

PRZEGLĄD UCZENIA NADZOROWANEGO

„The Elements of Statistical Learning”

Roz. 2.2-2.4

Trevor Hastie, Robert Tibshirani, Jerome Friedman

Dwa proste podejścia do predykcji

Najmniejsze kwadraty i najbliżsi sąsiedzi

Modele liniowe

Wzór na predykcję \hat{Y}
modelem liniowym

$$\hat{Y} = \hat{\beta}_0 + \sum_{j=1}^p X_j \hat{\beta}_j.$$

Postać macierzowa
wzoru na predykcję \hat{Y}

$$\hat{Y} = X^T \hat{\beta},$$

Minimalizujemy
rezidualną sumę
kwadratów

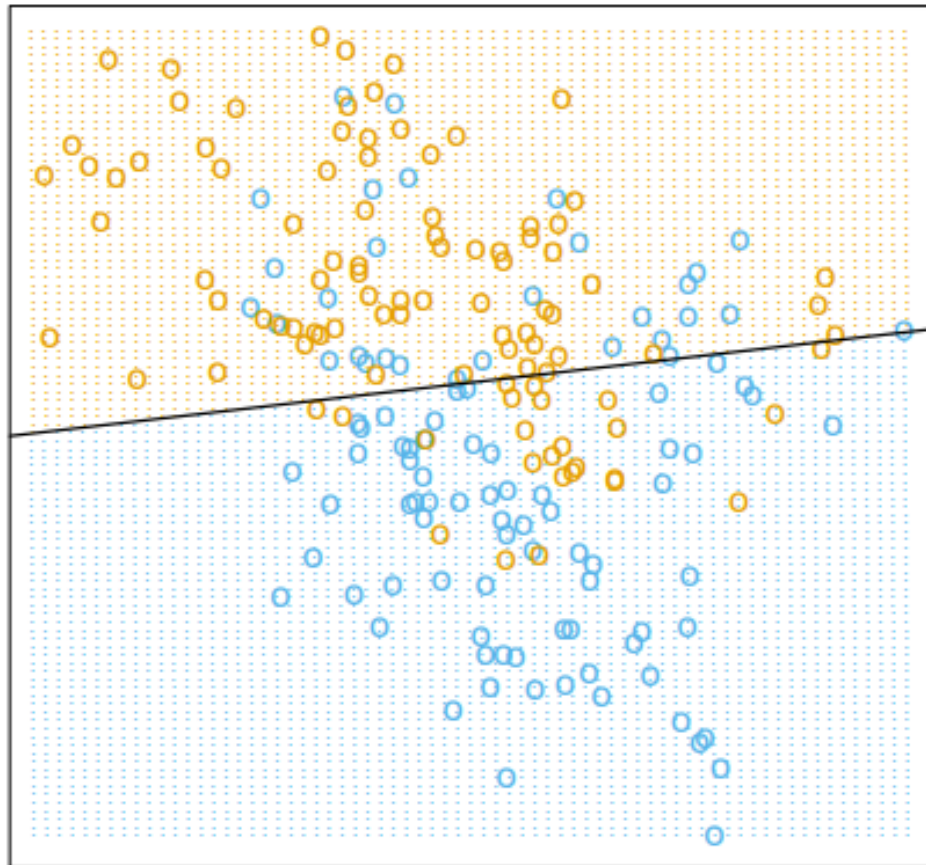
$$\text{RSS}(\beta) = \sum_{i=1}^N (y_i - x_i^T \beta)^2.$$

Jedyne rozwiązanie
wektora
współczynników

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

Przykład modelu liniowego w kontekście klasyfikacji

$$\hat{G} = \begin{cases} \text{ORANGE} & \text{if } \hat{Y} > 0.5, \\ \text{BLUE} & \text{if } \hat{Y} \leq 0.5. \end{cases}$$



$$\{x : x^T \hat{\beta} = 0.5\}$$

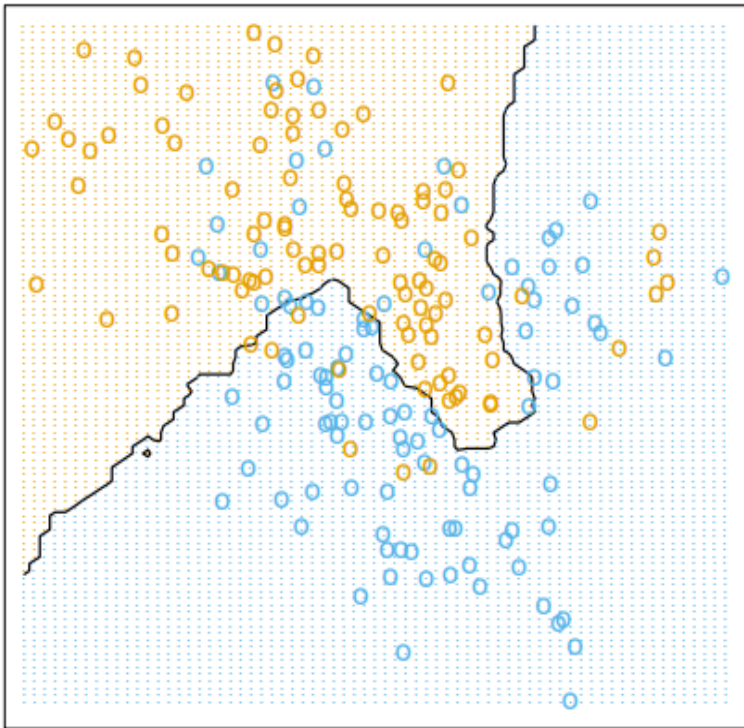
Metoda najbliższych sąsiadów

Dopasowanie k-
najbliższego sąsiada

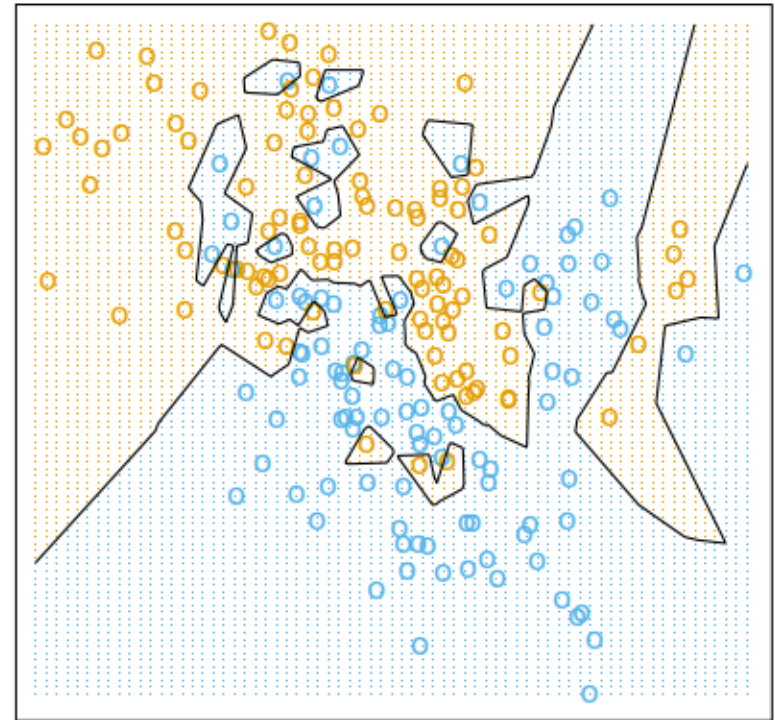
$$\hat{Y}(x) = \frac{1}{k} \sum_{x_i \in N_k(x)} y_i,$$

Porównanie liczby sąsiadów na przykładzie

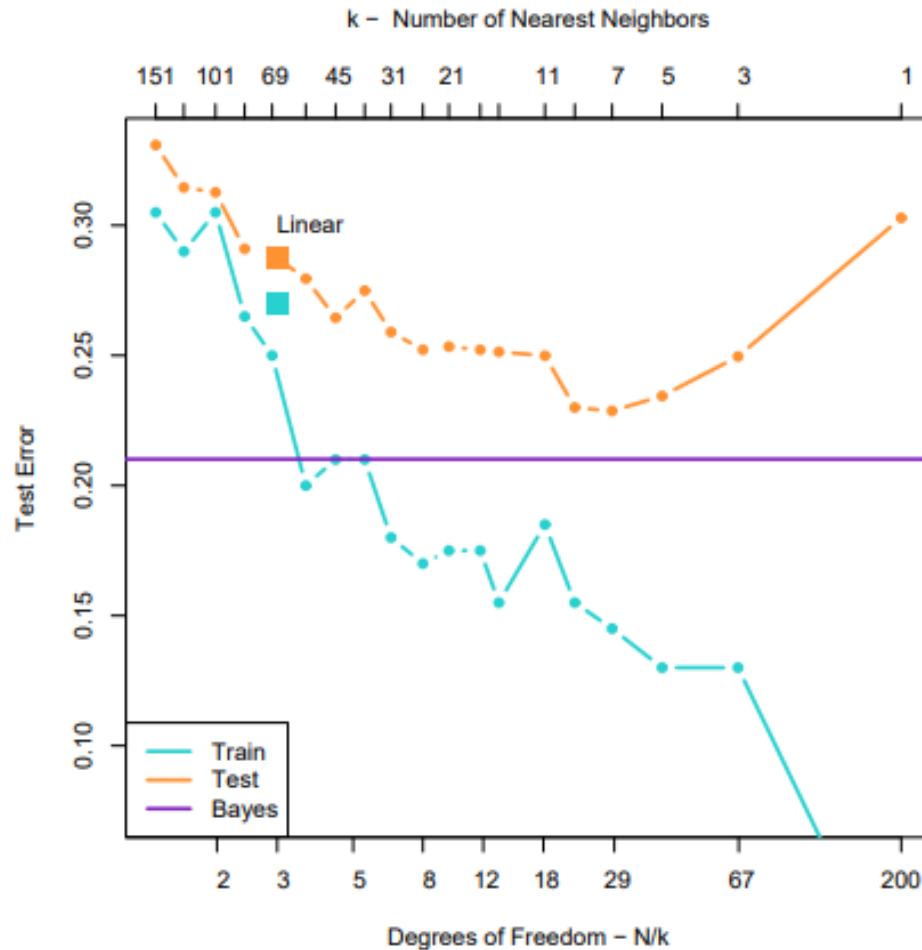
15 najbliższych sąsiadów



1 najbliższy sąsiad



Od najmniejszych kwadratów do najbliższych sąsiadów



Teoria decyzji statystycznych

Kwadratowa funkcja straty

$$L(Y, f(X)) = (Y - f(X))^2.$$

Oczekiwany błąd predykcji

$$\begin{aligned} \text{EPE}(f) &= \text{E}(Y - f(X))^2 \\ &= \int [y - f(x)]^2 \text{Pr}(dx, dy) \end{aligned}$$

Minimalizując EPE otrzymujemy

$$f(x) = \text{E}(Y|X = x)$$

Różnice pod względem założeń modelu

Funkcja regresji

$$f(x) \approx x^T \beta.$$

Wzór na wektor współczynników

$$\beta = [E(XX^T)]^{-1}E(XY).$$

Reguła k-najbliższych sąsiadów

Zakładamy, że $f(x)$ jest dobrze przybliżana przez lokalną funkcję stałą.

Rozwiązanie przy normie L1

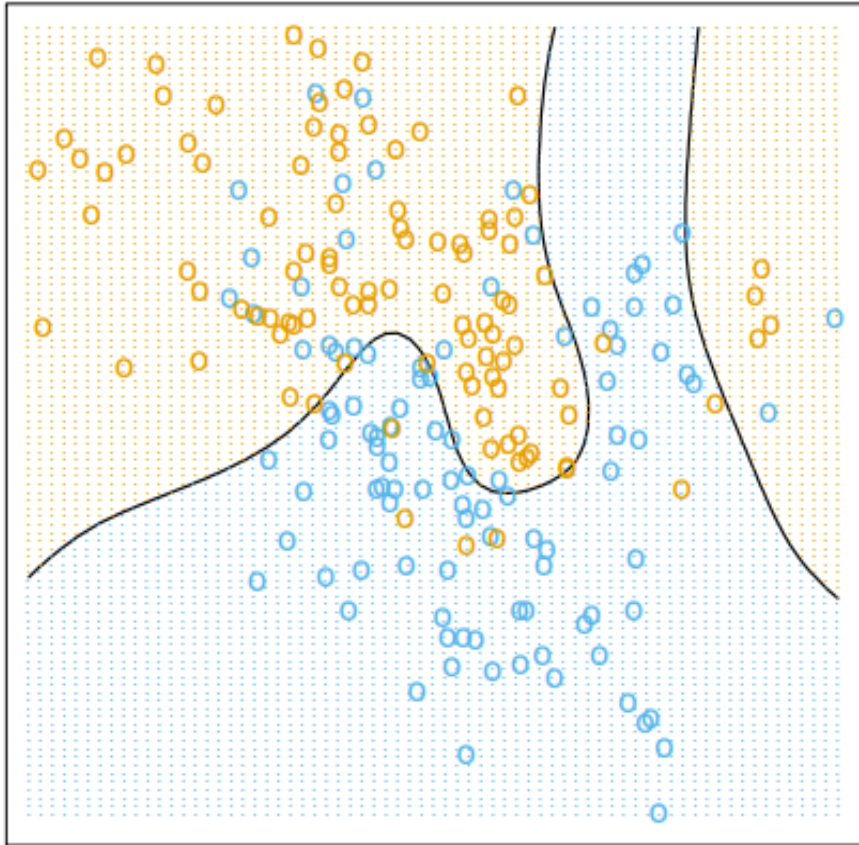
$$\hat{f}(x) = \text{median}(Y|X = x).$$

Metoda najmniejszych kwadratów

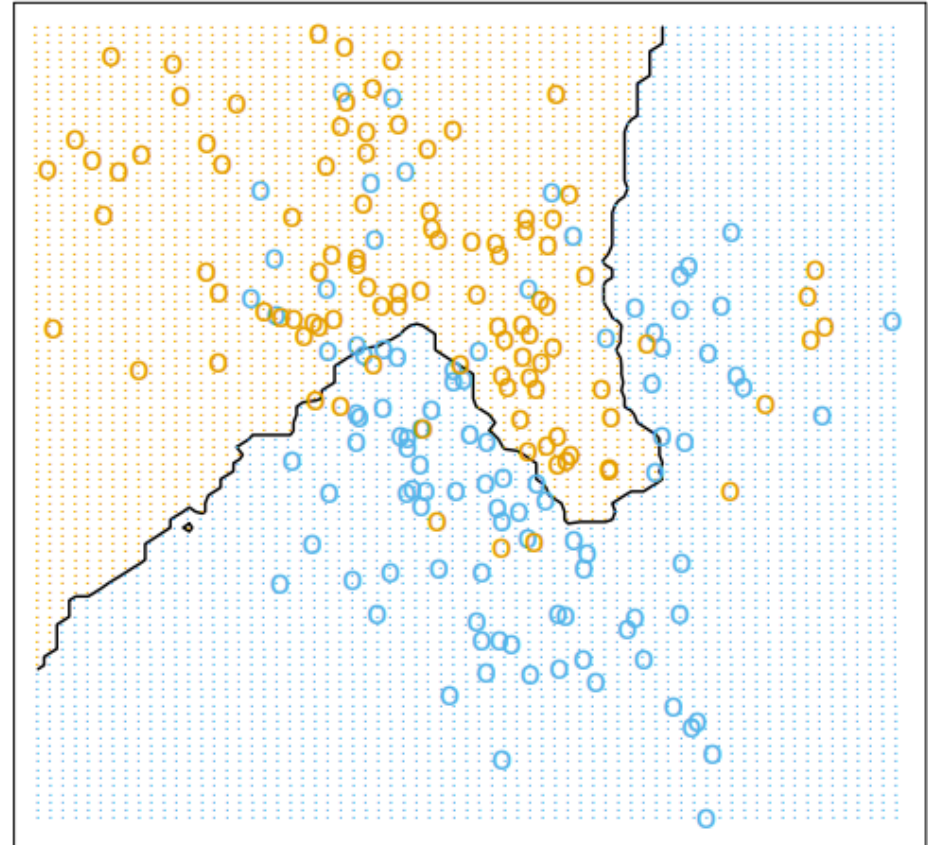
Zakładamy, że $f(x)$ jest dobrze przybliżana przez globalną funkcję liniową.

Klasifikator Bayesa

Bayes Optimal Classifier



15-Nearest Neighbor Classifier



Dziękuję za uwagę!