

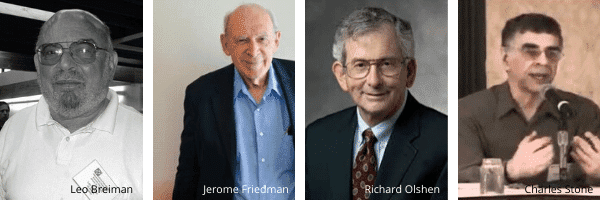
TRABAJO FINAL 2

***Qué es un árbol de decisión***

Un árbol de decisión es un modelo predictivo que divide el espacio de los predictores agrupando observaciones con valores similares para la variable respuesta o dependiente

* **Cuál es el origen de los árboles de decisión:**

Los creadores de la metodología del árbol de clasificación con aplicación al aprendizaje automático, también llamada metodología CART, fueron Leo Breiman, Jerome Friedman, Richard Olshen y Charles Stone. Su aplicación en el ámbito de la Estadística se inició en 1984.



* **Qué tipo de algoritmo es un árbol de decisión**

Los algoritmos de aprendizaje automático se clasifican en dos tipos:

* Supervisados.
* No supervisados.

Un árbol de decisión es un algoritmo supervisado de aprendizaje automático porque para que aprenda el modelo necesitamos una variable dependiente en el  
conjunto de entrenamiento.

* **Estructura básica de un árbol de decisión**

Los árboles de decisión **están formados por nodos y su lectura se realiza de arriba hacia abajo.**

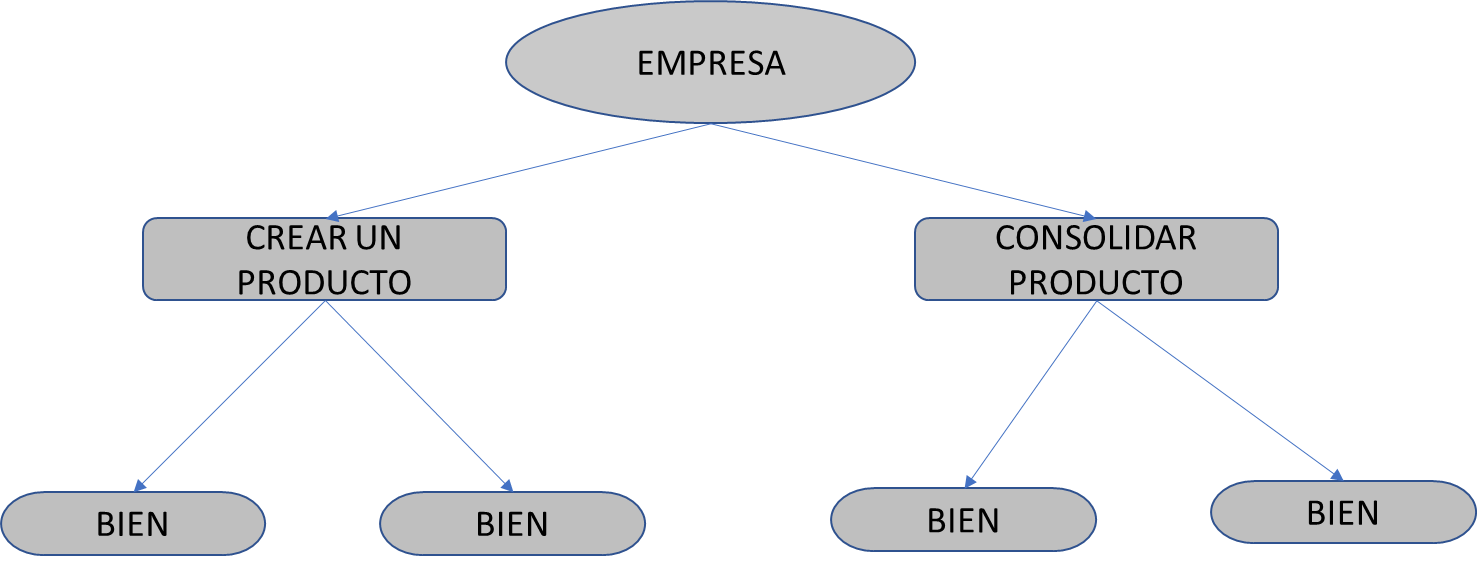
Dentro de un árbol de decisión distinguimos diferentes tipos de nodos:

* Primer nodo o nodo raíz: en él se produce la primera división en función de la variable más importante.
* Nodos internos o intermedios: tras la primera división encontramos estos nodos, que vuelven a dividir el conjunto de datos en función de las variables.
* Nodos terminales u hojas: se ubican en la parte inferior del esquema y su función es indicar la clasificación definitiva.
* ***Aplica entropía y ganancia de información.***

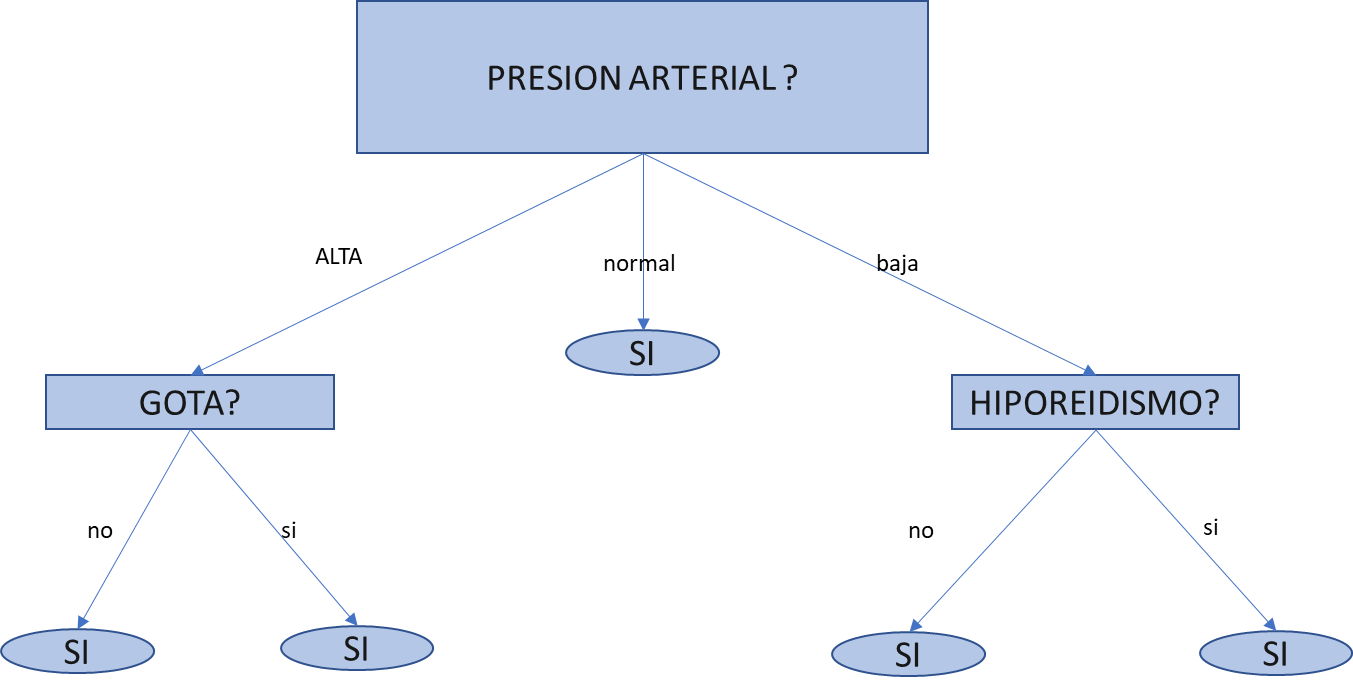
La entropía es una medida de la impureza de un conjunto de datos, mientras que la ganancia de información se utiliza para medir la reducción de la entropía al dividir los datos en subconjuntos. Estos conceptos son utilizados por los algoritmos de árboles de decisión para determinar la mejor forma de dividir los datos y construir el árbol.

* ***esuelve un problema con el algoritmo ID3.***
* El algoritmo ID3 es un algoritmo de árbol de decisión que se utiliza para resolver problemas de clasificación. Utiliza la entropía y la ganancia de información para construir un árbol de decisión que pueda clasificar nuevos ejemplos en función de sus características.

EJEMPLO 1



EJEMPLO 2



EJEMPLO 3

