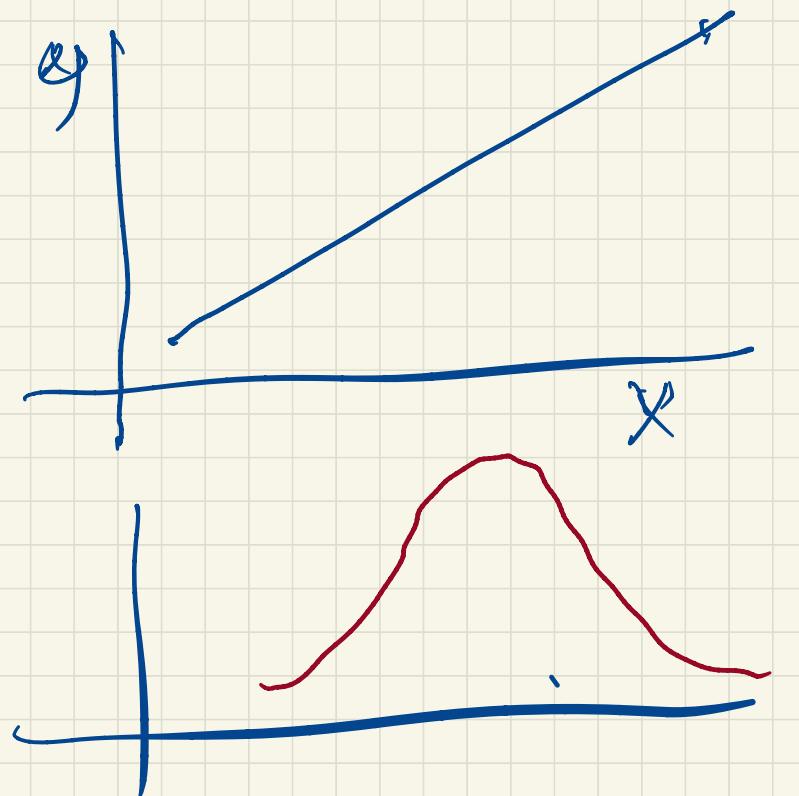


m^2	Precio	Predecir el Precio
150	450	
120	380	
170	480	
80	270	
11	11	VARIABLE OBJETIVO
12	12	

↳ $y = \alpha + \beta x$

OBJETIVO

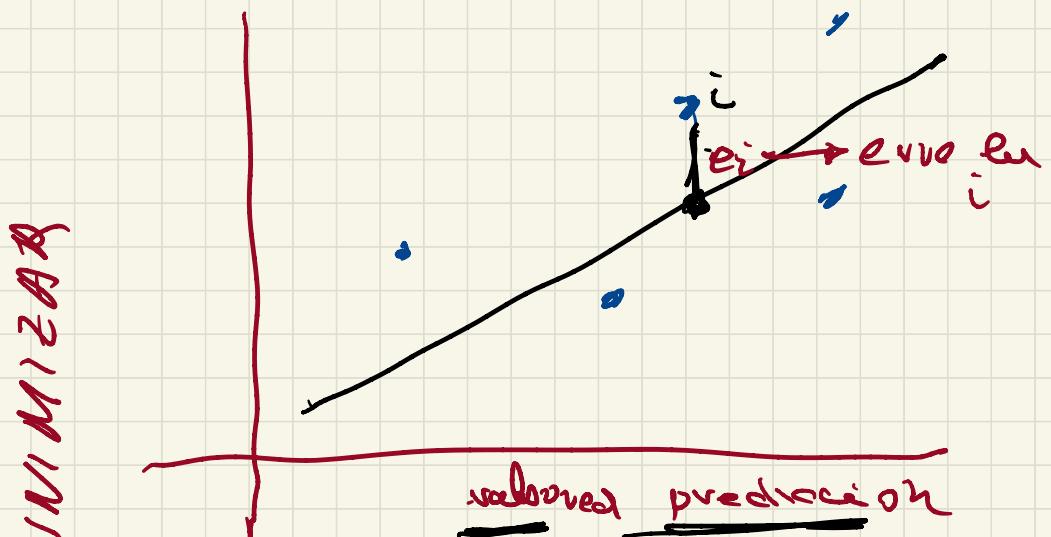
PREDICTOR



$$\text{ALQUILER} = \alpha + \beta * \text{SUPERFICIE}$$

U

$$y = 100 + 2,5 * \text{SUP}$$



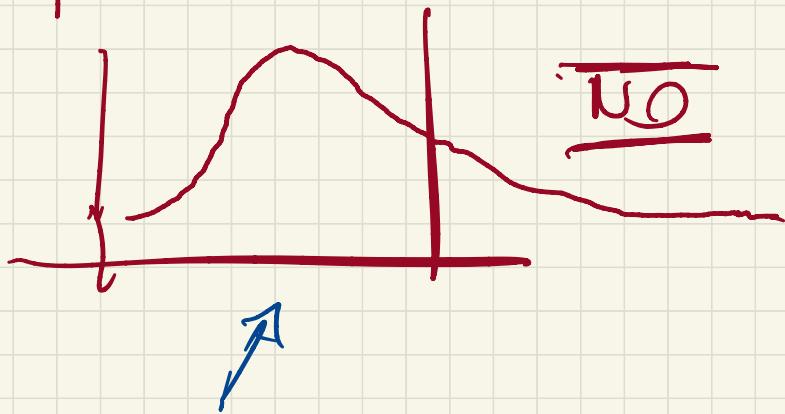
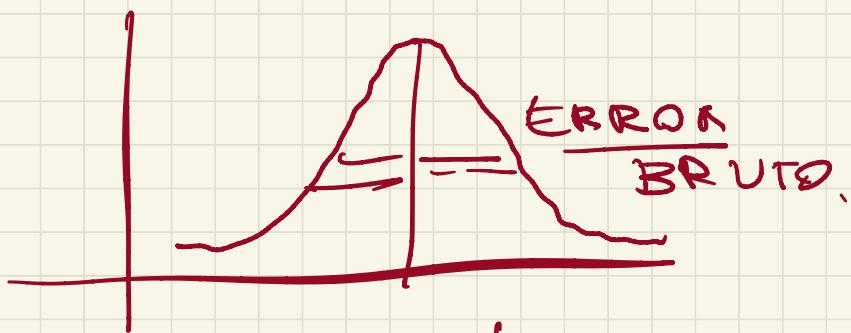
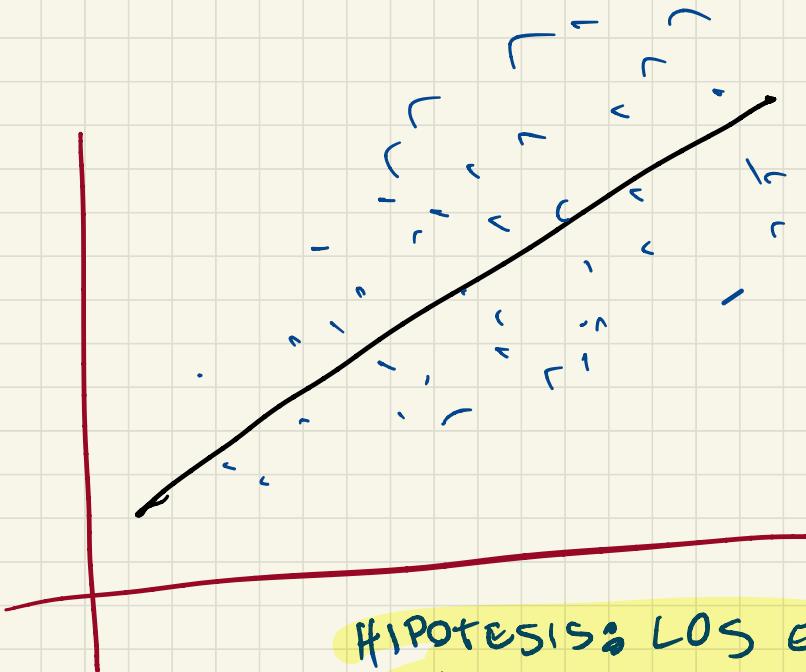
$$e_i = y_i - \hat{y}_o(x_i)$$

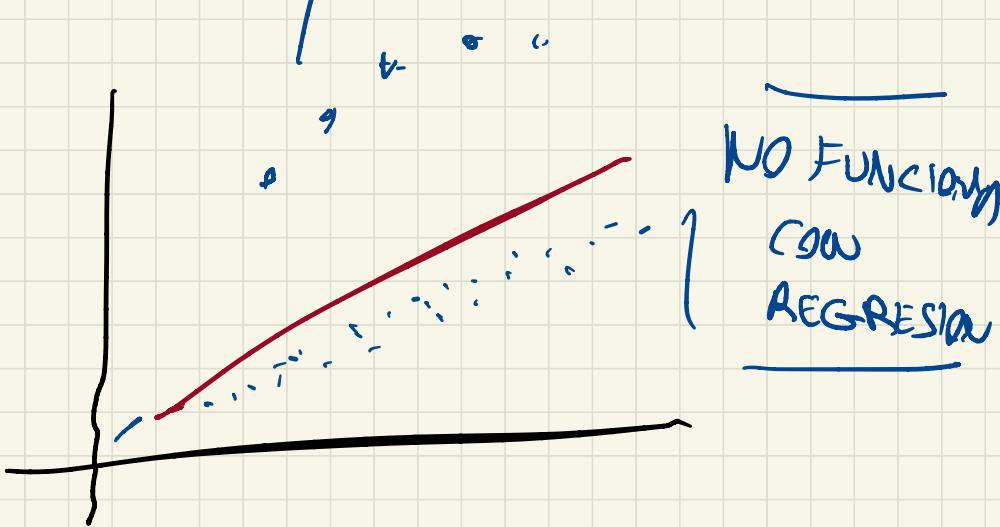
$$\sum e_i^2 = \sum_i^n (y_i - \hat{y}_o(x_i))^2 =$$

$$\hat{\sigma}^2 = \sum_i^n (y_i - (\alpha + \beta x_i))^2$$

$$\beta = \frac{\text{Cov}(x, y)}{\text{Var}(x)}$$

$$[\alpha = \bar{y} - \beta \cdot \bar{x}]$$

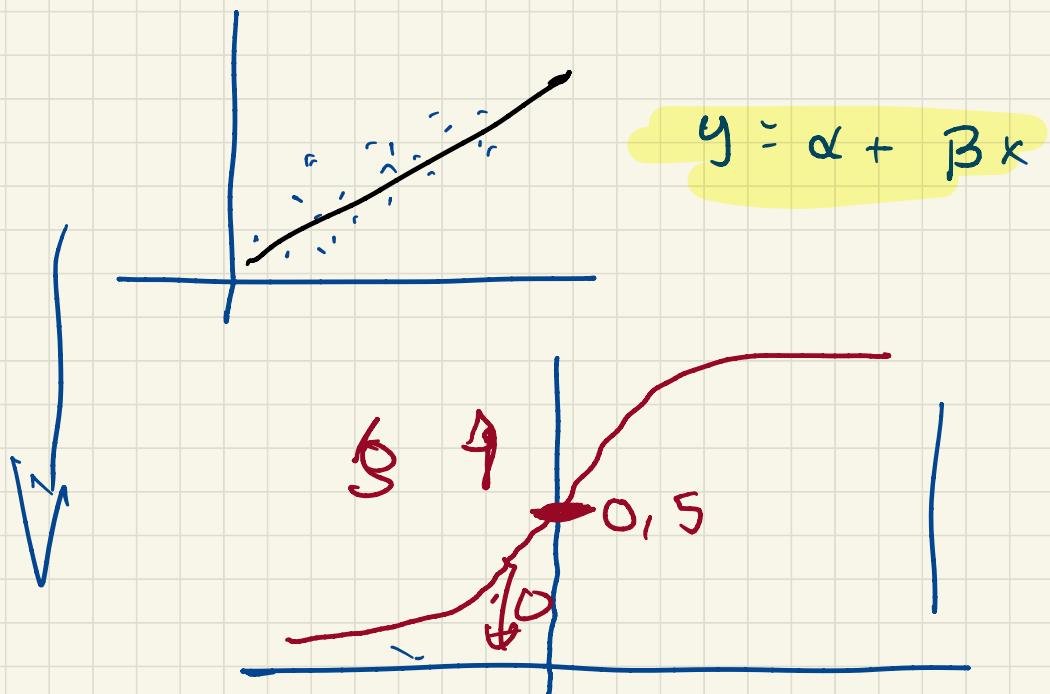




HAY QUE COMPROBAR LA
NORMALIDAD ERRORES

REGRESIÓN LOGÍSTICA

- Utiliza un modelo de regresión para clasificar



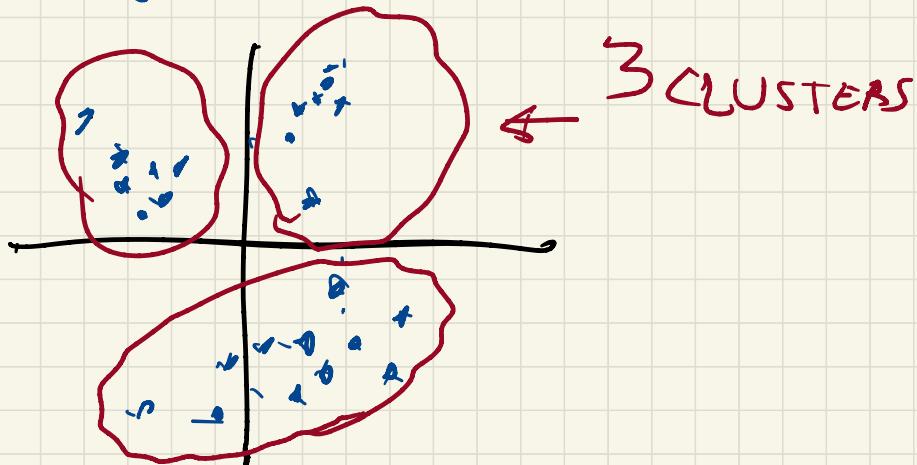
$$y = \frac{1}{1 + e^{-(\alpha + \beta x)}}$$

	R. LINEAL	R. LOGIST
PREDICTORA	- NUMERICAS - CATEGÓRICAS	- NUMERICAS - CATEGÓRICAS
PREDICTORES	- NUMERICAS CONTINUAS	- CATEGÓRICAS
RELACION	- LINEAL	- LINEAL (SE APLICAN TRANSFORMACIONES)



APRENDIZAJE NO SUPERVISADO

- Agrupación (CLUSTERING)



- Reducción de Dimensiones

$$X(a, b, c, d)$$


$$X_v(d_1, d_2, d_3)$$

a = edad

b = sex x 0

c = - - - - -

d =

d_v = dimensiones
de estudio

$$\underline{X} \rightarrow \underline{X_v} \rightarrow \text{Reducción Dimensión}$$

$$\underline{X_v} \rightarrow \underline{X} \rightarrow \text{Reconstrucción}$$

Si al reconstrucción

no sale bien

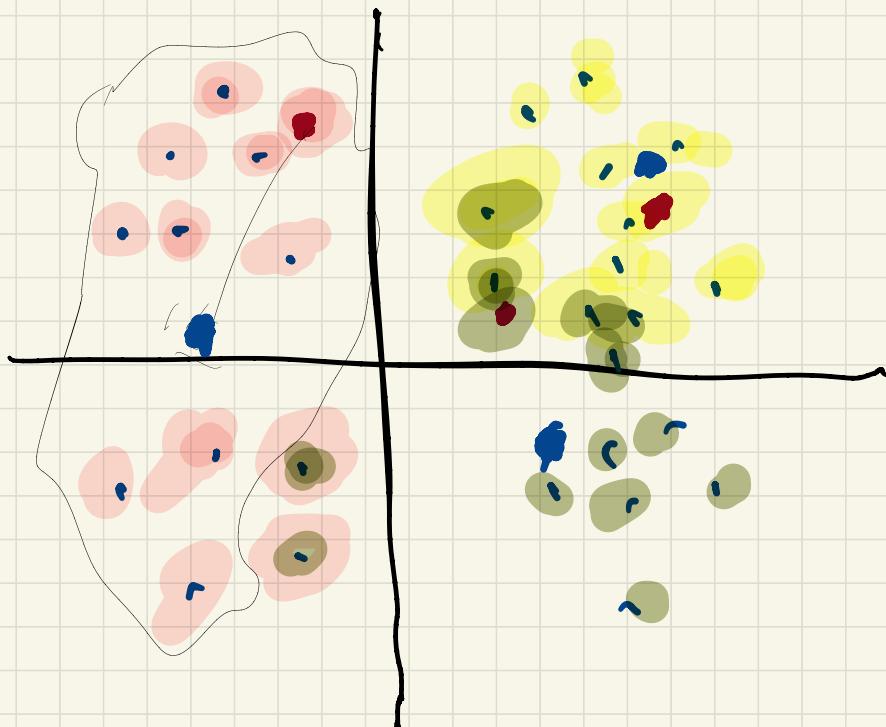
USAR PREPROCESO

DATO ANOMALO CON ALTA PROBABILIDAD

K - MEANS

$K =$ NÚMERO
DE
CLUSTERS

$K = 3$



1 → ASIGNAR LOS CENTROIDES DE FORMA ALEATORIA

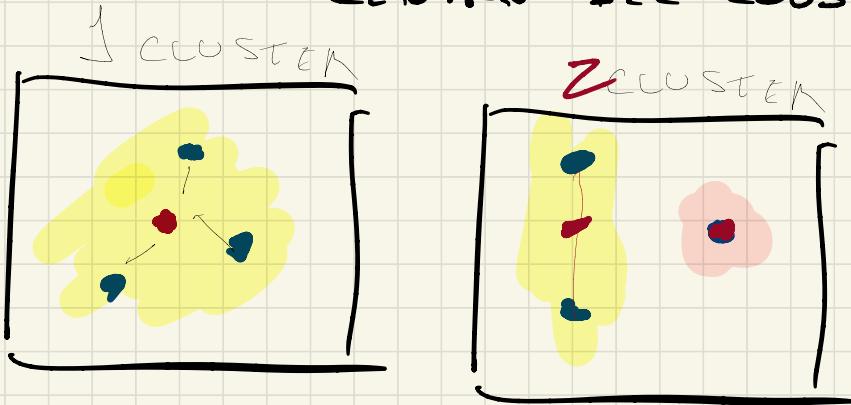
2 → a - ASIGNAR CADA ELEMENTO AL CENTROIDE MÁS CERCANO

b - BUSCAR EL CENTRO DE LOS ELEMENTOS DE CADA CLUSTER Y HACERLOS NUEVOS CENTROIDES

3 → PARAMOS CUANDO NO HAY NINGUNO QUE REASIGNAR

CLUSTER INERTIA

- INERCIA = SUMA CUADRADOS DE LA DISTANCIA ELEMENTO AL CENTRO DEL CLUSTER

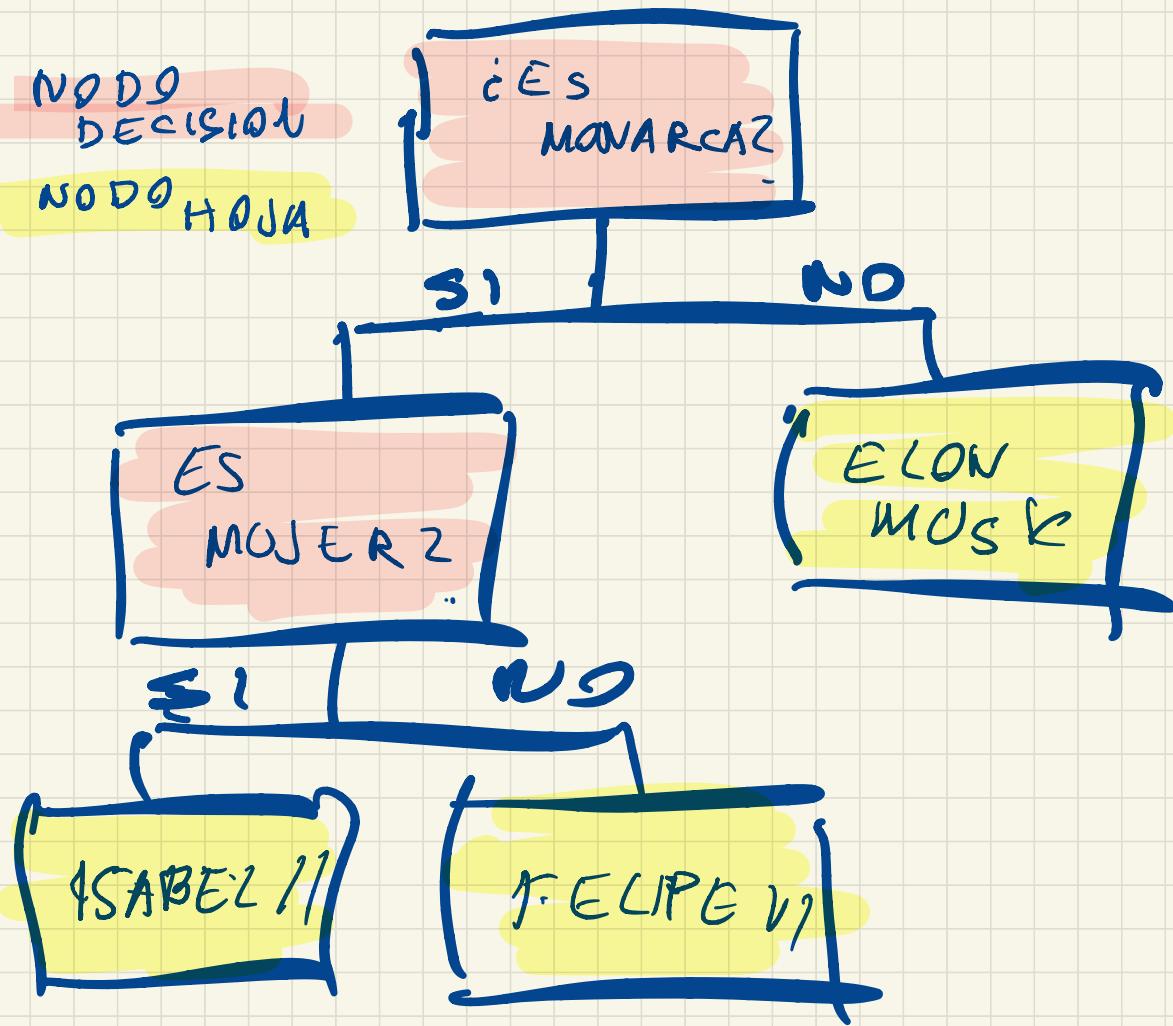


AKINATOR

- FELIPE VI
- ECON MUSK
- ISABEL II

NODO
DECISION

NODO
HOJA



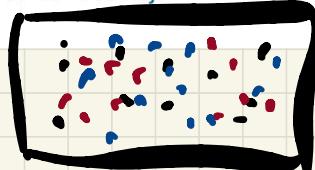
Paso 1: Comienza el árbol con el nodo raíz, llamado S, que contiene el conjunto de datos completo.

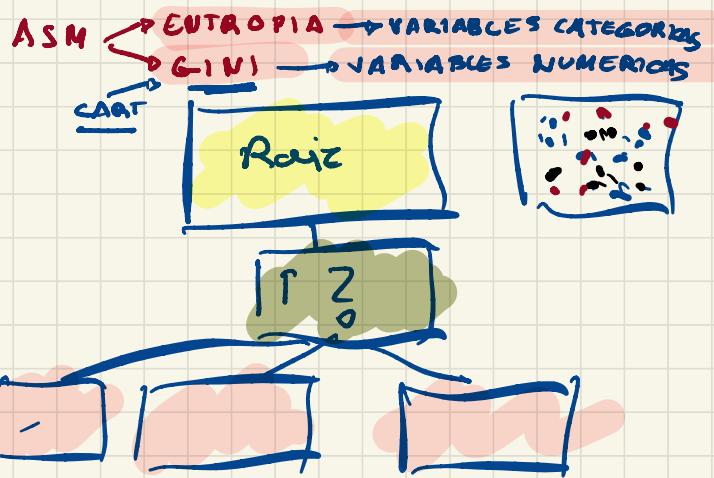
Paso 2: Encontrar el mejor atributo en el conjunto de datos utilizando la medida de selección de atributos (ASM)

Paso 3: Dividir S en subconjuntos que contengan los posibles valores de los mejores atributos

Paso 4: Generar el nodo del árbol de decisión que contiene el mejor atributo

Paso 5: Hacer recursivamente nuevos árboles de decisión utilizando los subconjuntos del conjunto de datos creados en el paso 3. Continúe este proceso hasta que se llegue a una etapa en la que no se puedan clasificar más los nodos y se llame al nodo final como nodo hoja



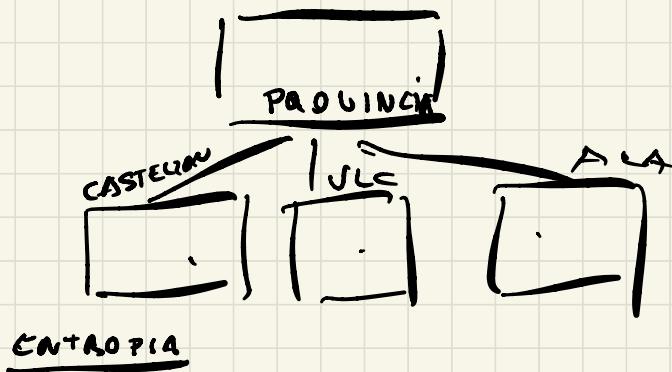


- PROVINCIA
- GÉNERO
- MAYOR DE EDAD

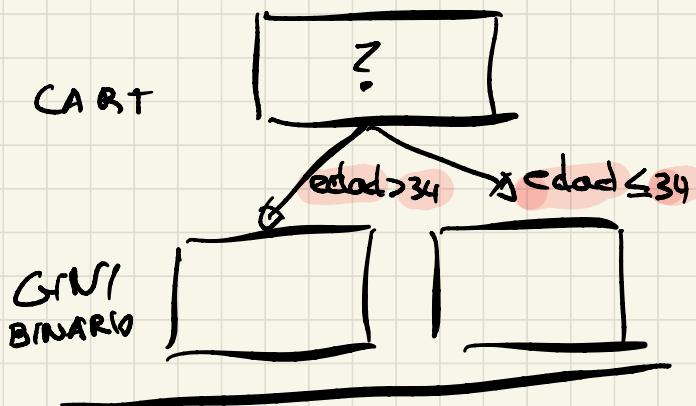
• ENTROPIA
• GANANCIA DE INFORMACIÓN



UN ATRIBUTO CON ALTA GANANCIA DE INFORMACIÓN DIVIDE EL DATASET EN CONJUNTOS MÁS HOMOGENEOS QUE UNO CON BAJA GANANCIA

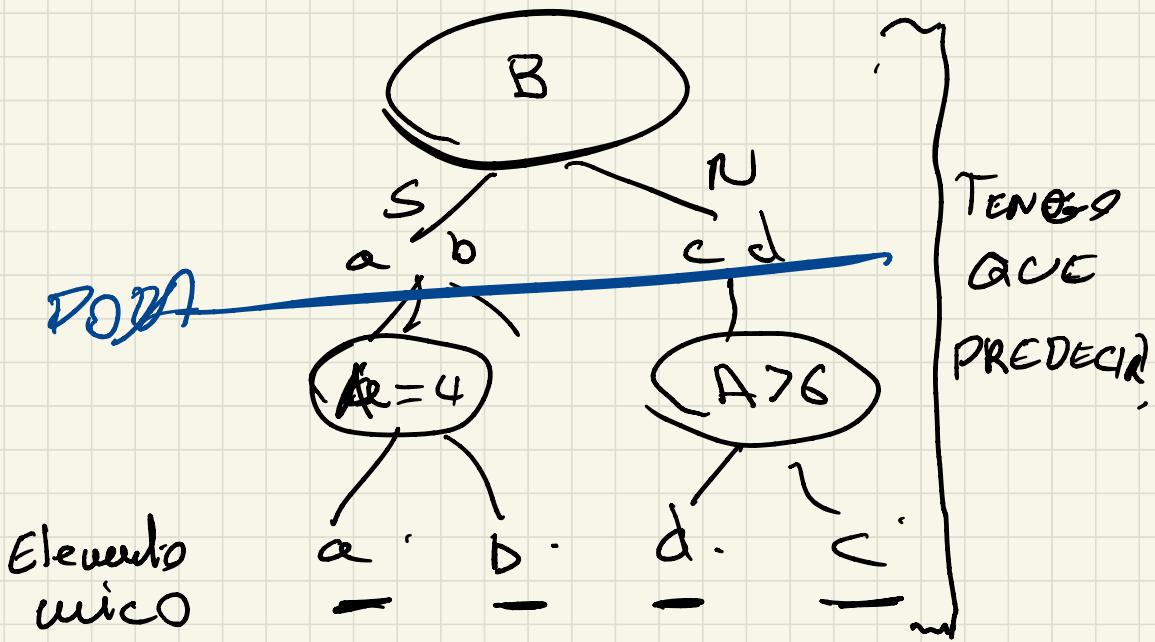


$$E(S) = \sum_{i=1}^k -p_i \cdot \log_2(p_i)$$



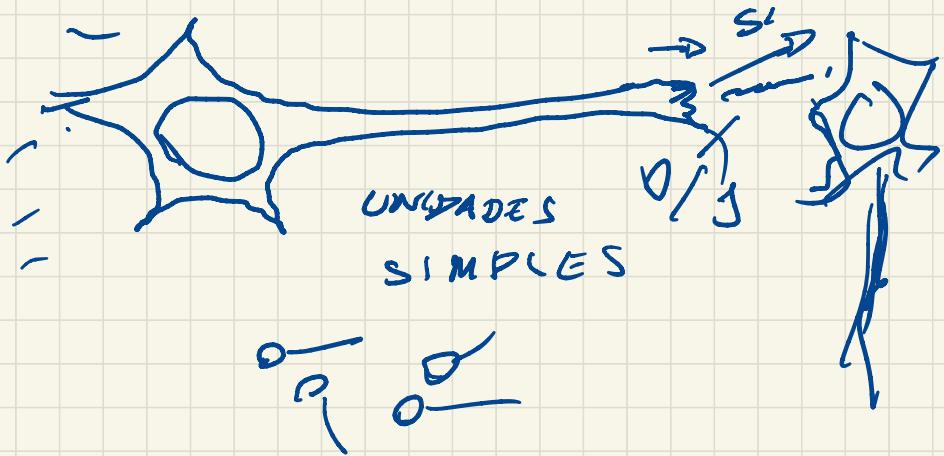
OVERFITTING - SOBREAJUSTE

	A	B	Y
a	4	S	S
b	6	S	N
c	8	N	S
d	10	N	N



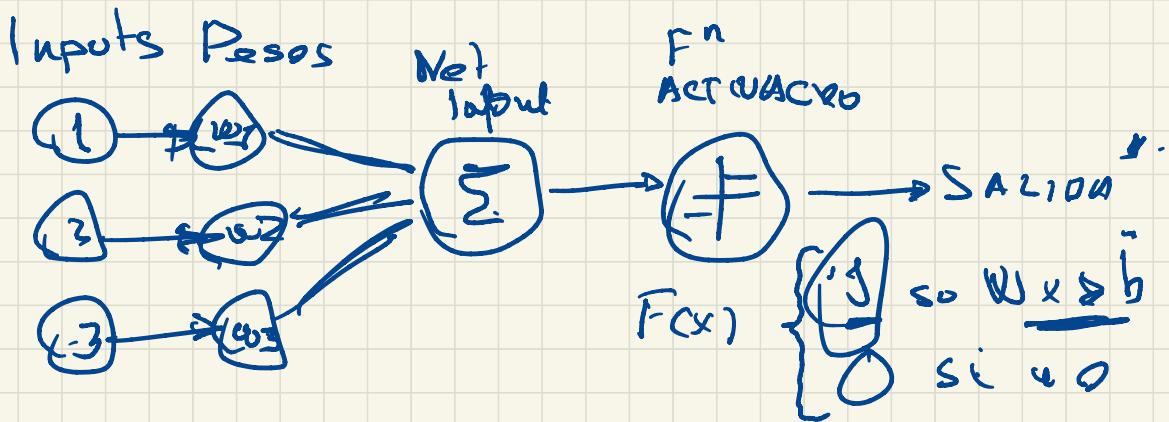
DEEP LEARNING

REDES NEURONAZGS



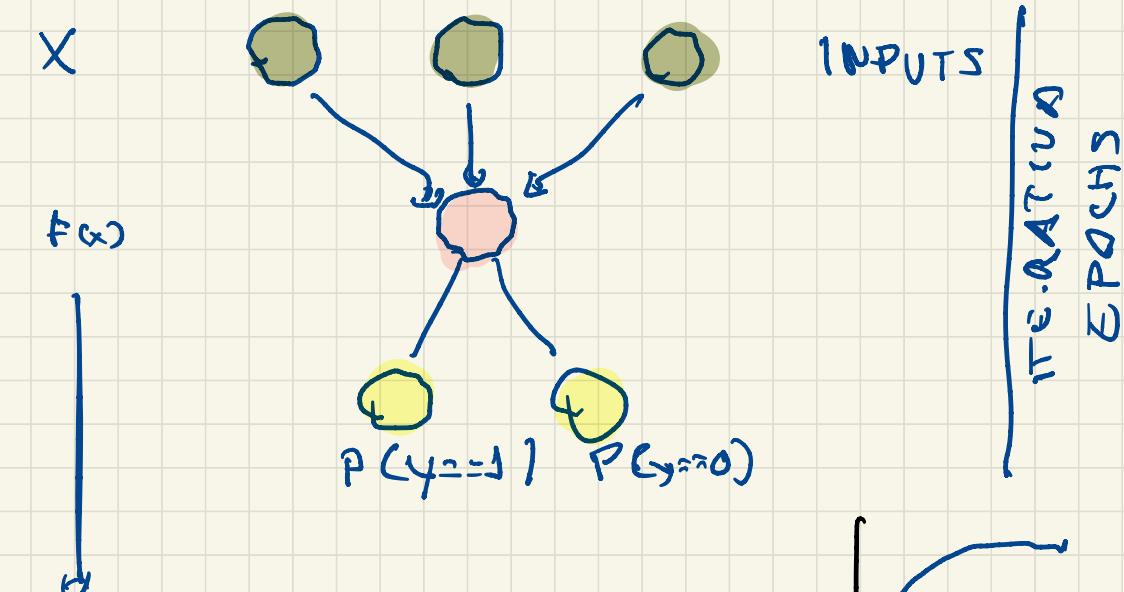
RED NEURONAL + SIMPLE

PERCEPTRON



- Seleccionar ω_i ALATORIO EN MUCHOS CASOS
 - Calculamos $F(x_i)$ para observar
 - Actualizamos los pesos,
- $$\hookrightarrow \Delta \omega = \alpha (y_i - F(x_i))x_i$$

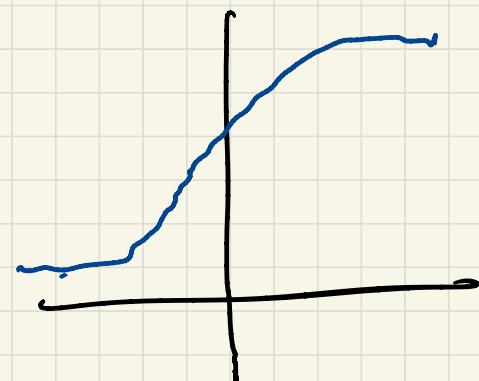
RED NEURONAL & NEURONA

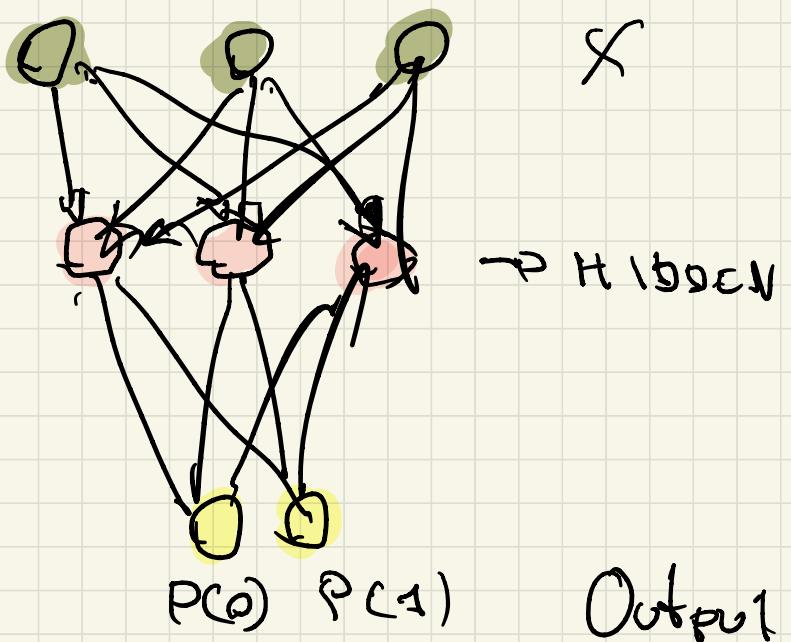


Fn Activación

SIGMOIDEA

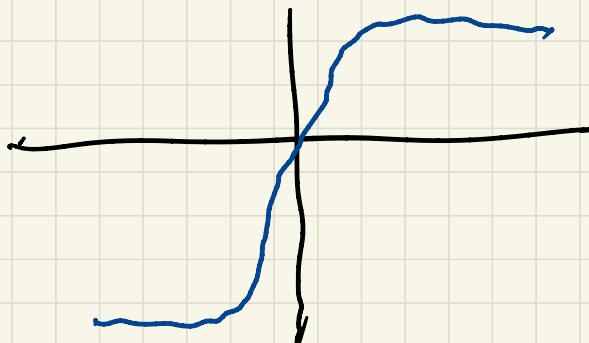
$$F(x) = \frac{1}{1 + e^{-x}}$$



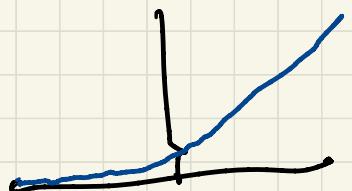


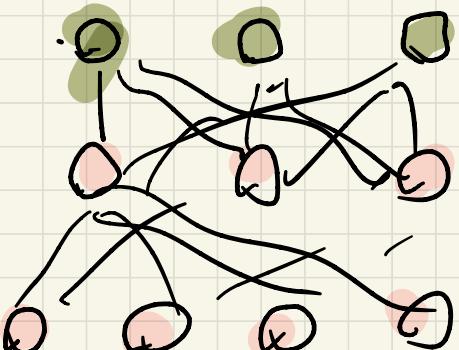
OTRAS FUNCIONES ACTIVACIÓN:

TANGENTE HIPERBÓLICA

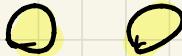


RECTIFICADORA





3 HIDDEN
LAYERS



TENSORFLOW

↳ DEDICADA A REDES
NEURONALES

↳ KERAS

↳ ACCELERAR LOS CALCULOS
DE GRANDES MATRICES
MULTIDIMENSIONALES
(TENSOR)