

# **TP02 Fouille de Données**

Guettouche Islem

Master 1 Génie de l'informatique logicielle (G.I.L)

20 Mai 2017

# Table des matières

Introduction : .....3

Question1 .....3

Question 2 .....3

Question3 .....3

Question4 .....3

Question5 .....4

Question6 .....4

Question7 .....4

Question8 .....4

Question9 .....4

Question10 .....4

Question11 .....4

Question12 .....5

Question13 .....5

Question14 .....5

Question15 .....6

Question16 .....6

## Introduction :

Le système Weka est développé à l'Université de Waikato, Nouvelle Zélande. Il permet de prétraiter des données, de les analyser à l'aide d'une méthode de data mining et d'afficher le modèle résultant et ses performances. Weka est entièrement développé en Java. Il est diffusé sous licence publique GNU. Les ressources nécessaires à l'installation et à l'utilisation du système sont disponibles à l'adresse suivante <http://www.cs.waikato.ac.nz/ml/weka/>. Dans Weka, chaque méthode de transformation, de sélection d'attributs, d'apprentissage, de clustering ou de découverte d'associations est implémentée par une classe Java. La documentation des classes est accessible à partir du fichier packages.html. Weka traite des données au format ARFF (Attribute Relation Format File) ou CSV (Comma Separated Values).

### Question1

**Les intervalles où il n'existe aucune instance de jeux de données sont:**

✚ L'intervalle [72.4-74.8]

✚ L'intervalle [76.6-78.7]

### Question 2

**Les différents intervalles générés sont :**

✚ '(-inf-69.25]' effectifs: 4

✚ '(69.25-74.5]' effectifs: 4

✚ '(74.5-79.75]' effectifs: 2

✚ '(79.5-inf]' effectifs: 4

### Question3

**Les différents intervalles générés sont :**

✚ '(-inf-69.5]' effectifs: 4

✚ '(69.5-73.5]' effectifs: 4

✚ '(73.5-80.5]' effectifs: 3

✚ '(80.5-inf]' effectifs: 3

### Question4





✚ '(-inf-70.5]' effectifs: 5

✚ '(70.5-77.5]' effectifs: 5

✚ '(77.5-80.5]' effectifs: 4



### Question5

Les valeurs numériques correspondant aux 4 premiers histogrammes sont:

-  65
-  70
-  75
-  80

### Question6

Le nom de ces valeurs nominales est:

-  neg\_5-10 occurrences 7
-  pos\_5-10 occurrences 7




### Question7

On observe sur l'histogramme de température qu'on peut jouer que si la température est faible soit "cool" ou "medium" alors que lorsque c'est hot on ne peut pas jouer.

On observe sur l'histogramme de l'humidité qu'on peut jouer que quand l'humidité est faible "low-medium" alors qu'on ne peut pas jouer si elle est sur "high".

### Question8

Les combinaisons de valeurs de température, et humidité pour que play = yes sont

-  cool-high
-  medium-low\_medium
-  hot-low\_medium

### Question9

- ❖ outlook pour que play=yes il faut que outlook=overcast et play=no il faut que outlook=sunny
- ❖ temperature pour que play=yes il faut que temperature=cool ou medium et play=no il faut que temperature=hot
- ❖ humidity pour que play=yes il faut que humidity=low\_medium et play=no il faut que humidity=high
- ❖ windy pour que play=yes il faut que windy=false et play=no il faut que windy=true

### Question10

Rien ne se passe on peut pas appliquer l'algorithme apriori.

### Question11

outlook=overcast 4 ==> play=yes 4 conf:(1)

outlook=rainy windy=FALSE 3 ==> play=yes 3 conf:(1)

temperature=cool windy=FALSE 3 ==> play=yes 3 conf:(1)

### Question12

temperature=hot play=no 2 ==> outlook=sunny 2    conf:(1)

outlook=sunny humidity=high 2 ==> play=no 2    conf:(1)

### Question13

If outlook = overcast then yes

If temperature = cool

    and windy = FALSE then yes

If temperature = temperate

    and humidity = low then yes

If humidity = medium

    and outlook = rainy then yes

If outlook = sunny

    and temperature = hot then no

If humidity = high

    and outlook = sunny then no

If outlook = rainy

    and windy = TRUE then no

### Question14

**temperature = temperate and humidity = low**

N'apparait pas dans les règles d'apriori.

### Question15

outlook=overcast 4 ==> play=yes 4    conf:(1)

outlook=rainy windy=FALSE 3 ==> play=yes 3    conf:(1)

temperature=cool windy=FALSE 3 ==> play=yes 3    conf:(1)

outlook=sunny humidity=low 2 ==> play=yes 2    conf:(1)

outlook=overcast temperature=hot 2 ==> play=yes 2    conf:(1)

outlook=sunny temperature=hot 2 ==> play=no 2    conf:(1)

outlook=sunny humidity=high 2 ==> play=no 2    conf:(1)

outlook=rainy windy=TRUE 2 ==> play=no 2    conf:(1)

### Question16

outlook=overcast 4 ==> play=yes 4    conf:(1)

outlook=rainy windy=FALSE 3 ==> play=yes 3    conf:(1)

temperature=cool windy=FALSE 3 ==> play=yes 3    conf:(1)

outlook=sunny humidity=low 2 ==> play=yes 2    conf:(1)

outlook=overcast temperature=hot 2 ==> play=yes 2    conf:(1)

outlook=sunny temperature=hot 2 ==> play=no 2    conf:(1)

outlook=sunny humidity=high 2 ==> play=no 2    conf:(1)

outlook=rainy windy=TRUE 2 ==> play=no 2    conf:(1)