

# TP4 Arbre de Décision

BOUZOURENE Madjid

2022-11-08

```
# Creation de la base de donnée
Meteo <- c('soleil','soleil','soleil','soleil','nuages','nuages','nuages','nuages','soleil')
Amis <- c('présents','absents','présents','absents','absents','présents',
          'absents','présents','absents')
Vent <- c('faible','fort','fort','faible','faible','fort','fort','faible','faible')
Jour <- c('week-end','semaine','semaine','semaine','week-end','week-end',
          'semaine','week-end','week-end')
Décision <- c('oui','non','non','oui','non','non','non','oui','non')

Table <- data.frame(Meteo, Amis, Vent, Jour, Décision)
```

## Affichage des données Table:

```
print(Table)
```

```
##      Meteo      Amis      Vent      Jour Décision
## 1 soleil présents faible week-end      oui
## 2 soleil absents   fort  semaine      non
## 3 soleil présents   fort  semaine      non
## 4 soleil absents faible semaine      oui
## 5 nuages absents faible week-end      non
## 6 nuages présents   fort week-end      non
## 7 nuages absents   fort  semaine      non
## 8 nuages présents faible week-end      oui
## 9 soleil absents faible week-end      non
```

```
# Pour l'arbre de décision
library(rpart)
# Pour la représentation de l'arbre de décision
library(rpart.plot)
```

```
str(Table)
```

```
## 'data.frame':   9 obs. of  5 variables:
## $ Meteo      : chr  "soleil" "soleil" "soleil" "soleil" ...
## $ Amis       : chr  "présents" "absents" "présents" "absents" ...
## $ Vent       : chr  "faible" "fort" "fort" "faible" ...
## $ Jour       : chr  "week-end" "semaine" "semaine" "semaine" ...
## $ Décision: chr  "oui" "non" "non" "oui" ...
```

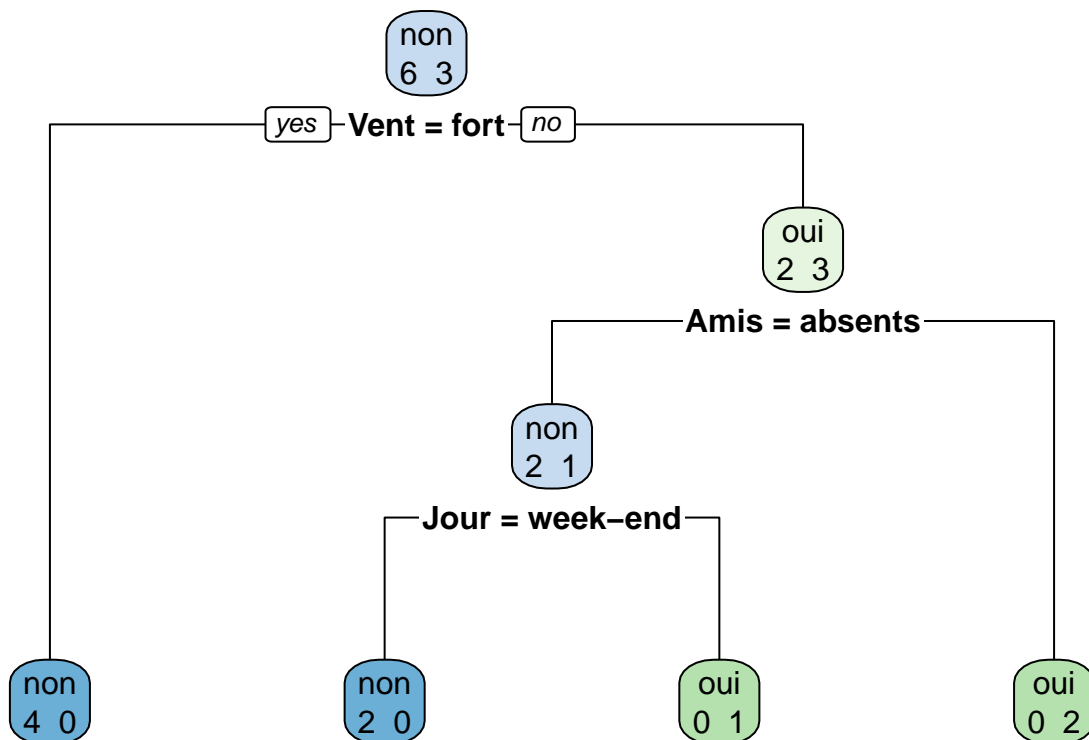
```
arbre <- rpart(formula = Décision ~ ., data = Table, method = "class", minsplit=2, cp=0)
```

```
print(arbre)
```

```
## n= 9
##
## node), split, n, loss, yval, (yprob)
##      * denotes terminal node
##
## 1) root 9 3 non (0.6666667 0.3333333)
##    2) Vent=fort 4 0 non (1.0000000 0.0000000) *
##    3) Vent=faible 5 2 oui (0.4000000 0.6000000)
##      6) Amis=absents 3 1 non (0.6666667 0.3333333)
##        12) Jour=week-end 2 0 non (1.0000000 0.0000000) *
##          13) Jour=semaine 1 0 oui (0.0000000 1.0000000) *
##        7) Amis=présents 2 0 oui (0.0000000 1.0000000) *
```

## Représentation de l'arbre

```
rpart.plot(arbre,extra = 1)
```



## Creation de l'arbre elagué :

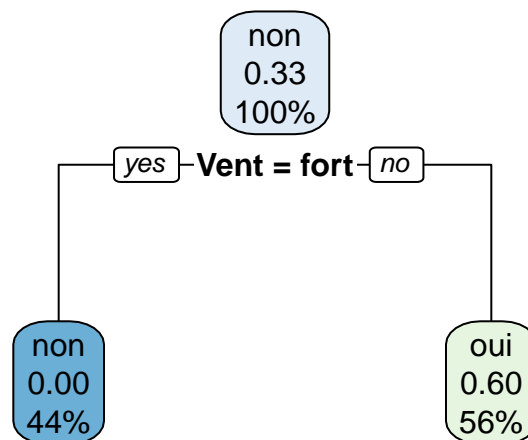
```
arbreElagué <- rpart(formula = Décision ~ ., data = Table, minsplit = 2, minbucket=3)
print(arbreElagué)
```

```
## n= 9
##
```

```
## node), split, n, loss, yval, (yprob)
##      * denotes terminal node
##
## 1) root 9 3 non (0.6666667 0.3333333)
##    2) Vent=fort 4 0 non (1.0000000 0.0000000) *
##    3) Vent=faible 5 2 oui (0.4000000 0.6000000) *
```

## Représentation de l'arbre elagué :

```
rpart.plot(arbreElagué)
```



## Données de deuxieme tableau pour le test:

```
Meteo <- c("soleil","nuages","soleil","soleil","nuages")
Amis <- c("présents","absents","absents","absents","présents")
Vent <- c("fort","faible","faible","faible","faible")
Jour <- c("week-end","semaine","week-end","semaine","semaine")
Décision <- c("non","non","oui","non","oui")
```

## Affichage de tableau test

```
Table_test <- data.frame(Meteo,Amis,Vent,Jour,Décision)
print(Table_test)
```

```
##      Meteo      Amis   Vent      Jour Décision
## 1 soleil présents   fort week-end      non
## 2 nuages absents faible semaine      non
## 3 soleil absents faible week-end      oui
## 4 soleil absents faible semaine      non
## 5 nuages présents faible semaine      oui

Table_predict <- predict(arbreElagué, newdata = Table_test, type = "class")
Table_predict

##      1      2      3      4      5
## non oui oui oui oui
## Levels: non oui
```

## Matrice de confusion :

```
matrix_c<-table(Table_test$Décision,Table_predict)
matrix_c

##      Table_predict
##      non oui
## non      1      2
## oui      0      2

erreur<-1.0-(matrix_c[1,1]+matrix_c[2,2])/sum(matrix_c)
cat("Erreur", erreur)

## Erreur 0.4
```