چه زمانی یک مدل مرکز است؟ و نیاز به

درمان دارد؟

چه زمانی مدل تصمیم گیرنده باشد! سه قدرت تخمین روی داده‌های

غیر از Train خود را

ندارد!

درمان :

مدل را پیچیده‌تر

under fitting $\longrightarrow$

Train-error : $\uparrow$

Test-error : $\times$

$\longrightarrow$ high bias

درمان : Data

over fitting $\longrightarrow$

Train-error : $\downarrow$

Test-error : $\uparrow$

$\longrightarrow$ high variance!

underfitting $\longrightarrow$ Reg $\begin{cases} D=1 \longrightarrow D=2 \\ D=2 \longrightarrow D=3 \end{cases}$

$\downarrow$ SVR

LR $\xrightarrow{?}$ SVM

DT

RF

RNN

Overfitting

Train

مدل به نویز داده‌ها چسبیده!

→ اصل و فیچرهای درست را رها کرده و سراغ فیچرهای خاص و بی‌اهمیت رفته!

مثال

آموزش

سیب سیب + سیب سیب سیب = سیب سیب سیب سیب سیب

آزمایش

Test

سیب سیب + پرتقال پرتقال پرتقال = سیب سیب سیب

داده پرت outlier

missing value رها شده

NaN

عدم پاک سازی داده ها

یادگیری نامناسب داده

پیچیدگی مشکل وزن

$$y = w_1 x_1 + w_2 x_2 + b$$

x_1	x_2	y
200	500	3

$w_1 = 10^6$

$w_2 = 10^{15}$

$b = 10^3$

$$\underset{3}{y} = \underbrace{10^6}\,\underbrace{x_1}_{500} + \underbrace{10^{15}}\,\underbrace{x_2}_{200} + \underbrace{3}$$

از حضور شما

۱. تمرین‌سازی داده!

Next Session!

label

features.

Normalization.

$$x = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Min max scaling!

$$X = \frac{X - X_{min}}{X_{max} - X_{min}}$$

کمینه ستون

بیشینه ستون

X_2	X_1	Y
	1000	
	100	
	50	

0/052

$$X = \frac{100 - 50}{1000 - 50} = \frac{50}{950} = \frac{5}{95} = 0/052$$

Minmax Scaling ⇒

$[0, 1]$

x_1	x_2	x_3
1	2	8
3	5	9
8	7	3

image!

128	0	1	18
20	13	130	3
201	80	19	87

$$\rightsquigarrow \frac{X}{X_{max}} = \frac{X}{255}$$

255

scaling

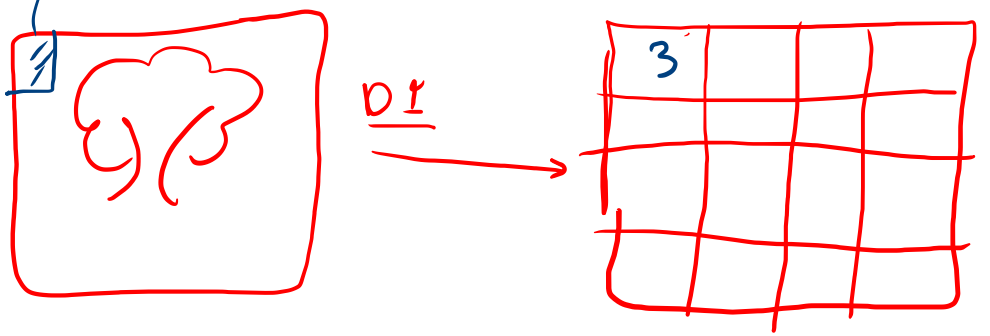
1,5	0	..
..	..	..
..	..	..

Image

سفید

سیاه

0 - 255



DF

دانش‌آموز

S1 → 1 → 0

S1 → 3 → ٪02

5 → ٪4

2 → ٪25

1 → 0

4 → ٪35

S2: 100

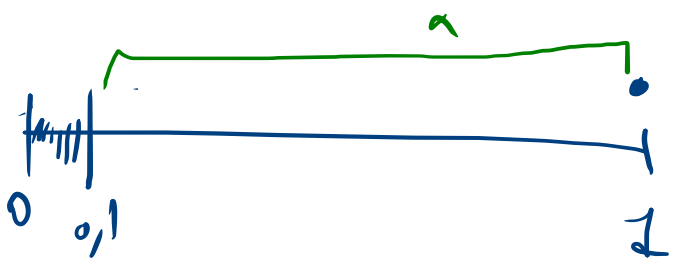
price

$$x_2 = \frac{3-1}{100-1} = \frac{2}{99} = ٪02$$

$$x_1 = \frac{1-1}{100-1} = 0$$

$$\alpha = \frac{x - x_{min}}{\underbrace{x_{max} - x_{min}}_{100}}$$

$$\frac{100-1}{100-1} = 1$$



۳. استاندارد کردن داده‌ها

Standardization.

$[-1, 1]$

$[0, 1]$

Z-Score!

$$X = \frac{X - \mu}{\sigma}$$

$\mu \rightarrow$ mean

$\sigma \rightarrow$ std dev.

ارتفاع خونه X_1	قیمت خونه X_2	قیمت خونه Y
	10	
	20	
	25	

$\mu = \frac{55}{3} = 18{,}33$

$\sigma = 6{,}24$

$$X = \frac{10 - 18{,}33}{6{,}24} = -1{,}33$$

$$\sigma = \sqrt{\left(\frac{(X_i - \mu)^2}{N-1}\right)} \rightarrow$$

X_2
10
20
25

$$\sigma = \cancel{\sqrt{}}$$

$$\sigma = \sqrt{\frac{(10-18{,}33)^2 + (20-18{,}33)^2 + (25-18{,}33)^2}{3-1}}$$

$$\sigma = \sqrt{\frac{69{,}38 + 2{,}78 + 44{,}4}{2}} = \sqrt{39} \approx 6{,}24$$