

Section B

Wednesday, February 1, 2023 9:15 AM

Def: Classification

$$D = \{(\vec{x}^{(i)}, y^{(i)})\}_{i=1}^N$$

$\forall i, \vec{x}^{(i)} \in \mathbb{R}^M$  ← features/attributes/inputs  
 $\forall i, y^{(i)} \in \{1, \dots, L\}$  ← label/class/output

$x$

$y$

$M = \# \text{ features}$

$N = \# \text{ training examples} = |D|$

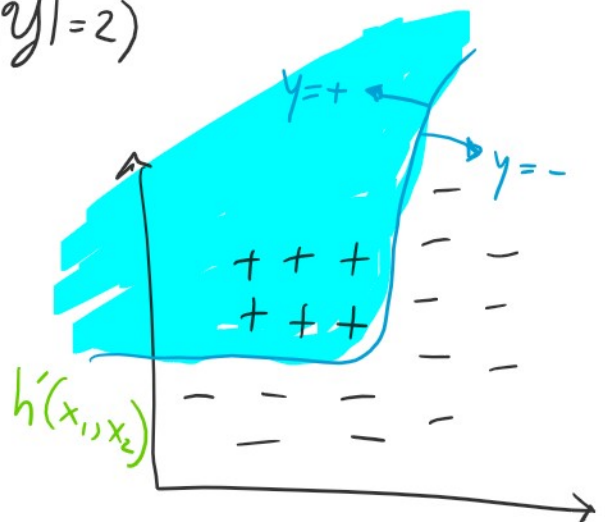
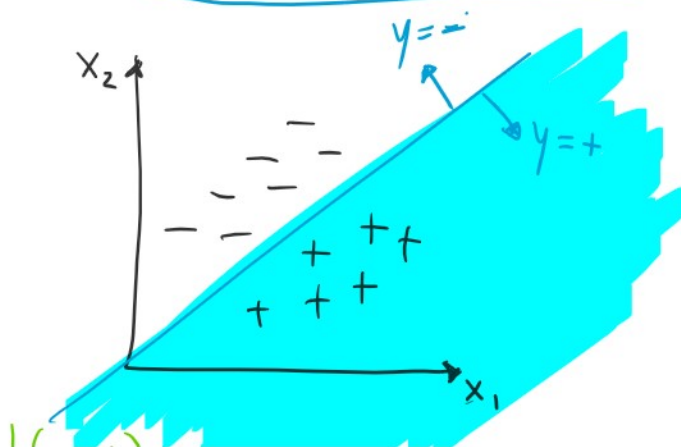
Def: Binary Classification  
 classification where  $|Y| = 2$   
 $\forall i, y^{(i)} \in \{+, -\}$   
 $\in \{0, 1\}$   
 $\in \{\text{red, blue}\}$

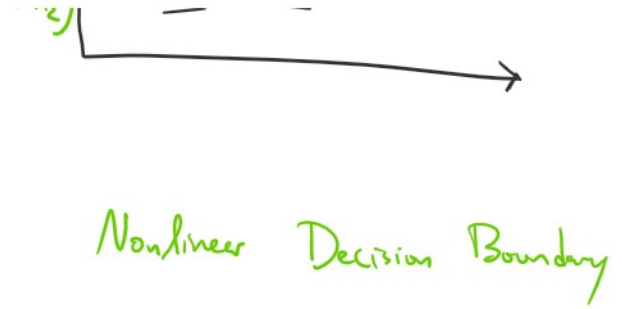
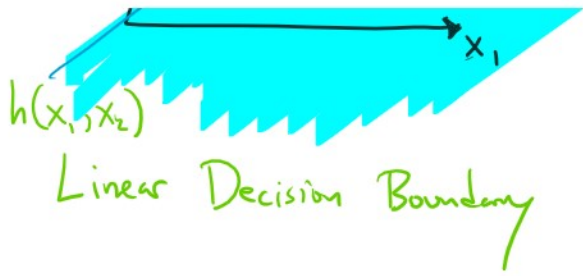
Def: Hypothesis (aka. Decision Rule) for Binary Class.

$$h: \mathbb{R}^M \rightarrow \{+, -\}$$

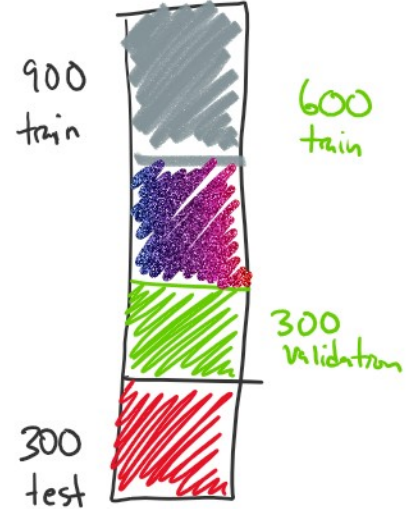
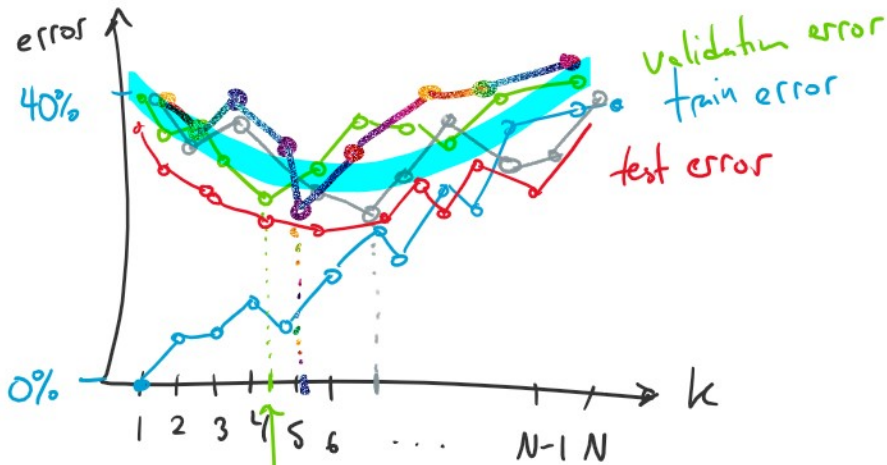
Train time learn  $h$   
Test time given  $\vec{x}$ , predict  $\hat{y} = h(\vec{x})$

Ex. 2D Binary Classification ( $M=2, |Y|=2$ )





Choosing  $k$  for  $kNN$



Assume  $D$  is 40%  $y^{(i)} = 0$  ←  
 60%  $y^{(i)} = 1$

best  $k$  on validation