

# Árboles de decisión

- 1844452 - Castillo Rodríguez, Leonardo Daniel
- 1851642 - Cedillo Charles, Hector
- 1877422 - Martinez Ortuño, Gustavo Juan
- 1869864 - Morales Salazar, César Alejandro
- 1851701 - Navarrete Rosales, José Eduardo



# ¿Qué son los árboles de decisión?

○ Un árbol de decisión es un diagrama en forma de árbol que muestra la probabilidad estadística o determina un curso de acción. Muestra a los tomadores de decisiones, qué pasos deben o pueden tomar y cómo las diferentes elecciones podrían afectar todo el proceso..



# Tipos de nodos

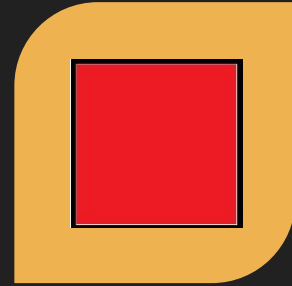
Un nodo es un punto de intersección, conexión o unión de varios elementos que confluyen en el mismo lugar. En los diagramas de árbol tenemos 3 tipos de nodos:



NODOS  
TERMINALES



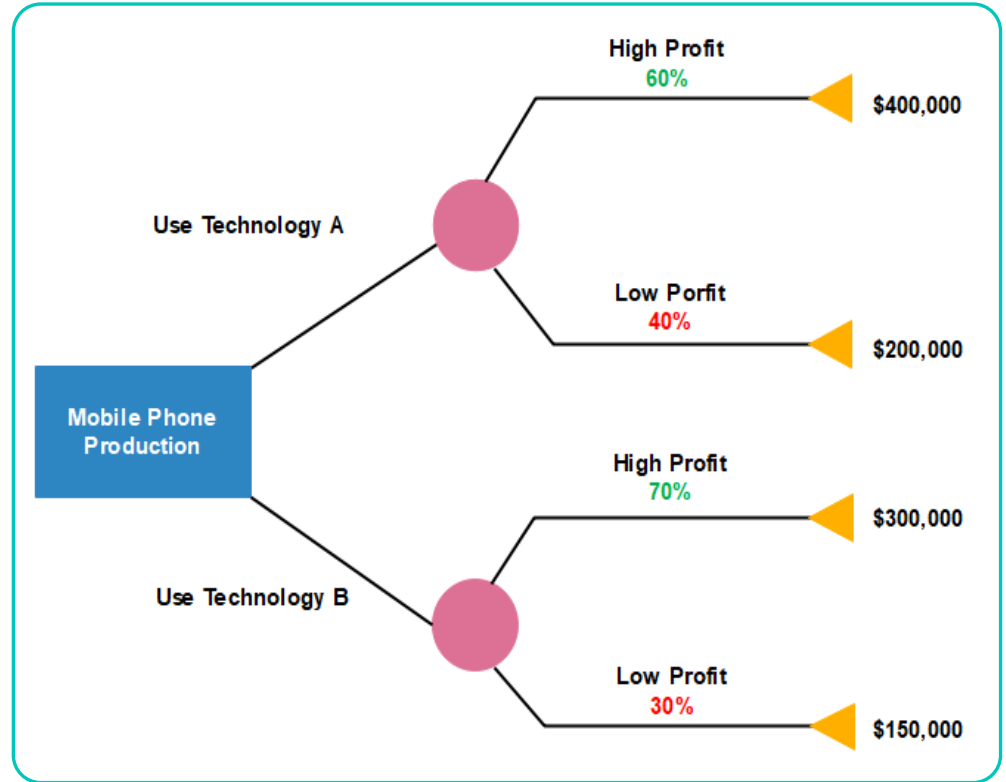
NODOS DE  
PROBABILIDAD

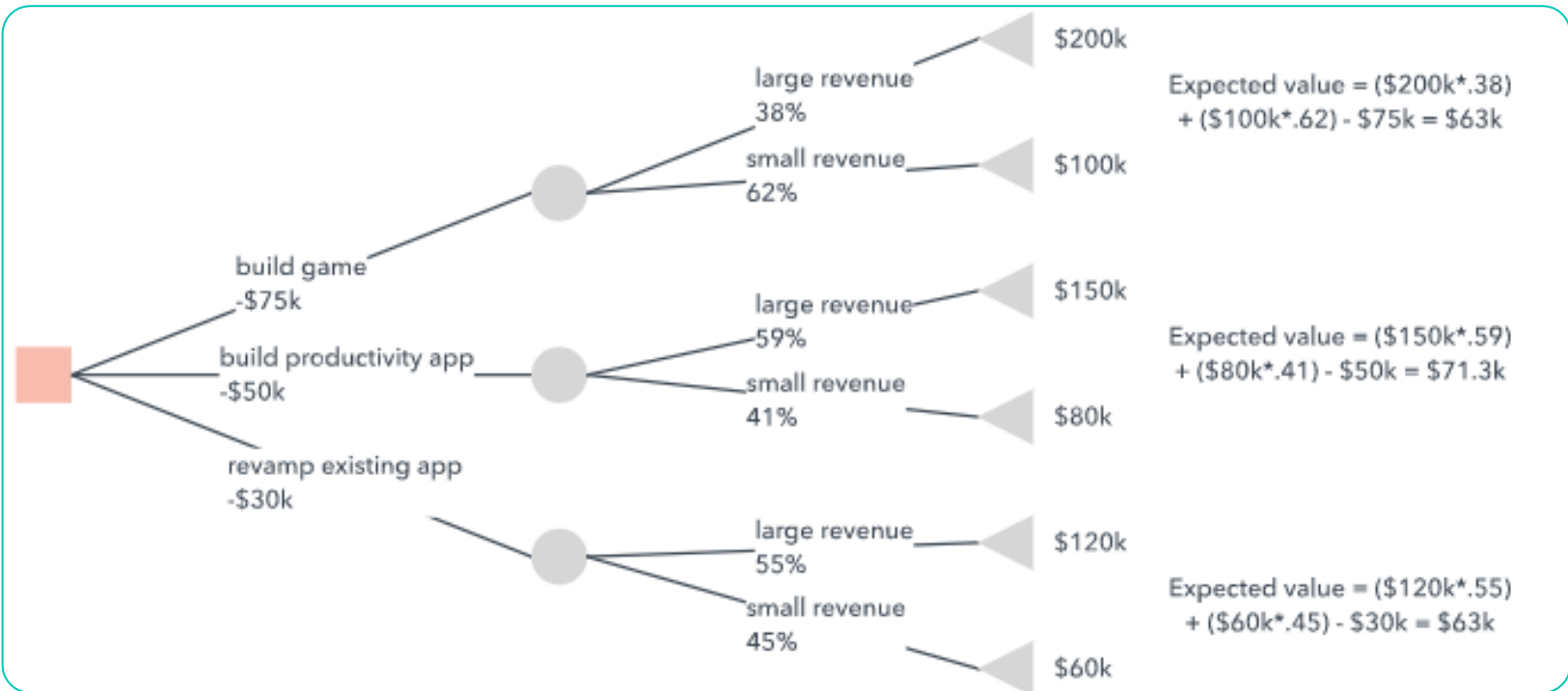


NODOS DE  
DECISIÓN

# Conceptos Generales

- Raíz
- Padre
- Hijo
- Hermanos
- Brazo
- Rama





## Interpretación de un árbol de decisión

# Como hacer un arbol de decision



Empieza por la decisión



Enumera tus opciones



Somételo a pruebas



Enumera las conclusiones

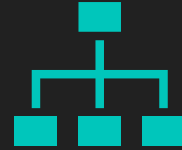
# Casos de uso/aplicaciones habituales de los árboles de decisiones



**Planificación  
de productos**



**Aprobación  
de préstamos**

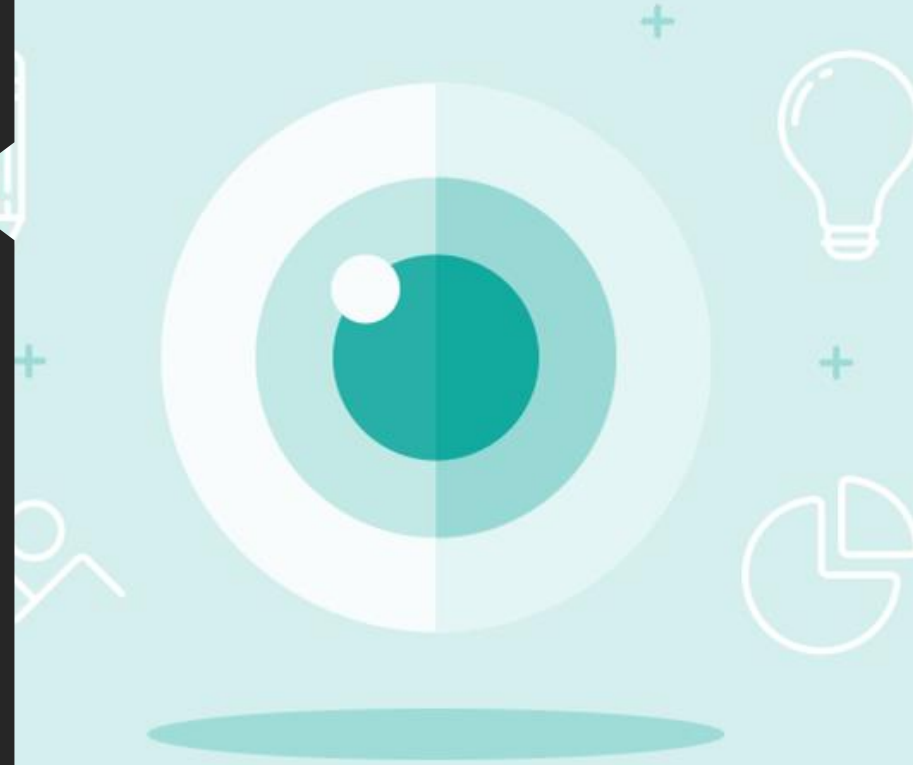


**Decisiones  
personales**

# Ventajas

Las ventajas del árbol de decisiones se reducen a los beneficios de la toma de decisiones basada en datos. Estas son algunas de las principales ventajas:

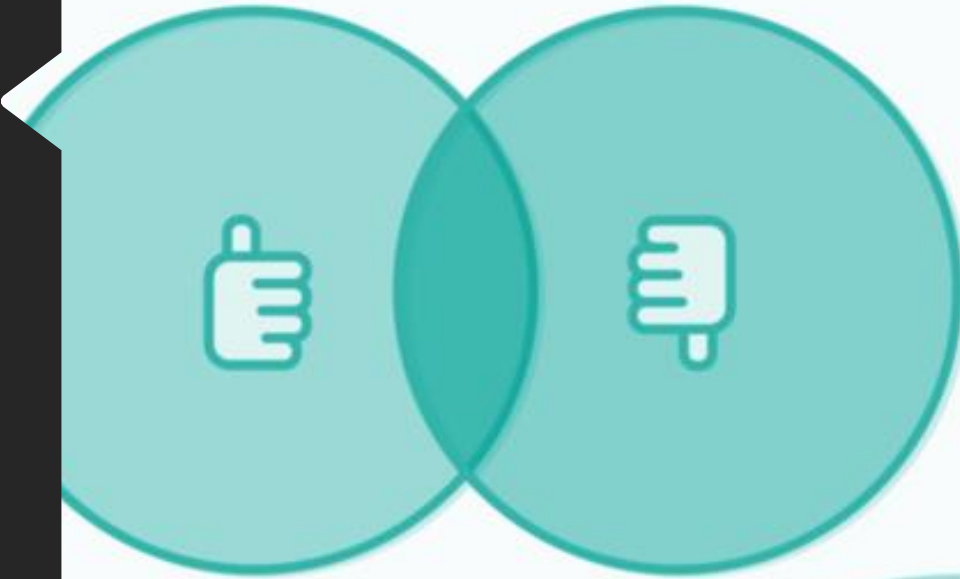
- Integral
- Gráfico
- Adaptable
- Reducen el sesgo
- Simple
- Bajo costo
- Ahorra tiempo
- Funciona para los datos numéricos o categóricos
- Puede modelar problemas con múltiples resultados





# Desventajas

- Son inestables
- No se puede garantizar que el árbol generado sea el óptimo
- Los cálculos pueden volverse complejos al lidiar con la falta de certezas y numerosos resultados relacionados
- Los principiantes crean árboles sesgados, sobre todo si una de las personas que interviene es dominante respecto al resto





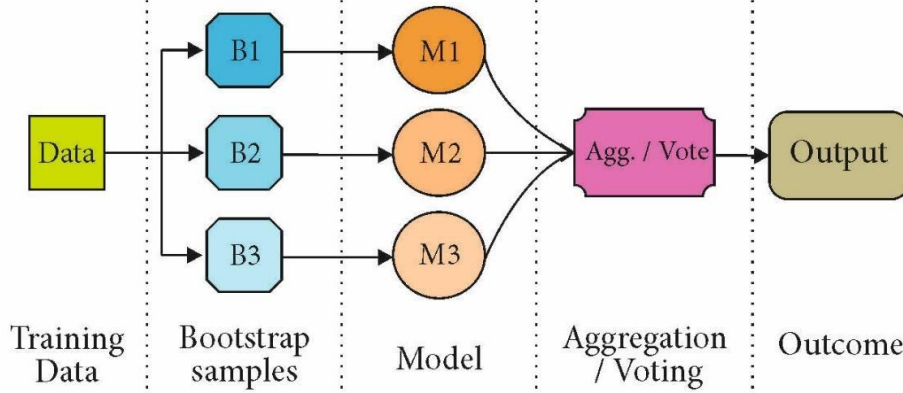
# DATA MINING

## Minería de datos

- Un árbol de decisión también puede usarse para ayudar a crear modelos predictivos automatizados, que puedan emplearse en el aprendizaje automático, la minería de datos y las estadísticas. Conocido como "el aprendizaje basado en árboles de decisión", este método toma en consideración las observaciones sobre un elemento para predecir su valor.
- Cada valor de datos o decisión forma una cláusula, de tal manera que, por ejemplo, "si las condiciones 1, 2 y 3 se cumplen, entonces el resultado X será el resultado definitivo con certeza Y".

# BAGGING Algorithm

*Bootstrap Aggrigating*



## Técnicas de construcción

- Baggin
- Árboles impulsados
- Rotation Forest
- Random Forest

# Aprendizaje basado en árboles de decisión

The background features a central brain shape composed of glowing blue circuitry. To the left, there is a circular gauge showing '100%' and a small globe icon. To the right, another circular gauge shows '78%' and a soccer ball icon. At the bottom, there are concentric circles and some faint, illegible text.

Es la construcción de un árbol de decisión a partir de tuplas de entrenamiento, cada una etiquetada con su correspondiente clase. Un árbol de decisión es similar a una estructura de diagrama de flujo, donde cada nodo interno (no hoja) denota una prueba en un atributo, cada rama representa el resultado de una prueba, y cada hoja (o terminal) nodo tiene una etiqueta de clase. El nodo superior en un árbol es el nodo raíz.

# Reglas para saber cuando detenerse



Pureza de nodo



Cota de  
profundidad



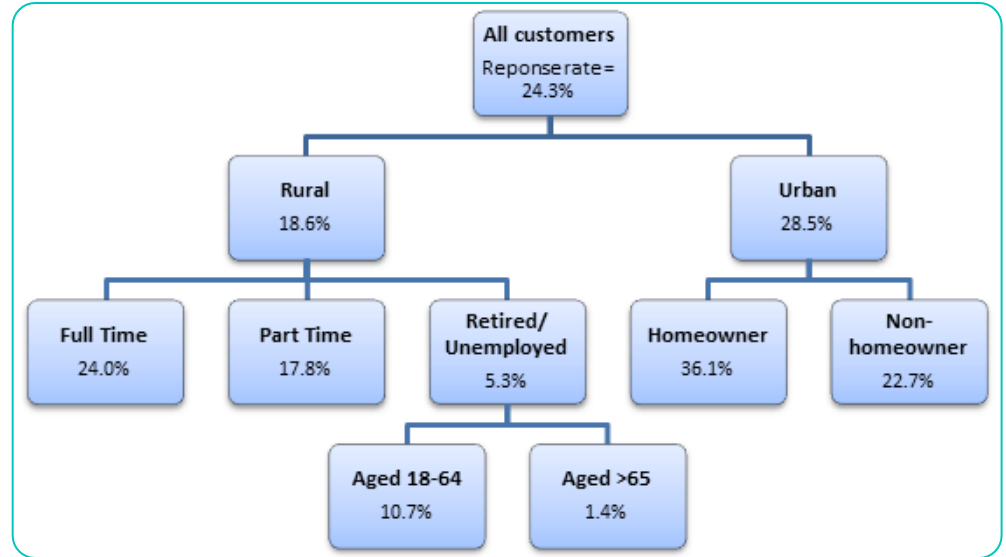
Umbral de soporte

# Propiedades en los algoritmos de arboles de decisiones

- El criterio para determinar la partición de cada nodo
- La regla que declara un nodo terminal
- La asignación de una clase a cada nodo terminal, lo que determina la regla de clasificación
- Fusión
- Partición
- Poda
- La evaluación de la bondad del clasificador obtenido

# ARBOL CHAID

- En el algoritmo AID las variables explicativas han de estar medidas en escalas nominales u ordinales y la variable a explicar, variable criterio o dependiente, puede medirse en una escala métrica.
- El análisis AID constituye un Análisis de la Varianza.
- La prueba estadística F la que se utiliza para seleccionar las mayores diferencias posibles.

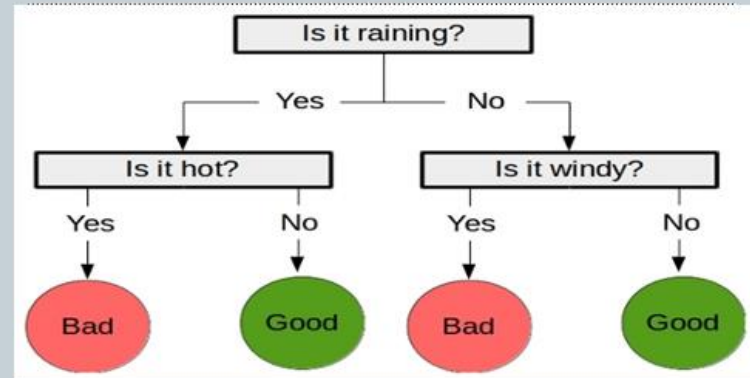




# Árbol CART

- Con este algoritmo, se generan árboles de decisión binario.
- Este modelo admite variables de entrada y de salida nominales, ordinales y continuas.
- El algoritmo utiliza el índice de Gini para calcular la medida de impureza.

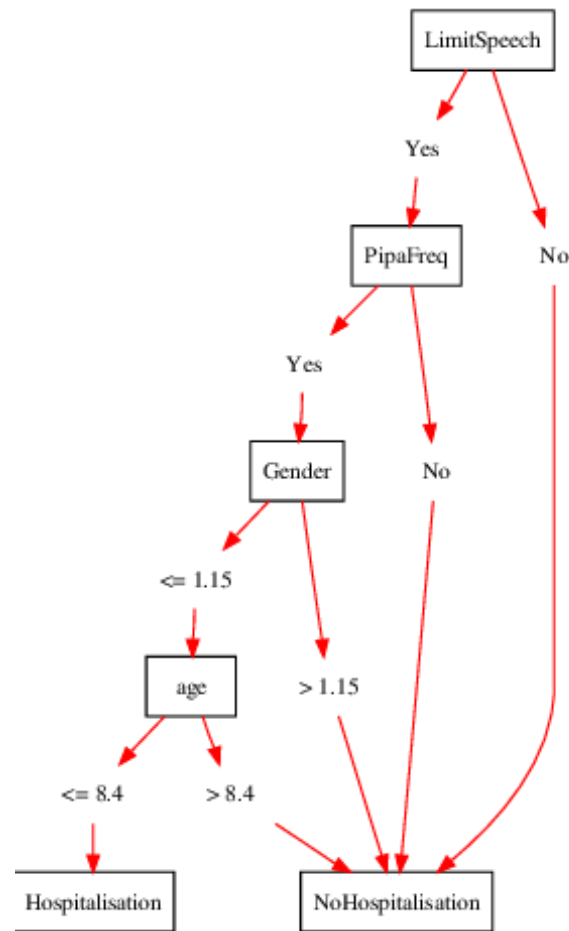
## Chapter 4 : Decision Tree(CART)





# ÁRBOL QUEST

- Este algoritmo trata de corregir y de restringir la exhaustiva búsqueda de particiones significativas que se generan tanto en los algoritmos AID y CHAID como en el CART.
- Sintetizando el procedimiento, primero se elige la mejor variable predictora cuyo objetivo es que el número de categorías que poseen las variables no afecte a la elección de la mejor variable, para realizar después la mejor segmentación de la variable que ha seleccionado.
- Este método QUEST sólo puede ser utilizado si la variable de salida es categórica nominal.



# Ejemplo

A continuación, se muestra un ejemplo que nos ayudará a comprender mejor cómo funcionan los árboles de decisión. Además, el ejemplo está hecho con el lenguaje de programación Python3.

El código es por medio de funciones y librerías propias de este lenguaje, lo cual hace que el código sea inteligente. De esta forma, el algoritmo aprende de la última vez que se corrió el programa para desarrollarlo más en la siguiente ocasión.

<https://colab.research.google.com/drive/1s4WT6ZiNQV2r51J7VluajH3Lg0jfHXz0?usp=sharing>



# Bibliografía

- <https://www.youtube.com/watch?v=269QJ5joMCc>
- <https://artyco.com/que-es-un-arbol-de-decision-y-su-importancia-en-el-data-driven/>
- <https://lucidspark.com/es/blog/como-hacer-arboles-de-decisiones>
- [https://www.lucidchart.com/pages/es/que-es-un-diagrama-de-arbol-de-decision/#section\\_5](https://www.lucidchart.com/pages/es/que-es-un-diagrama-de-arbol-de-decision/#section_5)
- [https://bookdown.org/content/2274/metodos-de-clasificacion.html#introduccion.](https://bookdown.org/content/2274/metodos-de-clasificacion.html#introduccion)