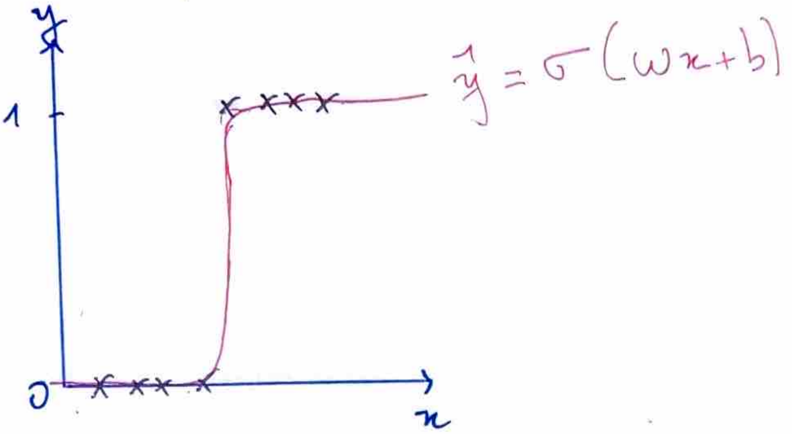
Logistic Regression Model

- Used for Binary Classification:

for one feature x:

$$\sigma(t) = \frac{1}{1-e^{-t}}$$

(Sigmoid)



In general, $\hat{y} = \sigma(w_n x_n + \dots + w_1 x_1 + b)$

x_i: features
w_i: parameters of the i-th feature
b : constant parameter.

$\Rightarrow$ If $x = [x_1 \dots x_n]$ and $w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$

$$\hat{y} = \sigma(xw + b)$$

Since we need a binary classification:

$$\begin{cases} \hat{y} = 1 & \text{if } \hat{y} \geq 0{,}5 \\ \hat{y} = 0 & \text{else.} \end{cases}$$

The cost function is: $Cost = -\frac{1}{m} \sum_{i=1}^{m} \left[y \log \hat{y} + (1-y) \log(1-\hat{y}) \right]$

After some calculus:

$$\frac{\partial cost}{\partial w} = \frac{(\hat{y}-y)^T . X}{m} \quad \text{and} \quad \frac{\partial cost}{\partial b} = \frac{\text{sum}(\hat{Y}-Y)}{m}$$

⚠ Note that $\frac{\partial cost}{\partial w}$ is $(1 \times m)$

$\boxed{\text{Accuracy}}$

If $\hat{y} = \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_m \end{pmatrix}$, we want $\hat{y}_i = 1$ iff $\hat{y}_i > 0,5$

then accuracy $= 1 - \dfrac{\sum_{i=1}^{m} |y_i - \hat{y}_i|}{m}$