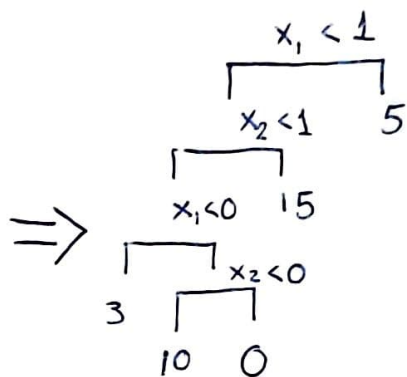
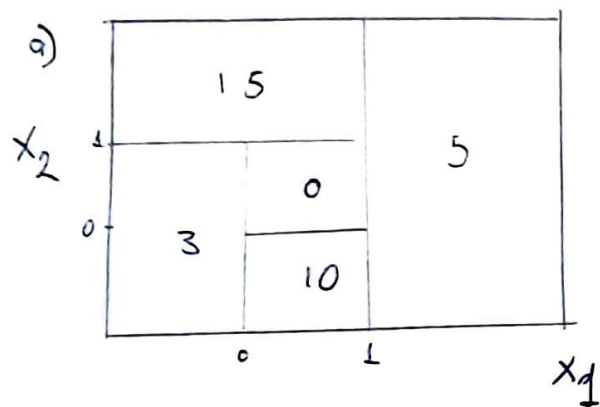
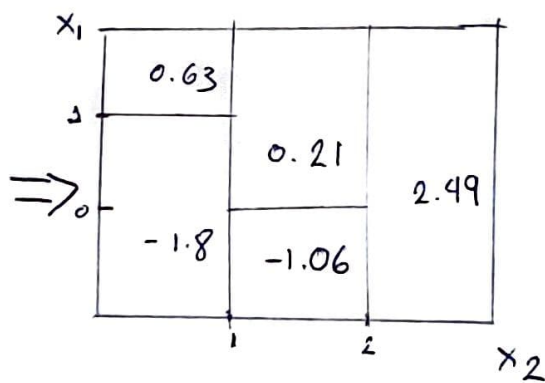
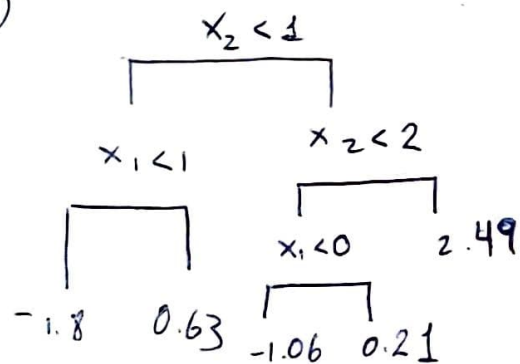


①



⑥ Son múltiples árboles de decisiones con diferentes variables predictoros

b)



⑤ Boosting
Es un método de ensemble que puede combinar muchos algoritmos débiles, convirtiéndolos en el algoritmo fuerte. La idea es entrenar secuencialmente y en cada nuevo intento corregir el intento precedente.

④ La variable más importante en el caso

a) x_1

b) x_2

Bagging

Por cada entrenamiento en diferentes subconjuntos del training set. cuando el muestreo de subconjuntos se utiliza el remplazamiento, es decir, permite entrenar con muestras muchas veces para el mismo predictor.

③

Hard Voting

Es posible que sea 0.6

ya que hay mayor

$$0.6, 0.6 \approx 0.65$$

peso entre ese rango $[0.55, 0.75]$

Soft Voting

$$= \frac{0.1 + 0.15 + 0.2 + 0.2 + 0.55 + 0.6 + 0.6 + 0.65 + 0.7 + 0.75}{10}$$

$$= 0.45;$$

⑦ verdadero negativo (-) (+) Falso positivo

(-)	0	0
(+)	4	6

FN TP Verdadero positivo

L	A	A	B	A	B	A	A	B	B	A
P	A	B	B	B	B	B	A	A	B	A
	V	X	V	X	V	X	V	X	V	V
	1	0	1	0	1	0	1	0	1	1

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$P = \frac{6}{6 + 0}$$

$$P = \underline{1}$$

$$F_1 = \frac{2}{\frac{1}{P} + \frac{1}{R}}$$

$$F_1 = \frac{2}{1 + \frac{5}{3}}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$R = \frac{6}{6 + 4}$$

$$R = \frac{6}{10} \rightarrow \underline{\frac{3}{5}}$$

$$F_1 = \underline{0.75}$$

8

DBSCAN no requiere un número de clusters

K-means funciona con centroides y DBSCAN funciona con distancias (ϵ) y con un número mínimo de muestras, si un punto cumple con esta condición es considerado un punto núcleo.

9

Generalmente es menor. Pero depende del número de muestras del nodo por ejemplo si solo hay una partición de igual número de muestras. La partición sería a la mitad.

$$G = \frac{1}{2} - \quad G = \frac{p_1}{\left(\frac{1}{2}\right)} \left(1 - \frac{1}{2}\right) + \frac{p_2}{\left(\frac{1}{2}\right)} \left(1 - \frac{1}{2}\right)$$

$$G = \frac{1}{2} \quad \frac{1}{2}$$

10

- Capa ocultas
- número de neuronas
- función de activación (step(z))



11

No, este algoritmo no es sensible a la presentación de los datos. Puede manejarlos en su escala original o en otra, pero seguro que tendrá el mismo desempeño.

12

A

$$G = 1 - \left[\left(\frac{3}{18} \right)^2 + \left(\frac{8}{18} \right)^2 + \left(\frac{7}{18} \right)^2 \right]$$

$$G = \underline{0.6234}$$

B

$$G = 1 - \left[\left(\frac{2}{20} \right)^2 + \left(\frac{2}{20} \right)^2 + \left(\frac{16}{20} \right)^2 \right]$$

$$G = \underline{0.34}$$