

DECISION TREE



DECISION TREES

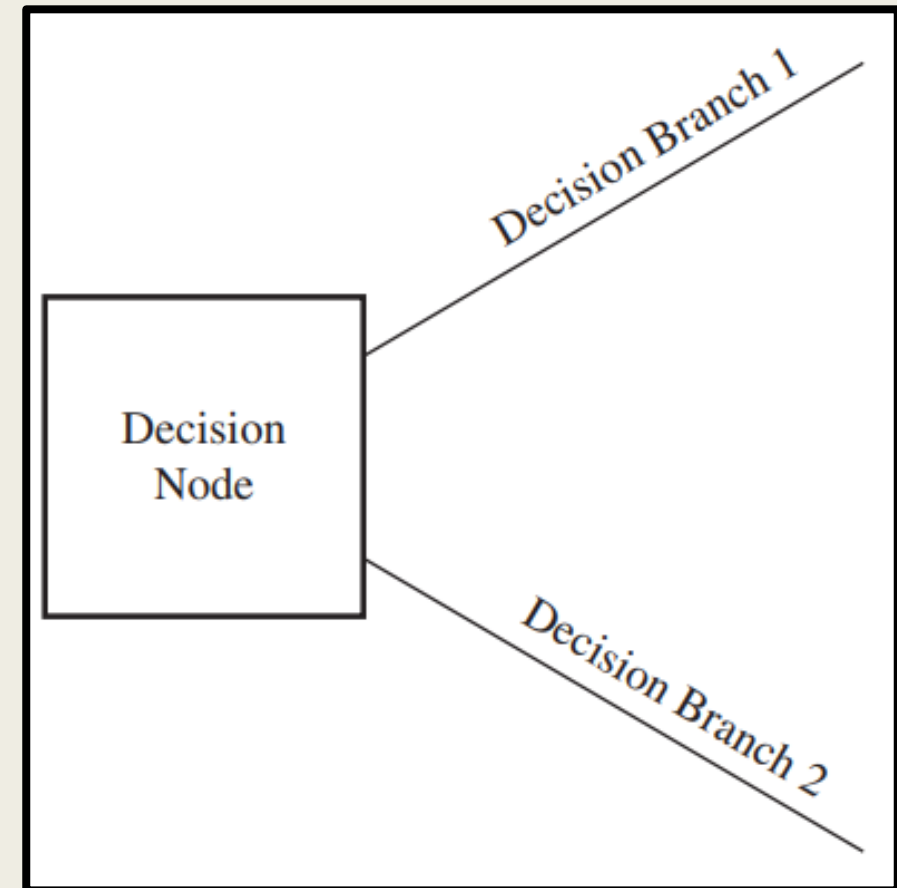
INTEGRANTES:

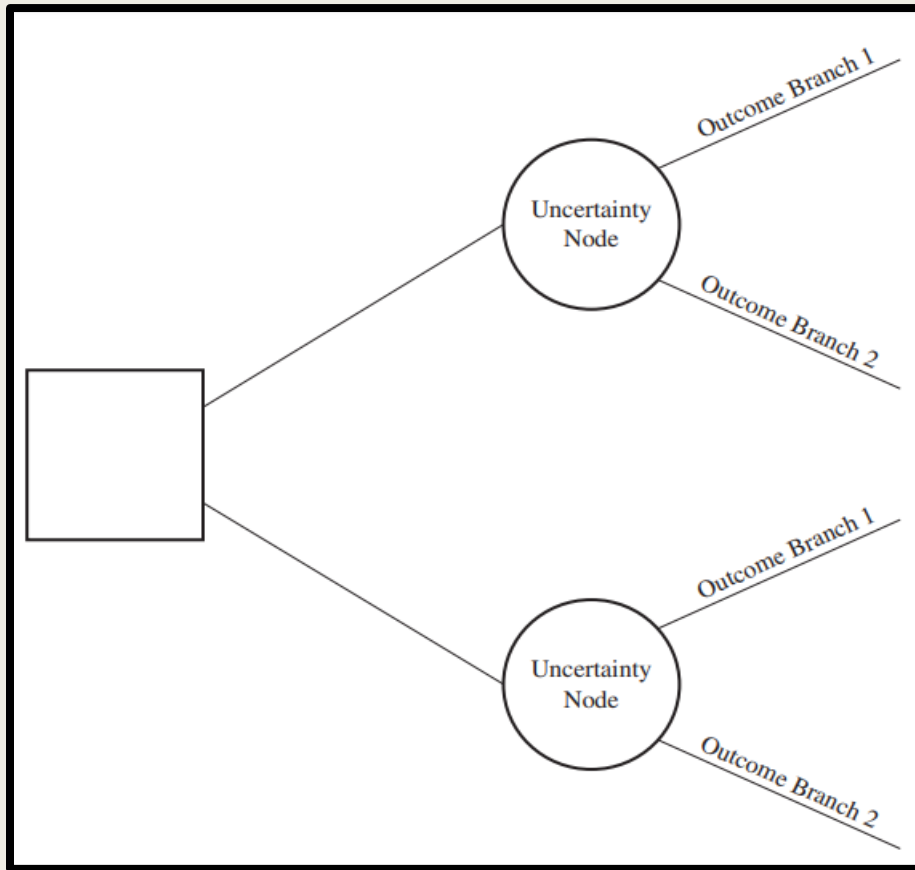
Alan René Cerda Serrato	1864878
Jorge Antonio Cruz Lerma	1867341
Jesús Eduardo Valencia González	1630606
Jesús Alfonso Galindo Vázquez	1941475

¿Qué es?

Un árbol de decisiones se utiliza a menudo para mostrar y analizar las opciones disponibles para la toma de una decisión. Los árboles de decisiones son especialmente informativos. Se compone por nodos en los que se muestran las diferentes consecuencias, alternativas y decisiones.

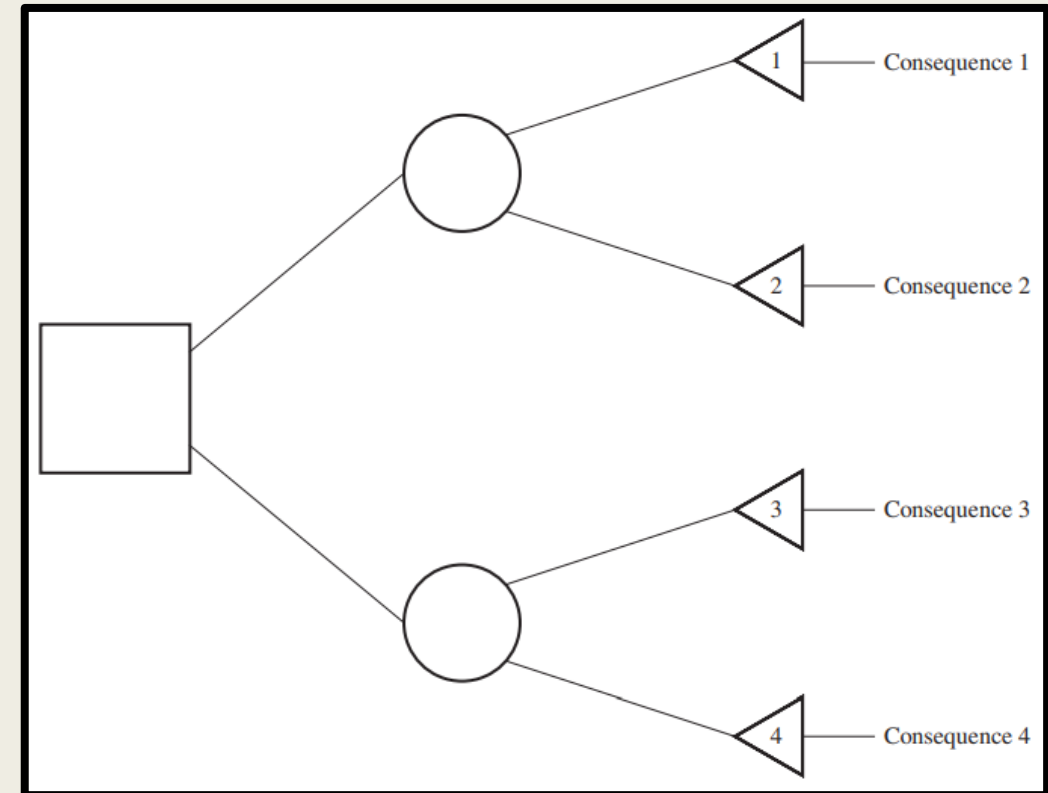
- *Un nodo de decisión tiene ramas de decisión que muestran cada curso de acción alternativo.*





- *Un nodo de Incertidumbre que refleja un resultado al azar con Ramas de Resultados que revelan los posibles resultados de esto.*

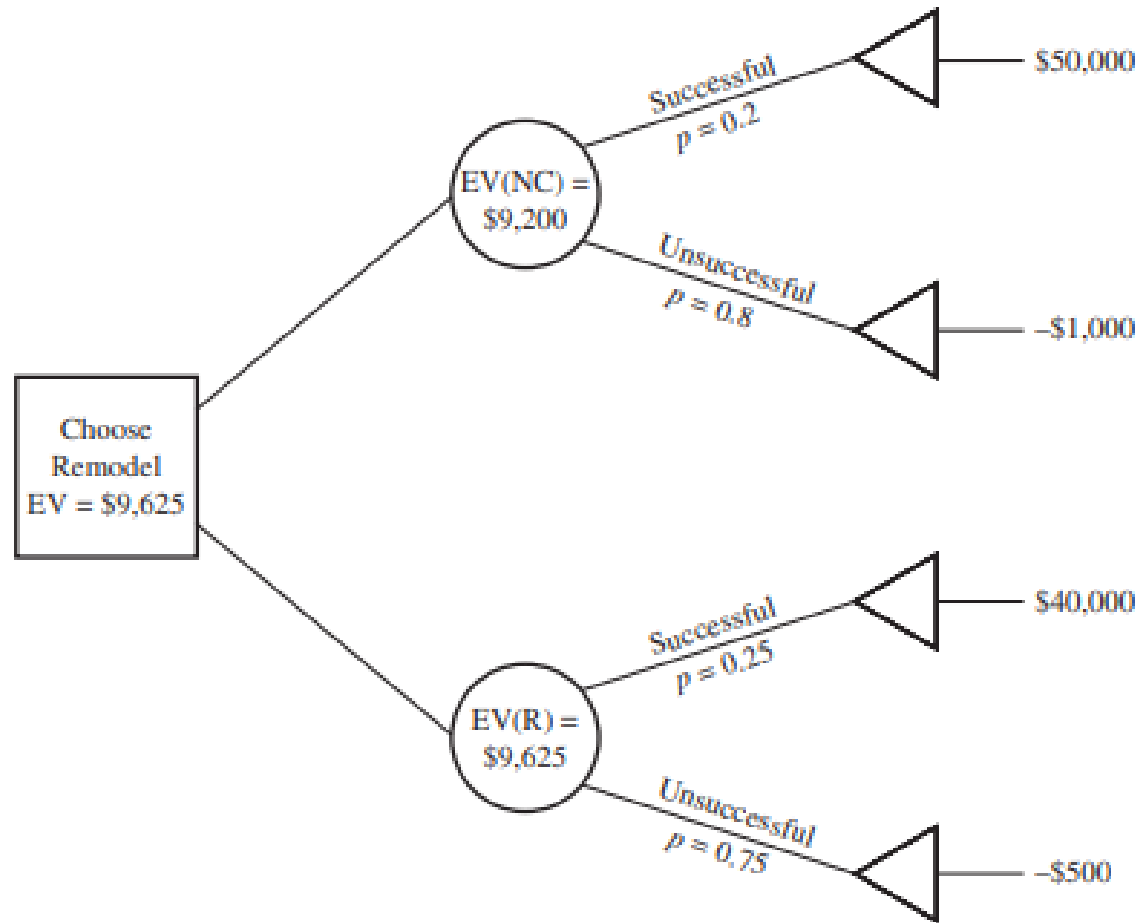
- *Un Nodo Terminal con una rama de consecuencia que muestra lo obtenido por ese resultado.*



	Construcción nueva (CN)	Remodelación (R)
Ganacias (neta)	\$50,000	\$40,000
Pérdidas	-\$1,000	-\$500
Probabilidad del contrato	20%	25%

Ejemplo

La decisión por hacer es si construir nuevos campos de golf o remodelar los ya existentes. El criterio de decisión es maximizar el beneficio de todo el recorrido maximizando el valor esperado.



© Cengage Learning 2013

1. Empezamos con un nodo de decisión con 2 ramas, una para cada acción posible.
2. Ahora añadimos los nodos de incertidumbre cada uno con las posibilidades de cada acción principal, con su probabilidad asociada.
3. Por último, se agregan los nodos terminales con la consecuencia de cada rama y su final.

Aplicaciones

- Para teoría de juegos: Se utilizan modelos para estudiar interacciones de estructuras formalizadas de incentivos.
- También se utiliza en las ramas de: Economía, política, biología, sociología, psicología, filosofía y ciencias de la computación.
- Un ejemplo en ciencias de la computación es en el diseño de aplicaciones informáticas, los arboles indican las acciones a realizar en función del valor de una o varias variables
- En economía los árboles de decisión se utilizan mayoritariamente para predecir las probabilidades de alcanzar un resultado en función de unas variables de entrada tales como edad, sexo, demografía o ingresos que indicarán por ejemplo, si el cliente es apto o no para recibir un préstamo.

Árbol de decisiones en la Minería de Datos

En minería de datos, un árbol de decisión sirve para abordar problemas tales como la clasificación, la predicción y la segmentación de datos con la finalidad de obtener información que pueda ser analizada para tomar decisiones futuras.

También se utilizan para evaluar las distintas alternativas o datos encontrados en nuestra base de información.



Ejemplo aplicado

Se presenta la siguiente problemática

¿Debemos ofrecer a un cliente determinado un producto concreto?

Primero, se debe definir el tipo de cliente, en este caso, si el cliente es VIP o no y, en función de la respuesta, habrá una serie de posibilidades.

Caso 1

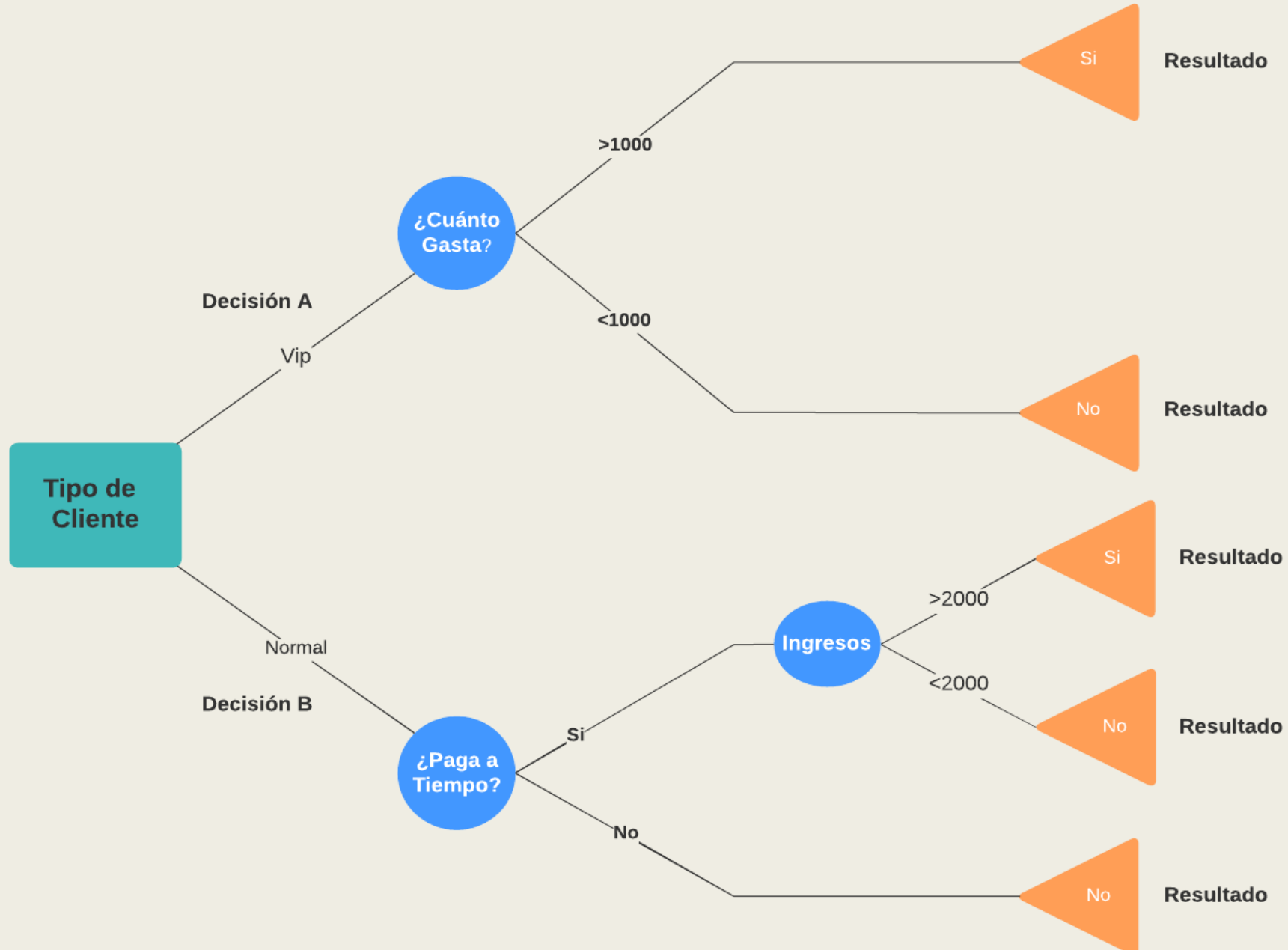
En el caso de que el cliente sea VIP, se debe tener en cuenta de cuánto gasta dicho cliente. En el caso de que el gasto sea mayor de 1000 sí se debe ofertar el producto, en caso contrario no.

Caso 2

En el caso de que el cliente no sea VIP, para poder ofrecer el producto habrá que considerar más posibilidades como si paga o no a tiempo o si sus ingresos son superiores a una cantidad determinada.

Árbol de Decisiones

September 8, 2021



Ejemplo con código

Este ejemplo se trata de clasificar los vinos con un árbol de decisión de clasificación.

1. Importo las librerías y datos que necesito.

```
from sklearn.tree import DecisionTreeClassifier #Para poder hacer un árbol de clasificación
from sklearn.model_selection import train_test_split #Para "entrenar" el árbol
from sklearn import tree #Para visualizar el árbol
from sklearn.datasets import load_wine #Para importar una base de datos
import matplotlib.pyplot as plt #Para poder acomodar la visualización
import pandas as pd #Para modificar la base de datos
```

2. Doy formato a los datos.

```
base_datos = load_wine() #Asigno la base de datos a una variable
vinos = pd.DataFrame(base_datos.data, columns = base_datos.feature_names) #Formato
vinos['Clasificación'] = base_datos.target #Clasifico
vinos.head(3) #Visualización de las primeras 3 filas de datos
```

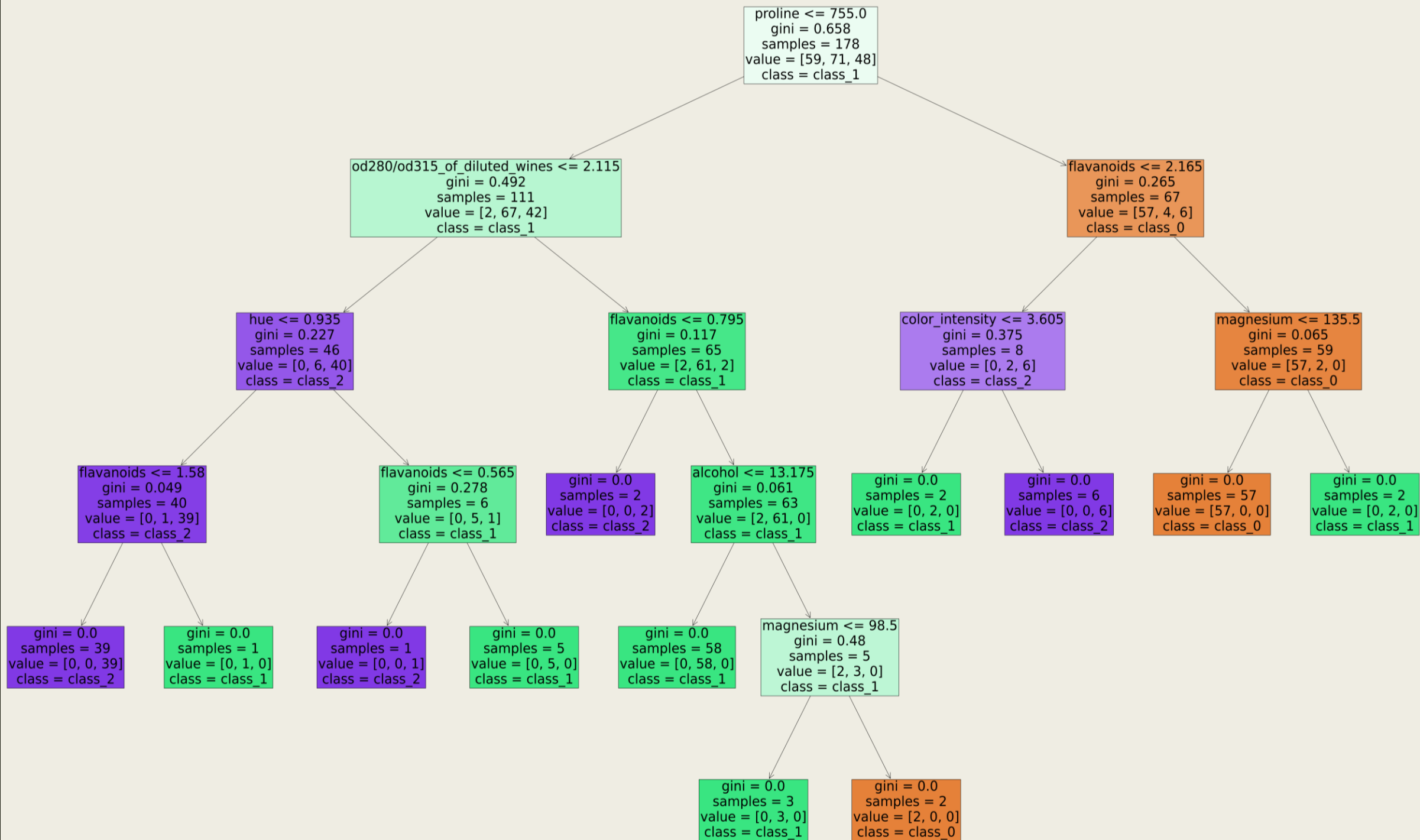
	alcohol	malic_acid	ash	alcalinity_of_ash	magnesium	total_phenols	flavanoids	nonflavanoid_phenols	proanthocyanins	color_intensity	hue	od280/od315_of_diluted_wines	proline	Clasificación
0	14.23	1.71	2.43	15.6	127.0	2.80	3.06	0.28	2.29	5.64	1.04	3.92	1065.0	0
1	13.20	1.78	2.14	11.2	100.0	2.65	2.76	0.26	1.28	4.38	1.05	3.40	1050.0	0
2	13.16	2.36	2.67	18.6	101.0	2.80	3.24	0.30	2.81	5.68	1.03	3.17	1185.0	0

3. Creo y lleno el árbol.

```
X = vinos[base_datos.feature_names] #Asigno criterios
Y = vinos['Clasificación'] #Asigno clasificación
arbol = DecisionTreeClassifier() #Creación el árbol
arbol.fit(X, Y) #Lleno el árbol
```

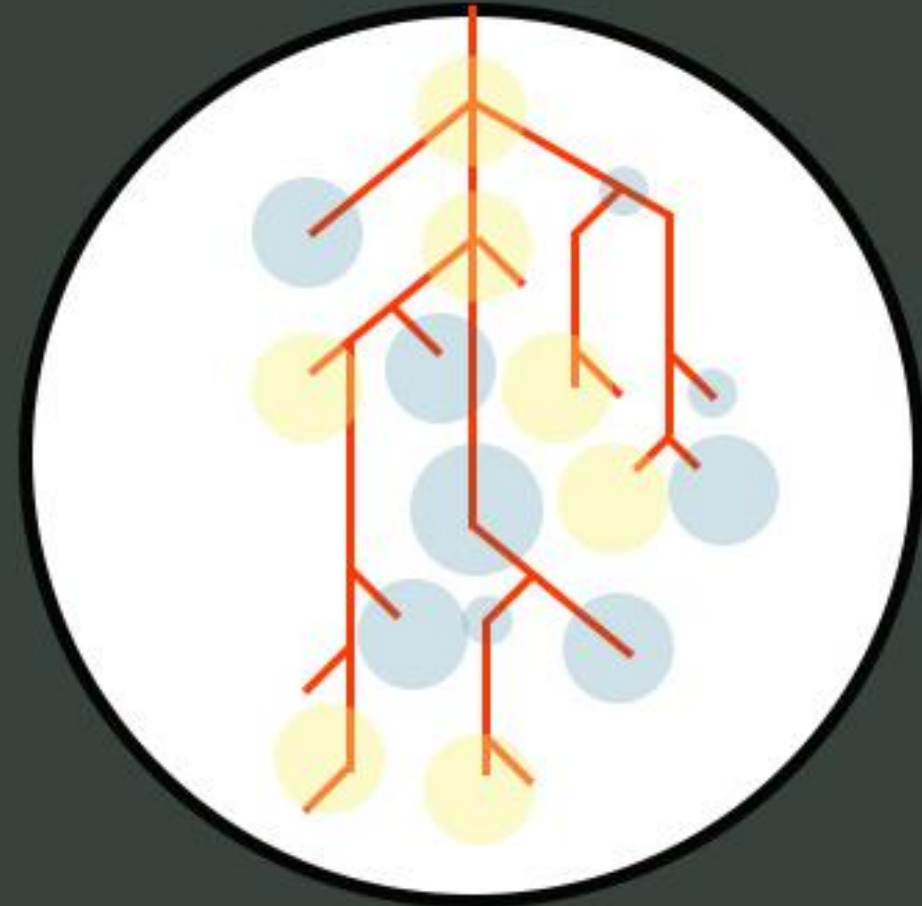
4. Visualizo el árbol.

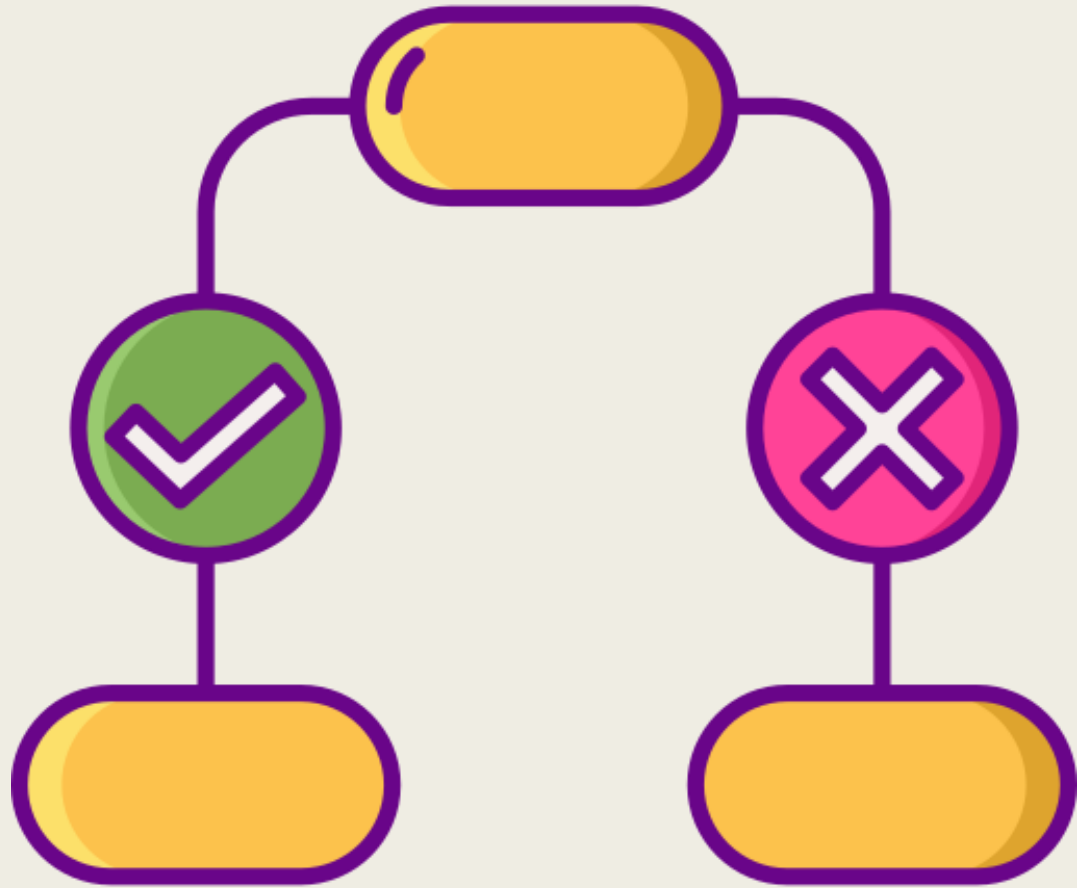
```
nombres = base_datos.feature_names #Criterios
clasif = base_datos.target_names #Clasificación
fig, axes = plt.subplots(nrows = 1,ncols = 1,figsize = (80,40), dpi=300) #Mejorar la visualización
tree.plot_tree(arbol, feature_names = nombres, class_names = clasif, filled = True, fontsize = 40) #Ver el árbol
```



Preguntas de interés

- ¿Para qué nos sirven los arboles de decisión?
- ¿Cuándo se utilizan arboles de decisión en la vida cotidiana?
- ¿Cuándo pueden ser utilizado en el campo de la actuaría o finanzas?
- ¿Cuáles son los componentes de un árbol de decisión?
- ¿En minería de datos, cuáles son las 3 principales funciones en los arboles de decisión?



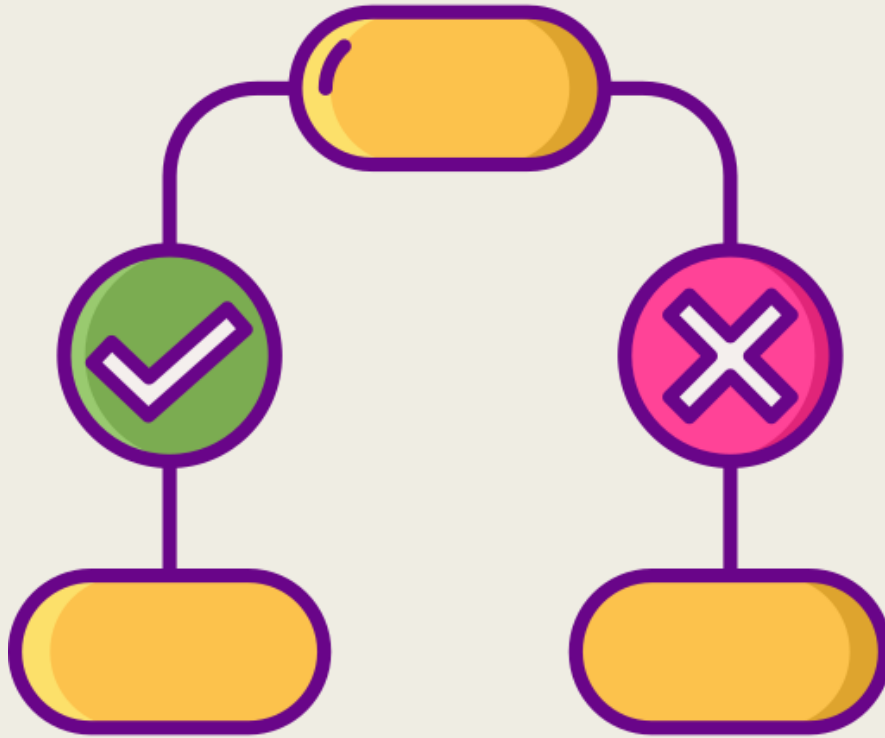


Bibliografía

- *A First Course in Mathematical Modeling, Fifth Edition.*
- Lior Rokach and Oded Maimon (2008). *Data mining with decision trees: theory and applications.* World Scientific.
- <https://conectasoftware.com/analytics/arboles-de-decisiones-en-la-mineria-de-datos/>
- <https://www.crehana.com/mx/blog/desarrollo-web/arboles-decision-mineria-datos/>
- <https://www.crehana.com/mx/blog/desarrollo-web/arboles-decision-mineria-datos/>

Bibliografía

(Para el código)



- <https://aprendeia.com/arboles-de-decision-clasificacion-scikit-learn-machine-learning/>
- <https://aprendeia.com/arboles-de-decision-clasificacion-practica-con-python-machine-learning/>
- <https://ichi.pro/es/visualizacion-de-arboles-de-decision-con-python-scikit-learn-graphviz-matplotlib-31132954028252>
- https://scikit-learn.org.translate.goog/stable/modules/generated/sklearn.tree.plot_tree.html?_x_tr_sl=auto&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=nv,elem#sklearn.tree.plot_tree
- https://scikit-learn.org/stable/datasets/toy_dataset.html
- <https://www.iartificial.net/arboles-de-decision-con-ejemplos-en-python/>