



Apprentissage Supervisé Classification

Arbres de décision





Content

- 1 Définition et exemple
- 2 Algorithme
- 3 Construction d'arbre de décision
- 4 Conclusion



Arbres de Décision

Motivation

Produire des classifications compréhensibles par l'utilisateur (versus les autres méthodes)



Arbres de Décision

Motivation

Produire des classifications compréhensibles par l'utilisateur (versus les autres méthodes)

Définition

Ensemble de règles de classification basant leur décision sur des tests associés aux attributs, organisés de manière arborescente.



Arbres de Décision

Motivation

Produire des classifications compréhensibles par l'utilisateur (versus les autres méthodes)

Définition

Ensemble de règles de classification basant leur décision sur des tests associés aux attributs, organisés de manière arborescente.

Objectif

L'objectif est de créer un modèle qui prédit les valeurs de la variable cible, en se basant sur un ensemble de séquences de règles de décision déduites à partir des données d'apprentissage.



Arbres de Décision

Motivation

Produire des classifications compréhensibles par l'utilisateur (versus les autres méthodes)

Définition

Ensemble de règles de classification basant leur décision sur des tests associés aux attributs, organisés de manière arborescente.

Objectif

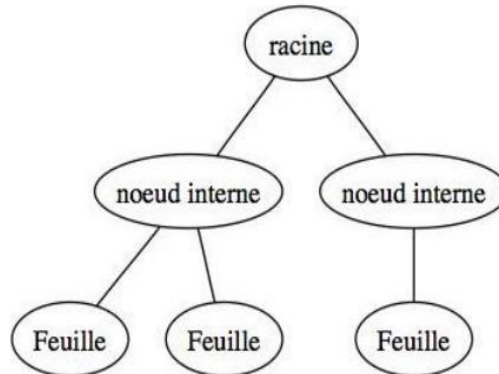
L'objectif est de créer un modèle qui prédit les valeurs de la variable cible, en se basant sur un ensemble de séquences de règles de décision déduites à partir des données d'apprentissage.

Remarque

Un arbre de décision (decision tree) est une structure très utilisée en Machine learning. Son fonctionnement repose sur des heuristiques construites selon des techniques d'apprentissage supervisées.

Arbre de Décision

- Les arbres de décision ont **une structure hiérarchique**.
- Les arbres de décision sont composés de **noeuds** et de **feuilles** reliés par des **branches**.
- Dans leur représentation graphique, la **racine** est placée en haut et **les feuilles** en bas.
- Les **noeuds internes** sont appelés des noeuds de décision.
- Les **noeuds internes** peuvent contenir une ou plusieurs règles (aussi appelées tests ou conditions).





Généralités sur les arbres de décision

- Les valeurs que peut prendre une variable dans un arbre de décision sont appelées **instances ou attributs**.
- Les **noeuds terminaux** contiennent la classe aussi appelée classe à prédire ou variable cible.
- Après sa construction, un arbre de décision peut être traduit sous la forme d'un **ensemble de règles de décision**.
- Lorsque la classe à prédire est une **variable qualitative**, nous avons un arbre de classification.
- Si la classe à prédire est une **variable quantitative**, nous avons un arbre de régression.



Exemple

Arbre de décision

Décider si une institution bancaire peut accorder ou non un prêt à un demandeur de prêt.



Exemple

Arbre de décision

Décider si une institution bancaire peut accorder ou non un prêt à un demandeur de prêt.

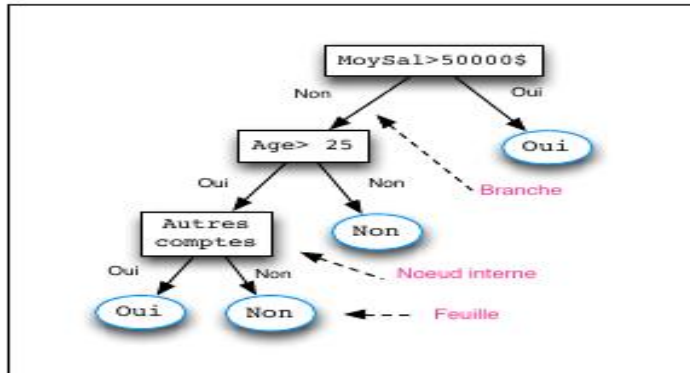
- La décision de la **variable cible Prêt** se base sur les **variables** moyenne du salaire (**MoySal**), âge (**âge**) et possession d'autres comptes (**Autres comptes**).
- Les **attributs des feuilles** de l'arbre: **Prêt**, sont des valeurs booléennes qui correspondent à une classification dans l'ensemble **{Oui, Non}**.

Exemple

Arbre de décision

Décider si une institution bancaire peut accorder ou non un prêt à un demandeur de prêt.

- La décision de la **variable cible Prêt** se base sur les **variables** moyenne du salaire (**MoySal**), âge (**âge**) et possession d'autres comptes (**Autres comptes**).
- Les **attributs des feuilles** de l'arbre: **Prêt**, sont des valeurs booléennes qui correspondent à une classification dans l'ensemble **{Oui, Non}**.





Exemple

Cet arbre de décision engendre les **règles de décisions** suivantes:

- Si **MoySal** > 50000\$ alors **Prêt = Oui**.
- Si **MoySal** < 50000\$ et **Âge** < 25 alors **Prêt = Non**.
- Si **MoySal** < 50000\$ et **Âge** > 25 et **AutresComptes** = **Oui** alors **Prêt = Oui**.
- Si **MoySal** < 50000\$ et **Âge** > 25 et **AutresComptes** = **Non** alors **Prêt = Non**.



Exemple

Cet arbre de décision engendre les **règles de décisions** suivantes:

- Si **MoySal** > 50000\$ alors **Prêt = Oui**.
- Si **MoySal** < 50000\$ et **Âge** < 25 alors **Prêt = Non**.
- Si **MoySal** < 50000\$ et **Âge** > 25 et **AutresComptes** = **Oui** alors **Prêt = Oui**.
- Si **MoySal** < 50000\$ et **Âge** > 25 et **AutresComptes** = **Non** alors **Prêt = Non**.

L'institution est en mesure de décider facilement si elle accorde ou non un prêt à un demandeur.



Algorithmes de construction d'arbres de décision.

Début Algorithme

Input :

- E : ensemble d'échantillons
- L : ensemble d'attributs.

Output:

- un arbre de décision

Faire :

- 1 Si tous les exemples sont de la même classe
 - ▶ Alors créer et retourner une feuille étiquetée par cette classe.
- 2 Sinon
 - 1 Répéter
 - 1 Trouver le meilleur attribut dans l'ensemble L : A
 - 2 Répartir les exemples de E en n sous-ensembles, selon la valeur de cet attribut.
 - 3 Allouer un noeud étiqueté par A
 - 4 Refaire le même processus pour toutes les valeurs de A et faire de ces noeuds
 - 5 les fils du noeud étiqueté par A
 - 2 Jusqu'à indice d'arrêt.

Fin Algorithme



Algorithme pour construire les arbres de décision

Il existe plusieurs algorithmes automatiques pour construire les arbres de décision:

- **ID3 (Iterative Dichotomiser 3):** développé en 1986 par Ross Quinlan.
 - ▶ Il peut être appliqué seulement sur les caractéristiques nominales.
 - ▶ Il est utilisé pour le classement.
- **C4.5:** une extension de ID3 par Ross Quinlan.
 - ▶ Il peut être appliqué sur tous les types de caractéristiques.
 - ▶ Il est utilisé pour le classement.
- **C5.0:** une extension commerciale de C4.5, toujours par Ross Quinlan.
- **CART (Classification and Regression Trees)**

Algorithme pour construire les arbres de décision

Algorithme CART

- Parmi les plus performants et plus répandus (Scikit-learn)
- Accepte tout type de variables
- Utilise le Critère de séparation : Indice de Gini ou l'entropie.
- L'algorithme récursif CART (Classification And Regression Trees) permet la construction d'un arbre de décision par la maximisation de l'indice de Gini.
- Cet algorithme génère des arbres de décision binaires.
- Construit un arbre de décision avec CART se fait en deux étapes:
 - ① Un arbre maximal (indice de Gini)
 - ② Élagage: on construit une suite de sous arbres optimaux élagués de l'arbre maximal (ccp)

Indice de Gini

- L'indice de Gini a été introduit par Breiman en 2001.
- Cet indice mesure l'impureté, qui est un concept très utile dans la construction des arbres de décision.
- La qualité d'un noeud et son pouvoir discriminant peuvent être évalués par son impureté.
- L'indice Gini est donné par la relation suivante:

$$\text{GINI}(T) = 1 - \sum_{j=1}^m \left(\frac{|T_j|}{|T|} \right)^2$$

avec T ensemble de donnée, m nombre de classes et $|T_j|$ désigne le cardinal de la classe j .

- Si l'ensemble de données T est partitionné en deux partitions (T_1, T_2) , alors:

Impureté moyenne pondérée après la division $(X, T) =$

$$\frac{|T_1|}{|T|} \text{GINI}(T_1) + \frac{|T_2|}{|T|} \text{GINI}(T_2)$$

Entropie d'un noeud

Définition générale :

L'entropie est une mesure de l'incertitude ou de l'imprévisibilité d'un système. En apprentissage automatique, elle mesure l'homogénéité d'un ensemble de données par rapport à la variable cible.

Dans le contexte de l'arbre CART :

- S représente l'ensemble des exemples (ou instances) présents dans un noeud.
- Ces exemples sont répartis en m groupes C_1, C_2, \dots, C_m selon les classes de la variable cible y .
- On note $p_i = \frac{|C_i|}{|S|}$ la proportion d'exemples appartenant à la classe C_i .

L'entropie du noeud S est donnée par :

$$H(S) = - \sum_{i=1}^m p_i \log(p_i)$$

Gain d'information sur un attribut catégoriel

- Supposons qu'on souhaite découper le noeud S selon un attribut catégoriel a , qui possède k valeurs distinctes.
- Cela génère k sous-groupes de données : S_1, S_2, \dots, S_k .
- On note $p_i = \frac{|S_i|}{|S|}$ la proportion d'exemples dans chaque sous-groupe.

Le gain d'information associé à a est :

$$GI(S, a) = H(S) - \sum_{i=1}^k p_i H(S_i)$$

- Il mesure la réduction d'entropie après le découpage selon l'attribut a .
- Plus le gain est élevé, plus l'attribut a est pertinent pour la séparation.

Construction d'arbre de décision: Exemple banque

client	M	A	R	E	I
01	moyen	moyen	village	oui	oui
02	élevé	moyen	bourg	non	non
03	faible	âgé	bourg	non	non
04	faible	moyen	bourg	oui	oui
05	moyen	jeune	ville	oui	oui
06	élevé	âgé	ville	oui	non
07	moyen	âgé	ville	oui	non
08	faible	moyen	village	non	non

Construction d'arbre de décision: Exemple banque

client	M	A	R	E	I
01	moyen	moyen	village	oui	oui
02	élevé	moyen	bourg	non	non
03	faible	âgé	bourg	non	non
04	faible	moyen	bourg	oui	oui
05	moyen	jeune	ville	oui	oui
06	élevé	âgé	ville	oui	non
07	moyen	âgé	ville	oui	non
08	faible	moyen	village	non	non

- M: moyenne des montants sur le compte client.
- A: tranche d'âge du client.
- R: localité du résidence du client.
- E: valeur oui si le client a un niveau d'études supérieures.
- I: classe oui correspond à un client qui effectue une consultation de ses comptes bancaires en utilisant internet.

Construction d'arbre de décision: Exemple banque

client	M	A	R	E	I
01	moyen	moyen	village	oui	oui
02	élevé	moyen	bourg	non	non
03	faible	âgé	bourg	non	non
04	faible	moyen	bourg	oui	oui
05	moyen	jeune	ville	oui	oui
06	élevé	âgé	ville	oui	non
07	moyen	âgé	ville	oui	non
08	faible	moyen	village	non	non

Indice de Gini avant séparation ($IG(as)$):

8 clients:

● $I = \text{oui}$: 3 client.

● $I = \text{non}$: 5 client.

$$IG(as) = 1 - \left(\left(\frac{3}{8}\right)^2 + \left(\frac{5}{8}\right)^2\right)$$

$$IG(as) = 0.46878$$

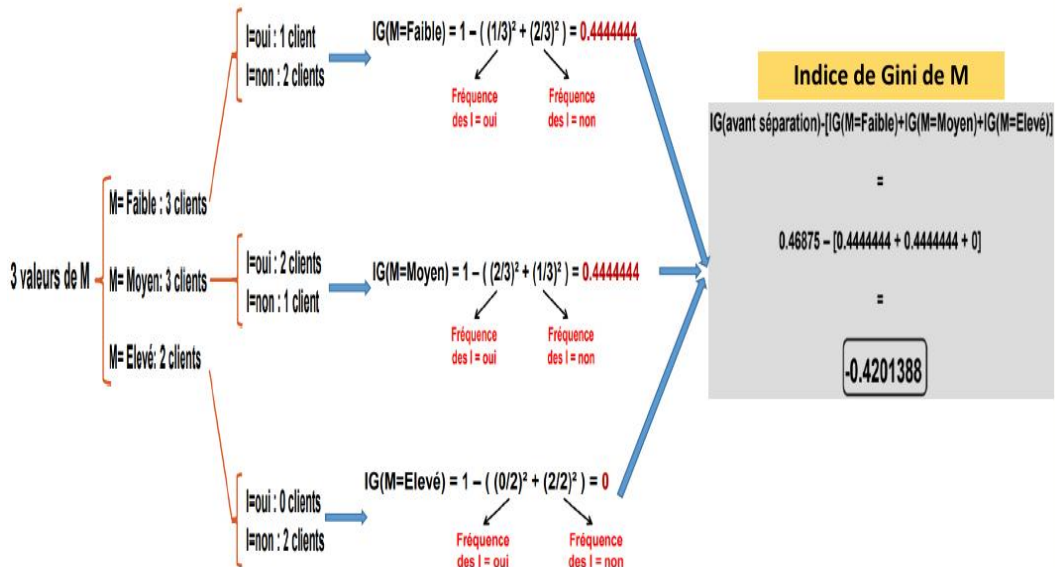
● $\frac{3}{8}$: Fréquence de $I = \text{oui}$

● $\frac{5}{8}$: Fréquence de $I = \text{non}$

- M: moyenne des montants sur le compte client.
- A: tranche d'âge du client.
- R: localité du résidence du client.
- E: valeur oui si le client a un niveau d'études supérieures.
- I: classe oui correspond à un client qui effectue une consultation de ses comptes bancaires en utilisant internet.

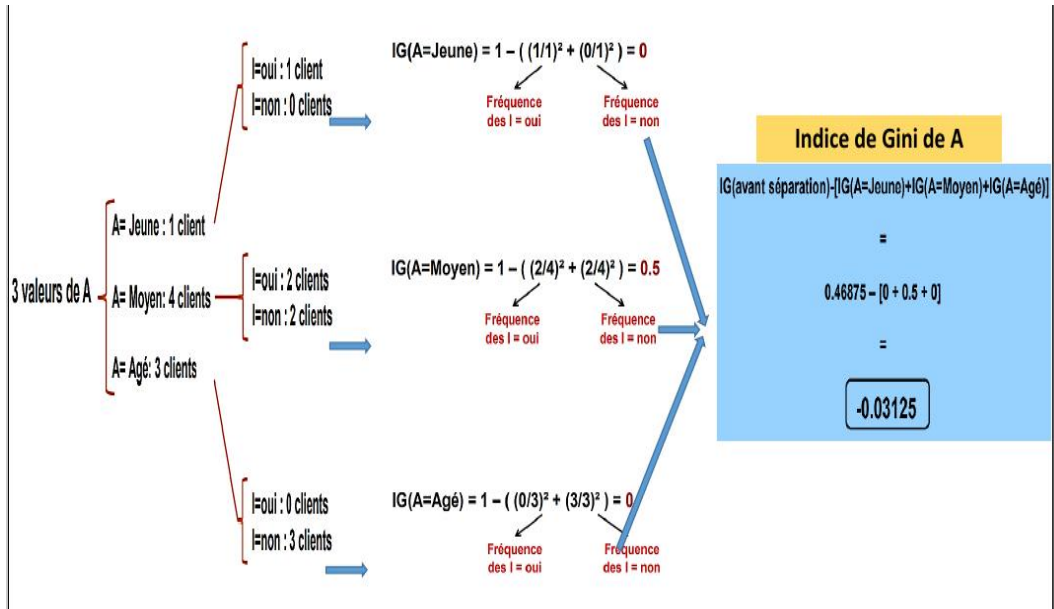
Indice de GINI

L'indice de GINI de la variable M: moyenne des montants sur le compte client.



Indice de GINI

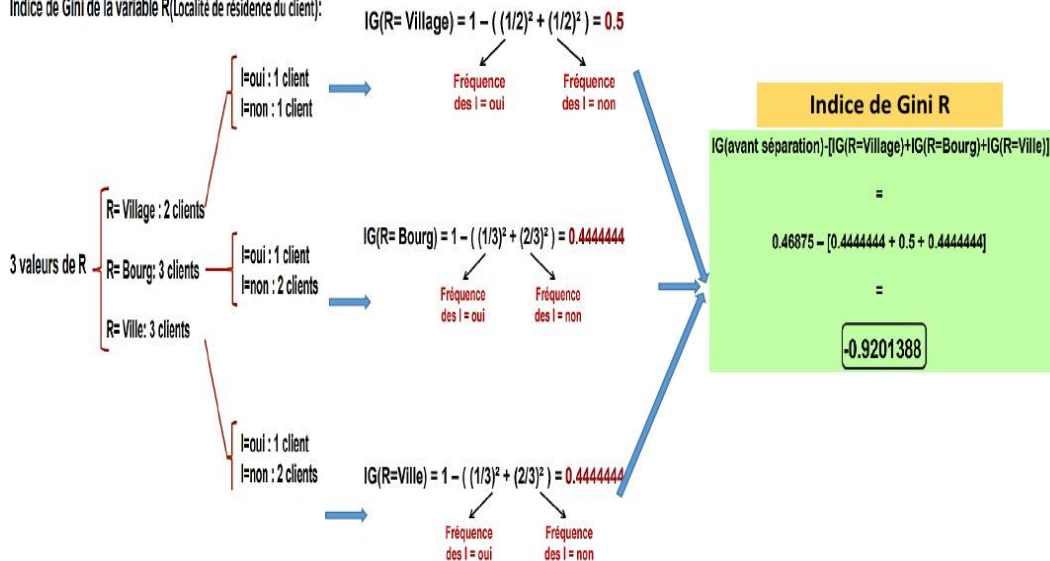
L'indice de GINI de la variable A: tranche d'âge du client.



Indice de GINI

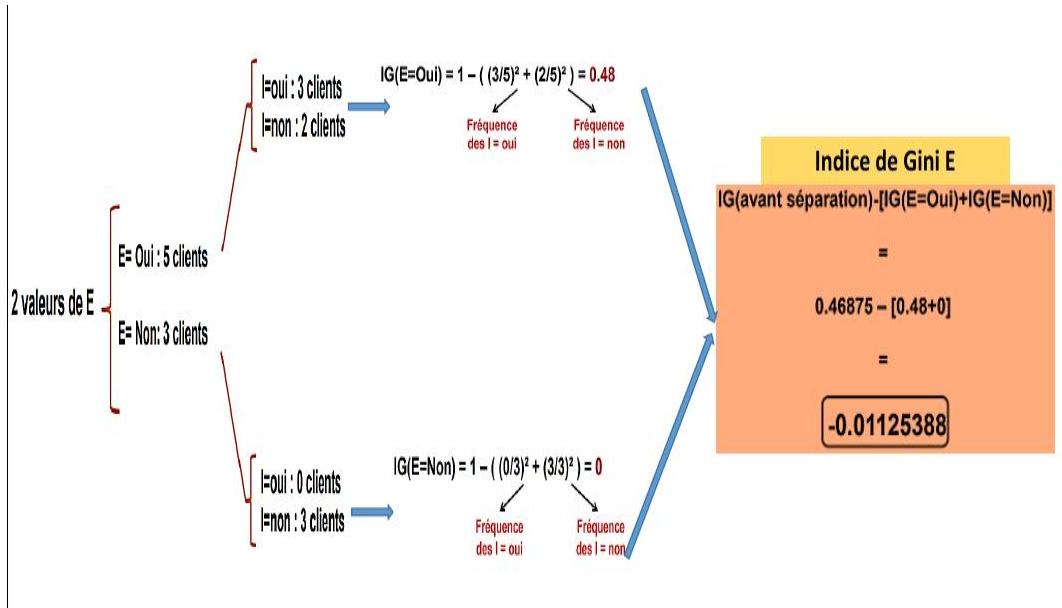
L'indice de GINI de la variable R: localité du résidence du client.

Indice de Gini de la variable R(Localité de résidence du client):



Indice de GINI

L'indice de GINI de la variable E: valeur oui si le client a un niveau d'études supérieures.



Indice de GINI

Indice de GINI de M

$$\begin{aligned} IG(as) &= (IG(M=facile) + IG(M=Moyene) + IG(M=Elevé)) \\ &= 0.46875 - (0.44444 + 0.44444 + 0) \\ IG(M) &= -0.4201388 \end{aligned}$$

Indice de GINI

Indice de GINI de M

$$\begin{aligned} IG(as) - (IG(M=facile) + IG(M=Moyene) + IG(M=Elevé)) \\ = 0.46875 - (0.44444 + 0.44444 + 0) \\ IG(M) = -0.4201388 \end{aligned}$$

Indice de GINI de R

$$\begin{aligned} IG(as) - (IG(R=Village) + IG(R=Bourg) + IG(R=Ville)) \\ = 0.46875 - (0.44444 + 0.5 + 0.4444444) \\ IG(R) = -0.9201388 \end{aligned}$$

Indice de GINI

Indice de GINI de M

$$\begin{aligned} IG(as) - (IG(M=facile) + IG(M=Moyene) + IG(M=Elevé)) \\ = 0.46875 - (0.44444 + 0.44444 + 0) \\ IG(M) = -0.4201388 \end{aligned}$$

Indice de GINI de R

$$\begin{aligned} IG(as) - (IG(R=Village) + IG(R=Bourg) + IG(R=Ville)) \\ = 0.46875 - (0.44444 + 0.5 + 0.4444444) \\ IG(R) = -0.9201388 \end{aligned}$$

Indice de GINI de A

$$\begin{aligned} IG(as) - (IG(A=jeune) + IG(A=Moyen) + IG(A=Agé)) \\ = 0.46875 - (0 + 0.5 + 0) \\ IG(A) = -0.03125 \end{aligned}$$

Indice de GINI

Indice de GINI de M

$$\begin{aligned} IG(as) &- (IG(M=facile) + IG(M=Moyene) + IG(M=Elevé)) \\ &= 0.46875 - (0.44444 + 0.44444 + 0) \\ IG(M) &= -0.4201388 \end{aligned}$$

Indice de GINI de R

$$\begin{aligned} IG(as) &- (IG(R=Village) + IG(R=Bourg) + IG(R=Ville)) \\ &= 0.46875 - (0.44444 + 0.5 + 0.4444444) \\ IG(R) &= -0.9201388 \end{aligned}$$

Indice de GINI de A

$$\begin{aligned} IG(as) &- (IG(A=jeune) + IG(A=Moyen) + IG(A=Agé)) \\ &= 0.46875 - (0 + 0.5 + 0) \\ IG(A) &= -0.03125 \end{aligned}$$

Indice de GINI de E

$$\begin{aligned} IG(as) &- (IG(E=Oui) + IG(E=Non)) \\ &= 0.46875 - (0.48 + 0) \\ IG(M) &= -0.01125388 \end{aligned}$$

PREMIER RESULTAT DE L'INDICE DE GINI

La variable la plus séparatrice est celle qui maximise:

$$IG(as) - (IG(fil_s_1) + IG(fil_s_2) + + IG(fil_s_n))$$



E

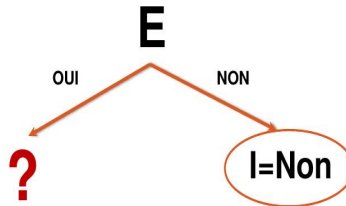
PREMIER RESULTAT DE L'INDICE DE GINI

La variable la plus séparatrice est celle qui maximise:

$$IG(as) - (IG(fil_s_1) + IG(fil_s_2) + + IG(fil_s_n))$$



E





CALCUL DE L'INDICE DE GINI : E=OUI

Indice de Gini avant séparation avec E = Oui :

Indice de Gini avant séparation ($IG(as_1)$):

5 clients:

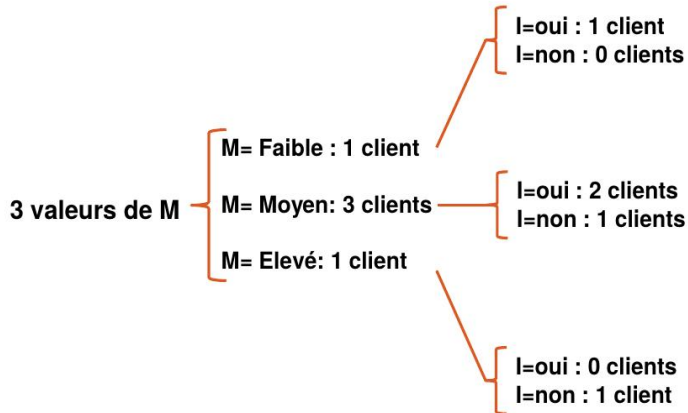
- $I = \text{oui}$: 3 client.
- $I = \text{non}$: 2 client.

$$IG(as_1) = 1 - \left(\left(\frac{3}{5} \right)^2 + \left(\frac{2}{5} \right)^2 \right) = 0.48$$

- $\frac{3}{5}$: Fréquence de $I = \text{oui}$ lorsque $E = \text{oui}$
- $\frac{2}{5}$: Fréquence de $I = \text{non}$ lorsque $E = \text{oui}$

Indice de GINI

Indice de Gini de la variable M (Moyenne des montants sur le compte client) avec E=Oui :



Indice de GINI

Indice de GINI de fils $M = \text{faible}$ & $E = \text{oui}$: 1 client.

● $I = \text{oui}$: 1 client.

● $I = \text{non}$: 0 client.

$$IG(M = \text{faible} \& E = \text{oui}) = 1 - \left(\left(\frac{1}{1}\right)^2 + \left(\frac{0}{1}\right)^2\right) = 0$$

Indice de GINI

Indice de GINI de fils $M = \text{faible}$ & $E = \text{oui}$: 1 client.

● $I = \text{oui}$: 1 client.

● $I = \text{non}$: 0 client.

$$IG(M = \text{faible} \& E = \text{oui}) = 1 - \left(\left(\frac{1}{1}\right)^2 + \left(\frac{0}{1}\right)^2\right) = 0$$

Indice de GINI de fils $M = \text{Moyen}$ & $E = \text{oui}$: 3 client.

● $I = \text{oui}$: 2 clients.

● $I = \text{non}$: 1 client.

$$IG(M = \text{Moyen} \& E = \text{oui}) = 1 - \left(\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2\right) = 0.4444444$$

Indice de GINI

Indice de GINI de fils $M = \text{faible}$ & $E = \text{oui}$: 1 client.

- $I = \text{oui}$: 1 client.
- $I = \text{non}$: 0 client.

$$IG(M = \text{faible} \& E = \text{oui}) = 1 - \left(\left(\frac{1}{1}\right)^2 + \left(\frac{0}{1}\right)^2\right) = 0$$

Indice de GINI de fils $M = \text{Moyen}$ & $E = \text{oui}$: 3 clients.

- $I = \text{oui}$: 2 clients.
- $I = \text{non}$: 1 client.

$$IG(M = \text{Moyen} \& E = \text{oui}) = 1 - \left(\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2\right) = 0.4444444$$

Indice de GINI de fils $M = \text{Elevé}$ & $E = \text{oui}$: 1 client.

- $I = \text{oui}$: 0 client.
- $I = \text{non}$: 1 client.

$$IG(M = \text{Elevé} \& E = \text{oui}) = 1 - \left(\left(\frac{0}{1}\right)^2 + \left(\frac{1}{1}\right)^2\right) = 0$$



Indice de GINI de M avec $E = \text{où}$

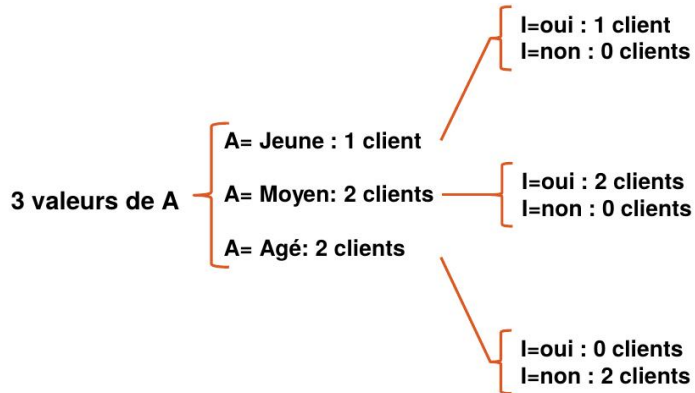
Indice de GINI de M avec $E = \text{où}$

$$\begin{aligned} IG(as_1) - (IG(M=\text{facile}) + IG(M=\text{Moyene}) + IG(M=\text{Elevé})) \\ = 0.48 - (0 + 0.44444 + 0) \end{aligned}$$

$$IG(M \& E = \text{où}) = 0.0355556$$

Indice de GINI

Indice de Gini de la variable A (Tranche d'âge du client) avec $E = \text{Oui}$:



Indice de GINI

Indice de GINI de fils $A = \text{jeune}$ & $E = \text{où}$: 1 client.

● $I = \text{où}$: 1 client.

● $I = \text{non}$: 0 client.

$$IG(A = \text{Jeune} \& E = \text{où}) = 1 - ((\frac{1}{1})^2 + (\frac{0}{1})^2) = 0$$

Indice de GINI

Indice de GINI de fils $A = \text{jeune}$ & $E = \text{oui}$: 1 client.

● $I = \text{oui}$: 1 client.

● $I = \text{non}$: 0 client.

$$IG(A = \text{Jeune} \& E = \text{oui}) = 1 - \left(\left(\frac{1}{1} \right)^2 + \left(\frac{0}{1} \right)^2 \right) = 0$$

Indice de GINI de fils $A = \text{jeune}$ & $E = \text{oui}$: 2 client.

● $I = \text{oui}$: 2 clients.

● $I = \text{non}$: 0 client.

$$IG(A = \text{jeune} \& E = \text{oui}) = 1 - \left(\left(\frac{2}{2} \right)^2 + \left(\frac{0}{2} \right)^2 \right) = 0$$

Indice de GINI

Indice de GINI de fils $A = \text{jeune}$ & $E = \text{oui}$: 1 client.

- $I = \text{oui}$: 1 client.
- $I = \text{non}$: 0 client.

$$IG(A = \text{Jeune} \& E = \text{oui}) = 1 - \left(\left(\frac{1}{1}\right)^2 + \left(\frac{0}{1}\right)^2\right) = 0$$

Indice de GINI de fils $A = \text{jeune}$ & $E = \text{oui}$: 2 client.

- $I = \text{oui}$: 2 clients.
- $I = \text{non}$: 0 client.

$$IG(A = \text{jeune} \& E = \text{oui}) = 1 - \left(\left(\frac{2}{2}\right)^2 + \left(\frac{0}{2}\right)^2\right) = 0$$

Indice de GINI de fils $M = \text{Agé}$ & $E = \text{oui}$: 2 client.

- $I = \text{oui}$: 0 client.
- $I = \text{non}$: 2 clients.

$$IG(M = \text{Agé} \& E = \text{oui}) = 1 - \left(\left(\frac{0}{2}\right)^2 + \left(\frac{2}{2}\right)^2\right) = 0$$

Indice de GINI de A avec $E = \text{oui}$

Indice de GINI de A avec $E = \text{oui}$

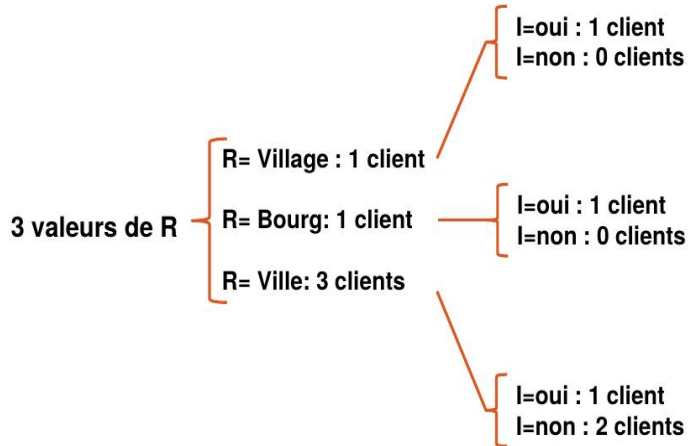
$$IG(as_1) - (IG(A=\text{jeune}) + IG(A=\text{Moyen}) + IG(A=\text{Agé}))$$

$$= 0.48 - (0 + 0 + 0)$$

$$IG(A \& E = \text{oui}) = 0.48$$

Indice de GINI

Indice de Gini de la variable R (Localité de résidence du client) avec $E = \text{Oui}$:



Indice de GINI

Indice de GINI de fils $R = \text{Village}$ & $E = \text{oui}$: 1 client.

● $I = \text{oui}$: 1 client.

● $I = \text{non}$: 0 client.

$$IG(R = \text{Village} \& E = \text{oui}) = 1 - \left(\left(\frac{1}{1} \right)^2 + \left(\frac{0}{1} \right)^2 \right) = 0$$

Indice de GINI

Indice de GINI de fils $R = \text{Village}$ & $E = \text{où}$: 1 client.

- $I = \text{où}$: 1 client.
- $I = \text{non}$: 0 client.

$$IG(R = \text{Village} \& E = \text{où}) = 1 - \left(\left(\frac{1}{1} \right)^2 + \left(\frac{0}{1} \right)^2 \right) = 0$$

Indice de GINI de fils $R = \text{Bourg}$ & $E = \text{où}$: 1 client.

- $I = \text{où}$: 1 client.
- $I = \text{non}$: 0 client.

$$IG(R = \text{Bourg} \& E = \text{où}) = 1 - \left(\left(\frac{1}{1} \right)^2 + \left(\frac{0}{1} \right)^2 \right) = 0$$



Indice de GINI

Indice de GINI de fils $R = \text{Village}$ & $E = \text{oui}$: 1 client.

- $I = \text{oui}$: 1 client.
- $I = \text{non}$: 0 client.

$$IG(R = \text{Village} \& E = \text{oui}) = 1 - \left(\left(\frac{1}{1} \right)^2 + \left(\frac{0}{1} \right)^2 \right) = 0$$

Indice de GINI de fils $R = \text{Bourg}$ & $E = \text{oui}$: 1 client.

- $I = \text{oui}$: 1 client.
- $I = \text{non}$: 0 client.

$$IG(R = \text{Bourg} \& E = \text{oui}) = 1 - \left(\left(\frac{1}{1} \right)^2 + \left(\frac{0}{1} \right)^2 \right) = 0$$

Indice de GINI de fils $R = \text{Ville}$ & $E = \text{oui}$: 3 client.

- $I = \text{oui}$: 1 client.
- $I = \text{non}$: 2 clients.

$$IG(R = \text{Ville} \& E = \text{oui}) = 1 - \left(\left(\frac{1}{3} \right)^2 + \left(\frac{2}{3} \right)^2 \right) = 0.44444444$$



Indice de GINI de R avec $E = \text{oui}$

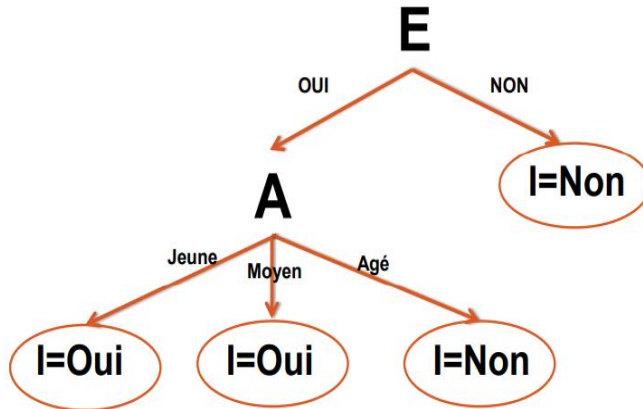
Indice de GINI de R avec $E = \text{oui}$

$$IG(as_1) - (IG(R=\text{Village}) + IG(R=\text{Bourg}) + IG(R=\text{Ville}))$$

$$= 0.48 - (0 + 0 + 0.444444)$$

$$IG(R \& E = \text{oui}) = 0.0355556$$

Arbre de décision



Avantages et Inconvénients

Avantages

- Simples à comprendre et à interpréter.
- Ils peuvent travailler sur des données avec peu de préparation.
- Ils acceptent les données numériques et nominales.
- Ils donnent de bonne performance même si leurs hypothèses sont un peu violées par le modèle réel à partir duquel les données ont été générées.

Avantages et Inconvénients

Avantages

- Simples à comprendre et à interpréter.
- Ils peuvent travailler sur des données avec peu de préparation.
- Ils acceptent les données numériques et nominales.
- Ils donnent de bonne performance même si leurs hypothèses sont un peu violées par le modèle réel à partir duquel les données ont été générées.

Inconvénients

- Ils peuvent être aussi complexes, ils ne généralisent pas bien (overfitting: surapprentissage).
- Ils peuvent être instable à cause des variations des données.
- Ils peuvent être biaisés à la classe dominante.
- Ce n'ai pas garanti de tomber sur l'arbre de décision optimal.



MERCI POUR VOTRE ATTENTION