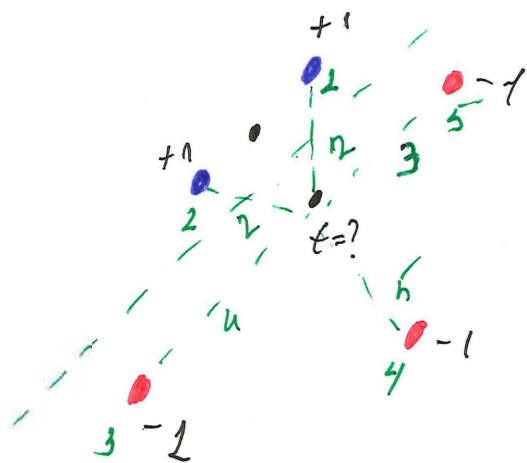


kNN, k=5

$$y(x) = \arg \max_{t \in \{b, r\}} \sum_{j=1}^5 \underbrace{\delta_t(t_j)}_{\mathbb{1}[t_j = t]}$$



Distance-weighted kNN

$$y(x) = \arg \max_{t \in \{r, b\}} \sum_{j=1}^5 \underbrace{\text{sim}(x_j, x)}_{\text{similarity between } x \text{ and } x_j} \cdot \delta_t(t_j)$$

$$\text{sim}(x, x_j) = \frac{1}{\|x - x_j\|^2}$$

similarity between x and x_j
test example

$$y(x) = \arg \max$$

$t = r$

$$1 \cdot \frac{1}{4^2} + 1 \cdot \frac{1}{5^2} + 1 \cdot \frac{1}{3^2}$$

$$\frac{1}{16} + \frac{1}{25} + \frac{1}{9} = 0.25$$

$t = b$

$$1 \cdot \frac{1}{2^2} + 1 \cdot \frac{1}{2^2}$$

$$\frac{1}{2} = 0.5$$

kNN for binary case, where labels are $\{+1, -1\}$

distance weighted kNN: $\hat{y}(x) = \arg \max_{t \in \{+1, -1\}} \sum_{j=1}^K \frac{1}{\|x - x_j\|} t_j [t_j = t]$

$$\hat{y}(x) = \text{sign} \left(\sum_{j=1}^K \frac{1}{\|x - x_j\|} t_j \right)$$

$$= \text{sign} \left(\frac{1}{2^2} \times (+1) + \frac{1}{2^2} \times (+1) + \frac{1}{4^2} \times (-1) + \frac{1}{5^2} \times (-1) + \frac{1}{3^2} \times (-1) \right)$$

$$= \text{sign} \left(\underbrace{\left(\frac{1}{2^2} + \frac{1}{2^2} \right)}_{S_+} - \underbrace{\left(\frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{3^2} \right)}_{S_-} \right)$$

$$= \begin{cases} +1, & \text{if } S_+ > S_- \\ -1, & \text{if } S_+ < S_- \end{cases}$$

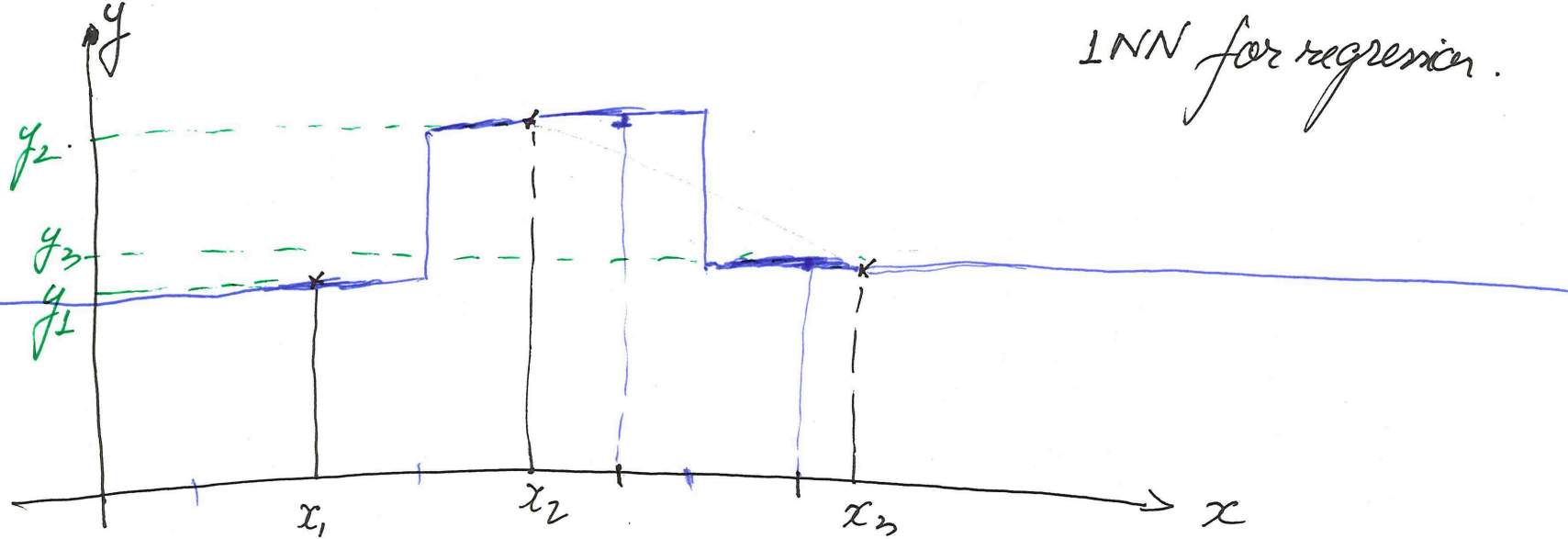
~~kNN~~ Kernel-based distance-weighted kNN

$$\hat{y}(x) = \text{sign} \sum_{j=1}^k \underbrace{\text{sim}(x, x_j)}_{k(x, x_j)} \cdot \cancel{y_j}$$

Gaussian kernel:

$$k(x, x_j) = e^{-\frac{\|x - x_j\|^2}{2\sigma^2}}$$

1NN for regression.



$$y(x) = \frac{1}{k} \sum_{j=1}^k t_j = \sum_{j=1}^k 1 \cdot t_j / k = \sum_{j=1}^k \underbrace{1}_{\text{sim}(x, x_j)} \cdot t_j / \sum_{j=1}^k \underbrace{1}_{\text{sim}(x, x_j)}$$

$$\text{sim}(x, x_j) = \frac{1}{\|x - x_j\|^2}$$