

CS 66: Machine Learning

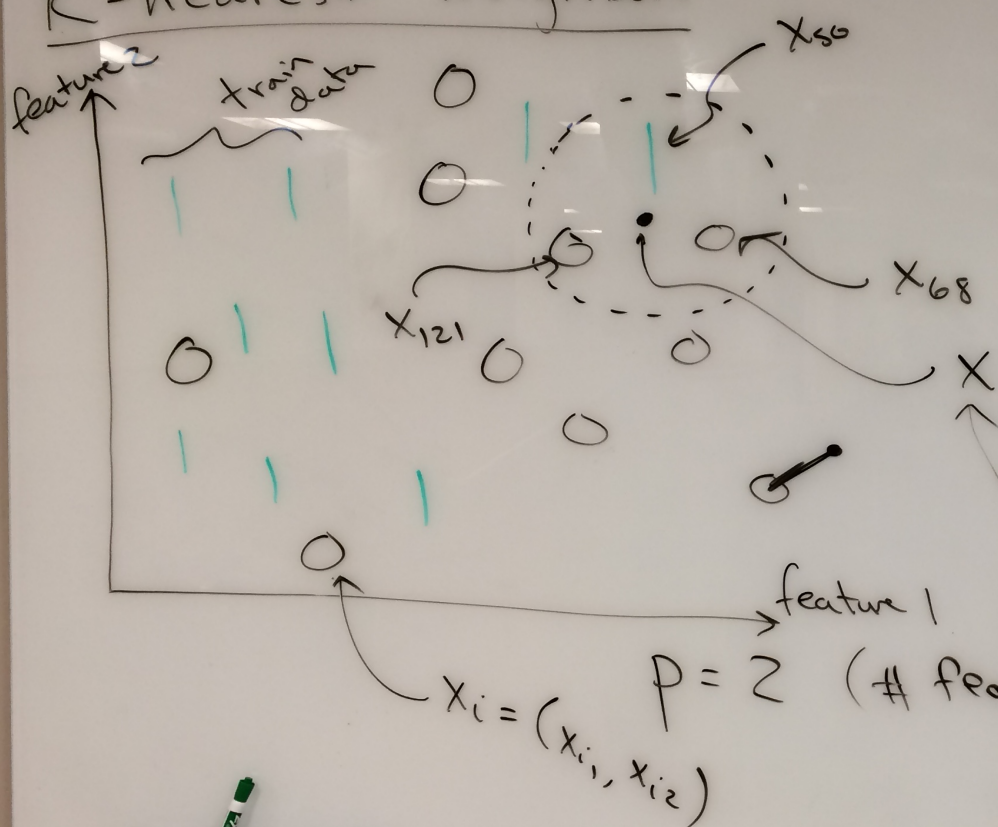
Prof. Sara Mathieson

Spring 2019



Lab 1: in-lab notes

K-nearest neighbors



k nearest neighbors.

$k=1 \Rightarrow$ nearest neighbor

$k=3 \Rightarrow$ 3 nearest neighbors

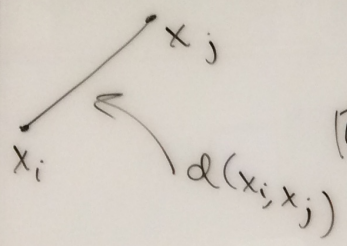
$y = 0$

test example.

Distance

Euclidean distance (what is close?)

$$d(x_i, x_j) = \sqrt{(x_{i1} - x_{j1})^2 + \dots + (x_{ip} - x_{jp})^2}$$



$N_k(x) = k$ nearest neighbors of x

EX $N_3(x) = \{x_{50}, x_{68}, x_{121}\}$

$$P(Y=c|x) = \frac{1}{k} \sum_{x \in N_k(x)} \mathbb{1}(y=c)$$

$$P(Y=0|x) = \frac{1}{3} \cdot 2 = \frac{2}{3}$$

$$P(Y=1|x) = \frac{1}{3} \cdot 1 = \frac{1}{3}$$

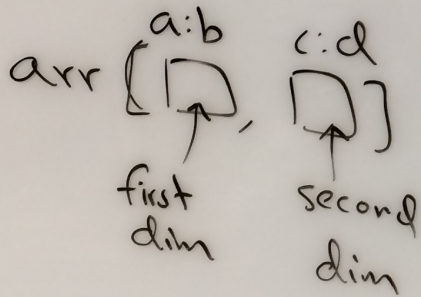
$$\hat{y} = \underset{c}{\operatorname{argmax}} (P(Y=c|x))$$

$\hat{y} = 0$

accuracy

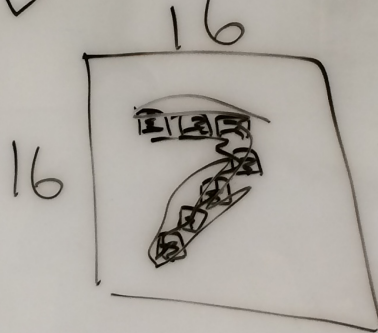
$$= \frac{1}{n} \sum_{i=1}^n \mathbb{1}(y_i = \hat{y}_i)$$

test
examples



Zipcode

✓



$$p = 256$$

