습니다.
 - 꾸준한 독서 및 다양한 연사와의 만남을 통해 산업이 나아가는 방향, 내가 나아가야 하는 방향을 늘 연구합니다.

About Me

- 목표

- Automative System : 사용자가 통계학, 머신러닝 등을 몰라도 잘 활용할 수 있는 산업 서비스를 만드는 것이 목표입니다.
- Efficient System : 기존 산업군에서 품질 보증 등을 이유로 비효율적인 실험 등이 많이 진행되었습니다. 이를 줄이고 생산성을 극대화 시키는 것이 목표입니다.
- User-friendly : 현장의 엔지니어들은 신기술의 어려움 및 본인들의 일자리 대체 등을 이유로 새로운 시스템 도입을 거절하는 경향이 있습니다. 이들도 좋아하는 서비스를 만드는 것이 목표입니다.

About Me

- 임태규란?
 - 취미 : 카페 탐방, 맛집 탐방, 운동
 - 관심사 : 투자, 트레이딩, 자동차