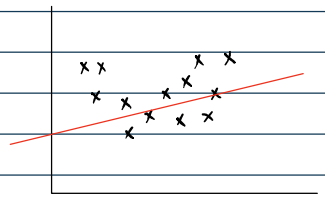
Polynomial Linear Regression

Monday 8 April 2024 10:25 PM

SLR → $y = \beta_0 + \beta_1 x$



MLR → $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n$

Polynomial Linear Regression $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \ldots + \beta_n x^n$

→ During the preprocessing stage we extract the polynomial features of every independent variable.

class sklearn.preprocessing.**PolynomialFeatures**(*degree=2*, ***, *interaction_only=False*, *include_bias=True*, *order='C'*) [source]

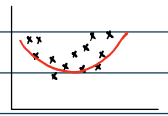
e.g. $x \mid y \longrightarrow x^0 \mid x^1 \mid (x^2) \mid y$ (degree = 2 → tries to extract the polynomial behaviour of the data)

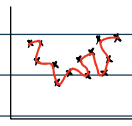
$y = \beta_0 + \beta_1 x_1 \longrightarrow y = \beta_0 + \beta_1 x^1 + \beta_2 x^2$

→ value of degree is a hyperparameter

↓ value, model tends to underfit

↑ value, model tends to overfit





* Despite having we have a higher degree relationship b/w y and X, why do we still call it Linear Regression?

→ When performing linear Regression we are looking at the relationship b/w y and the coefficients, degree of coefficients is 1, the relationship b/w y and the coefficients is still linear.

degree 2

X_1	X_2	y

→ $x_1^0 x_2^0$ $x_1^1 x_2^1$ $x_1^1 x_2^0$ $x_1^2 x_2^0$ $x_1^0 x_2^1$ $x_1^0 x_2^2$