

Daylight Entrance

Daniel Rybakken
danielrybakken.com



Photography by Kalle Sanner & Daniel Rybakken

32

이현경

shimmerlee13@gmail.com



이현경 Profile

수료교육

이미지 분석 기반의 인공지능 플랫폼 개발자 양성과정

정보통신기획평가원 주관

2020.06.24 ~ 2020.12.04



RDBMS JAVA Web

R분석 Python

머신러닝 딥러닝

- Oracle & SQL
- JSP/Java/JDBC
- HTML/CSS/JQuery
- Spring/ORM(Mybatis)
- R
- Python
- Numpy/Pandas
- Scikit-learn
- Tensorflow
- Keras

자격증

정보처리기사

한국산업인력공단

Link

Github

<https://github.com/shimmerlee13>

Kaggle

<https://www.kaggle.com/shimmerlee13/notebooks>

개요 Skill Inventory

<https://github.com/shimmerlee13/>

기간	프로젝트명	프로젝트 개요	기술	담당업무
2020.08.05 ~ 2020.08.13	[팀 프로젝트] 서울시 아파트 실거래 현황 및 가격지수 정보 시스템	<ul style="list-style-type: none"> · 아파트 데이터 분석 및 시각화 · Ajax를 사용한 비동기 통신 구현 · 데이터 크롤링을 통한 실시간 부동산 뉴스 제공 	<ul style="list-style-type: none"> · DB : Oracle · Language: Java/JSP, Spring, JQuery, UI/UX Template(Bootstrap) · Open API : Google APIs (map, gocoder) +Restful, Jsoup · Framework : Spring, MyBatis · CM : Git 	<ul style="list-style-type: none"> · 국토교통부 서울 아파트 거래 가격 이용 DB 설계/모델링 · 서울시 아파트 매매/전세 가격 변동 분석 및 시각화 · Spring Framework를 사용한 웹 서비스 개발
2020.09.17 ~ 2020.12.04	[분석모델 Kaggle] 빅데이터 분석을 위한 머신러닝 / 딥러닝 모델 개발	<ul style="list-style-type: none"> · 데이터 전처리 · 분석모델 개발 · 평가/검증/시각화 	<ul style="list-style-type: none"> · Python (+Pandas, +Numpy) · OpenAPI : Sklearn, Tensorflow, Keras, OpenCV · RDBMS : Oracle, SQLite3 	<ul style="list-style-type: none"> · LogisticRegression, DecisionTree, RandomForest, XGBoost, LightBGM, GradientBoostingRegressor, · PCA, K-means, t-SNE, DBSCAN · AlexNet, InceptionV3, ResNet, vgg16 · 오픈 데이터 활용 모델링, 학습, 예측, 평가
2020.10.28 ~ 2020.10.30	[팀 프로젝트] BC Card 주최 2020 금융 빅데이터 분석 챌린지	<ul style="list-style-type: none"> · 코로나, 금리 정책, 부동산 정책별 소비에 미친 영향도 분석 · 카드 매출액 변화, 소비 트렌드, 성장성 예측 	<ul style="list-style-type: none"> · Language: Python (+sklearn) · Open API : Json, Numpy, Pandas, Seaborn, Matplotlib 	<ul style="list-style-type: none"> · 유동인구 데이터 시각화 · 업종별/지역별 카드 사용량 분석 및 시각화
2020.11.02 ~ 2020.11.23	[팀 프로젝트] CNN 이미지 인식 기반 개인화 맞춤 레시피 추천 서비스	<ul style="list-style-type: none"> · 재료 이미지 인식 모델 개발 · 개인화 맞춤 레시피 추천 서비스 구현 	<ul style="list-style-type: none"> · DB : Oracle · Language: Python +sklearn +TensorFlow +Keras +OpenCV, UI/UX Template(Bootstrap) · Open API : Selenium, BeautifulSoup, Numpy, Pandas, Matplot · Framework : Flask 	<ul style="list-style-type: none"> · 공공데이터포털의 레시피 데이터 활용 DB 설계 및 구축 · Contents-based Filtering 추천 알고리즘 개발

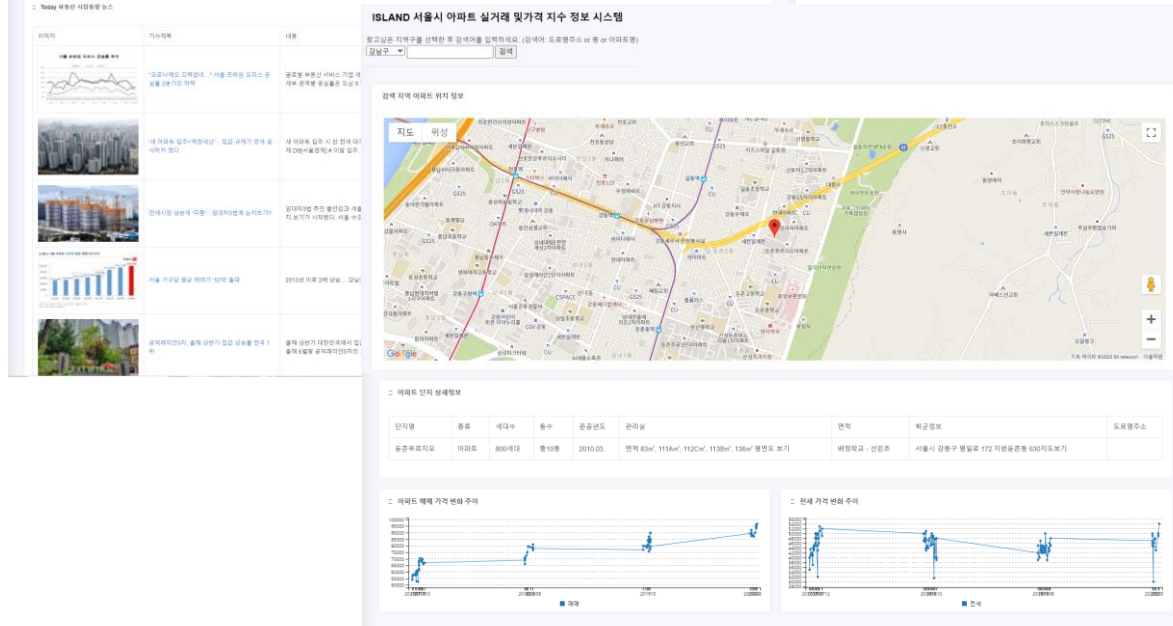
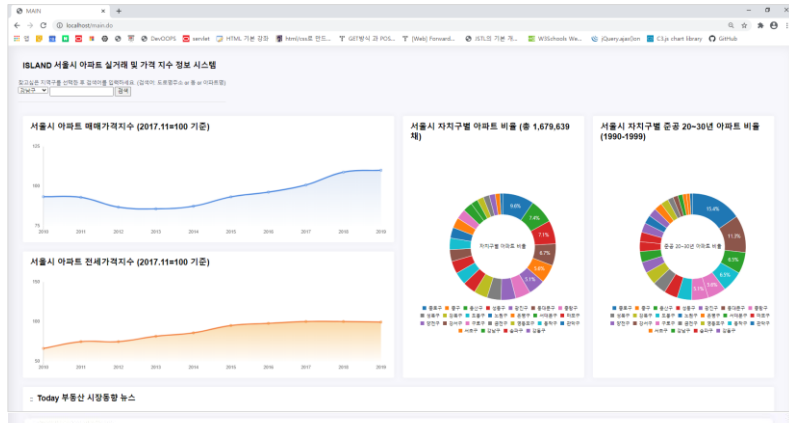
분석 모델 Kaggle

<https://github.com/shimmerlee13/mlidl>

Subject	Model	Metrics	Data set
Boston House Prices	Regression [Ridge / Lasso]	rmse : 4.772	506 rows , 13 cols
House Prices: Advanced Regression Techniques	Staking Ensemble [Ridge+ Lasso+XGBRegressor+LGBMRegressor]	rmse : 0.1164	1460 rows, 81 cols
Bike Sharing Demand	RandomForestRegressor GradientBoostingRegressor XGBRegressor LGBMRegressor	rmsle: 0.347, rmse: 48.226, mae: 30.400 rmsle: 0.335, rmse: 52.964, mae: 33.779 rmsle: 0.349, rmse: 49.941, mae: 31.688 rmsle: 0.318, rmse: 45.093, mae: 28.187	10886 rows, 12 cols
Santander Customer Satisfaction	XGBClassifier LGBMClassifier	roc-auc: 0.8453 roc-auc: 0.8417	76020 rows, 371 cols
Titanic: Machine Learning from Disaster	DecisionTreeClassifier RandomForestClassifier LogisticRegression	auc : 0.7877 auc : 0.8324 auc : 0.8659	891 rows, 12 cols
Credit Card Fraud Detection	LogisticRegression LGBMRegressor	f1: 0.7442, roc-auc:0.9643 f1: 0.8710, roc-auc:0.9833	284807 rows, 31 cols
Pima Indians Diabetes Database	LogisticRegression	f1 : 0.6931	768 rows, 9 cols
Digit Recognizer(MNIST)	Classification : RandomForestClassifier DNN / CNN[AlexNet, InceptionV3, ResNet]	oss : 0.0024 , acc : 0.9723	1797 rows, 64 cols
Breast Cancer Wisconsin (Diagnostic) Data Set	Classification XGBoost / LightGBM	f1 : 0.9677 f1 : 0.9615	357 rows, 212 cols
Custom Segmentation	Clustering [K-means/t-SNE/DBSCAN]	silhoutte : 0.303	541909 rows, 8 cols
TMDB 5000 Movie Dataset	cosine_similarity	CountVectorizer	4803 rows, 20 cols
MovieLens Latest Datasets	cosine_similarity	mse: 9.895	9742 rows, 3 cols
Cat and Dog	CNN[vgg16]	epoch:100 , batch-size:32	image : 217.88 MB

서울시 아파트 실거래 현황 및 가격지수 정보 시스템 Team Project

2020.08.05 ~ 2020.08.13



OVERVIEW

서울시 아파트 실거래 현황 및 가격지수 정보 시스템

- 서울시 아파트 매매/전세 가격 변동 분석 및 시각화
- 자치구별 아파트 가격 분석 차트 및 실시간 부동산 뉴스 제공
- <https://github.com/shimmerlee13/homes>

적용 기술

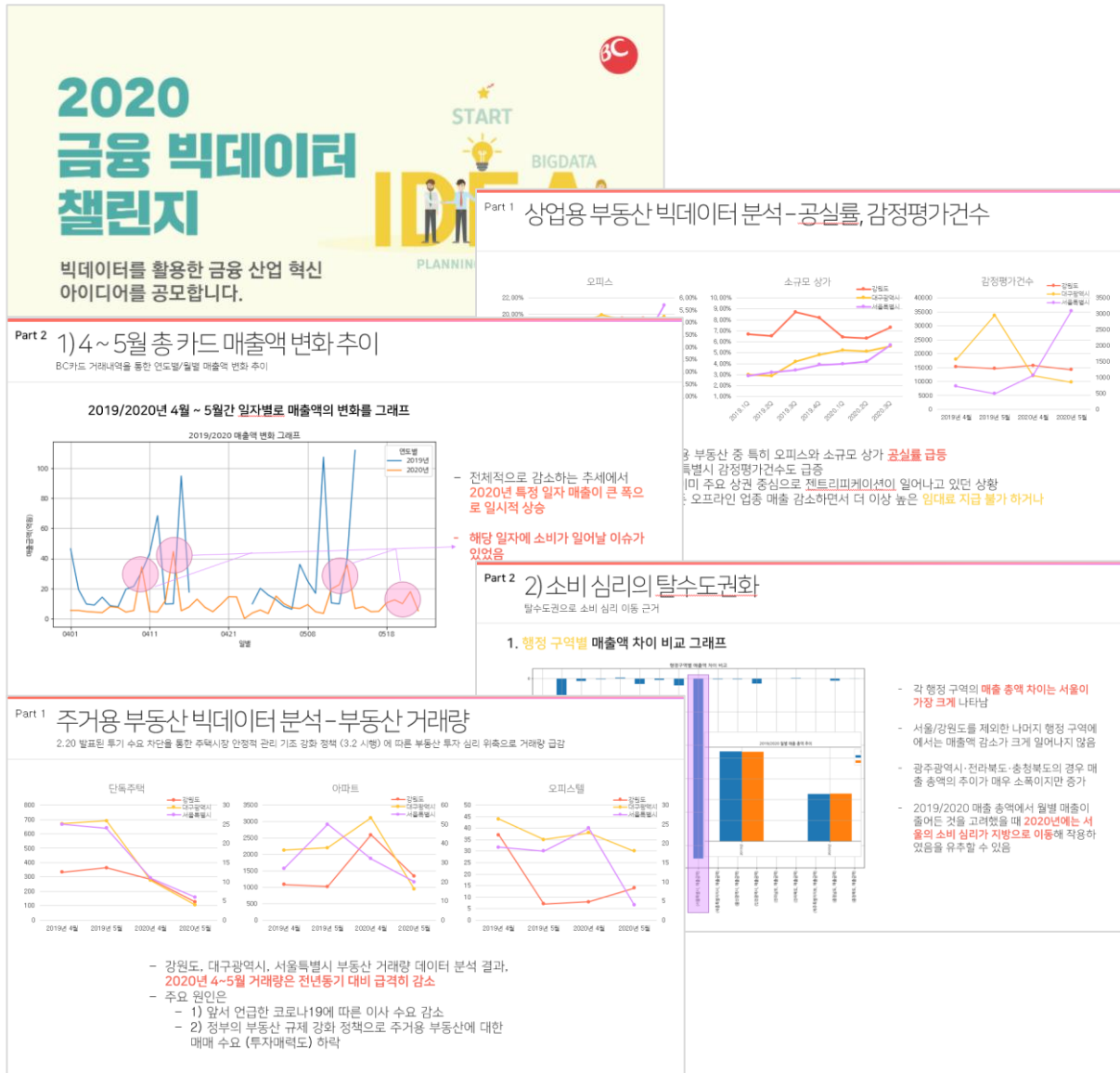
- DB : Oracle
- Language: Java/JSP, Spring, JQuery, UI/UX Template(Bootstrap)
- Open API : Google APIs (map, gocoder) +Restful, Jsoup
- Framework : Spring, MyBatis
- CM : Git

담당 업무

DB 설계 및 구축, 아파트 가격 분석, 차트 시각화, 프론트엔드 웹 구현

BC Card 주최 2020 금융 빅데이터 분석 챌린지 Team Project

2020.10.28 ~ 2020.10.30



OVERVIEW

BC Card 주최 2020 금융 빅데이터 분석 챌린지

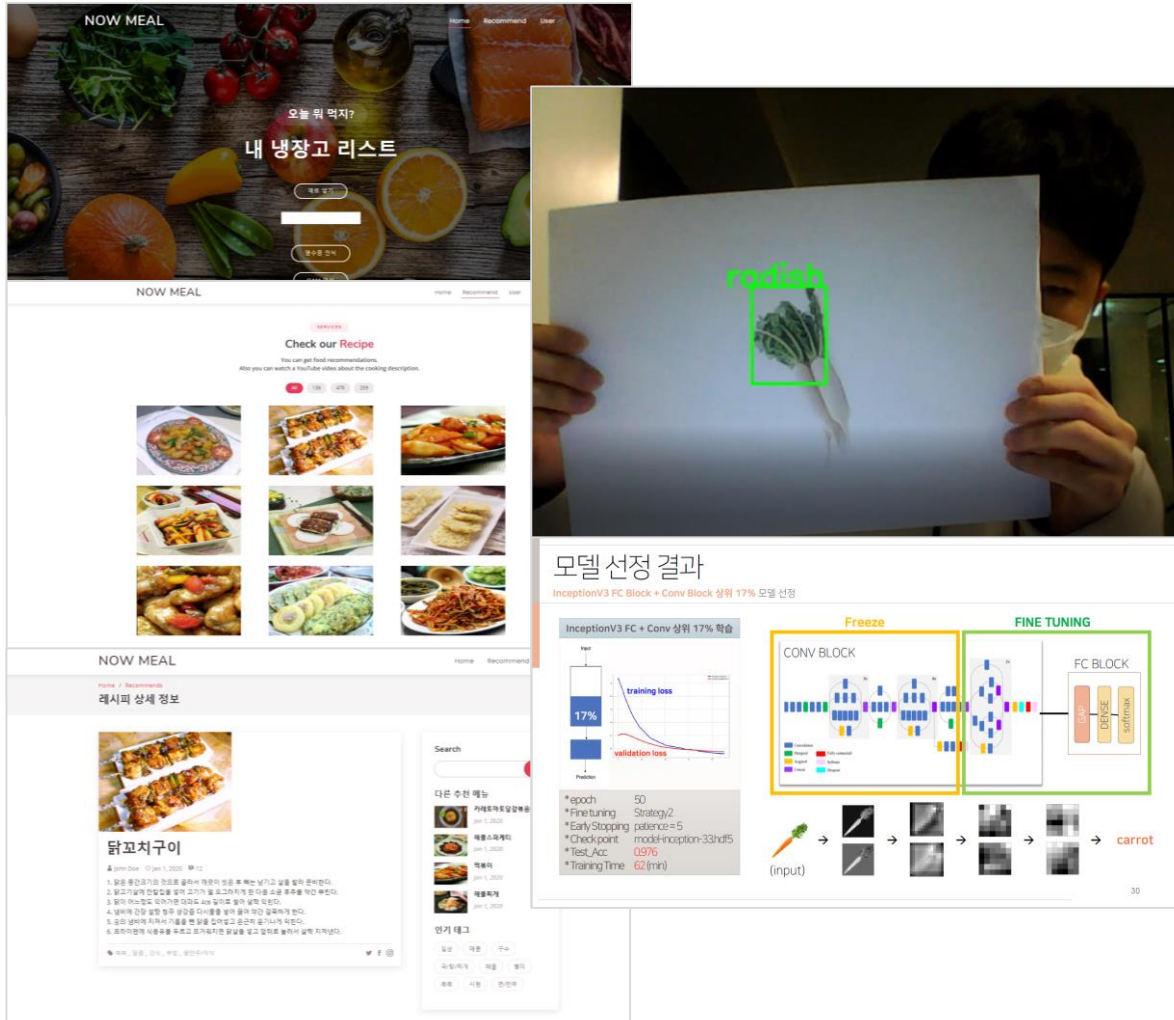
- 코로나, 금리정책, 부동산 정책 별 소비에 미친 영향도 분석
- 카드 매출액 변화, 소비 트렌드, 성장성 예측
- https://github.com/shimmerlee13/bc_card_2020

적용 기술

- Language: Python +sklearn
- Open API : Json, Numpy, Pandas, Seaborn, Matplotlib

데이터셋

- 다음소프트 키워드 버즈량
- 기웅 아파트 세대 거주기간/전출입/시세 정보
- 비씨카드 거래내역
- 망고플레이트 인기식당
- 케이티 유동인구일집계
- SBCN 주식일간기본지표
- 한국감정평가사협회 부동산 거래/평가이력



OVERVIEW

Now Meal : CNN 이미지 인식 기반 개인화 맞춤 레시피 추천 서비스

- 사용자 별 재료 이미지 인식 및 관리
- 취향/재료/요리 히스토리를 반영한 개인화 맞춤 레시피 추천 서비스 구현
- <https://github.com/shimmerlee13/nowmeal>

적용 기술

- DB : Oracle
- Language: Python +sklearn +TensorFlow +Keras +OpenCV, UI/UX Template(Bootstrap)
- Open API : Selenium, BeautifulSoup, Numpy, Pandas, Matplot
- Framework : Flask

담당 업무

Project Manager, Database 설계 및 구축, 추천 알고리즘 개발