

02/26 · Normalization (정규화) → 비율을 이용하여 data의 scale을 조정 (0 ~ 1)

Year	집(坪)	집가격
1	1000	
2	700	
5	500	
10	200	

Month	집(개월)	집가격
12	1000	
24	700	
60	500	
120	200	

집(坪)	가격
0	1
0.7	0.7
0.56	0.5

↓ 정규화를 하는 방법

- Min Max Normalization
- Z-score Normalization (Standardization)

✓ Min-Max Normalization

“데이터 정규화의 가장 일반적인 방법”

모든 feature의 값을 [최소값 0, 최대값 1] → 변환

✓

$$x_{\text{scaled}} = \frac{x - x_{\text{MIN}}}{x_{\text{MAX}} - x_{\text{MIN}}}$$

문제점?

이항치에 상응하는 평균치
scaling 문제

(outcome) 이상치처리를 반드시
미리 해야해요.

- Standardization

(Z-score Normalization)

* $x_{\text{scaled}} = \frac{x - \mu_i}{\sigma_i}$

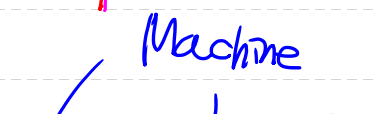
μ_i : 0이상의 크기 변동을
 나타냅니다.
 σ_i : 동일한 scale을 적용하는
 것입니다.

→ 권차리에 대해 알게됨으니

(온도가 파를 O_{250K} 정도 상승하면 역투과도
급격히 다시 구현

- ① CSU 파일 loading
- ② 볼록해져서
- ③ 이상치 처리 (처리)
- ④ Normalization (Min-Max)
- ⑤ Training data set
- ⑥ Tensorflow를 이용하는 Machine Learning + Python 구현도 포함
- ⑦ graph를 그려서 sklearn의 결과와 비교.
- ⑧ predict 해서 그 결과가 sklearn과 유사하게 나오는지 확인

★ Simple Linear Regression



Machine Learning

컨버퍼 구현

- 독립변수가 여러개인 **Multiple Linear Regression**에 대해서 알아보아 오~

"Classical Linear Regression Model"

$$\hat{y} = \beta_0 + \sum_{i=1}^p \beta_i x_i$$

"독립변수가 3개이면"

$$H = \overset{b}{\beta_0} + \overset{w_1}{\beta_1} x_1 + \overset{w_2}{\beta_2} x_2 + \overset{w_3}{\beta_3} x_3$$

$$H = w_1 x_1 + w_2 x_2 + w_3 x_3 + b$$

$$H = x_1 w_1 + x_2 w_2 + x_3 w_3 + b$$

$$\bullet H = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} + b = \begin{pmatrix} y_{11} \\ y_{21} \\ y_{31} \end{pmatrix}$$

$$x_{11} \cdot w_1 + x_{12} \cdot w_2 + x_{13} \cdot w_3 + b = y_{11}$$

$$H = \underset{\text{2차원}}{X} \cdot \underset{\text{2차원}}{W} + b$$