

데이터 분석 교육 / 참석자 (WebEX) – 2021/02 ~ 2021/04

파일(F) 편집(E) 공유(S) 보기(V) 오디오 및 비디오(A) 참가자(P) 미팅(M) 도움말(H)

발언 중: 안건이

서태민
나

CTO_이세열

webex 171

공성우

박종열

수영 심

임현수

안건이

류기선

이수현 선임

최재우s

조호습

레이아웃

참가자 (12)

검색

서태민
나

CTO_이세열

webex 171

공성우

류기선

박종열

수영 심

안건이

이수현 선임

임현수

조호습

최재우s

참석자 정보

채팅

조호습 님이 모두에게: 오후 2:03
네

저장 위치: --- 선택하십시오 ---

여기에 채팅 메시지를 입력하십시오.

음소거 해제

비디오 시작

공유

...

참가자

채팅

...

□ 전체 공유화면

파일(F) 편집(E) 공유(S) 보기(V) 오디오 및 비디오(A) 참가자(P) 미팅(M) 도움말(H)

서태민 나 CTO_이세열 webex 171 공성우 류기선

레이아웃

안건이 님의 응용프로그램 ...

안건이/연구원/CXT센터 빅데이터실 빅데이터MI Task(gonie.ahn@lge.com)

삽입 디자인 전환 애니메이션 슬라이드 쇼 검토 보기

칼라내기

복사

붙여넣기

슬라이드

클립보드

레이아웃

다시 설정

구역

슬라이드

글꼴

단락

SmartArt로 변환

그리기

정렬

스타일

도형 채우기

도형 윤곽선

도형 효과

찾기

바꾸기

선택

Loss Function Remind

LGE Internal Use Only

- Regression Loss Function

미분 불가능

$$MAE = \frac{1}{N} \sum_{i=1}^n |\hat{y}_i - y_i|$$

Range of predicted values: (-10,000 to 10,000) | True value: 100

미분 가능

$$MSE = \frac{1}{N} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$

Range of predicted values: (-10,000 to 10,000) | True value: 100

고객과 함께 하는 미래 더 나은 삶을 향한 도전

참가자 (12)

검색

서태민 나

안건이

CTO_이세열

webex 171

공성우

류기선

박종열

수영 심

이수현 선임

임현수

조호습

최재우s

채팅

최재우s님이 모두에게: 오후 2:04

SDQ

최재우s님이 모두에게: 오후 2:04

넵

저장 위치: ---선택하십시오---

여기에 채팅 메시지를 입력하십시오.

슬라이드 2/0 15 한국어

손 슬라이드 노트 메모

오후 2:05 2021-03-11

음소거 해제 비디오 시작 공유

□ 이론 설명 (발표자 도구 포함)

▼ 채팅

天々天々天々

이수현 선임님이 모두에게: 오후 2:06

안녕하세요 ^^;

수영 심 님이 모두에게: 오후 2:06

𠄎𠄎𠄎

박종열 님이 모두에게: 오후 2:06

$$\frac{\pi}{2} \quad \frac{\pi}{2} \quad \frac{\pi}{2}$$

서태민 님이 모두에게: 오후 2:10

[질문] Bias는 훈련데이터에서 확인되고, Variance는 테스트나검증데이터에서 확인되는건가요?

[illegible]

데이터 분석 교육 / Python 실습 - 2021/02 ~ 2021/04

파일(F) 편집(E) 공유(S) 보기(V) 오디오 및 비디오(A) 참가자(P) 미팅(M) 도움말(H)

서태민 나

안전이

CTO_이세열

webex 171

공성우

류기선

레이아웃

JupyterLab

notebook.lgebigdata.com/user/gonie.ahn/lab

업 Collab Slack Analysis Server Intellytics GitLab 기타

File Edit View Run Kernel Git Tabs Settings Help

Mem:0.00 B

Launcher

Regression_Test.ipynb

Code

git

Regression_Test.ipynb 4 hours ago

Regression.ipynb a day ago

Regularized Linear Mo... 4 hours ago

Regularized Linear Mo... a day ago

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

import warnings
warnings.filterwarnings("ignore")

from collections import Counter
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import MinMaxScaler
from regressors import stats

[2]: %%time
# Data Load
with open("/nas001/projects/Solar_Edu/Data/Edu_solar_hw_210302.pickle", "rb") as f:
    data = pickle.load(f)
print(">>> Data Shape : {}".format(data.shape))

>>> Data Shape : (38968, 588)
CPU times: user 248 ms, sys: 299 ms, total: 547 ms
Wall time: 546 ms

[5]: selc_col = ['Model', 'Eff', 'FF', 'UOC', 'ISC', 'RS', 'IREV2', 'TEMP']
data = data[selc_col]

data = data[data['Model'] == 'E6']
print(">>> Data Shape : {}".format(data.shape))

>>> Data Shape : (38968, 8)

[8]: data.drop(columns=['Model'], inplace=True)
print(">>> Data Shape : {}".format(data.shape))

>>> Data Shape : (38968, 7)

[7]: data.dropna(inplace=True)
data.reset_index(inplace=True, drop=True)
```

참가자 (12)

검색

서태민 나

안전이

CTO_이세열

W1 webex 171

공성우

류기선

박종열

수영 심

이수현 선임

임현수

조호습

최재우s

채팅

조호습님이 모두에게: 오후 2:52
솔라데이터는 뭘 써야하나요

조호습님이 모두에게: 오후 2:53
네..

조호습님이 모두에게: 오후 2:54
네

저장 위치: 모두

여기에 채팅 메시지를 입력하십시오.

0 1 GONIE | Idle

Mode: Command Ln 1, Col 1 Regression_Test.ipynb

오후 2:55 2021-03-11

제 20210429 호

상 장

Data Scientist 부문
개 근 상



조 호 습



임 광 영



최 재 우

위 학생들은 빅데이터실이 주관하는
데이터 학교에서 부지런히 3달간 개
근 하였으므로 이에 상장을 수여합니
다.

2021년 4월 29일

영원한 동반자 만 건 이



빅데이터 실전 교육 - 2021/07/05 ~ 2021/07/09

강사소개

gonie.ahn

👍 GIGO

@gonie.ahn · Member since February 11, 2020

gonie.ahn@lge.com · 📍 LG SciencePark, W04, 5F · 🏢 LGE, CTO, DXT, BigData ScienceTP

Data Scientist

<http://mod.lge.com/hub/smartdata/data-science-online-course-from-dstp>

Background

- 기존 데이터 분석에 관련된 사외 교육을 이수하고 현업 데이터에 바로 적용하기 어려움이 있었음
- LG전자 내 여러 사업부와 협업한 경험을 바탕으로 현업 상황에 맞고 바로 적용 가능한 데이터 분석 교육을 기획함
- 2019 - 2020년 [redacted] 연구개발팀과 데이터 분석을 진행한 내용을 바탕으로 강의를 진행함

8:32 79

빅데이터 실전 교육 - 2021/07/05 ~ 2021/07/09



Mathematical Expression

- Simple & Multi-Linear Regression Estimation

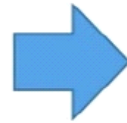
HSE Sim, Mat. β_0 β_1 $(y - \hat{y})$

$$Q(\beta_0, \beta_1) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$

Minimize $Q(\beta_0, \beta_1) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$

$$\left. \frac{\partial Q(\beta_0, \beta_1)}{\partial \beta_0} \right| = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\left. \frac{\partial Q(\beta_0, \beta_1)}{\partial \beta_1} \right| = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i = 0$$



$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$



신한카드 Data Science 심화 과정 – 2022.10.06



신한금융지주 Project Owner 과정 – 2022.10.06

