Lecture 9/17: Linear Classifiers and Perceptron

Tuesday, September 17, 2019 12:20 PM

Review 9/12:

Model Selection: pick parameters of learning algorithm
- Train/validation split
- k-fold cross validation

Model Assessment

Q1: What is the generalization error of h?
- Test null hypothesis $err_P(h) \geq \varepsilon$
  ↳ # of test errors is Binomial $e \sim Binom(p \geq \varepsilon, n)$
- Confidence intervals: $err_P(h) \in [l, u]$
  ↳ Hoeffding's inequality

Q2: Is h_1 better/worse than h_2?
- Test null hypothesis $err_P(h_1) = err_P(h_2)$
  ↳ McNemar's test / Binomial sign test

$[err_P(h_1) = err_P(h_2)] \rightarrow$ win/losses $d_1 \sim Binom(p=0.5, n=d_1+d_2)$

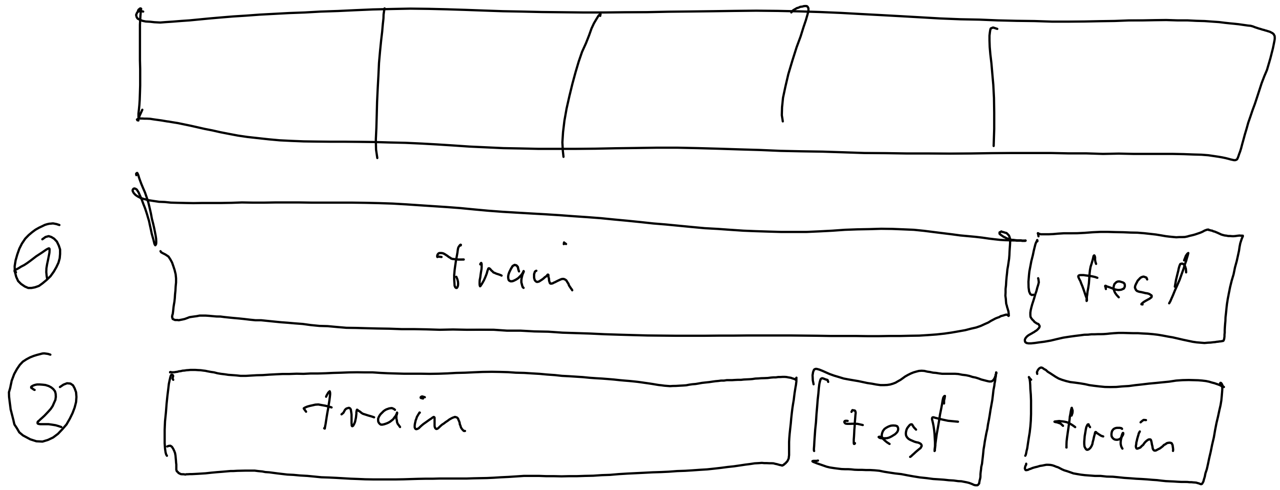
d_1 = number of wins by h_1

d_2 = number of wins by h_2

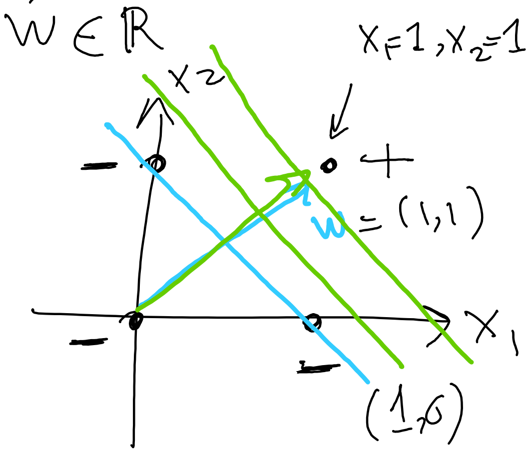
draws $h_1(x) = h_2(x)$

Q3: Is algorithm H_1 better/worse H_2?

H_1: DT $\qquad$ H_2: k-NN



$\vec{x} = (x_1, x_2) \in \{0,1\}^2$ $\qquad$ $\vec{w} \in \mathbb{R}^2$



1) $x_1 \wedge x_2$:

$\vec{w} = (1,1)$ $\qquad$ $b = -1 \to -1.99$

$h_{\vec{w},b}(x) = \text{Sign}(x_1 + x_2 - 1)$

$(1,0) \cdot \overset{w_1}{(1,1)} = 1$

$h(1,1) = \text{Sign}(\vec{w} \cdot \vec{x} - 1) = \text{Sign}(1) = +1$

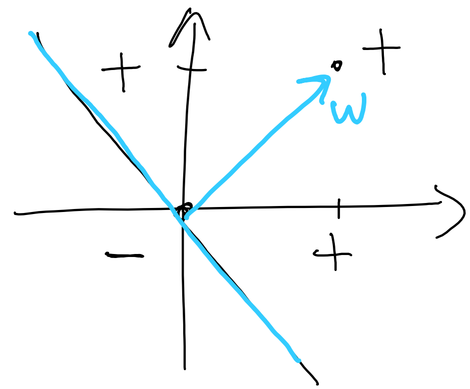
$(x_1, x_2) = (1,1)$ $\qquad$ $h_{\vec{w},b}(1,1) = \text{sign}\left(\underset{(1,1)\,(1,1)}{\vec{w}\cdot\vec{x}} - 1\right) = \text{sign}(1) = 1$

2) $x_1 \vee x_2$

$w = (1,1) \qquad b = 0$

$$h_{\vec{w},b}(\vec{x}) = \text{Sign}(x_1 + x_2)$$

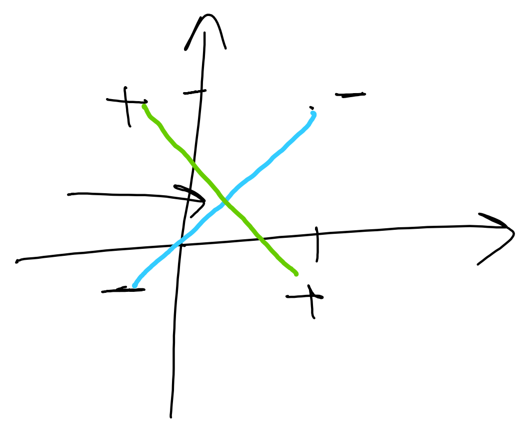
$$h_{\vec{w},b}((0,0)) = \text{Sign}(0) = -1$$



b=0

3) $x_1 \oplus x_2$

Cannot represent using a linear classifier



LP: linear Program $(\vec{u})$

$$\min \quad \vec{u} \cdot \vec{w}$$

$\vec{w}$

Subject to $\underbrace{\vec{a}_i \cdot \vec{w}}_{\text{linear}} \geq b_i$

We know

$y_i(\vec{w}^* \cdot \vec{x}_i) \geq \gamma$

$\Downarrow$

$y_i\left(\left(\frac{\vec{w}^*}{\gamma}\right) \cdot \vec{x}_i\right) \geq \underline{1}$

Wish:

$y_i(\vec{w} \cdot \vec{x}_i) > 0$

$\Downarrow$

Instead: $\min_{\vec{w}} \quad \vec{w} \cdot \vec{1} \; / \; \underline{1}$

Subject to $\forall \; i \in [m] \quad y_i(\vec{w} \cdot \vec{x}_i) \geq \underline{1}$

$h_{\vec{w}}(\vec{x}) =$ Consistent.