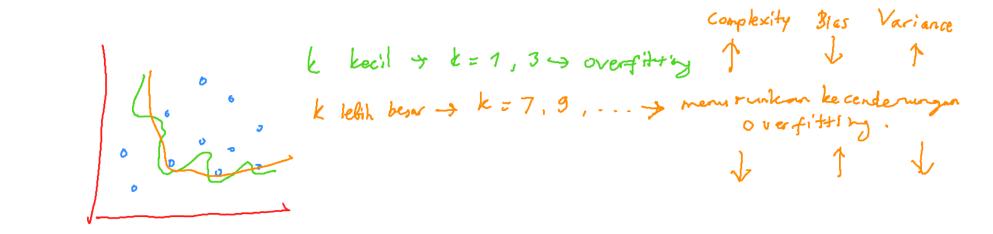
Generalization

1) KNN - S Number of k



Decision tree -- control the model's complexity

_ max - depth

_ min_ samples split / leaf

_ min - impurity - decrease

3) Linear models - Linney, logrey - Number of features - (Regularization) - mengatur weight / ni lai coef. - Polynomial -> mengubuh model linear menjadi non-linear B1×1 → ×1×2, ×72, ×2

mengatasi overfitting pada linear models. Kegularization-> 少= ゆの+ ら1×1+ら2×2····なの×の → Penalty rate 1) Ridge - membrat nilai over mendeketi Ø jika fitur dianggap trobak signifikan
fer hadap lube 1.

(2) Lasso - membrat nilai over menjadi O untuk fitur yang dianggap trobak signifikan.

(3) Elastic Net (3) Elastic Net

(1) Ridge Regression -> Lz regularization X \geq 0; 2 \rightarrow menambahkan yang didalum kotak ihi ke dalam cost function model linear. $J(\theta) = \frac{1}{h} \sum_{i=1}^{k} (y_i - \hat{y}_i)^2 + \lambda \sum_{i=1}^{k} \theta_i^2 \xrightarrow{\text{Starts from}} \beta_0 \text{ tidak terkena}$ regularization.

· Hyperparameter & mengontrol sebetapa besarregularisasi yang akan diterapkan pada model.

& sangat besur -> menjadikan nilai wef semakin mendekati),

· Wajib melder kan scaling pada duta sebelum melekukan regularization.

e.g. MPG _____ cylinder ___ 2, 4,6,8... ___ skala satuan e.g. MPG _____ mileage ___ ratusan, ribuan, ratus ribuan.

$$\frac{1}{y} = \beta \circ - \beta_1 \times 1 \quad \beta_2 \times 2$$

$$-0.7. \times 1 - 0.14. \times 2$$

$$\begin{array}{c} \text{hasilinga} \\ \text{hasilinga} \\ \text{he cil} \end{array}$$

(2) Lasso Regression (L1 regularization)

Lost Function $J(\theta) = MSE(\theta) + d \sum_{i=1}^{n} |\theta_i|$ Lisso

oleh &

dikontrol juga

Feature selection

- Semakih tinggi & , bisn juga semakih banyak fitur yang dihilangkan.
- Semakin kecil nilai coef, semakin tidak penting fitur tersebut.

(3) Elastic Net slasso ridge
$$J(\theta) = MSE(\theta) + r \times \frac{2}{2} |\theta| + \frac{1-r}{2} \times \frac{2}{1-1} |\theta|^{2}$$

$$r = m! \times ratio$$
 $r = 0 \longrightarrow ridge$

$$r=0.5?$$
 balanced?

- kalav mendapati fitur yang sebenarny a tidak signifikan, kita bis a menggunakan Lasso / Elastic Net