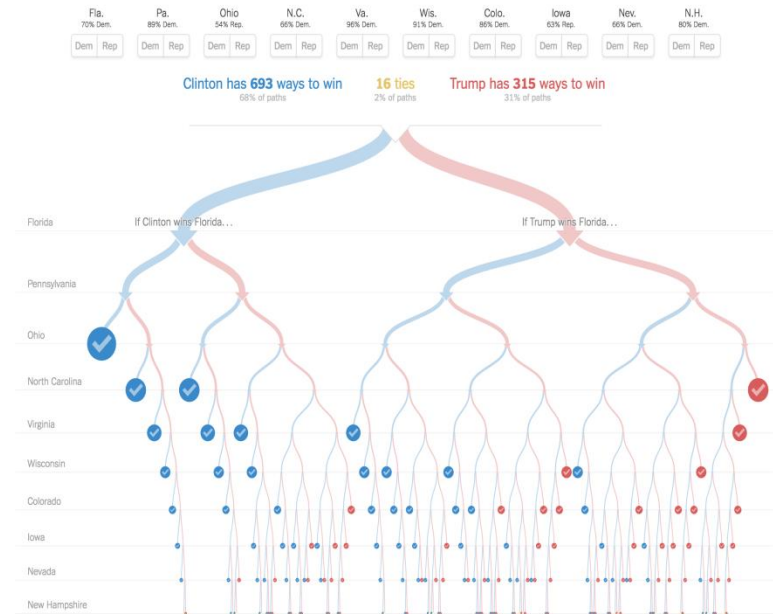
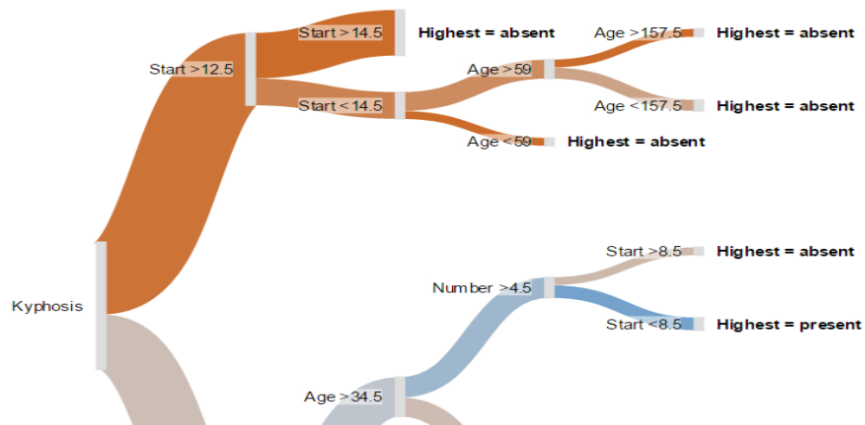
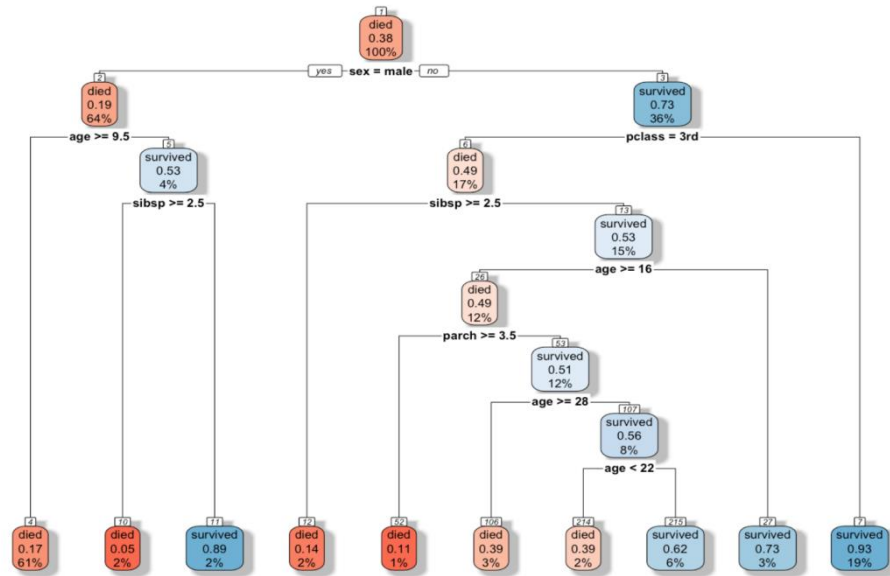


Árboles de decisión



Oscar Centeno Mora

Índice

1

Introducción

4

AC vs AR
ventajas y
desventajas

2

Principios del árbol
de decisión

5

Partición /
ramificación

3

Terminología

6

Estimación: el sobre
ajuste y la poda

Índice

7

Etapas de análisis

8

Árboles de decisión

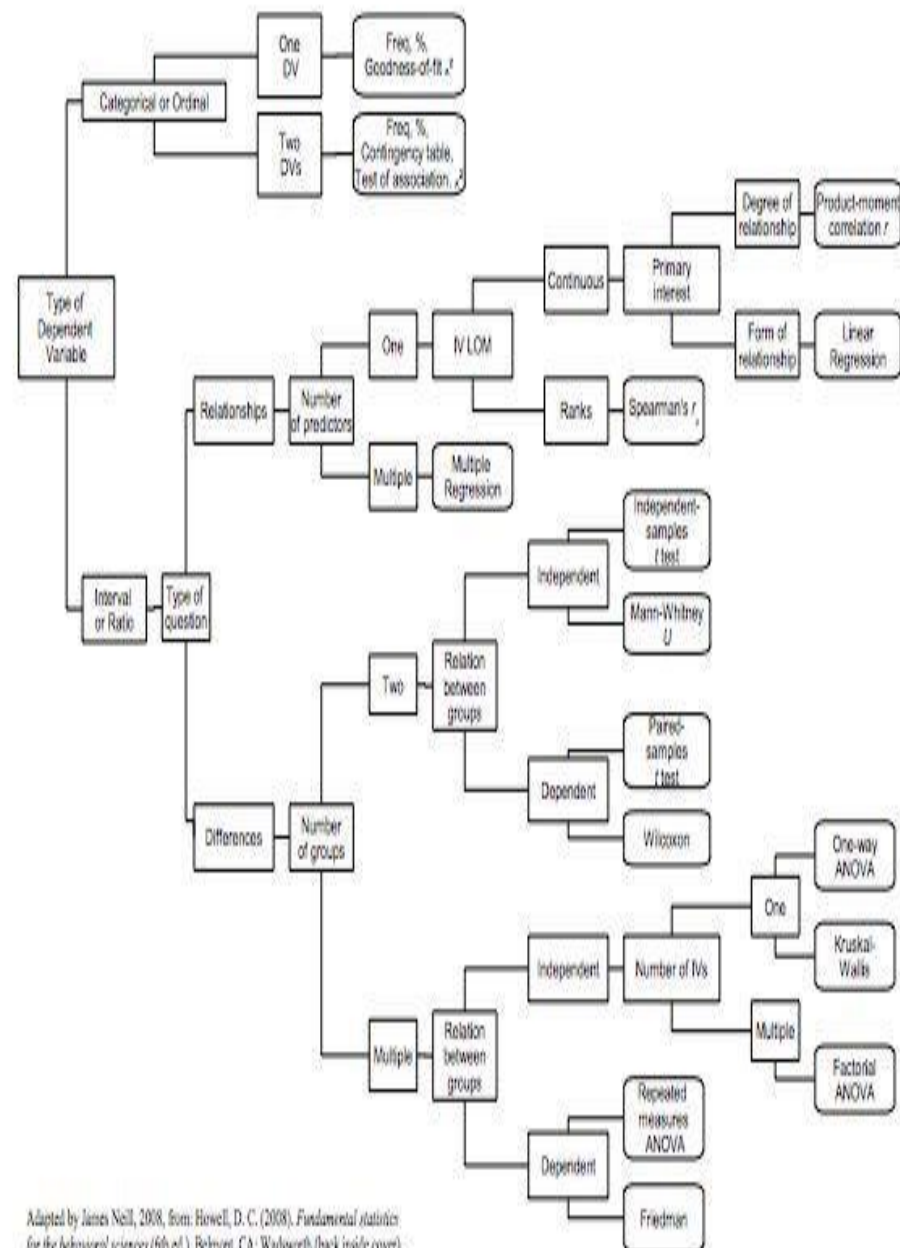
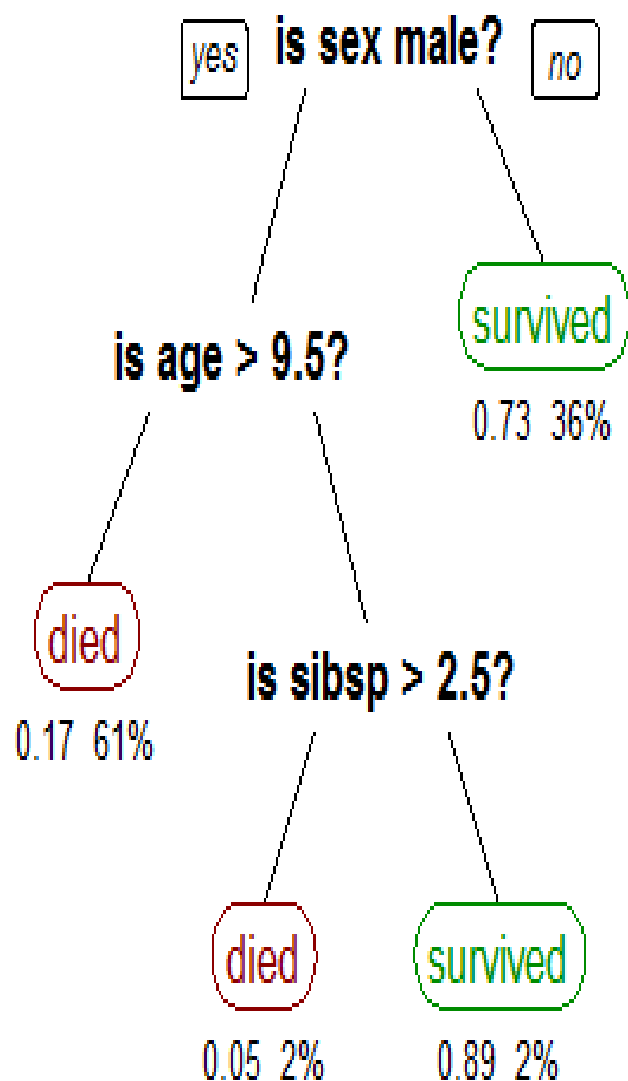
9

Árboles de regresión

Índice

1

Introducción



Adapted by James Neill, 2008, from: Howell, D. C. (2008). *Fundamental statistics for the behavioral sciences* (6th ed.). Belmont, CA: Wadsworth (back inside cover).

Fla. 70% Dem.	Pa. 89% Dem.	Ohio 54% Rep.	N.C. 66% Dem.	Va. 96% Dem.	Wis. 91% Dem.	Colo. 86% Dem.	Iowa 63% Rep.	Nev. 66% Dem.	N.H. 80% Dem.
Dem Rep	Dem Rep	Dem Rep	Dem Rep	Dem Rep	Dem Rep	Dem Rep	Dem Rep	Dem Rep	Dem Rep

Clinton has 693 ways to win

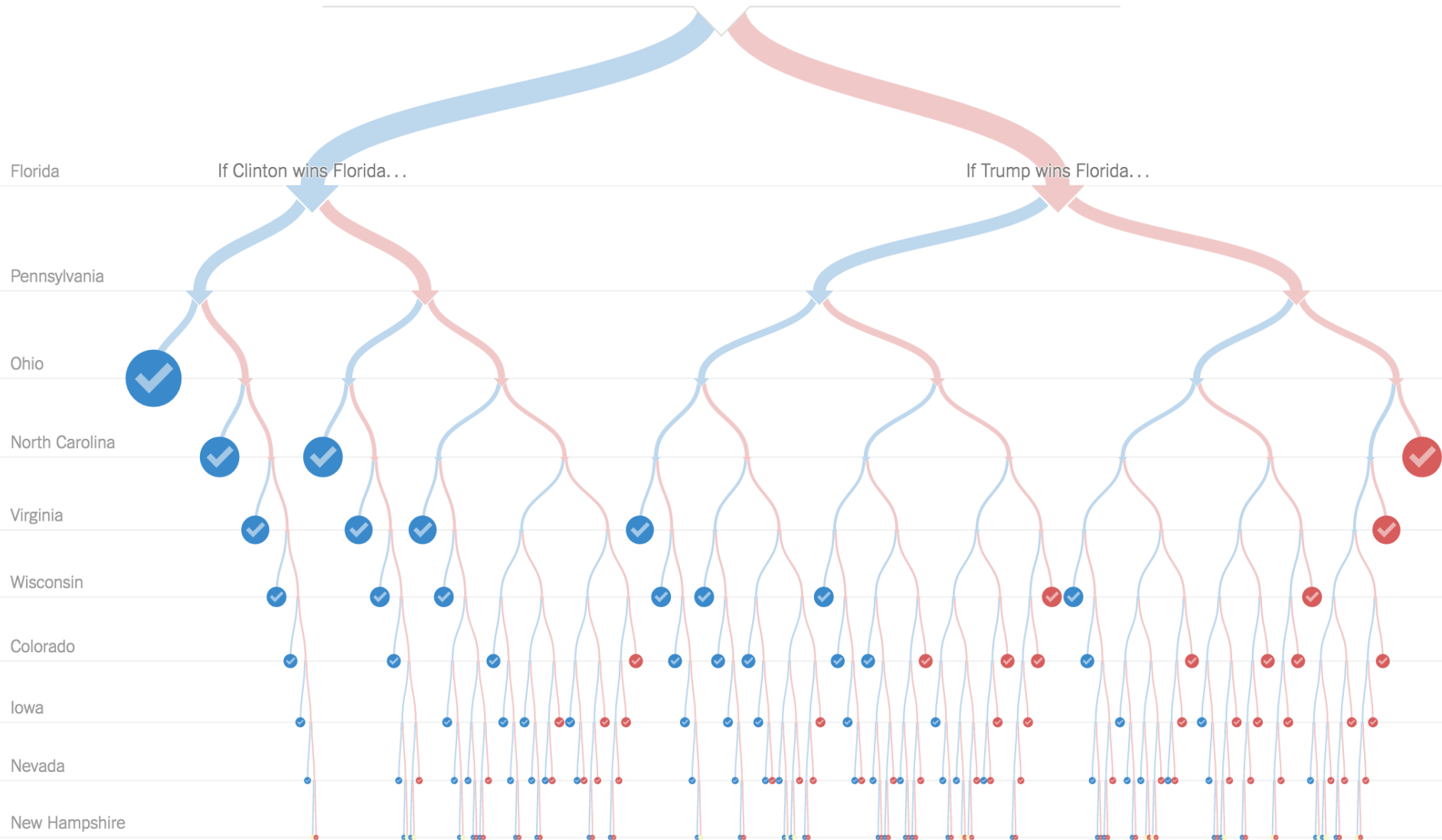
16 ties

Trump has 315 ways to win

68% of paths

2% of paths

31% of paths



Introducción

Los árboles de clasificación y regresión (CART= Classification and Regression Trees) son una alternativa al análisis tradicional de clasificación / discriminación o a la predicción tradicional (regresión).



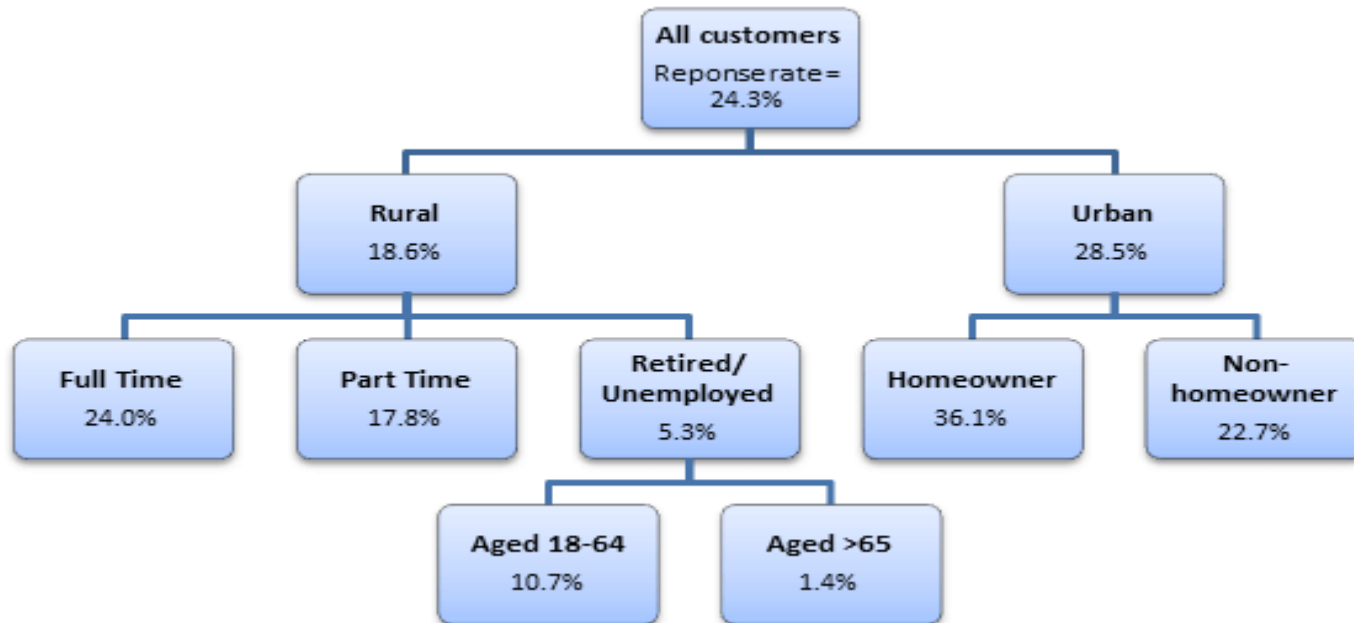
La ventaja de los árboles CART residen en su robustez a los outliers, la invariancia en la estructura de sus árboles de clasificación o de regresión a transformaciones monótonas de las variables independientes, y sobre todo, su interpretabilidad.

En esencia, se trata de dar con un esquema de múltiples dicotomías o bifurcaciones, anidadas en forma de árbol, de manera que siguiendo cada una de las ramas del árbol obtengamos, al final, una predicción para la clase de pertenencia (clasificación) o para el valor que toma (regresión) los individuos que cumplen con las propiedades que se han ido exigiendo en las distintas bifurcaciones.



Introducción

Los árboles de decisión se construyen mediante un algoritmo conocido como segmentación recursiva, que es el proceso paso a paso para dicha construcción. Existen principalmente tres procedimientos: CHAID (Chi-Squared Automatic Interaction Detector), QUEST (Quick Unbiased Efficient Statistical Tree) y el CART, el cual será utilizado en el presente curso.



Introducción

Historia de los principales algoritmos de árboles

C4.5. Introducido por Quinlan (1993) dentro de la comunidad de “Machine Learning”. Es descendiente del ID3 (Quinlan, 1986).

CHAID. Significa “Chi-square automatic interaction detection”, fue introducido por Kass (1980) y es un derivado del THAID: “A sequential search program for the analysis of nominal scale dependent variables” (Morgan and Messenger, 1973). El criterio para particionar está basado en χ^2 .

NewId. (Boswell, 1990). Es descendiente del ID3 (Quinlan, 1986)

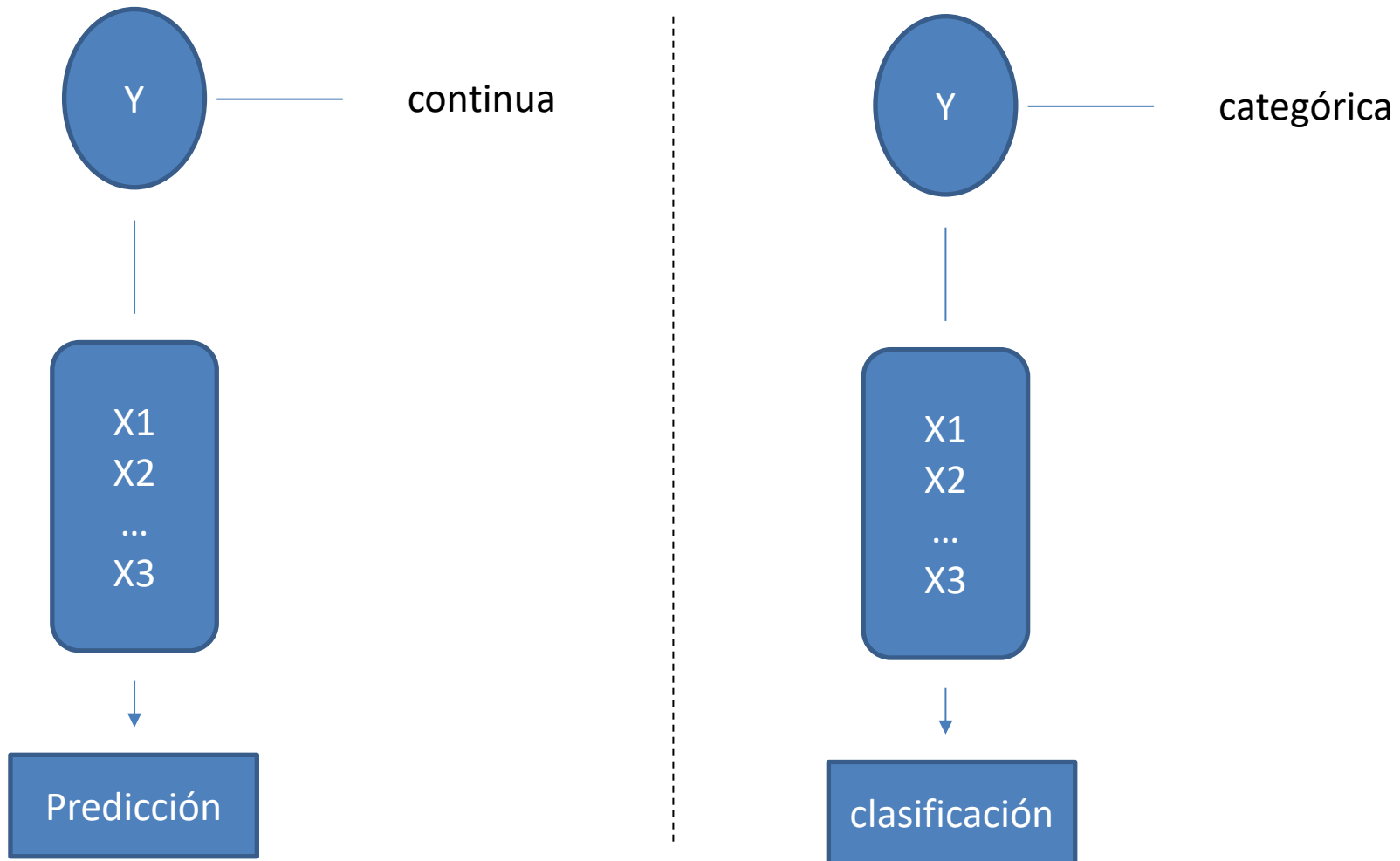
CART. Introducido por Breiman et al. (1984), propiamente es un algoritmo de árboles de decisión binario. Existe una versión similar llamada IndCART y que está disponible en el paquete IND distribuido por la NASA. RPART (PARTicionamiento Recursivo), una versión de CART esta disponible en R.

Arboles Bayesianos: Está basado en aplicación de métodos Bayesianos a arboles de decisión. Buntine (1992). Disponible en el paquete IND distribuido por la NASA.

CN2. Introducido por Clark and Niblett (1989).

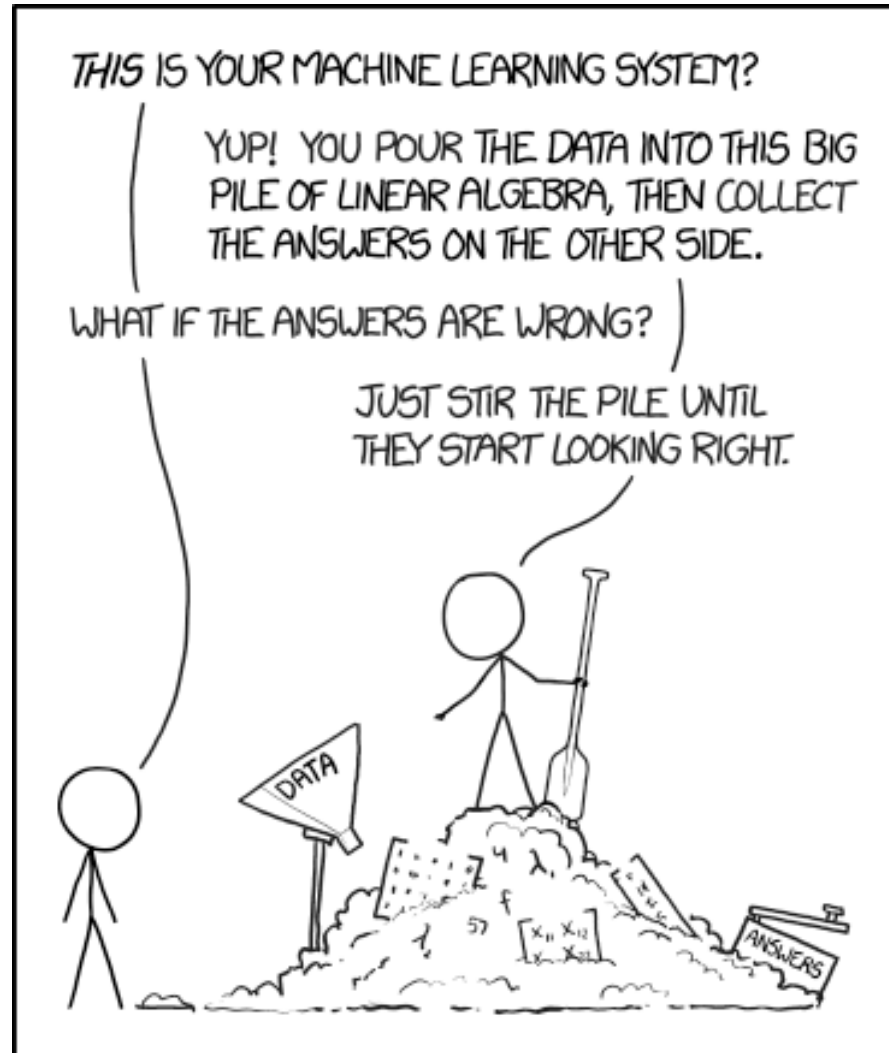
Introducción

- Recapitulando, los árboles son modelos supervisados de la forma:



Introducción

No olvidar: no se mete basura a un modelo... cuidado con las variables independiente que seleccionan...



Índice

1

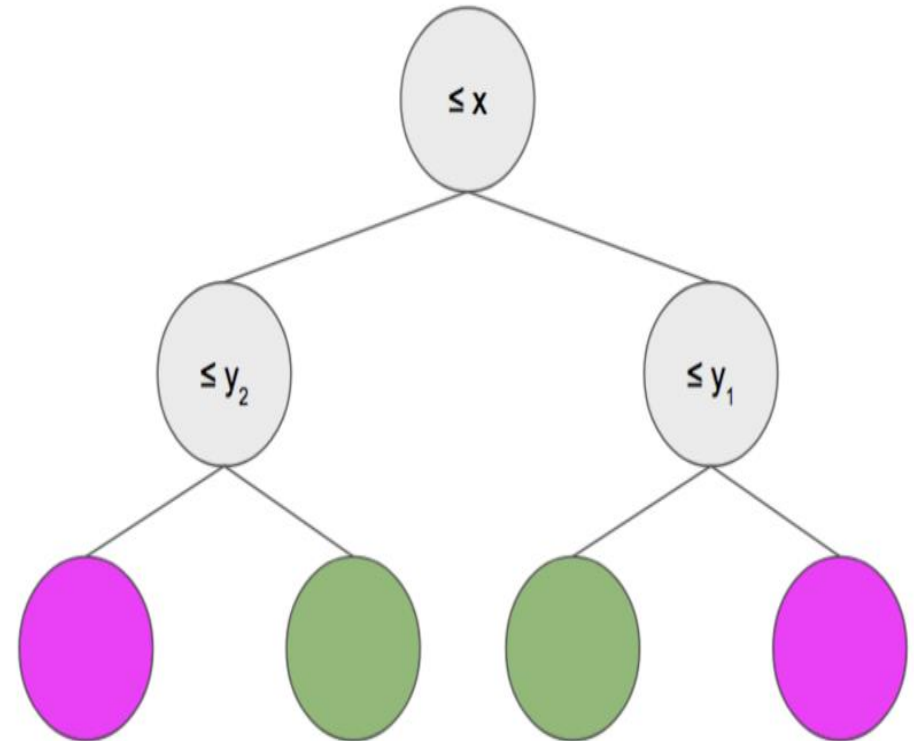
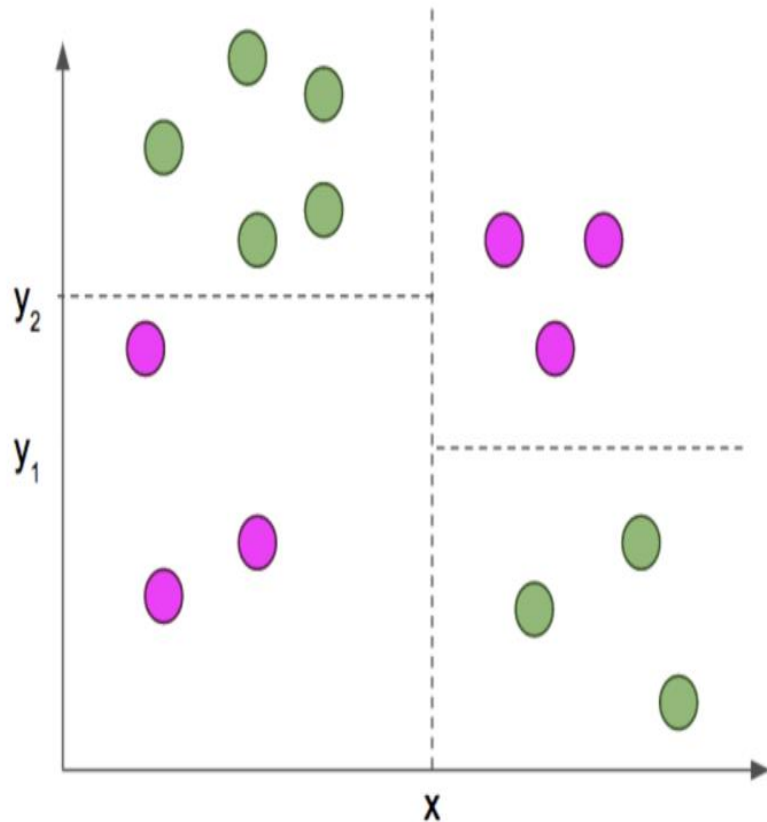
Introducción

2

Principios del árbol
de decisión

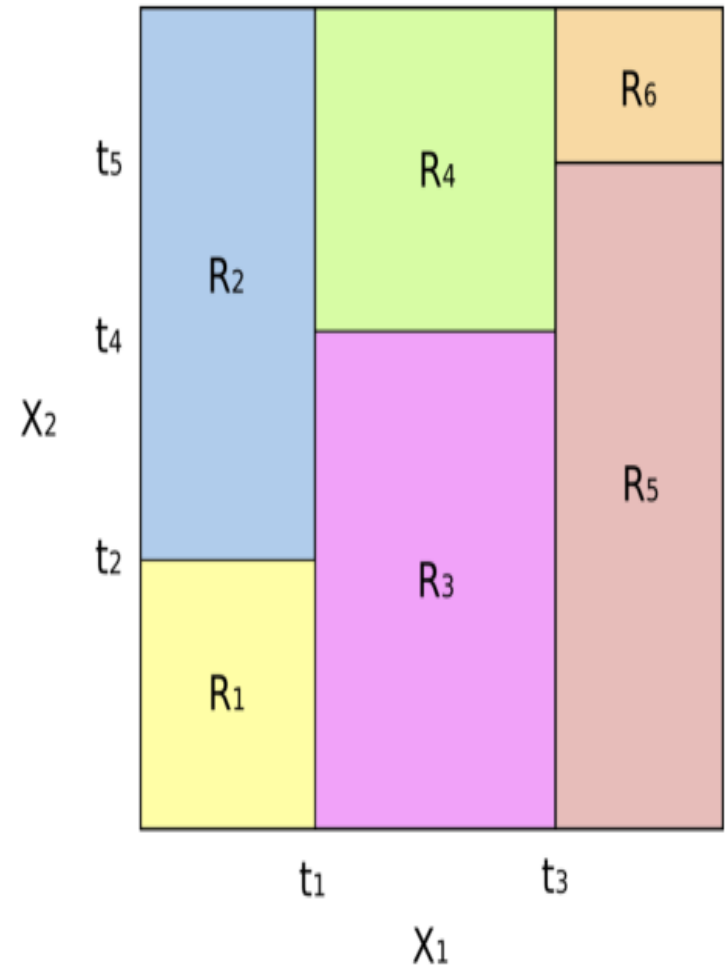
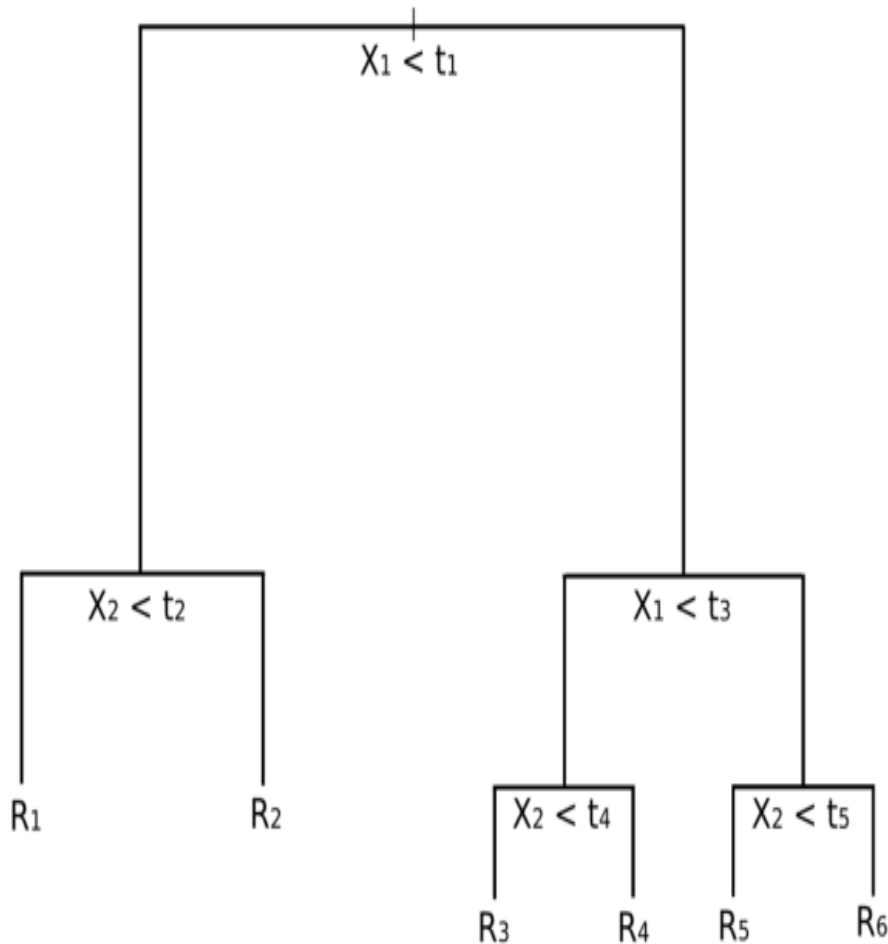
Principios de los árboles de decisión

Los árboles de decisión dividen el espacio de predictores (variables independientes) en regiones distintas y no superpuestas, para así poder asignar a las unidades de estudio.



Principios de los árboles de decisión

De forma general:



Principios de los árboles de decisión

Ilustremos el funcionamiento de los árboles de decisión con el siguiente ejemplo:

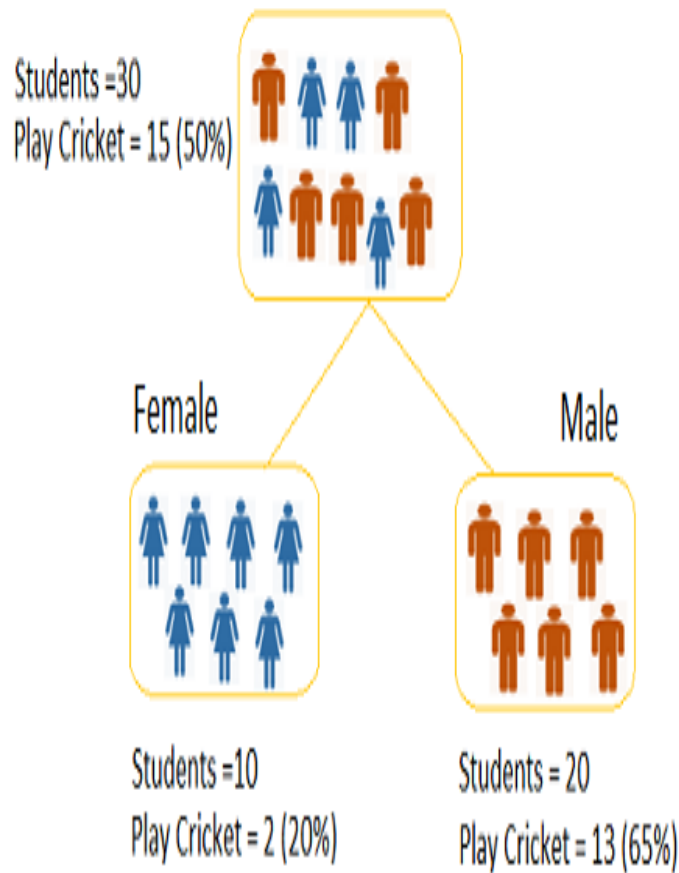
Digamos que tenemos una muestra de 30 estudiantes con tres variables: Género (Niño / Niña), Clase (IX / X) y Altura (5 a 6 pies). 15 de estos 30 jugarán cricket en tiempo libre. Ahora, quiero crear un modelo para predecir quién jugará cricket durante el período de ocio.

En este problema, necesitamos segregar a los estudiantes que juegan cricket en su tiempo de ocio basado en la variable de entrada altamente significativa entre los tres. Aquí es donde el árbol de decisiones ayuda.

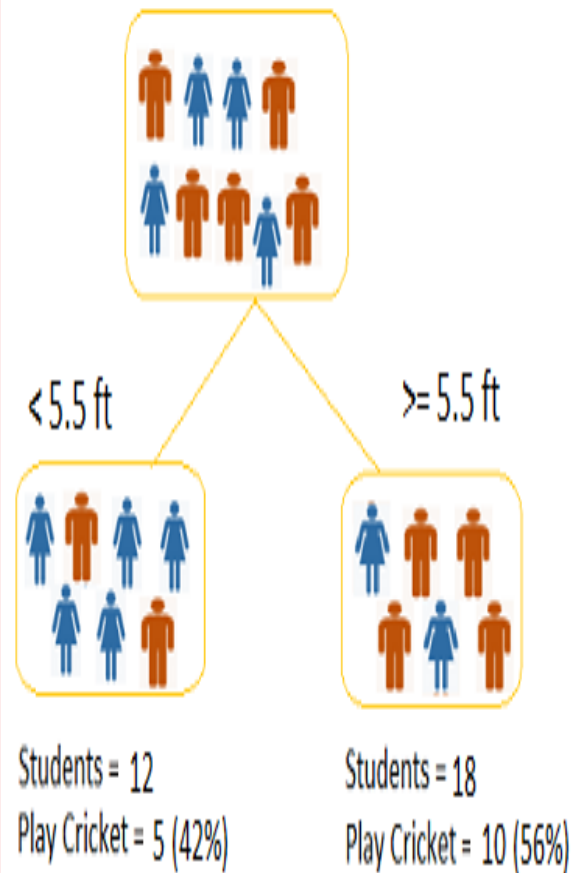
Segregará a los estudiantes en base a todos los valores de tres variables e identificará la variable, lo que crea los mejores conjuntos homogéneos de estudiantes (que son heterogéneos entre sí). En la instantánea a continuación, puede ver que la variable Género es capaz de identificar los mejores conjuntos homogéneos en comparación con las otras dos variables.

Principios de los árboles de decisión

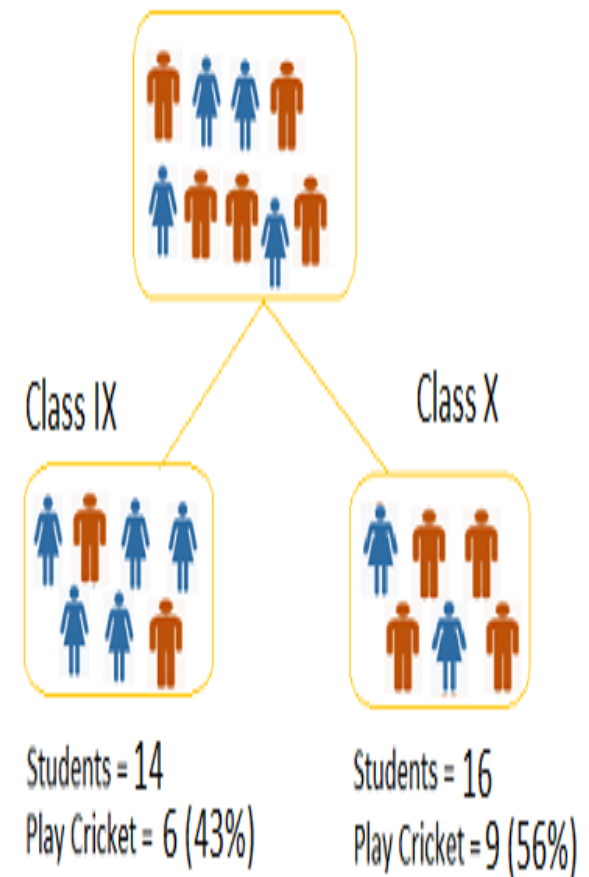
Split on Gender



Split on Height



Split on Class



Principios de los árboles de decisión

El árbol de decisiones identifica la variable más significativa y su valor que da los mejores conjuntos homogéneos de población. De esta forma va asignando (clasificando o prediciendo), a las unidades de estudio.

Como se ve, es un principio de bifurcación en la mejor asignación de los unidades. Tanto para una medida descriptiva, predictiva como de clasificación suele resultar muy conveniente.

Ahora la pregunta que surge es: ¿cómo identifica la variable y la división? Para ello, el árbol de decisiones utiliza varios algoritmos, que discutiremos en una siguiente sección.

Índice

1

Introducción

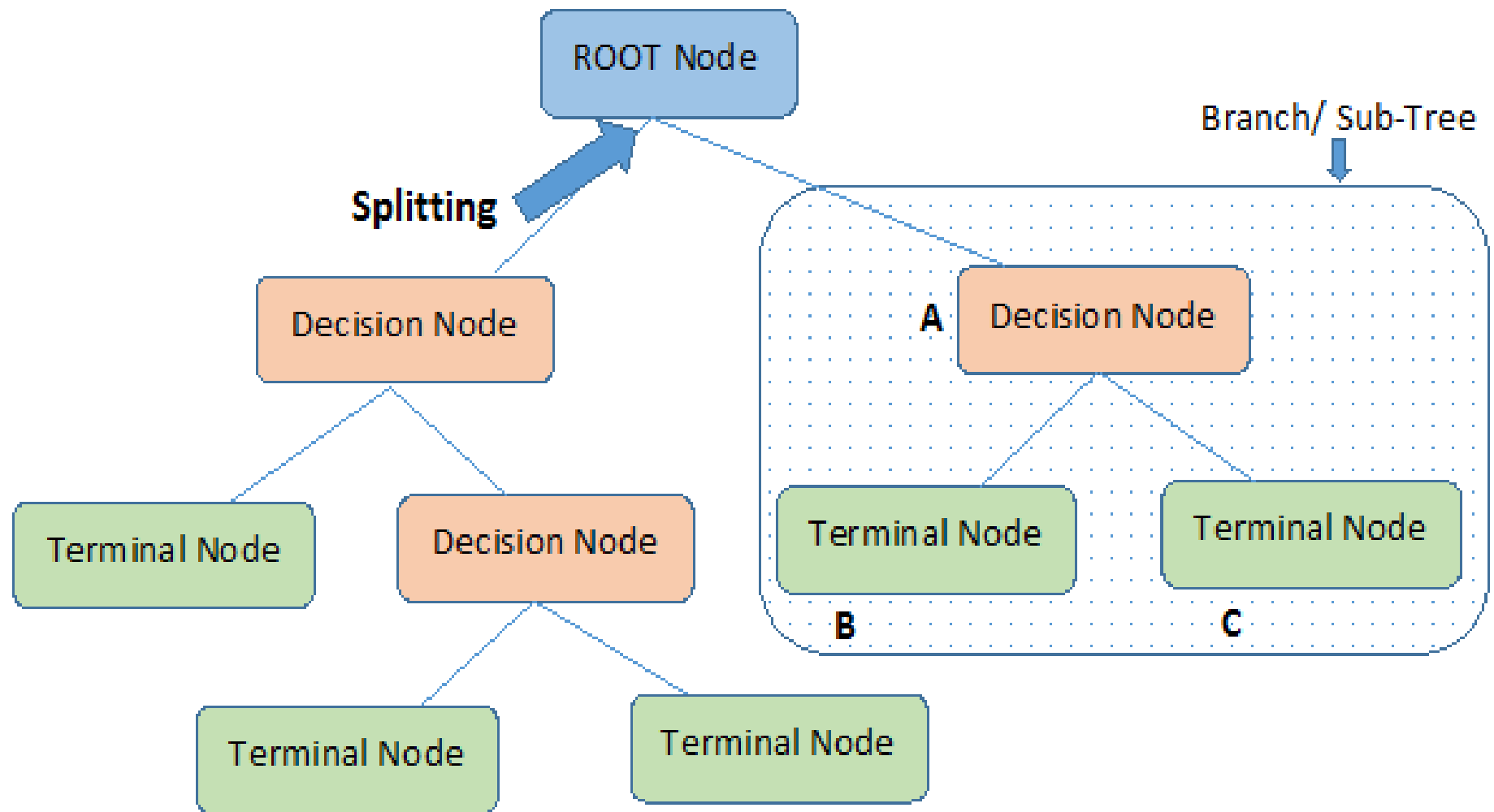
2

Principios del árbol
de decisión

3

Terminología

Terminología



Note:- A is parent node of B and C.

Terminología

La terminología básica utilizada para los árboles de decisión sería: nodo raíz, ramificación, nodo de decisión, nodo terminal u hoja, poda, rama (sub hoja), nodos padres e hijo.

Nodo raíz: representa toda la población o muestra y esto se divide en dos o más conjuntos homogéneos.

División / ramificación: es un proceso de dividir un nodo en dos o más subnodos.

Nodo de decisión: cuando un sub-nodo se divide en subnodos adicionales, se llama nodo de decisión.

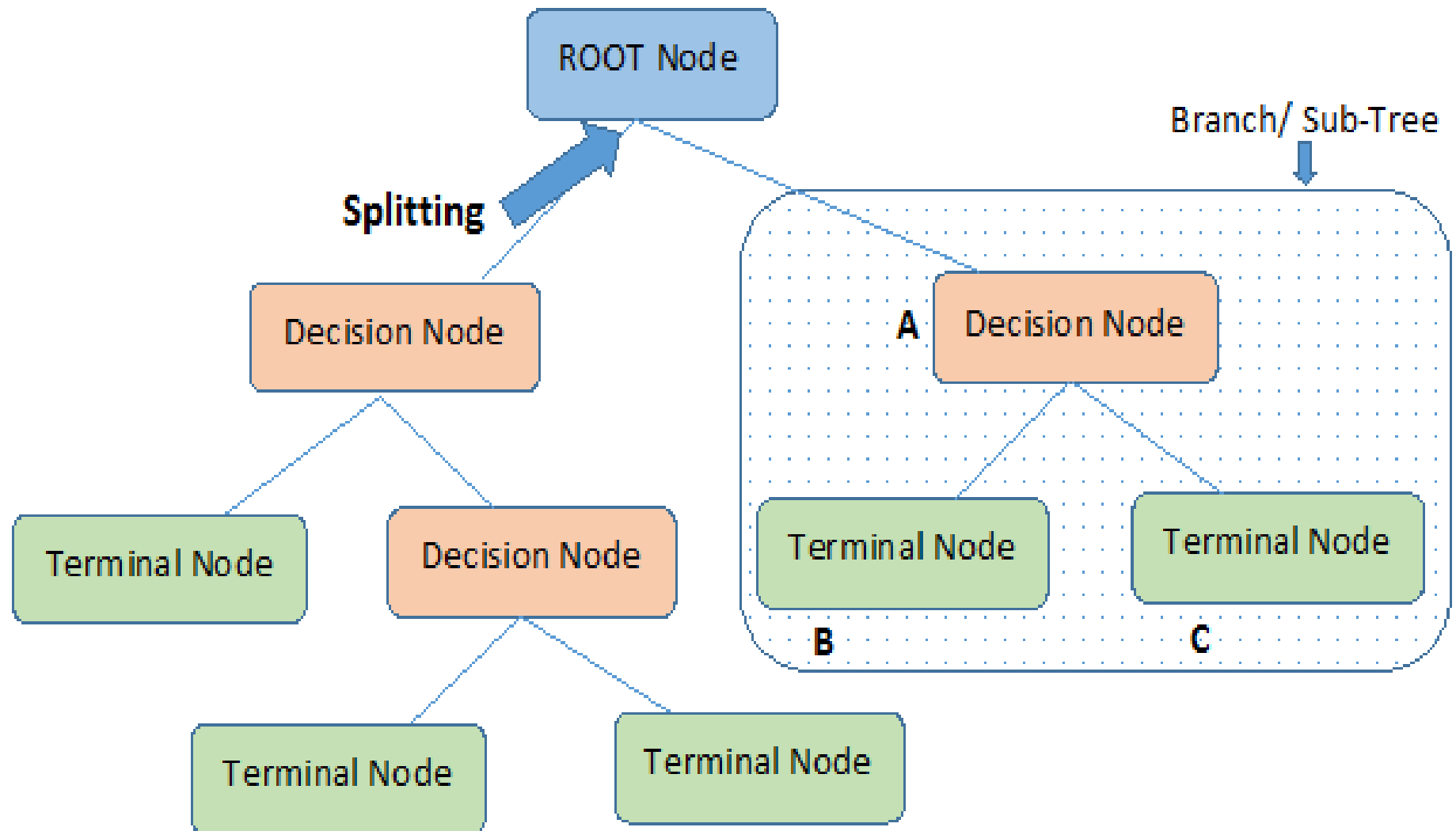
Nodo Leaf / Terminal: los nodos que no se dividen se llaman Leaf o nodo terminal.

Poda: Cuando eliminamos sub-nodos de un nodo de decisión, este proceso se llama poda. Se puede decir un proceso opuesto de división.

Subdivisión / Sub-Árbol: Una sub sección de todo el árbol se llama rama o sub-árbol.

Nodo padre e hijo: Un nodo, que se divide en sub-nodos, se llama nodo padre de sub-nodos donde como sub-nodos son el nodo secundario del nodo padre.

Terminología



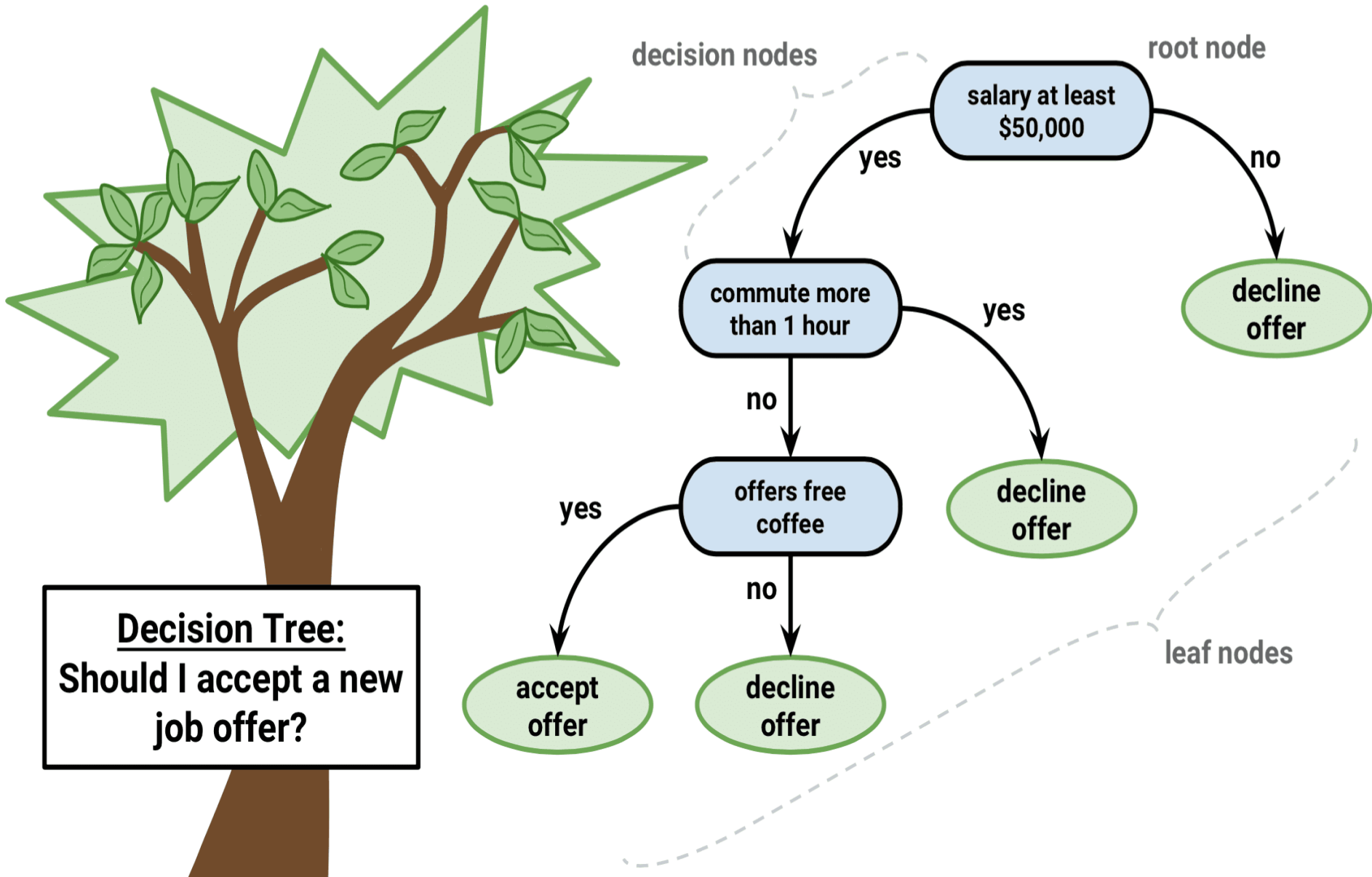
Note:- A is parent node of B and C.

Terminología

Vale la pena resaltar que.... sabemos que los nodos terminales (o hojas) se encuentra en la parte inferior del árbol de decisión. Esto significa que los árboles de decisión suelen ser dibujados al revés de manera que las hojas son el fondo y las raíces son las cimas (se muestra a continuación).



Terminología



Índice

1

Introducción

4

AC vs AR
ventajas y
desventajas

2

Principios del árbol
de decisión

3

Terminología

Árbol de regresión vs árbol de clasificación

Los tipos de árbol de decisión se basan en el tipo de variable objetivo que tenemos. Puede ser de dos tipos:

Árbol de decisión

Regresión

Variable dependiente es continua

Valores de los nodos terminales se reducen a la media de las observaciones en esa región.

Clasificación

Variable dependiente es categórica

El valor en el nodo terminal se reduce a la moda de las observaciones del conjunto de entrenamiento que han “caído” en esa región.

Ejemplo de árboles de regresión: el estudiante jugará al grillo o no, es decir, SI o NO.

Ejemplo de árboles de clasificación: estimación del salario.

Árbol de regresión vs árbol de clasificación

Ventajas

- Fácil de entender
- Util en exploración de datos: identificar importancia de variables a partir de cientos de variables.
- Menos limpieza de datos: outliers y valores faltantes no influyen el modelo (A un cierto grado)
- El tipo de datos no es una restricción
- Es un método no paramétrico (i.e., no hay suposición acerca del espacio de distribución y la estructura del clasificador)

Desventajas

- Sobreajuste
- Pérdida de información al categorizar variables continuas
- Precisión: métodos como SVM y clasificadores tipo ensamblador a menudo tienen tasas de error 30% más bajas que CART (Classification and Regression Trees)
- Inestabilidad: un pequeño cambio en los datos puede modificar ampliamente la estructura del árbol. Por lo tanto la interpretación no es tan directa como parece.

Índice

1

Introducción

4

AC vs AR
ventajas y
desventajas

2

Principios del árbol
de decisión

5

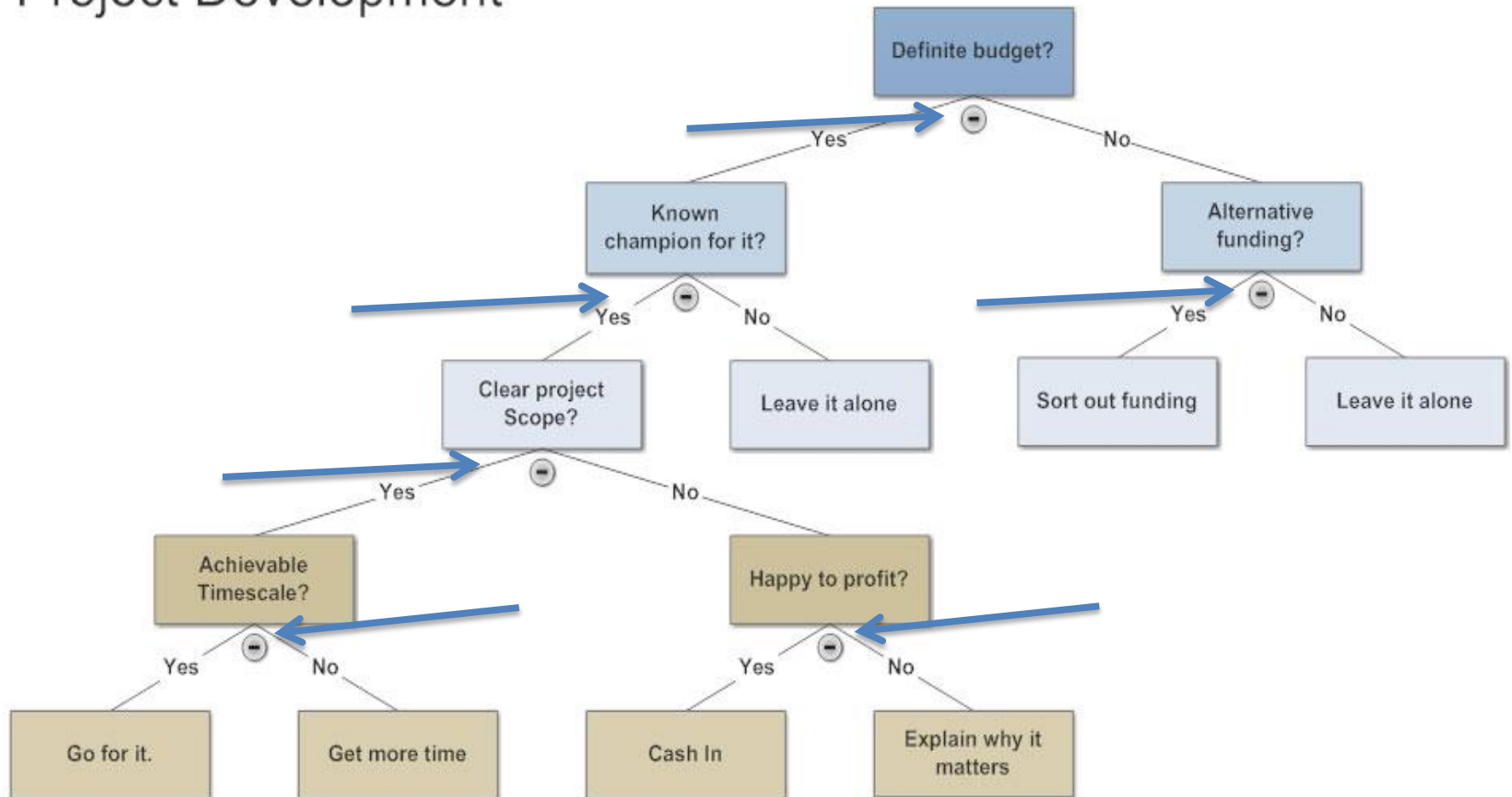
Partición /
ramificación

3

Terminología

Partición / ramificación

Project Development



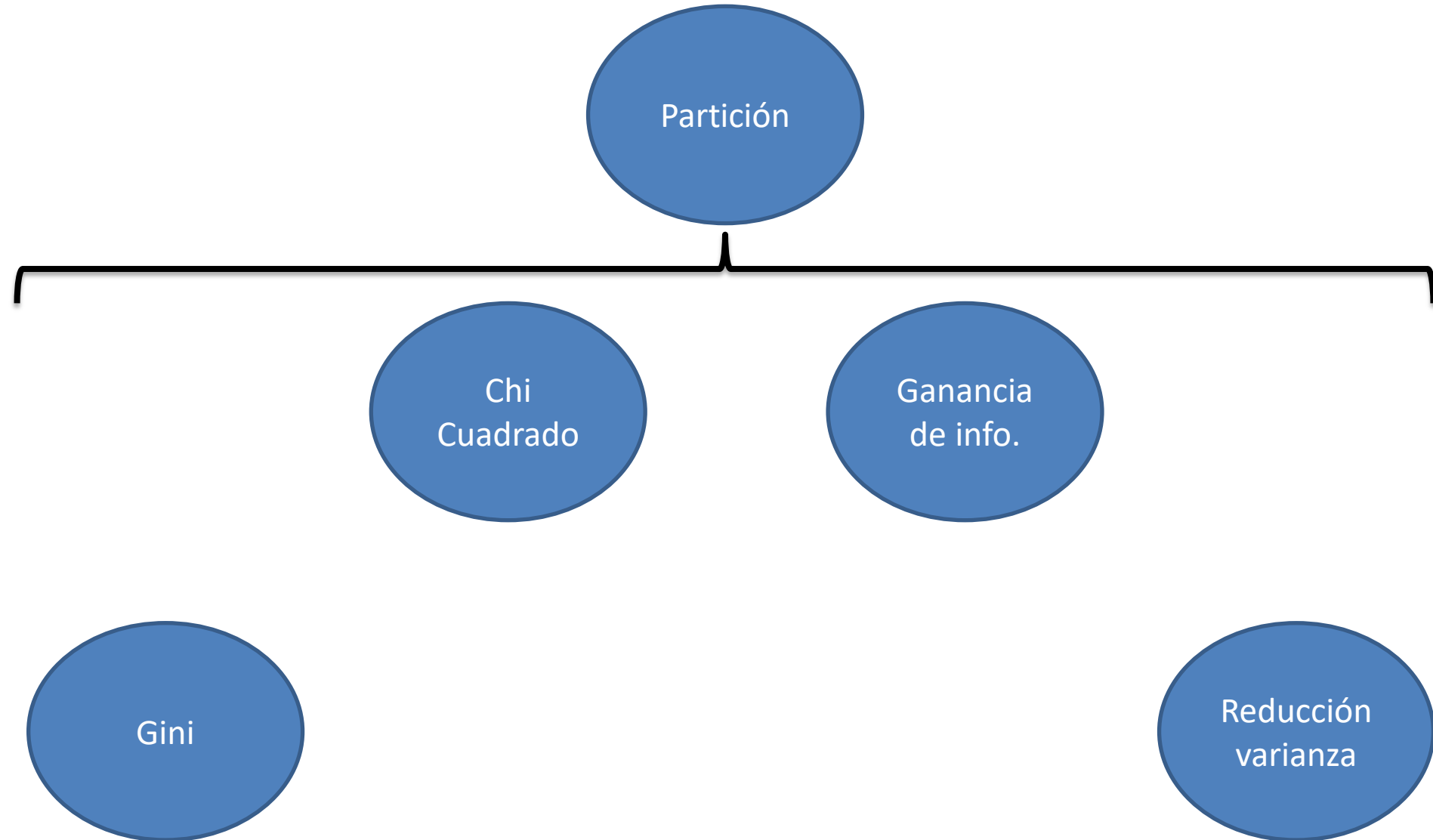
Partición / ramificación

La decisión de hacer divisiones estratégicas afecta en gran medida la precisión de un árbol. Los criterios de decisión son diferentes para árboles de clasificación y regresión.

Los árboles de decisión utilizan múltiples algoritmos para decidir dividir un nodo en dos o más subnodos. La creación de subnodos aumenta la homogeneidad de los subnodos resultantes. En otras palabras, podemos decir que la pureza del nodo aumenta con respecto a la variable objetivo. El árbol de decisiones divide los nodos en todas las variables disponibles y, a continuación, selecciona la división que da lugar a los subnodos más homogéneos.

La selección del algoritmo también se basa en el tipo de variables objetivo. Veamos los cuatro algoritmos más utilizados en el árbol de decisiones: Gini, Chi-Cuadrado, Ganancia de Información, Reducción de variancia.

Partición / ramificación



Partición / ramificación

Índice Gini

El índice de Gini establece como principio de funcionalidad: “Si seleccionamos aleatoriamente dos ítems de una población, entonces estos deben ser de la misma clase y la probabilidad de esto es 1 si la población es pura”. Así es un poco raro de entender... ya lo veremos de forma matemática.

Gini trabaja como:

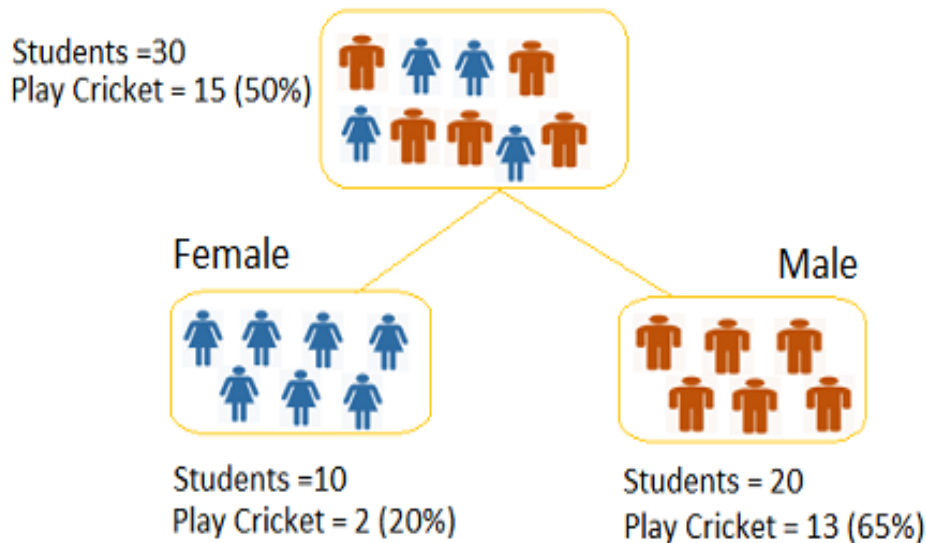
1. Funciona con la variable objetivo categórica "Éxito" o "Fracaso".
2. Realiza solamente divisiones binarias
3. Mayor el valor de Gini mayor la homogeneidad.
4. El método CART (Árbol de clasificación y de regresión) utiliza el método Gini para crear divisiones binarias... el que utilizamos en el presente curso

Partición / ramificación

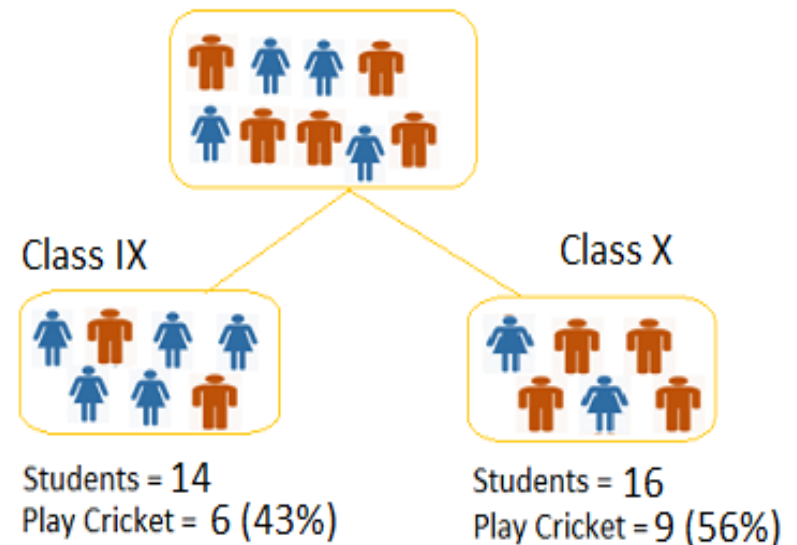
Para obtener el índice de Gini, hacemos lo siguiente:

1. Calcular Gini para subnodos, usando la fórmula suma de cuadrado de probabilidad de éxito y fracaso ($p^2 + q^2$).
2. Calcular el Gini para la división mediante el puntaje de Gini ponderado de cada nodo de esa división

Split on Gender



Split on Class



Partición / ramificación

Según la siguiente tabla:

Género		Clase	
Mujer	$(0.2)^2 + (0.8)^2 = 0.68$	IX	$(0.43)^2 + (0.57)^2 = 0.51$
Hombre	$(0.65)^2 + (0.35)^2 = 0.55$	X	$(0.56)^2 + (0.44)^2 = 0.51$
Pond.	$(10/30)0.68 + (20/30)0.55 = \mathbf{0.59}$	Pond.	$(14/30)0.51 + (16/30)0.51 = \mathbf{0.51}$

- ¿Cuál es el resultado para la variable “Altura”?
- ¿Cuál es el variable que se debe escoger para la primera división?

Partición / ramificación

Chi-Cuadrado

Es un algoritmo para descubrir la significación estadística entre las diferencias entre los subnodos y el nodo padre. Lo medimos por suma de cuadrados de diferencias estandarizadas entre las frecuencias observadas y esperadas de la variable objetivo. Funciona como sigue:

1. Funciona con la variable objetivo categórica "Éxito" o "Fracaso".
2. Puede realizar dos o más divisiones.
3. Mayor el valor de Chi-Cuadrado mayor la significación estadística de las diferencias entre el sub-nodo y el nodo padre.
4. Chi-cuadrado de cada nodo se calcula mediante la fórmula,
5. $\text{Chi-cuadrado} = ((\text{Actual} - \text{Esperado})^2 / \text{Esperado})^{1/2}$
6. Genera un árbol llamado CHAID (Chi-square Automatic Interaction Detector)

Partición / ramificación

Los Pasos para calcular el Chi cuadrado serían:

- 1. Calcular Chi-cuadrado para el nodo individual calculando la desviación para el éxito y el fracaso
- 2. Calcular el Chi-cuadrado de Split usando la Suma de todo Chi-cuadrado de éxito y Fracaso de cada nodo de la división

Node	Play Cricket	Not Play Cricket	Total	Expected Play Cricket	Expected Not Play Cricket	Deviation Play Cricket	Deviation Not Play Cricket	Chi-Square	
								Play Cricket	Not Play Cricket
Female	2	8	10	5	5	-3	3	1.34	1.34
Male	13	7	20	10	10	3	-3	0.95	0.95
Total Chi-Square								4.58	

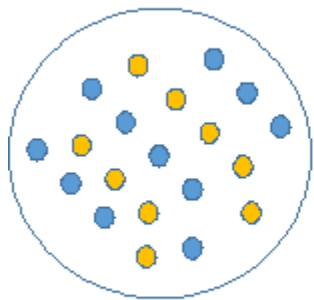


Node	Play Cricket	Not Play Cricket	Total	Expected Play Cricket	Expected Not Play Cricket	Deviation Play Cricket	Deviation Not Play Cricket	Chi-Square	
								Play Cricket	Not Play Cricket
IX	6	8	14	7	7	-1	1	0.38	0.38
X	9	7	16	8	8	1	-1	0.35	0.35
Total Chi-Square								1.46	

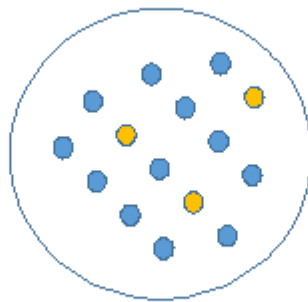
Partición / ramificación

Ganancia de información

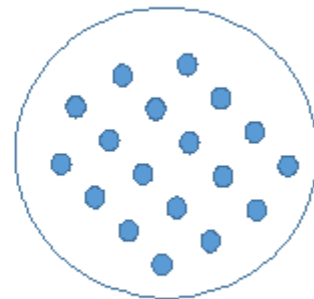
Si se observa la imagen de abajo y se piensa qué nodo se puede describir fácilmente. Estoy seguro, su respuesta es C porque requiere menos información ya que todos los valores son similares. Por otro lado, B requiere más información para describirlo y A requiere la máxima información. En otras palabras, podemos decir que C es un nodo Puro, B es menos impuro y A es más impuro.



A



B



C

Partición / ramificación

En la ganancia de información utilizamos el indicador de Entropía, y este se suele definir como:

$$Entropia = -p * \log_2(p) - q * \log_2(q)$$

Aquí, p y q es la probabilidad de “éxito” y “fracaso” respectivamente en el nodo. La entropía es también usada con variable objetivo categórica. Se escoje la división con la entropía más baja comparada con el nodo padre y otras variables de división. Mientras menor la entropía, mejor.

Sigamos con el mismo ejemplo de antes...

Partición / ramificación

Criterio	Entropy
Nodo Padre	$-(15/30) \log_2 (15/30) - (15/30) \log_2 (15/30) = 1$
Nodo Mujer	$-(2/10) \log_2 (2/10) - (8/10) \log_2 (8/10) = 0.72$
Nodo Hombre	$-(13/20) \log_2 (13/20) - (7/20) \log_2 (7/20) = 0.93$
Género ponderado	$(10/30)0.72 + (20/30)0.93 = \mathbf{0.86}$
Nodo Clase IX	$-(6/14) \log_2 (6/14) - (8/14) \log_2 (8/14) = 0.99$
Nodo Clase X	$-(9/16) \log_2 (9/16) - (7/16) \log_2 (7/16) = 0.99.$
Clase ponderado	$(14/30)0.99 + (16/30)0.99 = \mathbf{0.99}$

La entropía para la división por la variable "Género" es la más baja. Se puede derivar la ganancia de información a partir de la entropía así: $1 - \text{Entropía}$.

Partición / ramificación

Reducción en la varianza (regresión)

Los algoritmos anteriores se aplicaban para problemas de clasificación con variables objetivo categóricas. La reducción en la varianza es un algoritmo usado para variables objetivo continuas (problemas de regresión). Este algoritmo usa la fórmula estándar de la varianza para escoger el criterio de división. La división con la varianza más baja se elige para dividir la población:

$$Varianza = \frac{\sum (x_i - \bar{x})^2}{n}$$

Para el cálculo:

1. Se calcula la varianza para cada nodo.
2. Se calcula la varianza para cada división como un promedio ponderado de las varianzas de cada nodo.

Partición / ramificación

Criterio	Mean	Varianza
Nodo Padre	$(15 \times 1 + 15 \times 0) / 30 = 0.5$	$(15 \times (1 - 0.5)^2 + 15 \times (0 - 0.5)^2) / 30 = \mathbf{0.25}$
Nodo Mujer	$(2 \times 1 + 8 \times 0) / 10 = 0.2$	$(2 \times (1 - 0.2)^2 + 8 \times (0 - 0.2)^2) / 10 = 0.16$
Nodo Hombre	$(13 \times 1 + 7 \times 0) / 20 = 0.65$	$(13 \times (1 - 0.65)^2 + 7 \times (0 - 0.65)^2) / 20 = 0.23$
Género ponderado		$(10/30) \times 0.16 + (20/30) \times 0.23 = \mathbf{0.21}$
Nodo Clase IX	$(6 \times 1 + 8 \times 0) / 14 = 0.43$	$(6 \times (1 - 0.43)^2 + 8 \times (0 - 0.43)^2) / 14 = 0.24$
Nodo Clase X	$(9 \times 1 + 7 \times 0) / 16 = 0.56$	$(9 \times (1 - 0.56)^2 + 7 \times (0 - 0.56)^2) / 16 = 0.25$
Clase ponderado		$(14/30) \times 0.24 + (16/30) \times 0.25 = \mathbf{0.25}$

Índice

1

Introducción

4

AC vs AR
ventajas y
desventajas

2

Principios del árbol
de decisión

5

Partición /
ramificación

3

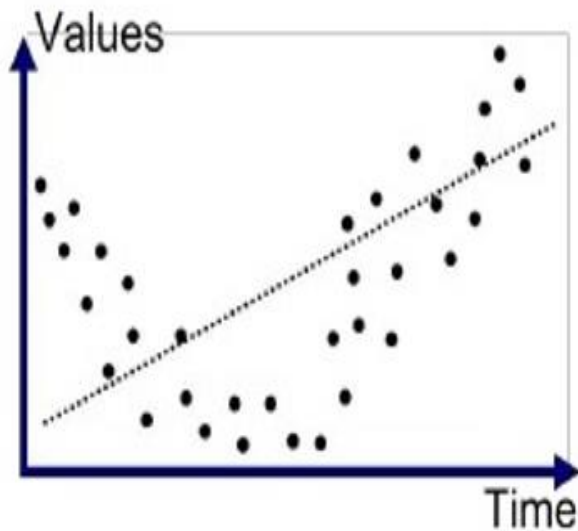
Terminología

6

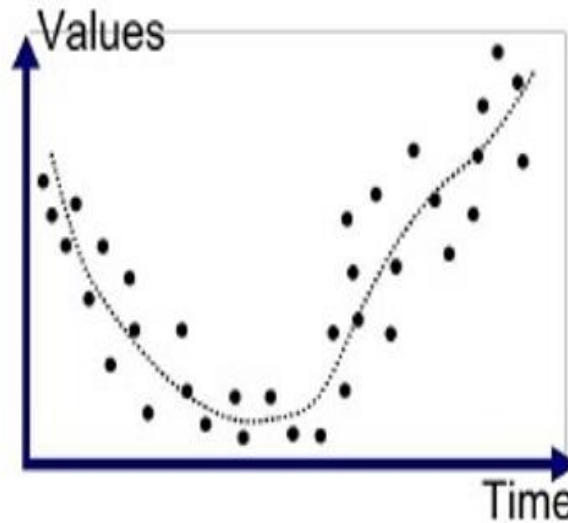
Estimación, el sobre
ajuste y la poda

Estimación, sobre ajuste y la poda

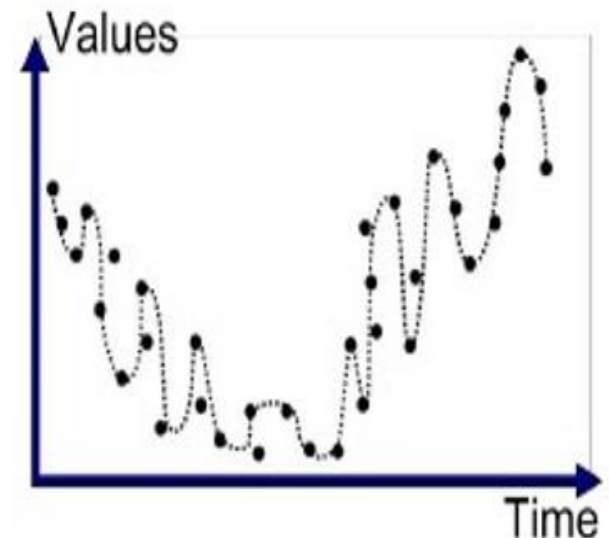
En cualquier tipo de modelización, se tiene el gran problema de estimar de forma efectiva el conjunto de datos: no ser ni tan específico, y no ser tan “general”. Se busca el punto medio.



Underfitted

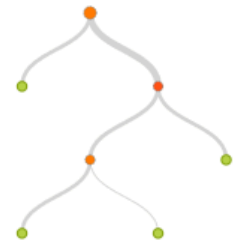


Good Fit/Robust



Overfitted

Estimación, sobre ajuste y la poda



- En la estimación de los parámetros de los árboles de decisión, el sobreajuste es uno de los desafíos más importantes en el proceso de modelación.
- Si no se definen límites, el árbol tendrá un 100% de precisión en el conjunto de datos de entrenamiento. En el peor caso tendrá una hoja por cada observación.
- Dos formas de evitar el sobreajuste:
 - (a) Definir restricciones sobre el tamaño del árbol y
 - (b) Podar el árbol.



Estimación, sobre ajuste y la poda

Restricciones sobre el tamaño del árbol

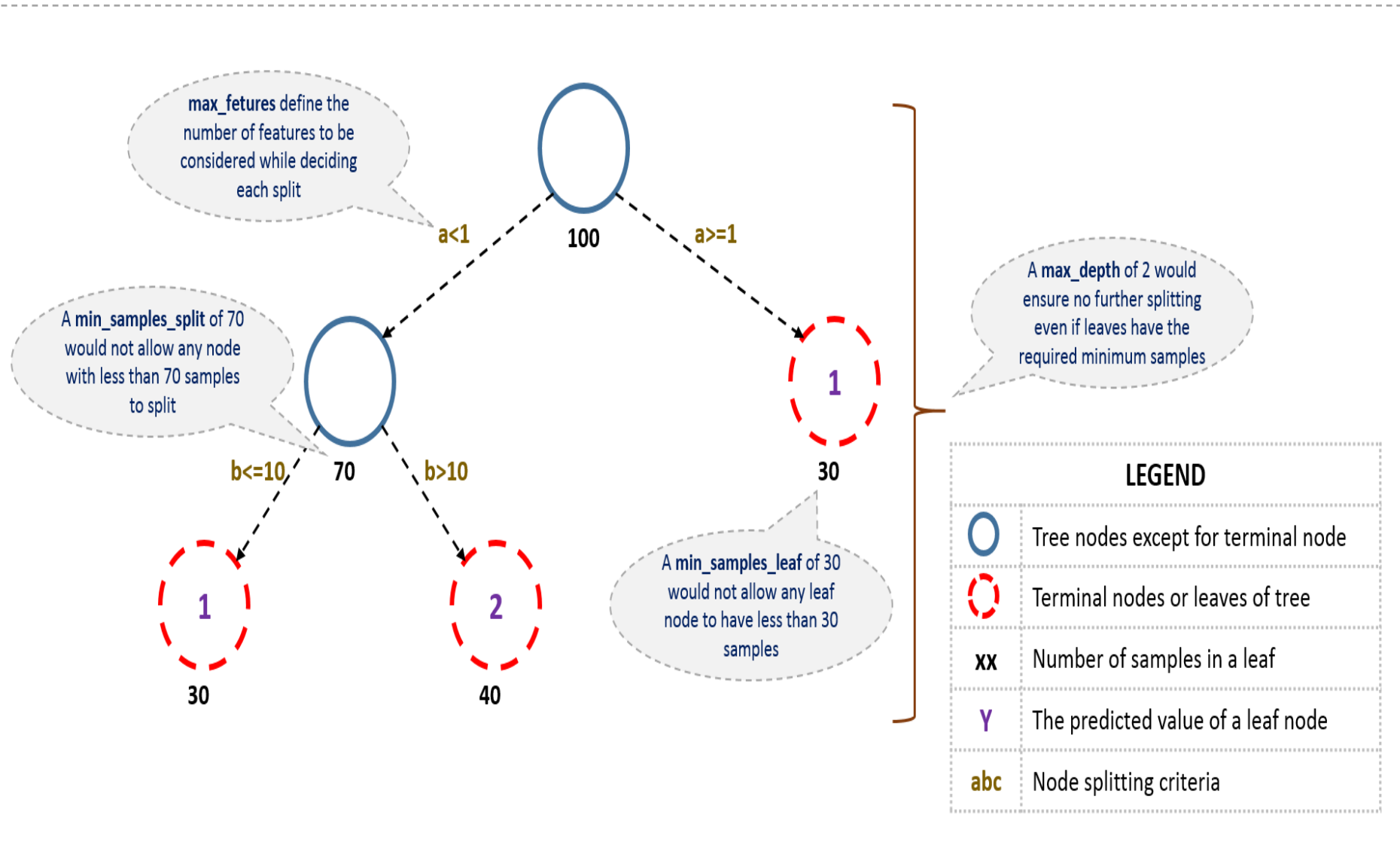
Las restricciones son independientes de la herramienta de programación, son formas de “acomodar” el árbol.

- 1. Mínimo de observaciones para dividir un nodo**
- 2. Mínimo número de observaciones para un nodo terminal**
- 3. Máxima profundidad del árbol (vertical)**
- 4. Máximo número de nodos hoja**
- 5. Máximo número de atributos a considerar para la ramificación**

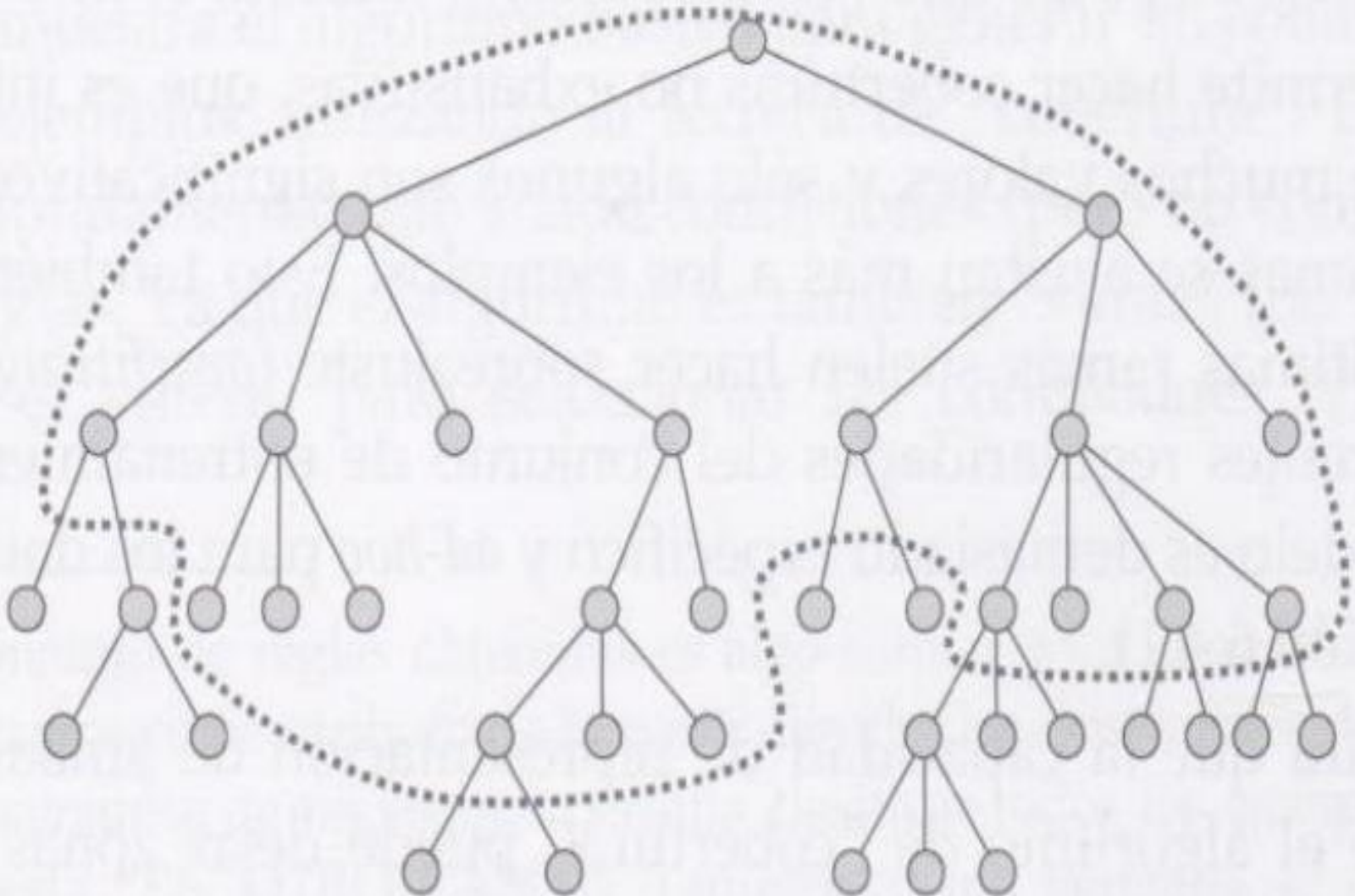
Refiérase a la siguiente página:

<https://www.analyticsvidhya.com/blog/2016/04/complete-tutorial-tree-based-modeling-scratch-in-python/>

Estimación, sobre ajuste y la poda



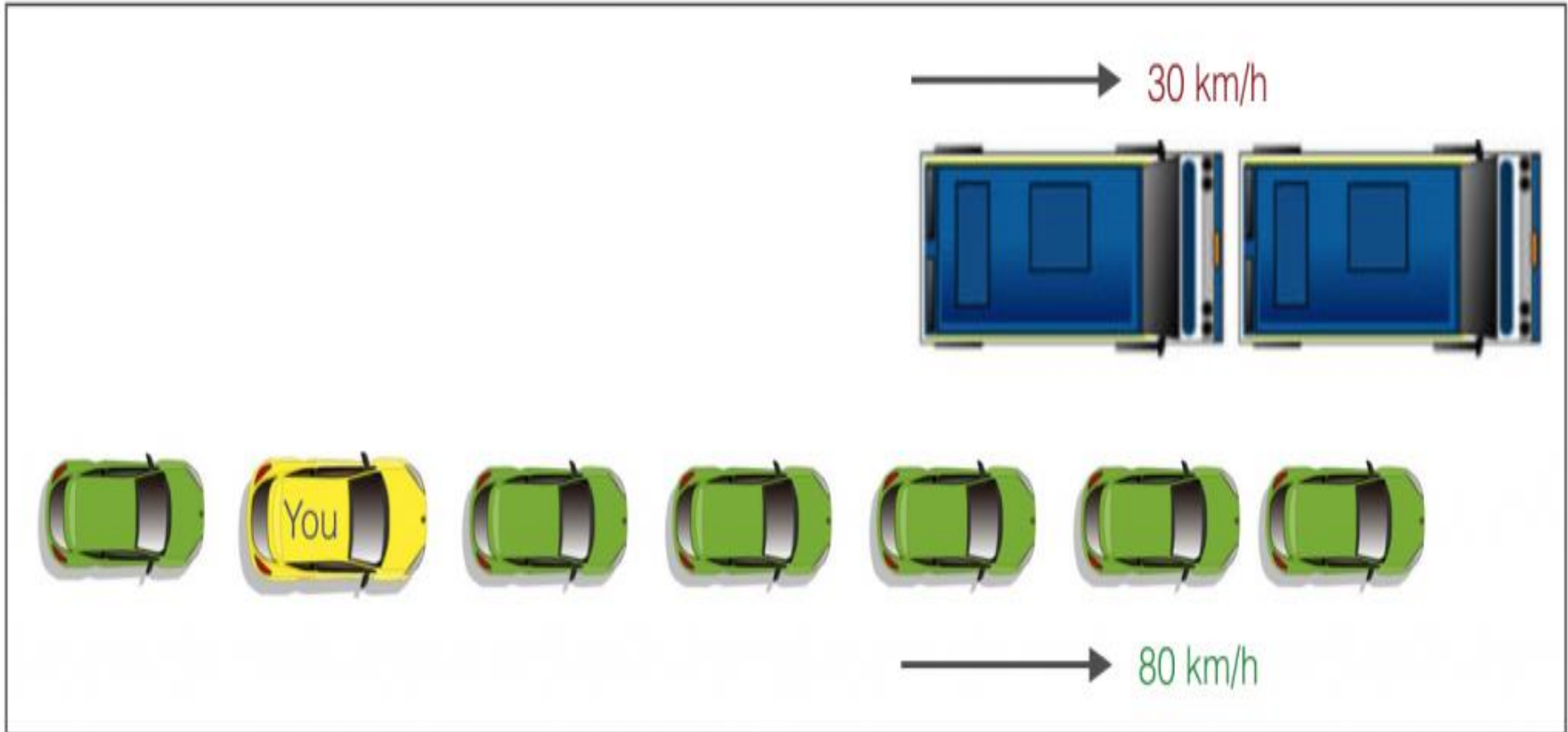
Estimación, sobre ajuste y la poda



Estimación, sobre ajuste y la poda

La poda

Para la explicación de la poda, veamos el siguiente dibujo.



Estimación, sobre ajuste y la poda

- Hay 2 carriles:

1. Un carril con coches que se mueven a 80km / h
2. Un carril con camiones que se mueven a 30km / h



En este instante, usted es el coche amarillo y tiene 2 opciones:

1. Tome a la izquierda y adelantar a los otros 2 coches rápidamente
2. Manténgase en movimiento en el carril actual

¿Cuál sería la mejor opción?



Estimación, sobre ajuste y la poda

Vamos a analizar estas opciones. En la primera opción, adelantará inmediatamente el coche y llegará detrás del camión y empezará a moverte a 30 km / h, buscando una oportunidad de retroceder. Todos los coches originalmente detrás de usted avanzan en el mientras tanto. Esta sería la opción óptima si su objetivo es maximizar la distancia recorrida en los siguientes 10 segundos. En la última opción, la venta a través de la misma velocidad, camiones cruzados y luego superar tal vez dependiendo de la situación por delante. Codicioso!

Esta es exactamente la diferencia entre el árbol de decisión normal y poda. Un árbol de decisión con restricciones no verá el camión por delante y adoptará un enfoque codicioso tomando una izquierda. Por otro lado si usamos la poda, en efecto miramos algunos pasos adelante y tomamos una decisión más acertada.

Estimación, sobre ajuste y la poda

Así que sabemos que la poda es mejor. Pero, ¿cómo implementarlo en el árbol de decisión? La idea es simple.

1. Primero hacemos el árbol de la decisión a una profundidad grande.
2. Luego empezamos en la parte inferior y empezamos a quitar las hojas que nos están dando resultados negativos cuando se comparan desde la parte superior.
3. Supongamos que una división nos está dando una ganancia de -10 (pérdida de 10) y luego la siguiente división en que nos da una ganancia de 20. Un árbol de decisión simple se detendrá en el paso 1, pero en la poda, vamos a ver que La ganancia total es de +10 y mantener ambas hojas.



Estimación, sobre ajuste y la poda

Importante

No profundizaremos en la poda, es un método más de machine learning que estadística, por lo que no está contemplado en los objetivos del curso.



Estimación, sobre ajuste y la poda

Algunas preguntas sobre los árboles pueden surgir:

"Si puedo usar la regresión logística para problemas de clasificación y regresión lineal para problemas de regresión, ¿por qué hay necesidad de usar árboles"?

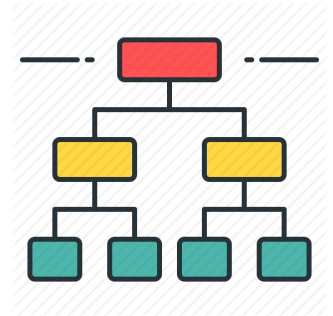


- Muchos de nosotros tenemos esta pregunta, y, esto es válido también.
- En realidad, puede utilizar cualquier algoritmo. Depende del tipo de problema que esté resolviendo. Echemos un vistazo a algunos factores clave que le ayudarán a decidir qué algoritmo utilizar.

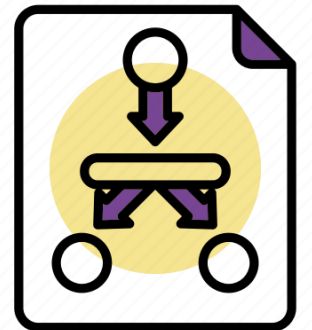


Estimación, sobre ajuste y la poda

1. Si la relación entre variable dependiente y independiente es bien aproximada por un modelo lineal, la regresión lineal superará al modelo basado en árboles.

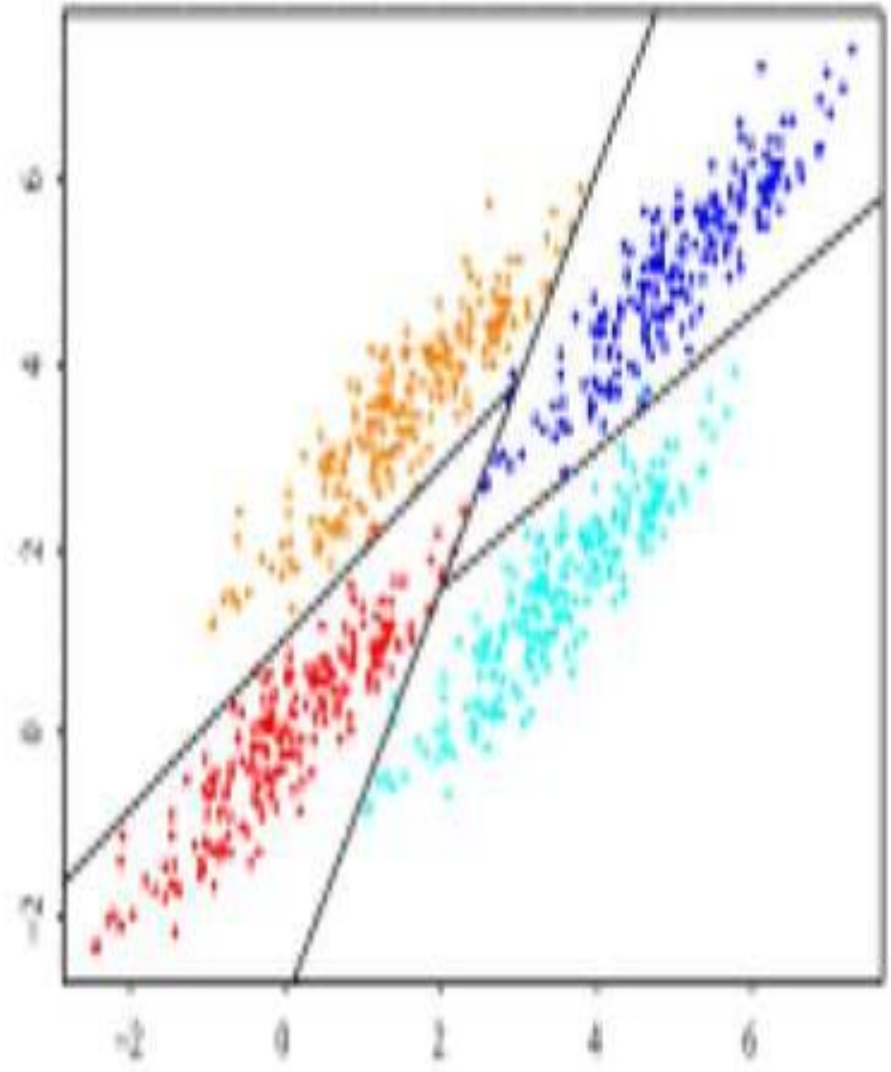
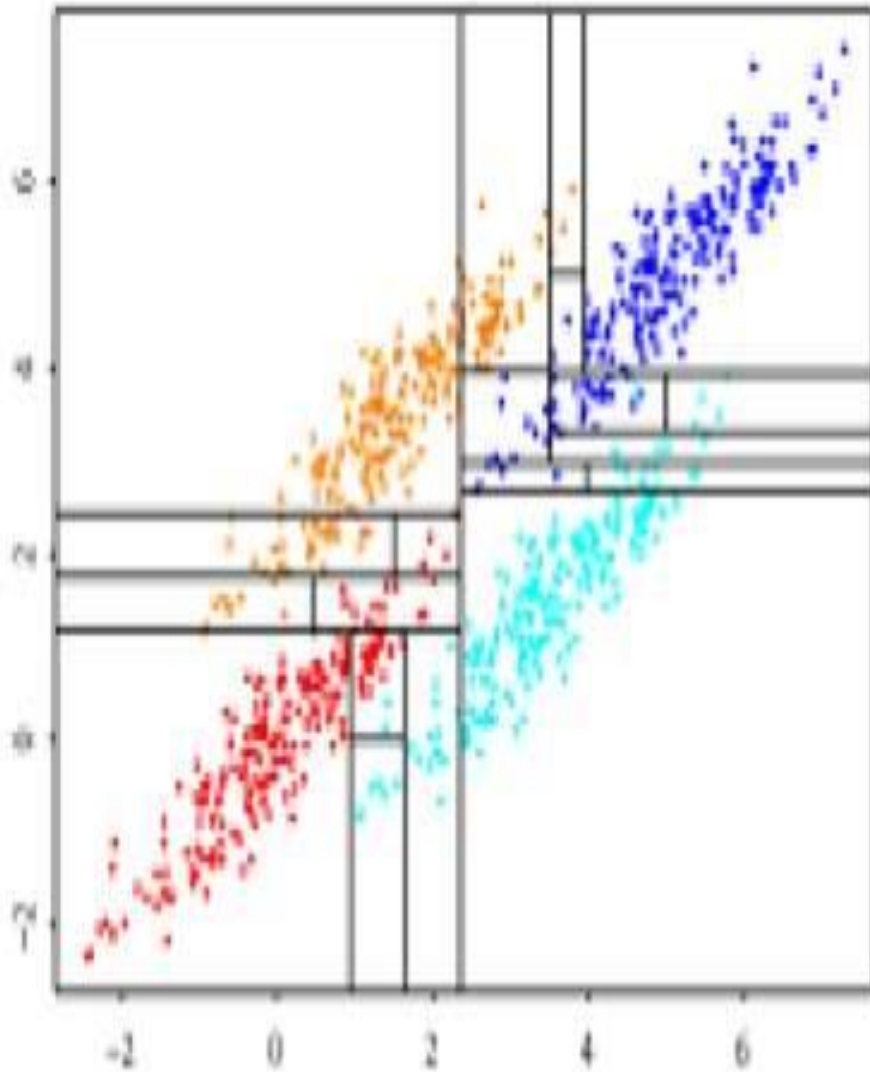


2. Si existe una alta no linealidad y relación compleja entre variables dependientes e independientes, un modelo de árbol superará a un método de regresión clásico.



3. Si necesita construir un modelo que sea fácil de explicar a la gente, un modelo de árbol de decisión siempre será mejor que un modelo lineal. Los modelos de árboles de decisión son aún más fáciles de interpretar que la regresión lineal.

Estimación, sobre ajuste y la poda



Índice

7

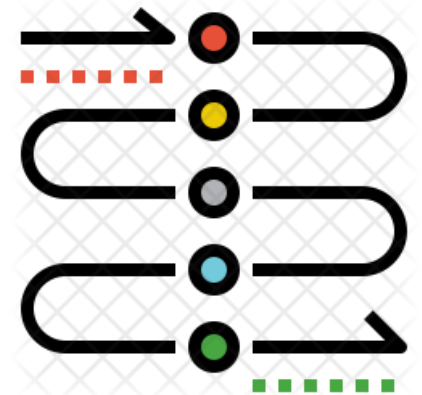
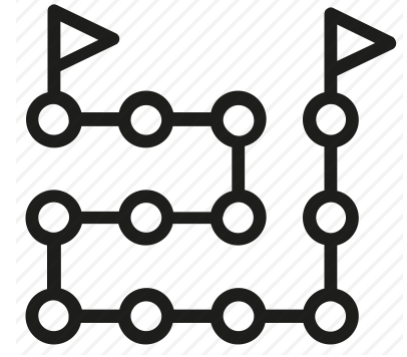
Etapas de análisis

Árboles de clasificación

Un árbol de clasificación posee los siguientes etapas

1. Estadísticas descriptivas
2. Particional datos: entrenamiento y test
3. Construir el modelo
4. Hacer la predicción
5. Medir el rendimiento del modelo
6. Ajustar los hyper-parámetros

Veamos el caso para un Árbol de clasificación y uno de regresión.



Índice

7

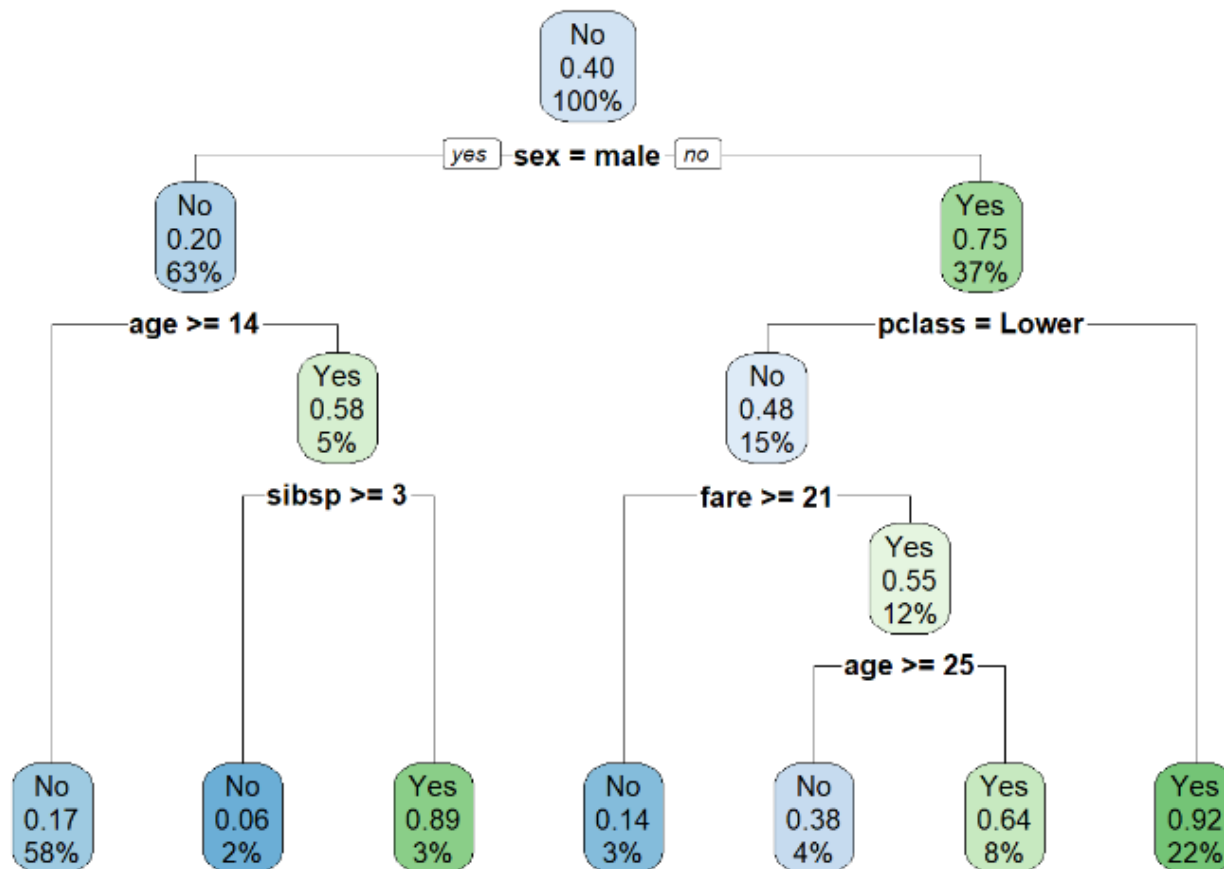
Etapas de análisis

8

Árboles de
clasificación

Árboles de regresión

Desde un punto descriptivo, estimaremos el árbol de clasificación, explicaremos dicha constitución, y veremos como implementar ciertas restricciones para mejorar las estimaciones.



Índice

7

Etapas de análisis

8

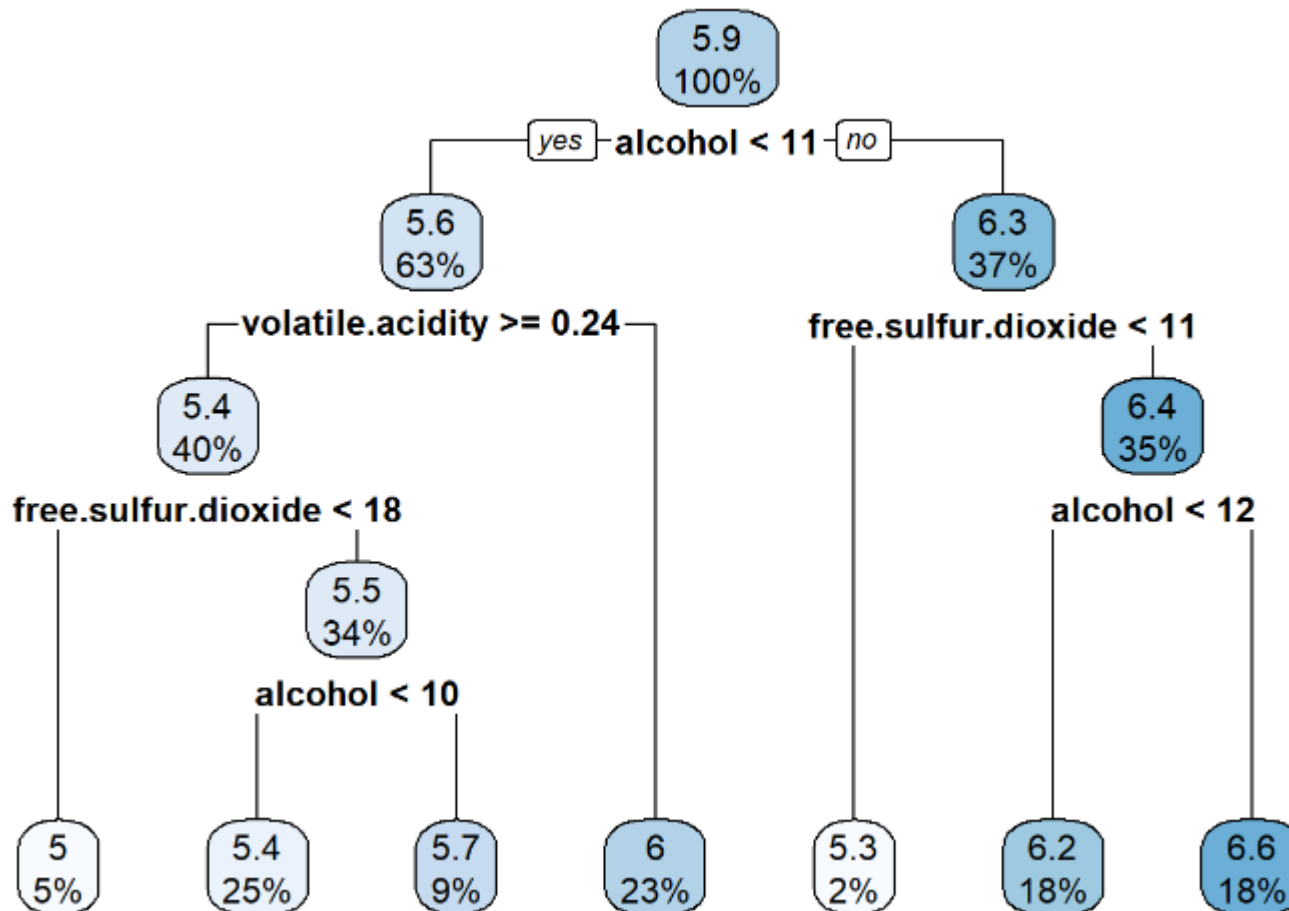
Árboles de decisión

9

Árboles de regresión

Árboles de regresión

De igual forma, veremos el análisis de un árbol de regresión, enfatizando el árbol y evaluando sus resultados.



Etapas en el análisis de un árbol

De forma general, las etapas son:

1. Análisis descriptivo de las variables.
2. Estimación del modelo (predicción o clasificación).
3. La poda o las restricciones de los árboles (en este curso esto no lo realizaremos).
4. La predicción o clasificación de las variables.
5. Evaluación del modelo
6. Cambiar ciertos hyper-parámetros.



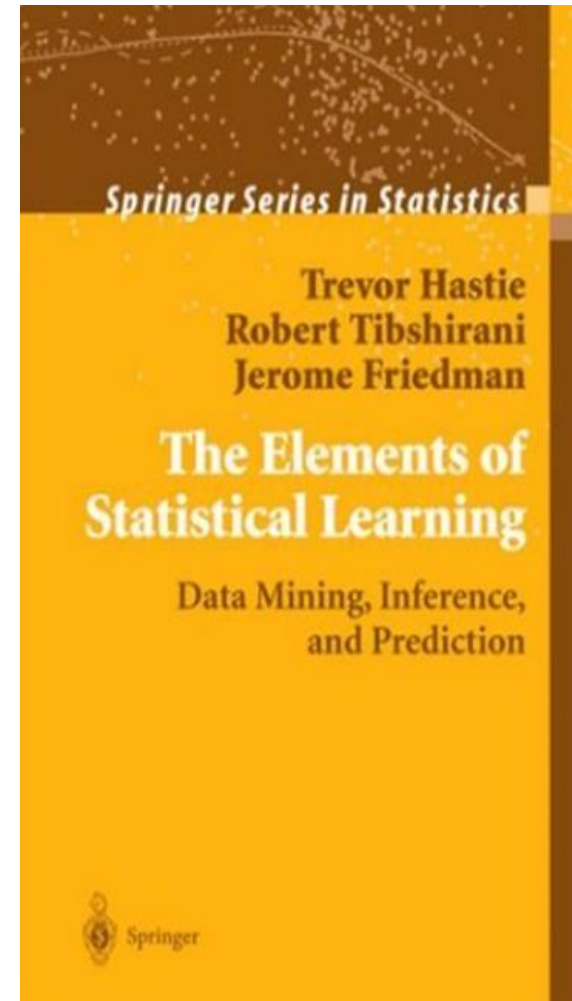
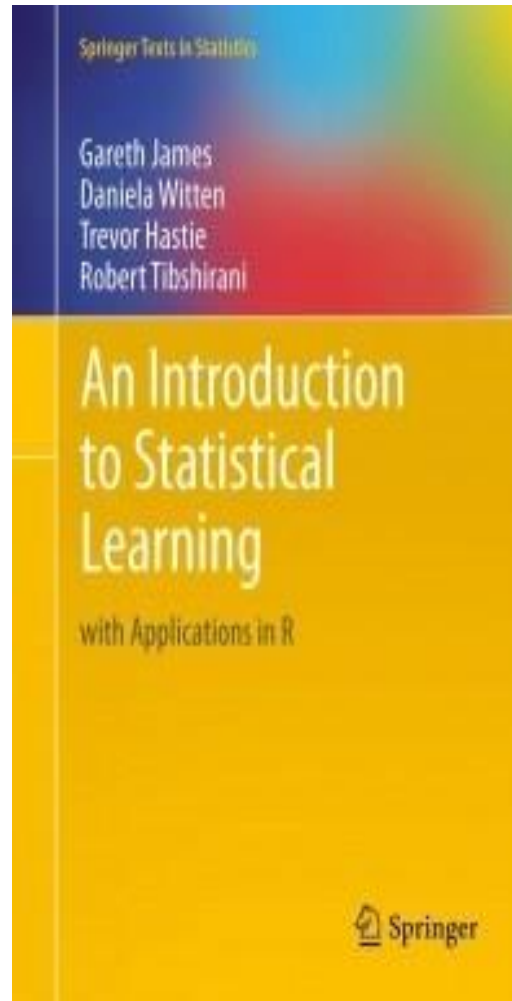
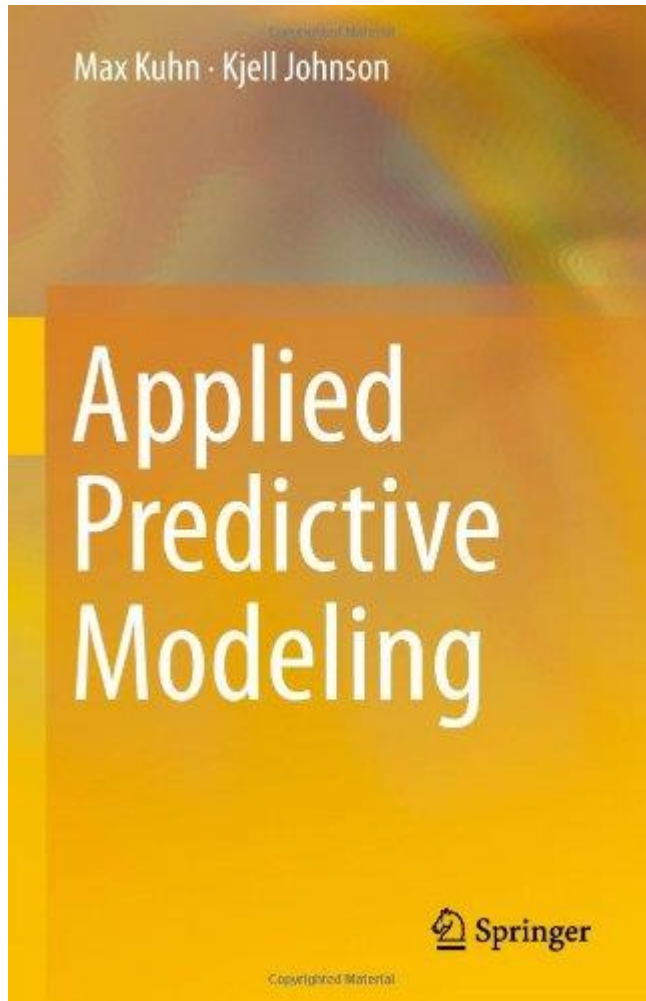
Conclusión: consulta de enlaces

Para ampliar el tempa del Cluster Analysis, pueden consultar los siguientes enlaces:

<https://bookdown.org/content/2031/arboles-de-decision-parte-i.html> <http://academic.uprm.edu/eacuna/treesacuna.pdf>
[http://www.oldemarrodriguez.com/yahoo_site_admin/assets/docs/Presentaci%C3%B3n - Arboles de Decisi%C3%B3n.234135906.pdf](http://www.oldemarrodriguez.com/yahoo_site_admin/assets/docs/Presentaci%C3%B3n_-_Arboles_de_Decisi%C3%B3n.234135906.pdf)
https://rstudio-pubs-static.s3.amazonaws.com/356615_c67d8544df654a68be06a7c914304b25.html
<https://www.uv.es/mlejarza/actuariales/tam/arbolesdecision.pdf>
https://rpubs.com/jboscomendoza/arboles_decision_clasificacion
https://rpubs.com/Cristina_Gil/arboles_ensemble
<https://rpubs.com/elfenixsoy/arbol-veronica>
<https://rpubs.com/amaurandi/ejemploTree>
https://www.datanalytics.com/libro_r/arboles-de-decision.html
https://www.datanalytics.com/libro_r/arboles-de-decision.html
<https://rpubs.com/ryankelly/dtrees>
https://rpubs.com/minma/cart_with_rpart
<https://rpubs.com/anish20/decisionTreesinR>
<https://rpubs.com/coleeagland/decisiontreesislr832>
https://rpubs.com/superseer/decision_trees_for_regression
<https://rpubs.com/cyobero/regression-tree>
<https://rpubs.com/blkdynamite/STAT6620HW5>
<https://rpubs.com/Loy/decisiontree-regression>
<https://rpubs.com/maulikpatel/218396>

Conclusión: consulta de libros

Se recomienda la siguiente bibliografía:



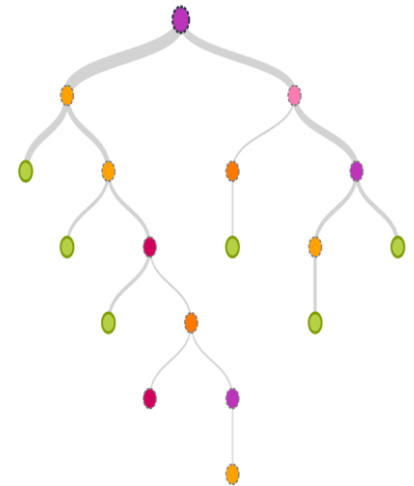
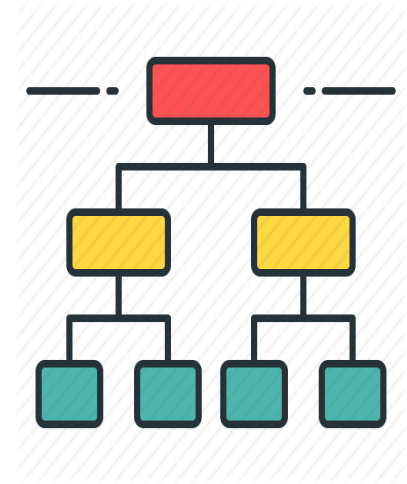
Conclusión

El presente capítulo abordo el análisis por árboles de decisión, desde un enfoque descriptivo y evaluativo del modelo.

Se vieron los principios de los árboles de decisión, los dos tipos de árboles, las etapas del análisis, entre otros.

Los árboles de decisión son utilizados para introducir a las personas en el mundo del machine learning.

No olvidar que este curso toco un enfoque descriptivo de la estadística, sabiendo que falta cubrir el enfoque predictivo.



*The
End*