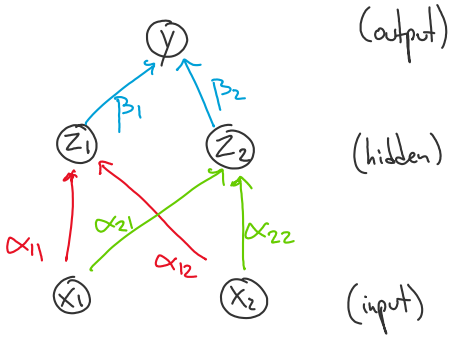
Section B: Neural Networks

Monday, October 3, 2022 10:04 AM

Ex#1: NN w/ 1 hidden and 2 hidden units



(output)

(hidden)

(input)

$x_m \in \mathbb{R}$

$z_i \in [0,1]$ if σ is sigmoid

$z_i \in \mathbb{R}$ more generally

$y \in [0,1]$ if σ sigmoid

Let σ be the activation function

If σ is sigmoid: $\sigma(a) = \frac{1}{1+\exp(-a)}$

$$z_1 = \sigma(\alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{10})$$
$$z_2 = \sigma(\alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{20})$$

each is like a "Log. Reg. Model" function

$$y = \sigma(\beta_1 z_1 + \beta_2 z_2 + \beta_0)$$
$$= \sigma(\beta_1 \sigma(\alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{10}) + \beta_2 \sigma(\alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{20}) + \beta_0)$$

don't forget the intercept/bias term

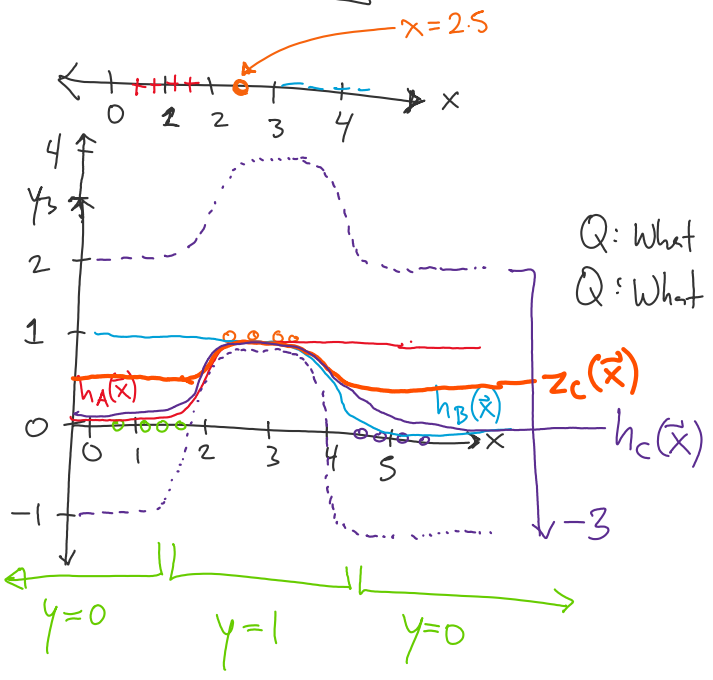
treat this as $\Pr[Y=1 \mid \vec{x}, \alpha, \beta]$

Use the Bayes Optimal Classifier

$$\hat{y} = h(\vec{x}) = \underset{\hat{y}}{\operatorname{argmax}} \Pr[Y=\hat{y} \mid \vec{x}, \alpha, \beta]$$

$$= \begin{cases} 1 & \text{if } y \geq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

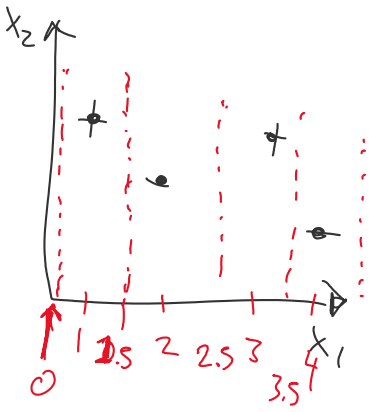
1D Face Recognition

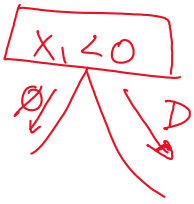


Q: What is $z_C(\vec{x}) = \frac{h_A(\vec{x}) + h_B(\vec{x})}{2}$

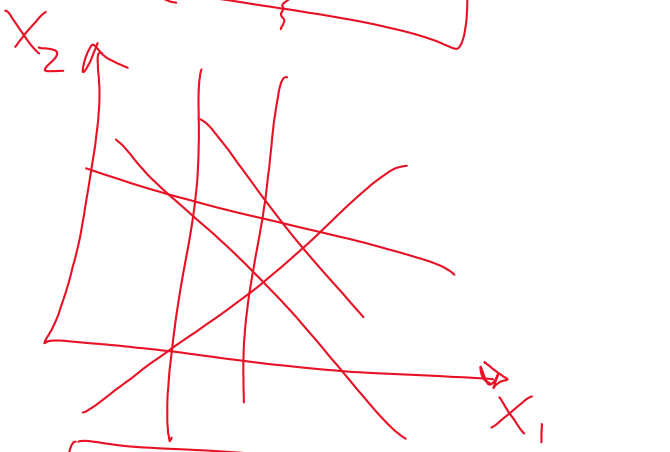
Q: What is $h_C(\vec{x}) = \sigma(w_A h_A(x) + w_B h_B(x) + w_C)$ where $w_A = 2$ $w_B = 2$ $w_C = -3$

OH





$$w^T x + b = 0$$



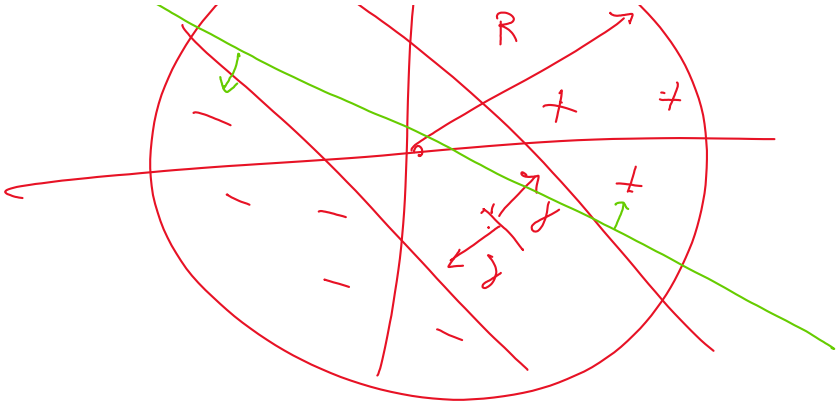
$$k \leq \left(\frac{R}{\gamma}\right)^2$$

radius

margin

$$\vec{x} \in \{0,1\}^M$$

$$\vec{x} = [1,1,0,1,1,1,0]$$


R
+
+
+
−
−
−
−
−
x
y
y