Lecture 4: 2/7/18

Supervised Learning $\mathbb{D}, \mathcal{H}, \mathcal{A}$

$g = \mathcal{A}(\mathbb{D}, \mathcal{H})$ → this is a model.

How to get predict: or fits?

$\mathbb{D} = \{ \langle \vec{x}_i, y_i \rangle \}$ $i = 1..n$

$g = \mathcal{A}(\mathbb{D}, \mathcal{H})$

$\hat{y}_i = g(\vec{x}_i)$ is called sample fits — "fit"

$\{\hat{y}_1, \ldots, y_n\}$ is sample fits/predictions

$\approx$

$\{y_1, \ldots, y_n\}$

How to predict for new data/observat: $\vec{x}^*$?

$\hat{y}^* = g(\vec{x}^*)$

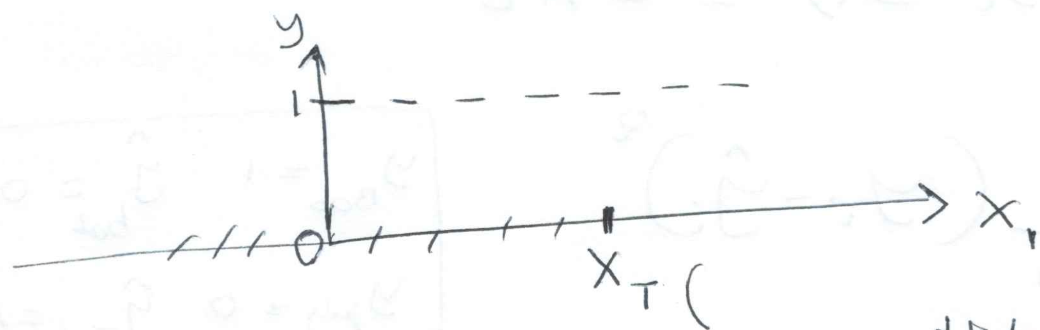
$y \in \{0, 1\}$ is binary.

let's use only

x_1 (salary, con

$\mathcal{H} = \{$

Let's graph.



$$\mathcal{H} = \{ \mathbb{1}_{x > x_T} : x_T \in \mathbb{R} \}$$

is every possible functions that look like

This is called parameter

Note

$$\mathbb{1}_A = \begin{cases} 1 & \text{if } A \\ 0 & \text{if } A^c \end{cases}$$

$$\mathbb{1}_{x > 10} = \begin{cases} 1 & \text{of } x > 10 \\ 0 & \text{of } x \leq 10 \end{cases}$$

Now

$$\mathcal{H} = \{ \mathbb{1}_{x > x_T} : x_T \in \{x_1, \ldots, x_n\} \}$$

Some error function.

$$Err(\vec{y}, \vec{\hat{y}}) > 0$$

sum of absolute error: $SAE = \sum_{i=1}^{n} |y_i - \hat{y}_i|$

Mean Abs Error.

$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$: called misscharacterization error.

sum of Square error: SSE ~~SSE~~ $SSE = \sum_{n=1}^{n} (y_i - \hat{y}_i)^2 = SA$

Mean of Square error: $MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

$$SSE = \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2 = SAE$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2$$

$$\boxed{\begin{array}{ll} y_{Bob} = 1 & \hat{y}_{bob} = 0 \\ y_{July} = 0 & \hat{y}_{July} = 1 \end{array}}$$

$$MAE = \frac{1}{n} \sum \left| y_i - \hat{y}_i \right|$$

$$\leq \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{y_i \neq \hat{y}_i}$$

$$SSE(h) = \sum_{i=1}^{n} \left(y_i - h(\vec{x}_i)\right)^2$$

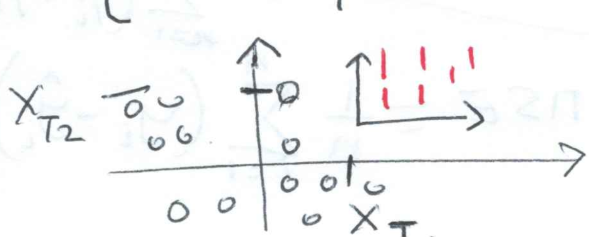
~~$g = argment\{SS$~~

$$g = avgmin\{SSE(h)\} \iff X_T^* = avgmin\Big\{\sum \big(y_i - \mathbb{1}_{\vec{x}_i >}$$

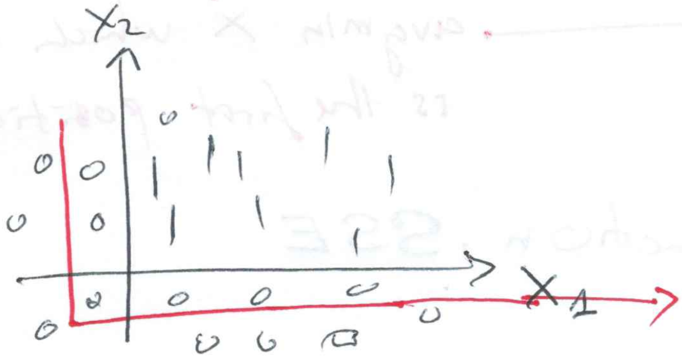
A would be a "great research" ± a very possible.

Now we have X_1 & X_2 both continuous.

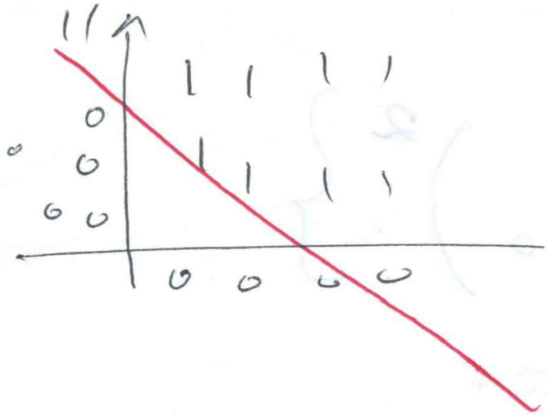
$$\mathcal{H} = \left\{ \mathbb{1}_{x_1 > x_T} \, \mathbb{1}_{x_2 > x_{T_2}} : \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \right\}$$



Let say your data looks like this:



arbitrary bad.



Best to do.

a+bx -line.

<u>Now</u>:

$$H = \left\{ \mathbb{1}_{x_2 > a\cancel{x} + bx} : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 \right\}$$

$$= \left\{ \mathbb{1}_{a+bx - x_2 < 0} : \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 \right\}$$

$$= \left\{ \mathbb{1}_{-a-bx + x_2 > 0} ; \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R} \right\}$$

Linear Classific

$$= \left\{ \mathbb{1}_{w_0 + w_1 x_1 + w_2 x_2 > 0} : \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} \in \mathbb{R}^3 \right\}$$

$$= \left\{ \mathbb{1}_{w_0 + \vec{w} \cdot \vec{x} > 0} : w_0 \in \mathbb{R}, \vec{w} \in \mathbb{R}^2 \right\}$$

let $\vec{x} = [1, x_1, x_2]$

argmin $\vec{x}$ which a 1 is the first position

$$= \left\{ \mathbb{1}_{\vec{w} \cdot \vec{x} > 0} : \vec{w} \in \mathbb{R}^3 \right\}$$

Use same error function, SSE

$$g = \text{argm} \left\{ SSE(h) \right\}$$

$$\Updownarrow$$

$$\vec{w} = \text{argm} \left\{ \sum_{i=1}^{n} \left(\overset{y_i}{} \mathbb{1}_{\vec{w} \cdot \vec{x}_i > 0} \right)^2 \right\}$$

This is Hard Problem.