

# Règles d'associations

*Azat Aleksanyan, Lucas Chabeau, Matthieu François et Etienne Hamard*

*11 février 2019*

## Contents

<b>Objectif</b>	<b>1</b>
<b>Fonctionnement des règles d'associations</b>	<b>1</b>
Support . . . . .	1
Confiance (confidence) . . . . .	1
Lift . . . . .	2
<b>Règles d'associations que nous avons trouvées</b>	<b>2</b>
Avec toutes les questions . . . . .	2
Sans les questions 8 . . . . .	3

## Objectif

L'objectif est ici de trouver de voir au travers des règles d'associations si des réponses à une question ou une pédagogie implique (avec assez de confiance) la réponse à une autre question (ou l'appartenance à une pédagogie). Ainsi nous avons un autre moyen de rechercher de l'information dans nos données.

## Fonctionnement des règles d'associations

Nous présenterons les règles sous la forme suivante : " $A \Rightarrow B$ " ce qui signifie "A implique B". Par exemple prenons la règle :  $\{T22, T61\} \Rightarrow \{T21\}$  : ça implique qu'un élève qui a bien répondu aux questions T22 et T61 aura tendance à avoir bien répondu à la question T21 (normal puisque 100% des élèves qui ont bien répondu à la question T22 ont répondu correctement à la question T21). Nous présenterons trois indicateurs pour quantifier la qualité et la force d'une règle d'association : - Le support - La confiance - Le lift

### Support

Le support est la proportion d'apparition de la règle sur l'ensemble de nos individus. (Par exemple, la règle  $\{T22, T61\} \Rightarrow \{T21\}$  a un support de 0.5192, c'est à dire que 51.92% des élèves interrogés ont réussi les questions T22 et T61 puis ont réussi la question T21).

### Confiance (confidence)

La confiance est un indicateur de fiabilité de la règle. Elle est la proportion des cas où la règle est observée parmi les individus qui suivent les antécédants de la règle. Pour expliquer de manière plus claire, revoyons la règle sous la forme  $A \Rightarrow B$  ou A est l'antécédant et B la conséquence de la règle. La confiance de la règle

sera la proportion des individus pour qui la règle  $A \Rightarrow B$  parmi ceux qui comportent les caractéristiques de l'ensemble A. (Par exemple, la règle  $\{T22, T61\} \Rightarrow \{T21\}$  a une confiance de 1, c'est à dire que 100% des élèves ayant réussi les questions T22 et T61 ont réussi la question T21).

## Lift

Le lift indique la force de la règle d'association. C'est le résultat du rapport entre le support de la règle et le produit des supports de l'antécédant et de la conséquence de la règle. Autrement dit, c'est le rapport entre la probabilité que la règle soit vraie si les antécédents sont respectés et la probabilité d'observer la règle sans forcément suivre la règle ( $lift = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A) * \mathbb{P}(B)}$ ) (Par exemple, la règle  $\{T22, T61\} \Rightarrow \{T21\}$  a un lift de 1.155, c'est à dire que si on demande à un élève qui a réussi à répondre aux questions T22 et T61, on a 15% de chances en plus qu'il ait réussi à répondre à la question T21 que si nous avions demandé à n'importe quel élève de la classe).

## Règles d'associations que nous avons trouvées

```
## Loading required package: arules
```

```
## Loading required package: Matrix
```

```
##
```

```
## Attaching package: 'arules'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      abbreviate, write
```

Certaines règles qui vont apparaître peuvent paraître triviales (du genre  $T22 \Rightarrow T21$ ) mais en faisant abstraction de ces “règles” qui ne nous servent à rien, nous pouvons trouver quelques règles qui pourraient peut-être vous intéresser.

## Avec toutes les questions

Voici toutes les règles que nous avons trouvées (avec un support minimum de 0.7 et une confiance minimum de 0.9)

```
## Apriori
```

```
##
```

```
## Parameter specification:
```

```
## confidence minval smax arem aval originalSupport maxtime support minlen
```

```
##      0.9      0.1      1 none FALSE                TRUE          5      0.7      2
```

```
## maxlen target  ext
```

```
##      10 rules FALSE
```

```
##
```

```
## Algorithmic control:
```

```
## filter tree heap memopt load sort verbose
```

```
##      0.1 TRUE TRUE  FALSE TRUE      2      TRUE
```

```
##
```

```
## Absolute minimum support count: 109
```

```
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[30 item(s), 156 transaction(s)] done [0.00s].
## sorting and recoding items ... [6 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 done [0.00s].
## writing ... [19 rule(s)] done [0.00s].
## creating S4 object ... done [0.00s].
```

	lhs	rhs	support	confidence	lift	count
## [1]	{T83}	=> {T21}	0.7115385	0.9327731	1.077871	111
## [2]	{T61,T83}	=> {T81}	0.7179487	0.9824561	1.071770	112
## [3]	{T83}	=> {T81}	0.7435897	0.9747899	1.063407	116
## [4]	{T61,T81}	=> {T21}	0.8012821	0.9124088	1.054339	125
## [5]	{T82}	=> {T81}	0.7115385	0.9652174	1.052964	111
## [6]	{T61,T62}	=> {T81}	0.7051282	0.9649123	1.052632	110
## [7]	{T62,T81}	=> {T61}	0.7051282	0.9909910	1.051664	110
## [8]	{T21}	=> {T61}	0.8525641	0.9851852	1.045503	133
## [9]	{T61}	=> {T21}	0.8525641	0.9047619	1.045503	133
## [10]	{T21,T81}	=> {T61}	0.8012821	0.9842520	1.044512	125
## [11]	{T62}	=> {T81}	0.7115385	0.9487179	1.034965	111
## [12]	{T62}	=> {T61}	0.7307692	0.9743590	1.034014	114
## [13]	{T21}	=> {T81}	0.8141026	0.9407407	1.026263	127
## [14]	{T21,T61}	=> {T81}	0.8012821	0.9398496	1.025290	125
## [15]	{T81,T83}	=> {T61}	0.7179487	0.9655172	1.024631	112
## [16]	{T82}	=> {T61}	0.7115385	0.9652174	1.024312	111
## [17]	{T81}	=> {T61}	0.8782051	0.9580420	1.016698	137
## [18]	{T61}	=> {T81}	0.8782051	0.9319728	1.016698	137
## [19]	{T83}	=> {T61}	0.7307692	0.9579832	1.016635	114

## Sans les questions 8

Comme nous avons beaucoup de règles qui impliquent la question 8, nous avons choisi de l'écarter pour voir les autres règles. Les règles restent les mêmes, ça nous permet juste d'afficher les autres règles avec un support moindre (sans avoir trop de règles, pour que ça reste lisible.) Nous avons maintenant un support minimum de 0.5 et une confiance minimum de 0.9

```
## Apriori
##
## Parameter specification:
## confidence minval smax arem aval originalSupport maxtime support minlen
##          0.9    0.1    1 none FALSE                TRUE         5     0.5     2
## maxlen target  ext
##          10  rules FALSE
##
## Algorithmic control:
## filter tree heap memopt load sort verbose
##       0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 78
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[21 item(s), 156 transaction(s)] done [0.00s].
```

```

## sorting and recoding items ... [6 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 done [0.00s].
## writing ... [15 rule(s)] done [0.00s].
## creating S4 object ... done [0.00s].

```

	lhs	rhs	support	confidence	lift	count
## [1]	{T22}	=> {T21}	0.5256410	1.0000000	1.155556	82
## [2]	{T22,T61}	=> {T21}	0.5192308	1.0000000	1.155556	81
## [3]	{T31,T61}	=> {T21}	0.5576923	0.9560440	1.104762	87
## [4]	{T31}	=> {T21}	0.5641026	0.9462366	1.093429	88
## [5]	{T21,T62}	=> {T61}	0.6538462	1.0000000	1.061224	102
## [6]	{T21,T31}	=> {T61}	0.5576923	0.9886364	1.049165	87
## [7]	{p1,T21}	=> {T61}	0.5256410	0.9879518	1.048439	82
## [8]	{T22}	=> {T61}	0.5192308	0.9878049	1.048283	81
## [9]	{T21,T22}	=> {T61}	0.5192308	0.9878049	1.048283	81
## [10]	{T21}	=> {T61}	0.8525641	0.9851852	1.045503	133
## [11]	{T61}	=> {T21}	0.8525641	0.9047619	1.045503	133
## [12]	{p1,T61}	=> {T21}	0.5256410	0.9010989	1.041270	82
## [13]	{T31}	=> {T61}	0.5833333	0.9784946	1.038402	91
## [14]	{T62}	=> {T61}	0.7307692	0.9743590	1.034014	114
## [15]	{p1}	=> {T61}	0.5833333	0.9479167	1.005952	91