

# Correction ADD

Préparé par Sarra Zouari

## Exercice1 (examen 2015 session principale)

### 1. Donner un pseudocode du fonctionnement de l'algorithme des arbres de décision. [2pts]

Procédure construire-arbre(X)

**SI** tous les individus I appartiennent à la même modalité de la variable décisionnelle

**ALORS** créer un nœud feuille portant le nom de cette classe : Décision

**SINON**

choisir le meilleur attribut pour créer un nœud // l'attribut qui sépare le mieux, le test associé à ce nœud sépare X en des branches

construire-arbre(Xd), ..., construire-arbre(Xg)

**FIN**

### 2. En déduire la variable la plus décisive par rapport à l'appartenance d'un individu à l'origine orientale. (Donner le calcul complet et la formule utilisée)

[3pts=2pts(formule)+1pt(calcul)]

### 3.

**Indice de Gini :**

1-somme(fréquence de la classe décisionnelle dans le nœud)

- 11 individus : Oriental = Oui : 6, Oriental = Non : 5 donc

### • **Indice de Gini avant séparation au NIVEAU DE LA RACINE :**

$$IG(\text{Oriental}) = 1 - ((5/11)^2 + (6/11)^2) = 0.4958678$$

### • **Indice de Gini de la variable Yeux :**

$$4 \text{ Noir} : 4 \text{ Oui}, 0 \text{ Non } IG(Y=\text{Noir}) = 1 - ((4/4)^2 + (0/4)^2) = 0$$

$$3 \text{ Brun} : 2 \text{ Oui}, 1 \text{ Non } IG(Y=\text{Brun}) = 1 - ((2/3)^2 + (1/3)^2) = 0.8888889$$

$$4 \text{ Bleu} : 0 \text{ Oui}, 4 \text{ Non } IG(Y=\text{Bleu}) = 1 - ((0/4)^2 + (4/4)^2) = 0$$

### • **Indice de Gini de la variable Cheveux :**

$$4 \text{ Noir} : 3 \text{ Oui}, 1 \text{ Non } IG(\text{Ch}=\text{Noir}) = 1 - ((3/4)^2 + (1/4)^2) = 0.75$$

$$4 \text{ Blanc} : 3 \text{ Oui}, 1 \text{ Non } IG(\text{Ch}=\text{Blanc}) = 1 - ((3/4)^2 + (1/4)^2) = 0.75$$

$$3 \text{ Blond} : 0 \text{ Oui}, 3 \text{ Non } IG(\text{Ch}=\text{Blond}) = 1 - ((0/3)^2 + (3/3)^2) = 0$$

### • **Indice de Gini de la variable Taille :**

$$6 \text{ Petit} : 3 \text{ Oui}, 3 \text{ Non } IG(T=\text{Petit}) = 1 - ((3/6)^2 + (3/6)^2) = 0.5$$

$$5 \text{ Grand} : 3 \text{ Oui}, 2 \text{ Non } IG(T=\text{Grand}) = 1 - ((3/5)^2 + (2/5)^2) = 0.48$$

- La variable la plus décisive est celle qui maximise  $IG(\text{avant séparation}) - *IG(\text{fils1}) + \dots + IG(\text{filsn})$ , donc la couleur des Yeux est la variable la plus décisive par rapport à l'appartenance d'un individu à l'origine orientale

## Exercice2 (examen 2016 session principale)

$$1/. IG [\text{age}, 18 ; 35] = 1 - ((2/4)^2 + (2/4)^2) = 0,5$$

$$IG [\text{age}, 36 ; +] = 1 - ((2/6)^2 + (4/6)^2) = 0,44$$

$$\rightarrow IG[\text{age}] = 0,94$$

$$IG[\text{sexe}, H] = 1 - ((4/6)^2 + (2/6)^2) = 0,44$$

$$IG[\text{sexe}, F] = 1 - ((0/4)^2 + (4/4)^2) = 0$$

$$\rightarrow IG[\text{sexe}] = 0,44$$

$$IG [\text{Propriétaire}, N] = 1 - ((1/3)^2 + (2/3)^2) = 0,44$$

$$IG [\text{Propriétaire}, O] = 1 - ((3/7)^2 + (4/7)^2) = 0,49$$

$$\rightarrow IG[\text{Propriétaire}] = 0,93$$

**->La racine de l'arbre est la variable SEXE**

$$IG [\text{age}, 18 ; 35, \text{sexe}=H] = 1 - ((2/3)^2 + (1/3)^2) = 0,44$$

$$IG [\text{age}, 36 ; +, \text{sexe}=H] = 1 - ((2/3)^2 + (1/3)^2) = 0,44$$

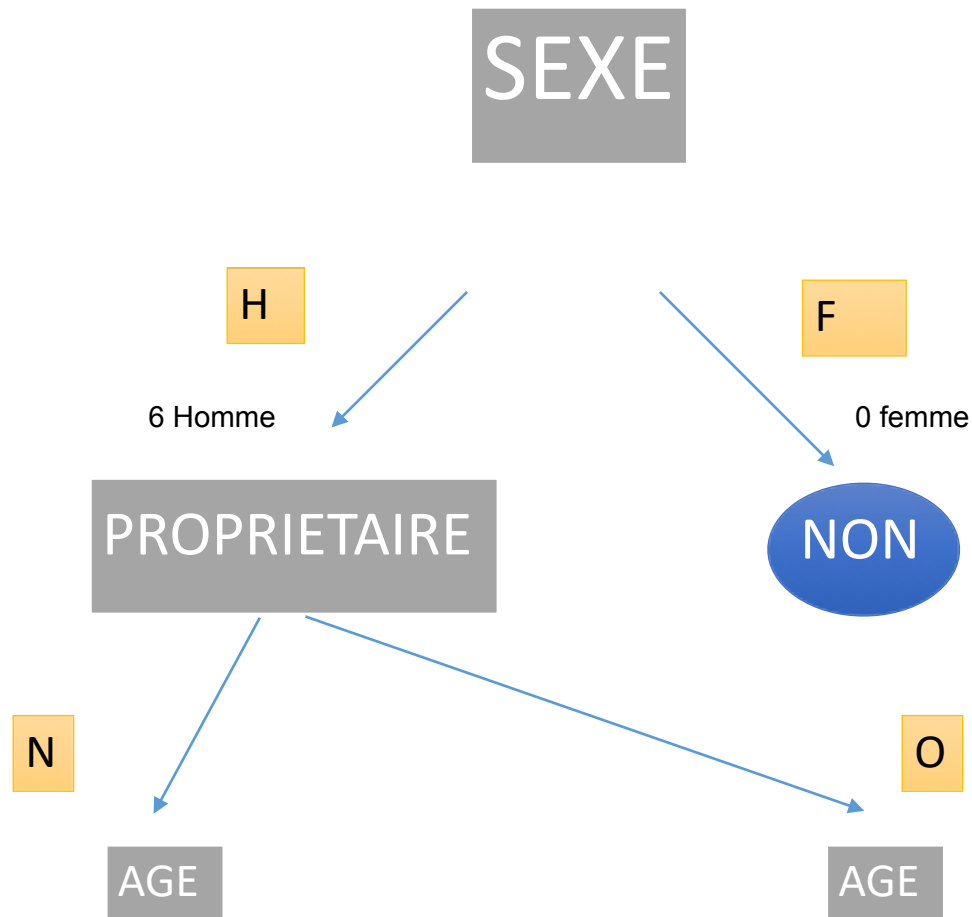
$$\rightarrow IG[\text{age}, \text{sexe}=H] = 0,888$$

$$IG [\text{Propriétaire}, N, \text{sexe}=H] = 1 - ((1/2)^2 + (1/2)^2) = 0,5$$

$$IG [\text{Propriétaire}, O, \text{sexe}=H] = 1 - ((3/4)^2 + (1/4)^2) = 0,375$$

$$\rightarrow IG [\text{Propriétaire}, \text{sexe}=H] = 0,875$$

**->La racine de l'arbre est la variable PROPRIETAIRE**



Sexe H et Prop **N = 2** (1->18:35 = **N** / 1-> 35 plus=**O** )     
 Sexe H et Prop **O = 4** ( 2->18:35= **O** / 2->35 plus=**O et N**)

