

• 03/12

→ Tensorflow 2.x 을 이용해 Linear Regression 구현

⇒ One data를 이용해 구현

↓
결측치, 이상치, 전처리

↓
deletion
Imputation

↓
평균치기법 (mean, median)

↓
여측치법 (KNN)

(종속변수가 결측치일때 사용)

(독립변수)가 결측치일때 사용.

mode

→ 어떤 알고리즘인지에 대해 설명

(강함) hyperparameter [k (?)
거리계산 방법

① 독립변수 (Solar, Wind, Temp)에 대한 결측치를 찾아서 → **median**으로 채우기 (mean을 이용하지 않는 이유는 이상치 때문에 그래요!!!)

② 독립변수에 대한 이상치를 검출해서 **mean**으로 채우기

③ 종속변수에 대해 이상치를 검출해서 **mean**으로 채우기

④ 전처리 진행

→ ⑤ 종속변수에 대한 결측치 채우기 **KNN**을 이용해서 Imputation.

→ "DataSet" 인식

- MNIST를 예로 들면,

Image data \Rightarrow 3차원 data

(가로, 세로, 색) \rightarrow color의 경우

(가로, 세로, 색) \leftarrow 흑백의 경우

grey scale의 경우 2차원으로 표현이 가능

\swarrow

0	0	0	0	0
0	0	70	50	0
0	0	0	0	0
0				

1차 2차원

(4)

1차원

\checkmark

x_1	x_2	x_3	x_4

\checkmark

1차원