Machine-learning Wednesday July 21

(y, x)

output, input

$$y(x) = f(x) + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2)$$

$$y(x) \simeq X\beta + \varepsilon$$

$$\mathbb{E}[y] = X\beta$$

$$MSE(\beta) = \frac{1}{n}\left((y - X\beta)^T (y - X\beta)\right)$$

$$\beta^{opt} = \hat{\beta} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}}\, MSE(\beta)$$

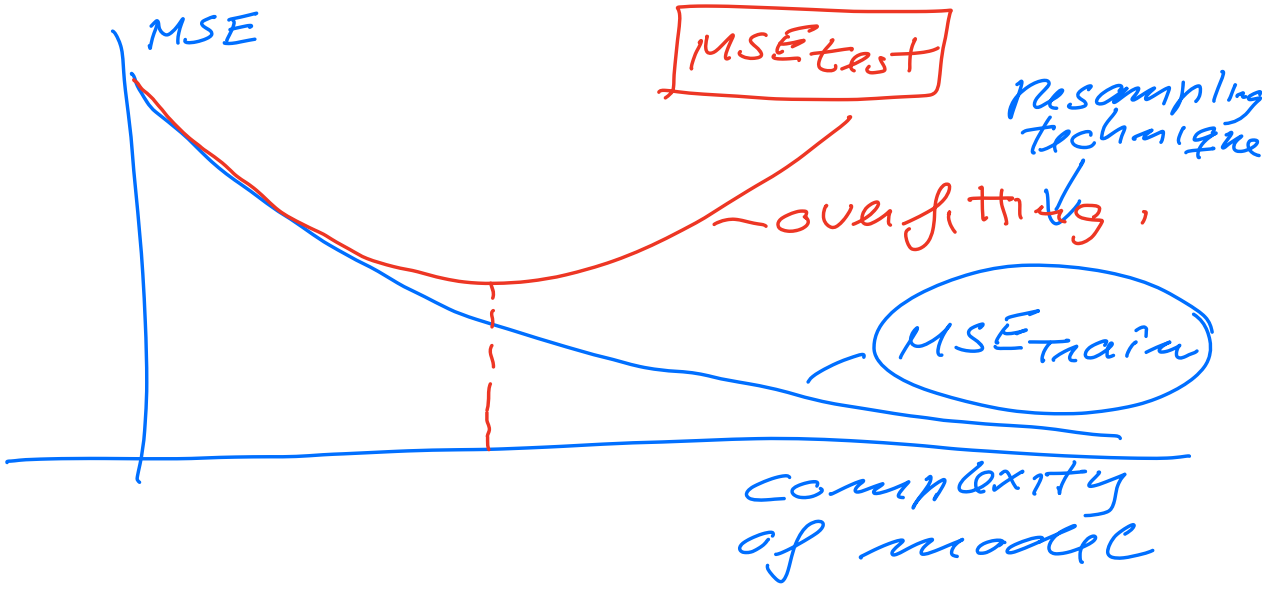
$$\hat{\beta} = (X^T X)^{-1} X^T y$$

$$MSE(\beta) = \text{Bias} + \text{Variance} + \sigma^2$$

$$\text{Bias} = \frac{1}{n}\sum_{i=0}^{n-1} (y_i - \mathbb{E}[\tilde{y}])^2$$

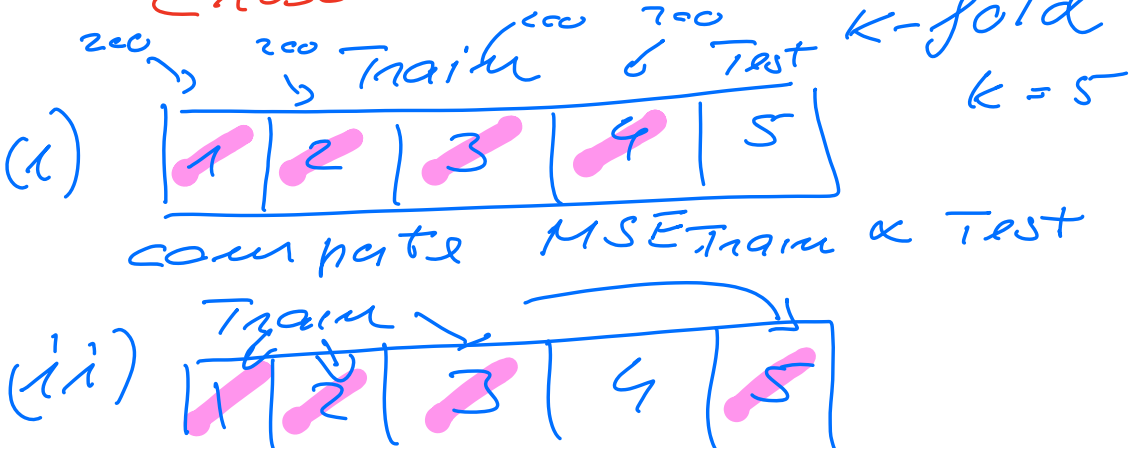
$$\text{Variance} = \frac{1}{n}\sum_{i=0}^{n-1}(\tilde{y}_i - \mathbb{E}[\tilde{y}])^2$$

Split data in train & Test (vault, and untouched)

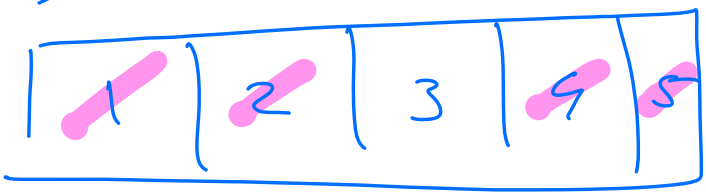


Resampling methods

Cross-validation

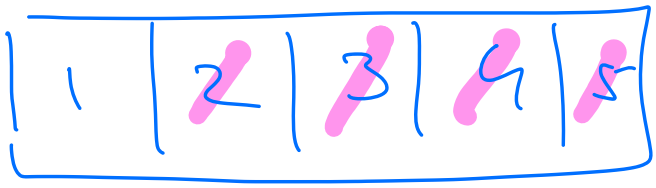


(iii)



↓

(v)



OLS

$$\beta^{opt} = (X^T X)^{-1} X^T y$$

A

$$A = X^T X = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}$$

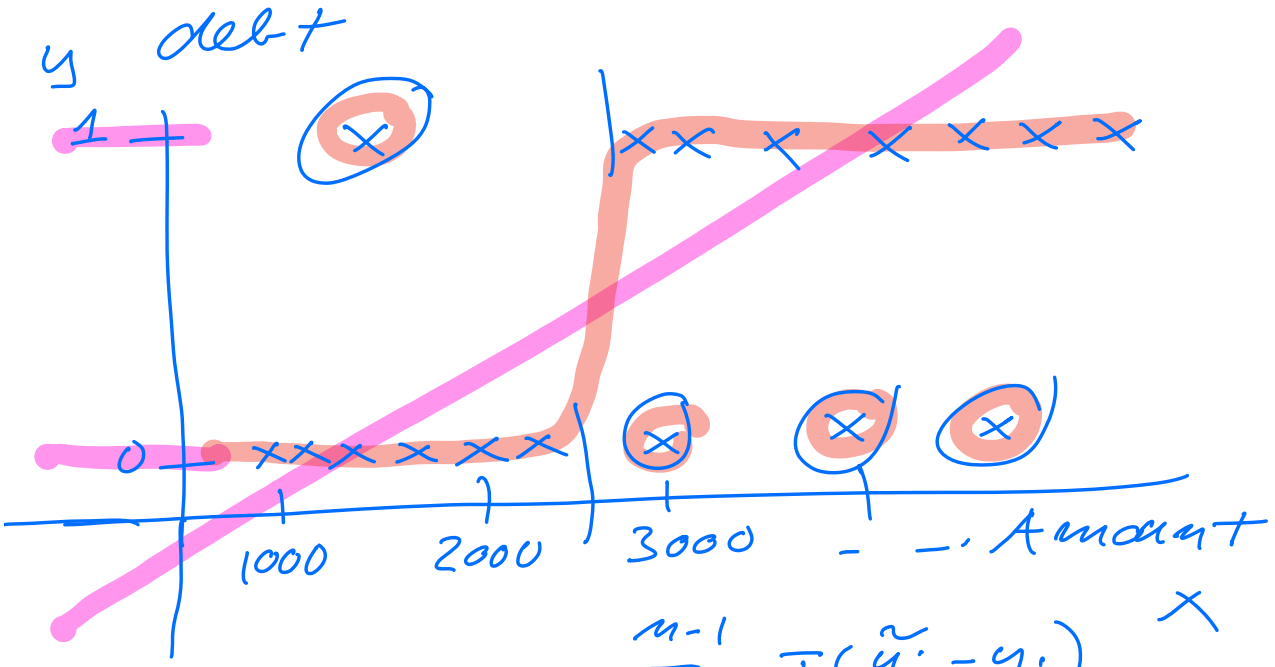
$a_{ii} + \lambda$

Classification

Binary : 0, 1, True, False

-1, 1, healthy, sick ...

Credit card (kaggle.com)



$$\text{accuracy} = \frac{\sum_{i=0}^{n-1} I(\tilde{y}_i = y_i)}{n}$$