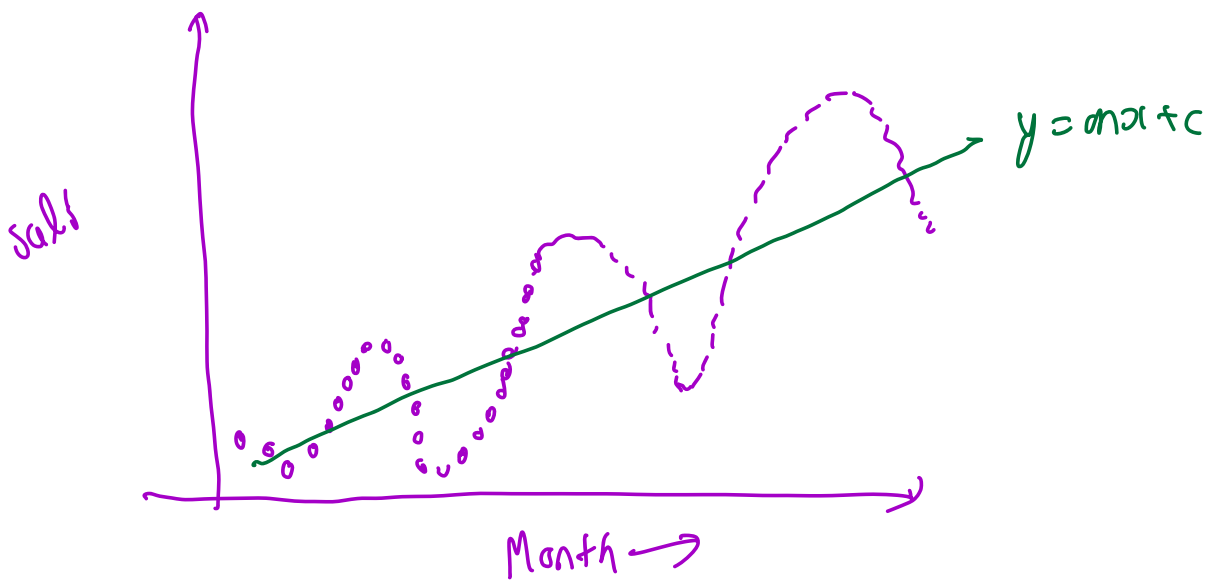
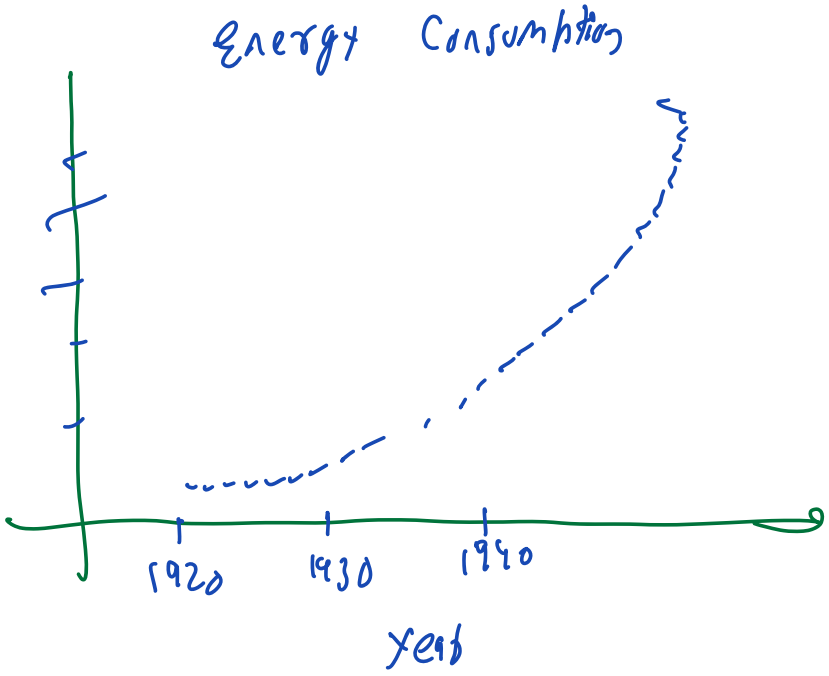

⇒ All topics
⇒ fired ⇒ stop

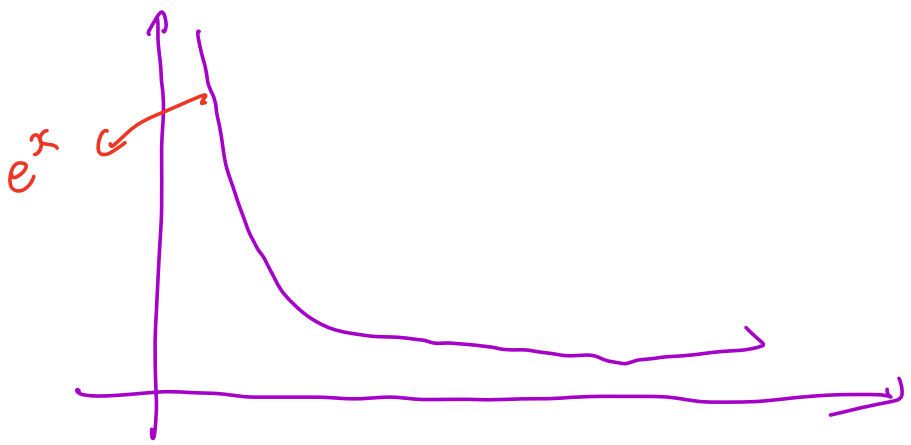
## Polynomial Regression

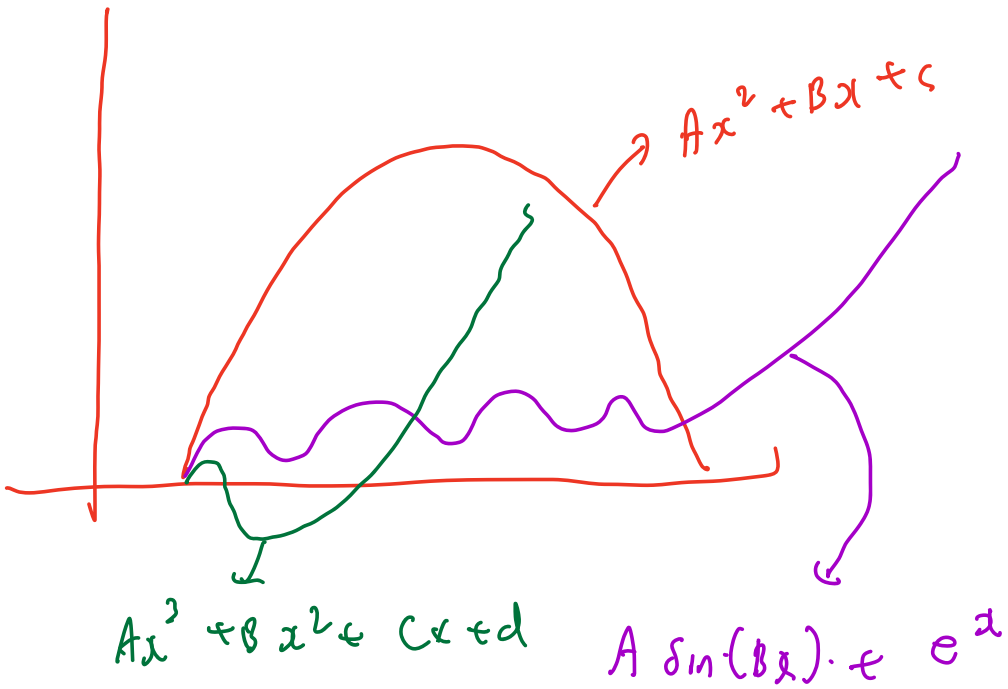



⇒ $30 * \sqrt{m} \cdot \cos(0.5 * m) + 5.03 * m + 70$

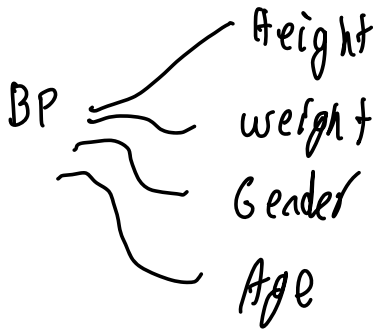



$$C = A + By + Cy^2 + Dy^3$$



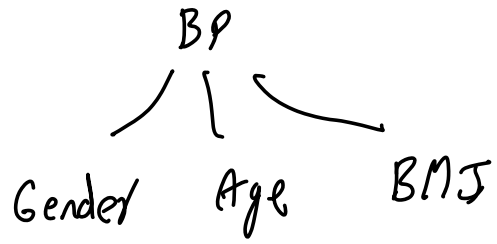



Height | weight | Gender | Age | BP




BMI = $\frac{\text{weight}}{\text{Height}}$

non-linear feature




Data Points $(x_1, x_2, x_3, \ldots, x_d, y)$ → Respon

raw Attributp

$\phi_1(\vec{x}) \rightarrow x_1$

$\phi_2(\vec{x}) \rightarrow \frac{x_1}{x_2}$

$\phi_3(\vec{x}) = x_3^3$

$y = a + b_1 x_1 + b_2 x_2 + \ldots b_n x_n$

$y = a + b_1 \phi_1(x) + b_2 \phi_2(x) + \ldots + b_n \phi_n(x_n)$

age

$R^2$ → overfit (train)

$R^2$ → test data →

Occam Razor

$y = a \cdot x_1 + c$
79%

$y = a x_1^3 + b x^2 + c x + d$
70%

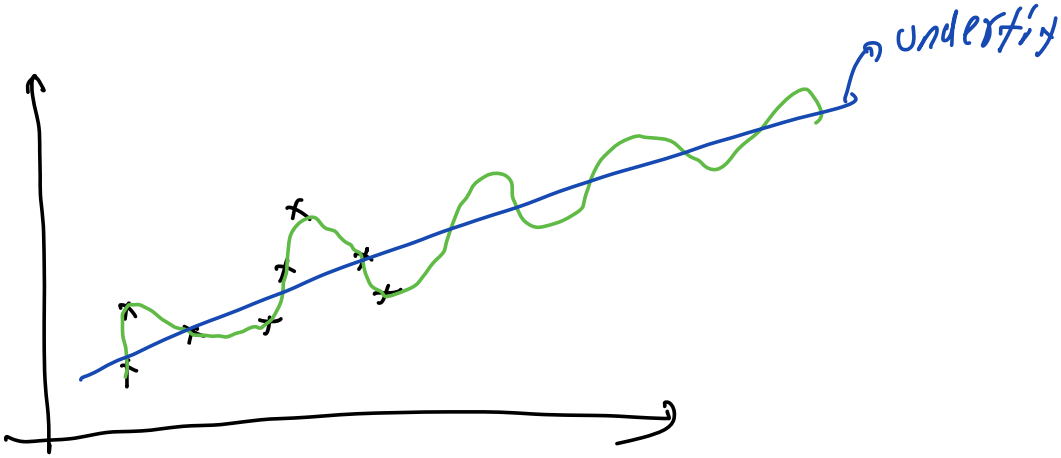
.63 .79

$x \; x^2 \; x^3$ Train $\rightarrow R^2 \uparrow$ $\uparrow$ $\uparrow$
Test $\rightarrow R^2 \uparrow$ $\uparrow$

Grid

L Quad cubic

Test (H) L L



Underfit: Train : $R^2$ ↓
Test : $R^2$ ↓

Overfit : Train data : $R^2$ : ✓ ↑
test : $R^2$ ↓

Perfect : 


$$y_i = w_0 + w_1 x_{i1}^1 + w_2 x_{i2}^2 + w_3 x_{i1}^3 + w_4 x_{i1}^4$$

original

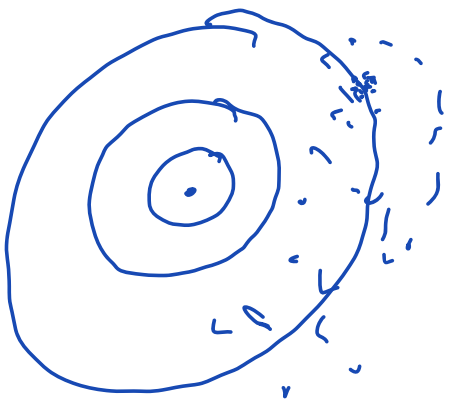
transformed features

Case 1: $w_1, w_2, w_3, w_4 \neq 0$ → overfit

Case 2: $w_3 = w_4 = 0$ ∴ Perfect
$w_1, w_2 \neq 0$

Case 3: $w_1 = w_2 = w_3 = w_4 = 0$ → underfit

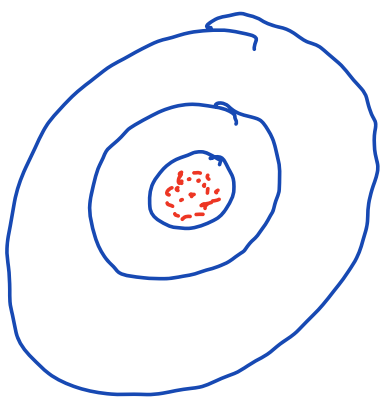
# Bias Variance tradeoff

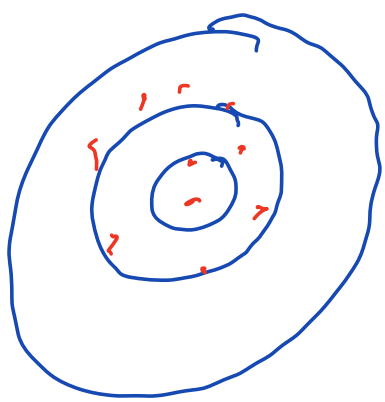


Low Variance

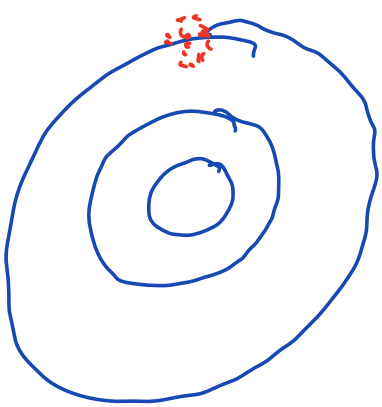
High Variance

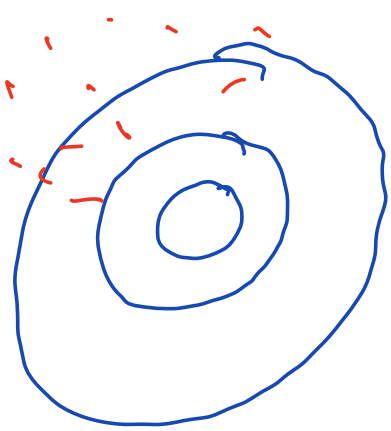
Low Bias



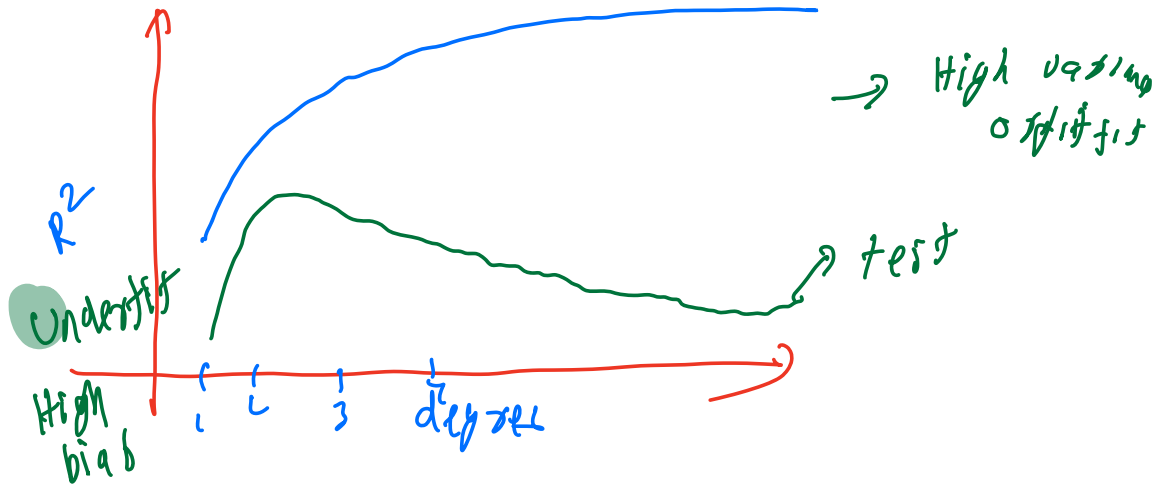


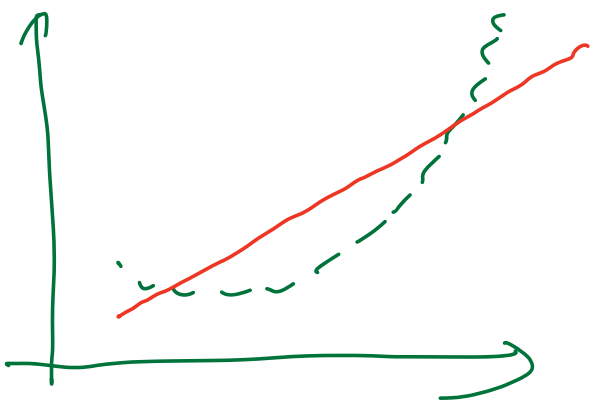
High bias



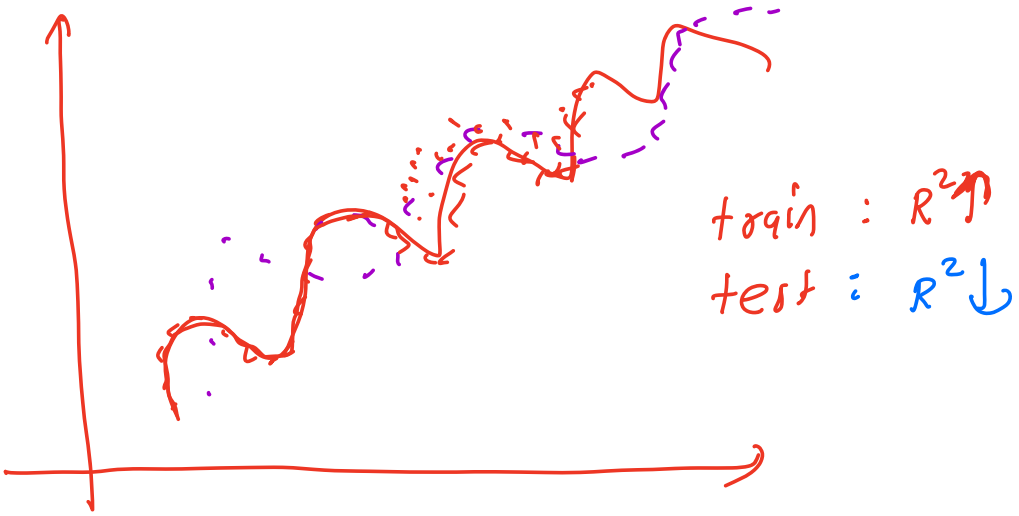


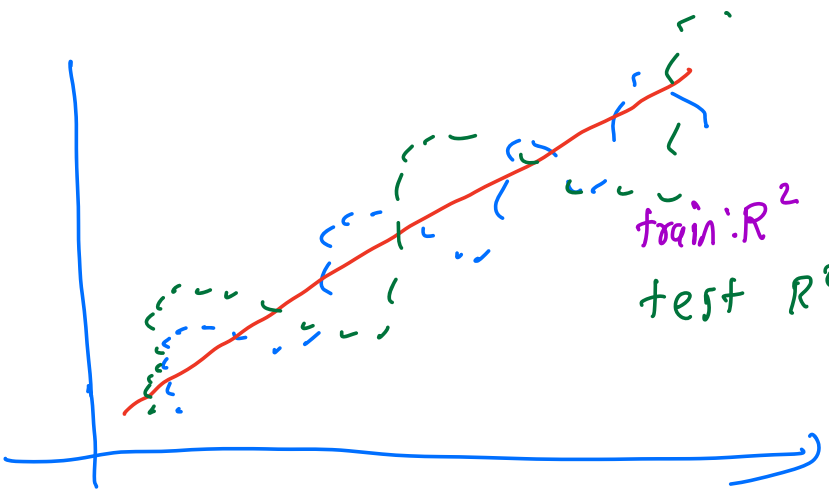
target


High variance
overfit
$R^2$
test
Underfit
High bias
1
2
3
degree



Break: 10:30




$$\text{loss} \Rightarrow \sum (y_i - \hat{y}_i)^2$$

(Ridge Reg)
L2-regularization

$$\sum (y_i - (w_1 x_1 + w_2 x_2 + w_0))^2 + \lambda \sum w_i^2$$

Loss — Regularization

$$\min \; (\text{Leakage} + \text{Budget})$$

Loss  Regularization

Case 1  $\lambda = 0 \Rightarrow$ overfit

Loss + $\lambda$ Reg

Case 2  $\lambda \rightarrow$ large $\Rightarrow$ underfit

min (Loss term + $\lambda |w_i|$)

↳ L1 regularization
Lasso regression

Elastic net reg

min ( loss + $\lambda_1 L_2$ reg + $\lambda_2 L_1$ reg )