



분석으로 유익한 사회를 !

김동우입니다

✉ zxcvbnm1997@hanmail.net

☎ 010 3106 7593

🐙 <https://github.com/ddooom>

About 김동우

목차



데이터 기반의 분석

데이터 기반의 분석
EDA, 데이터 시각화 경험



높은 러닝 레이트

새로운 지식을 빠르게 습득
자신만의 학습법



다양한 프로젝트 경험

정형 데이터부터 이미지, 텍스트까지
의사결정부터 예측 분류까지
머신러닝부터 딥러닝까지



그 외 강점들

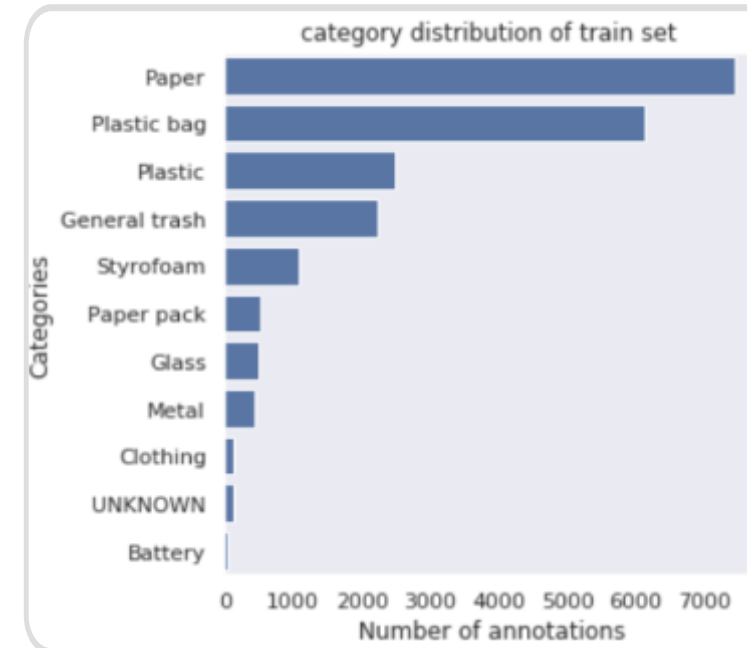
R, SQL 경험
통계기반 컴퓨팅 사고

Part 1

데이터 기반의 분석

"Garbage in, garbage out" 라는 말처럼 분석에는 데이터가 가장 중요합니다. 저는 제한된 데이터로 최상의 결과를 도출하기 위해 EDA와 시각화를 통해 데이터 이해, 구조 파악을 우선시했습니다.

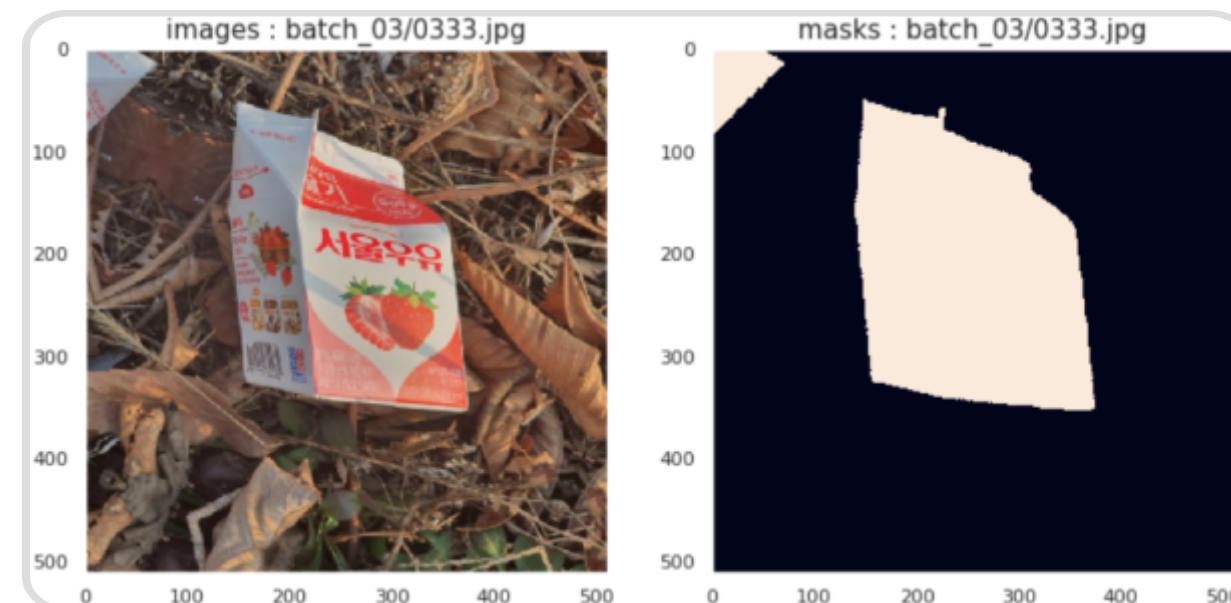
모델과 파라미터 기반이 아닌 데이터와 통계적 사고를 기반으로 분석합니다.



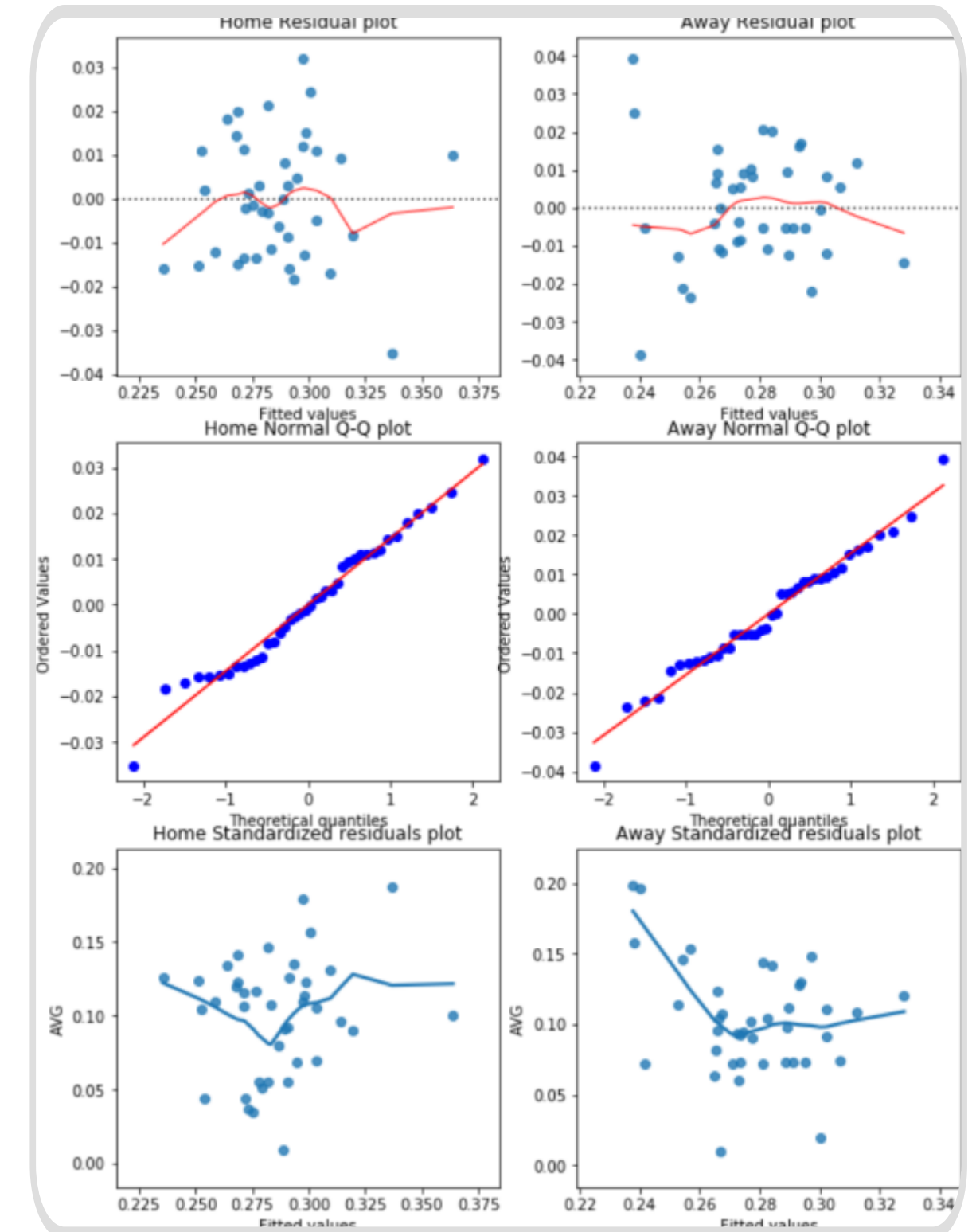
Class Imbalance

	width	height	ratio	area
count	21116.000000	21116.000000	21116.000000	21116.000000
mean	108.645776	112.751208	1.277337	18051.901280
std	90.702375	92.089715	0.999388	28614.262521
min	3.400000	2.000000	0.016807	27.740000
25%	40.400000	40.500000	0.690845	1751.712500
50%	83.000000	85.000000	1.052390	6854.155000
75%	147.500000	161.500000	1.564431	21905.692500
max	511.900000	511.900000	26.128571	262041.610000

Bounding Box Statistics



Data Augmentation



Data Assumption for Regression

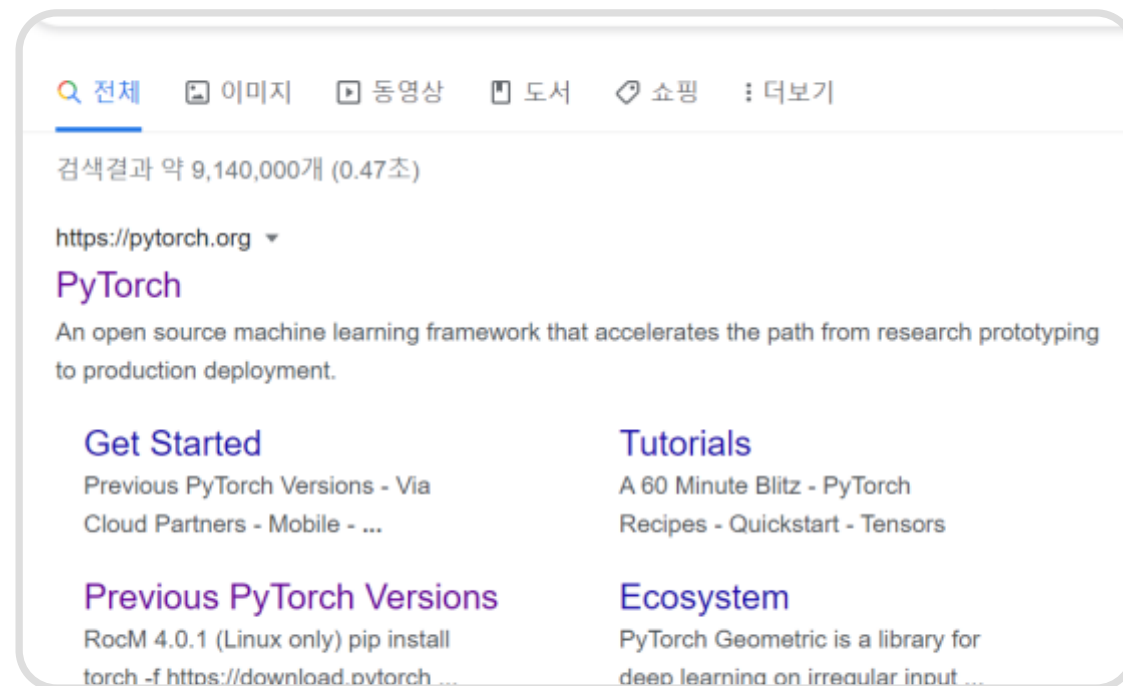
Part 2

높은 러닝 레이트

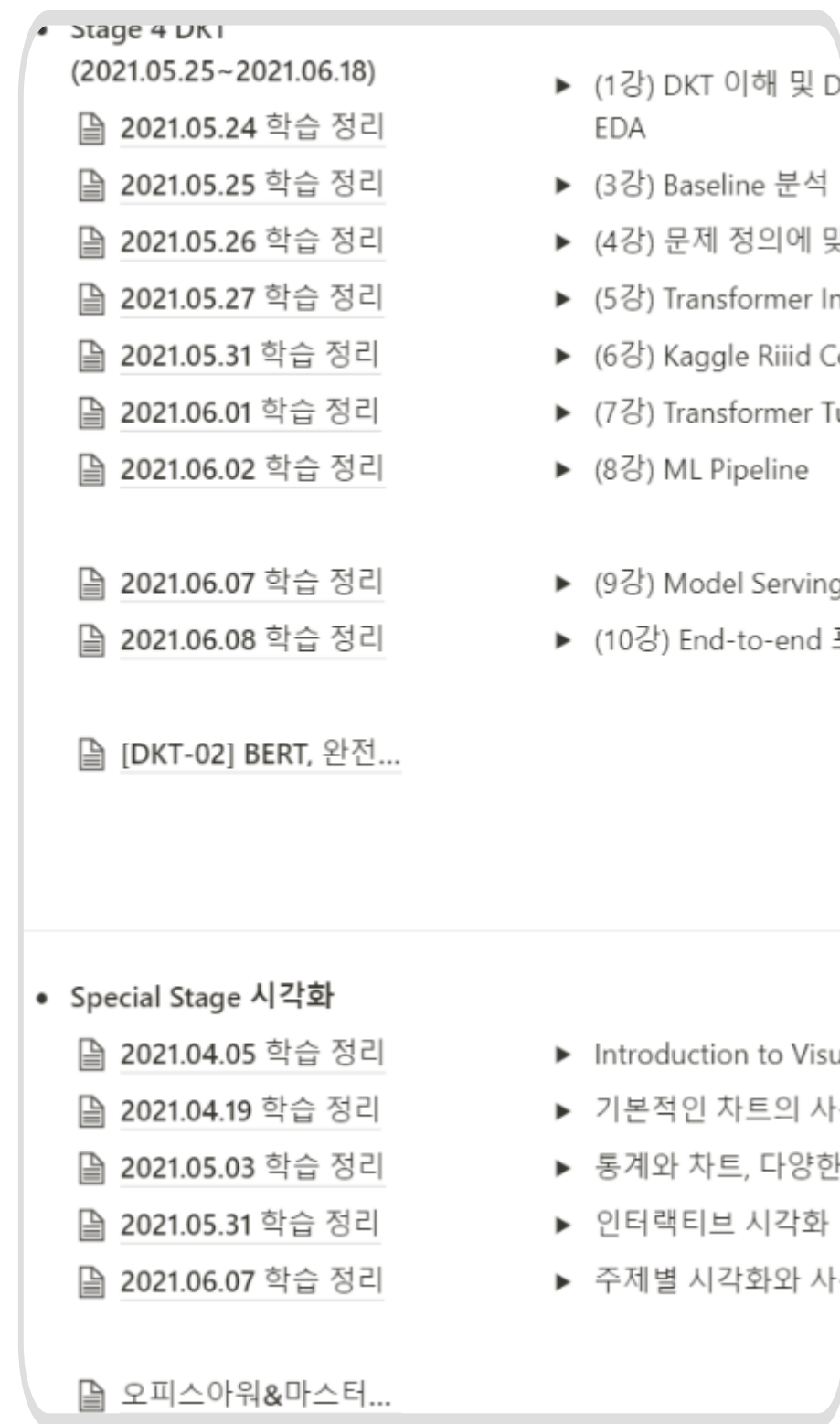
네이버 부스트캠프 AI Tech에서 파이썬부터 딥러닝 프레임워크까지
매주 새로운 것을 접하였습니다.

이를 통해 새로운 지식에 대한 두려움을 극복하고 빠르게 흡수하는
저만의 학습법을 만들었습니다.

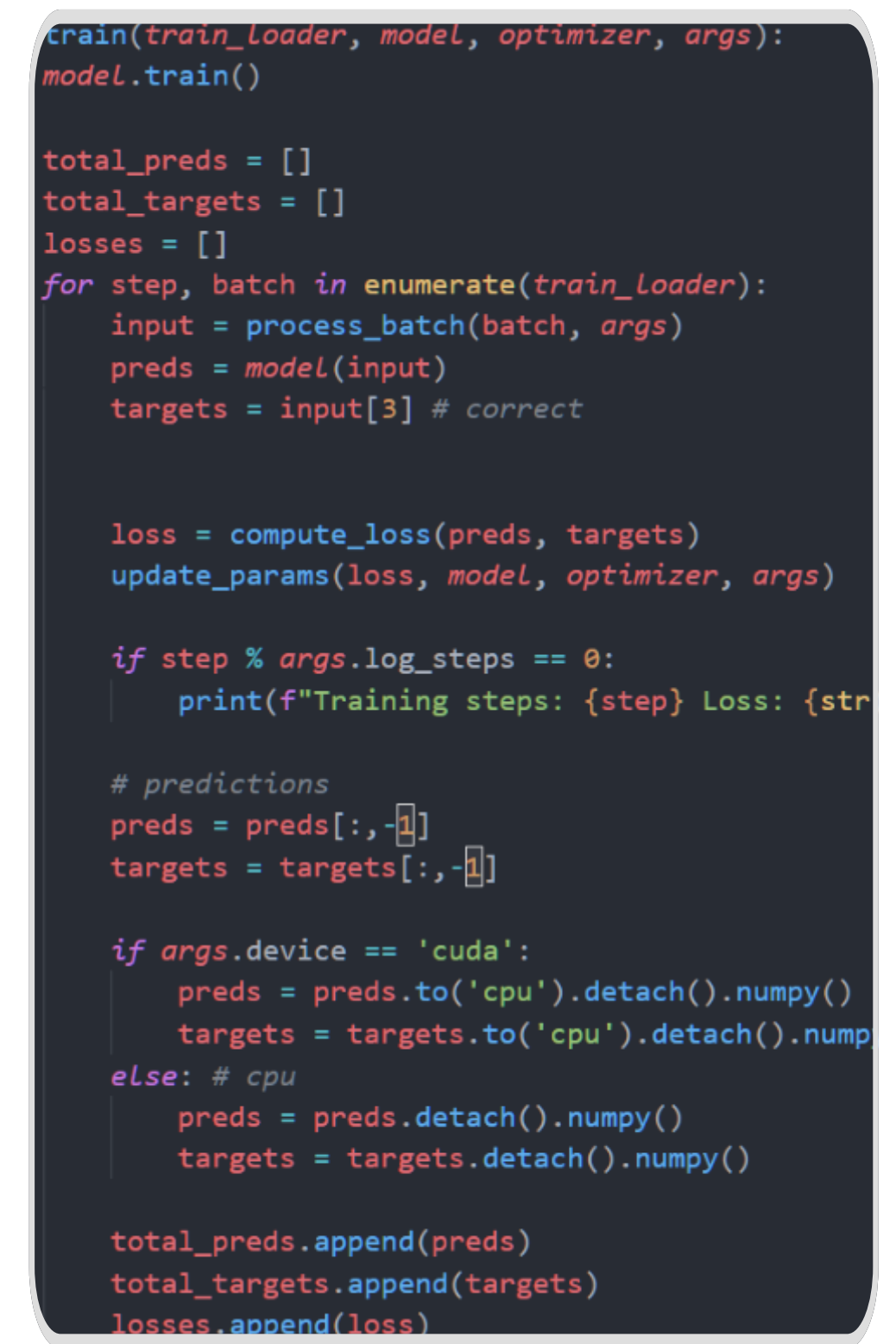
시행착오를 겪으며 높아진 러닝 레이트를 통해 낯선 프로젝트도 빠
르게 이행할 수 있습니다.



1. 탐색을 통한 개념, 툴의 이해



2. 학습한 내용과 예제 기록



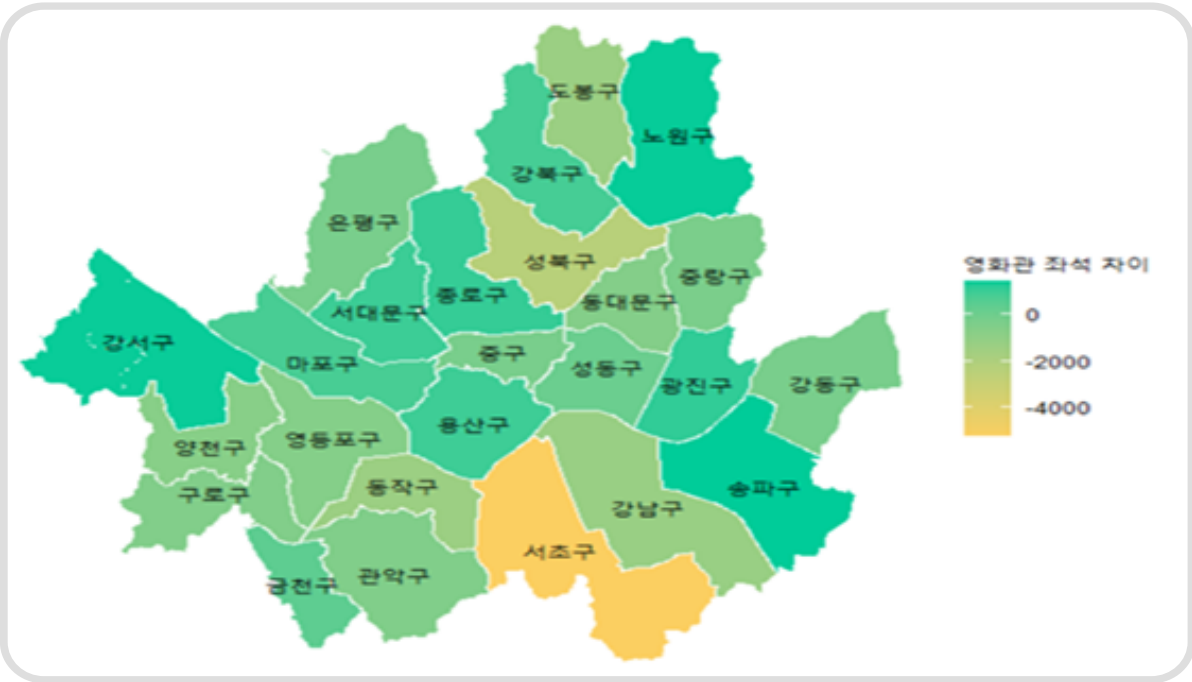
3. 학습 내용을 바탕으로 실습, 구현

Part 3

다양한 프로젝트 경험

회귀 분석을 통한 의사 결정과 머신러닝 모델을 이용한 정형 데이터 예측부터 딥러닝을 이용한 Computer Vision, NLP, Sequential Data 분석까지 다양한 분석 경험이 있습니다.

회귀, 머신러닝, 딥러닝 등 다양한 분석 경험을 바탕으로 다양한 시각에서 효율적인 분석을 진행할 수 있습니다.



회귀 분석을 통한 자동차 극장 입지 추천



딥러닝을 이용한 Object Detection

	pred_wp_rf	pred_wp_wgb
T_ID		
HH	0.379882	0.495114
HT	0.422385	0.497180
KT	0.504173	0.518345
LG	0.446562	0.509468
LT	0.452328	0.509172
NC	0.579986	0.539819
OB	0.570334	0.474821
SK	0.401247	0.460703
SS	0.475578	0.500070
WO	0.549309	0.552092

머신러닝을 이용한 KBL 승률 예측

Part 4

그 외 강점들

R을 사용한 분석 경험이 있습니다.

소프트웨어 학부 복수전공으로 SQL을 사용하여 간단한 DB를 구성

하고 다양한 SQL 쿼리문을 작성해봤습니다.

복수전공으로 통계적 사고뿐만 아니라 컴퓨팅 사고를 기반으로 효율

적인 분석을 할 수 있습니다.

```
# 변수선택법
plot(final)

reg = lm(영화관좌석 ~ ., data=final)
step(reg, direction = 'both', scope=~1))

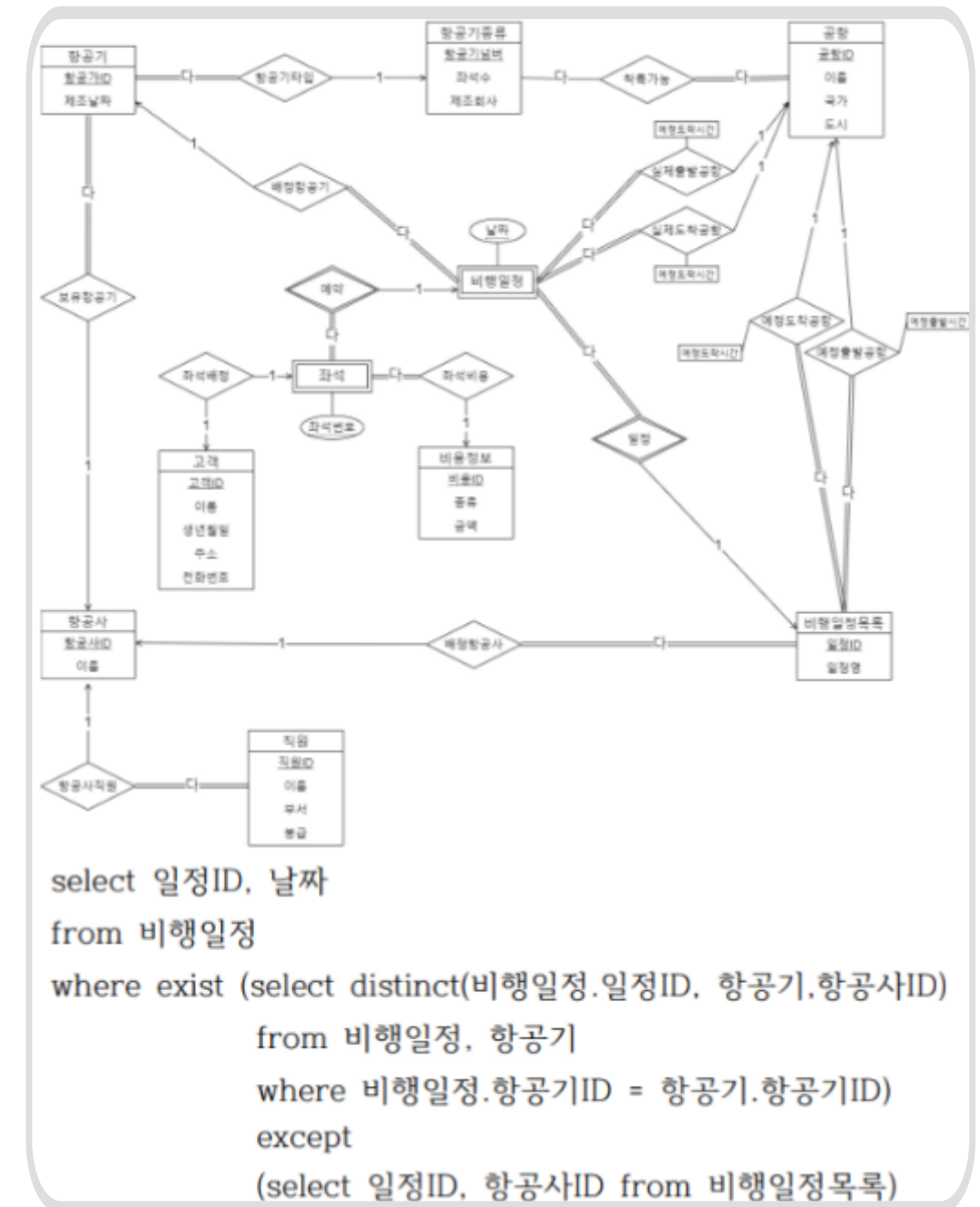
reg = lm(영화관좌석 ~ 도로면적 + 실거래가, data=final)
summary(reg)

standardized_residual = rstandard(reg)
plot(standardized_residual)
plot(data[, c('영화관좌석', '실거래가', '도로면적')])
```

R을 사용한 분석

- augmentation
- dataset
- inference
- loss
- model
- scheduler
- train
- utils

효율적 분석을 위한 모듈화



DB 설계와 SQL 쿼리문

저의 강점을 발휘하여
기업에, 사회에 도움이 되는
분석가로 성장하겠습니다.

감사합니다. 김동우였습니다!