ASI-MGK: IMPLICATIVE STATISTICAL ANALYSIS TOOL BASED ON M_{GK}

Bruno Bakys RALAHADY and André TOTOHASINA

Department of Mathematics and Computer Science
Ecole Normale Supérieure Pour l'Enseignement Technique (ENSET), University of Antsiranana BP.0 – Madagascar.

ABSTRACT

We propose in this paper to present a statistical analysis tool integrating some well-known algorithms for extracting optimized patterns and algorithm for extracting rules based on about forty existing interest measures in the literature, including the pruning threshold. is based on a critical value of the measure of interest M_{GK}. The tool is equipped with two results presentation windows, one in a table that can be exported in LaTeX and the other interactive dynamic display window for an implicit graph.

KEYWORDS

Association rule mining, itemset mining, ISA, implicative graphs.

1. INTRODUCTION

Association Rule Extraction is a very important task in data mining, usually consisting of four iterative and interactive steps, namely: data processing; generation of patterns and set of candidate patterns, extraction of valid rules and visualization and interpretation of results. In the two middle steps, many existing algorithms and the properties of these algorithms have given rise to a large number of works. In the first processing step, currently there are several data formats: raw text data, csv data, araff data. These data can be coded in binary or quantitative or qualitative data and can be represented in sequences or transactions.

DOI: 10.5121/ijcst.2019.01201

According to the coding and the representation of the data, the methodologies used in the second extraction step diverges; hence the existence of several pattern generation algorithms or set of patterns in the literature, each of them have the ambition to reduce the processing time and memory capacity used or even the access time to the Data Base .

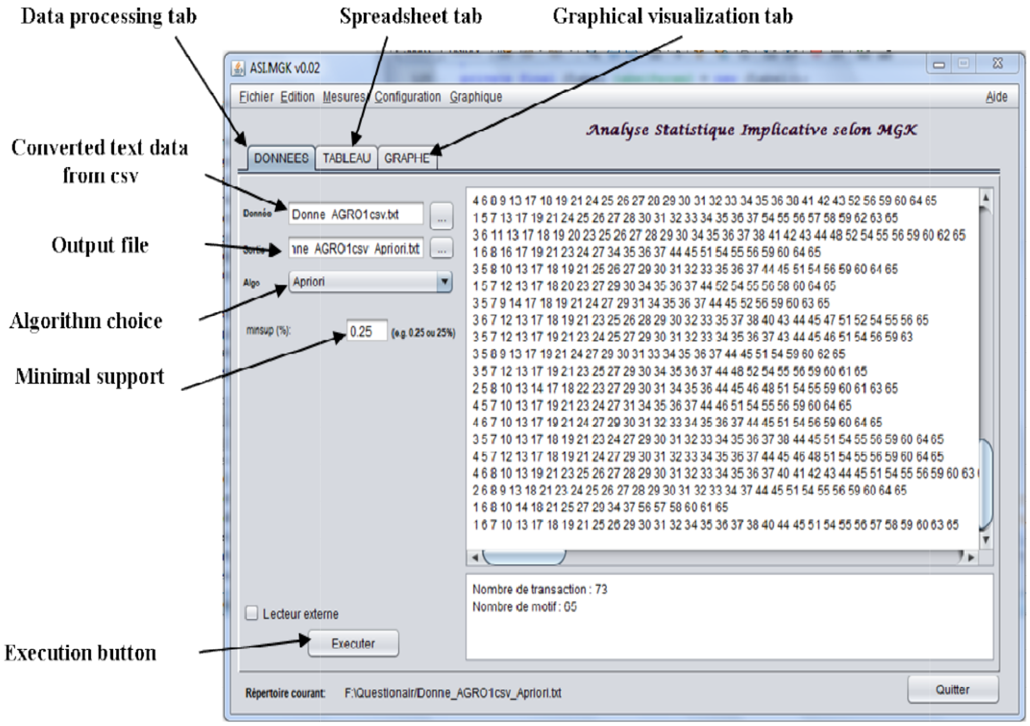
The third step, is also a very important steps. Several research centers are interested in this, it is the extraction of valid rules where the validation of the rules, often conditioned by the discriminant, the relevant, the robustness and the non-redundant. The existence of several measures of interest in the literature and the use of overdraft and optimization of rules extraction algorithm explain the divergence of ideas and the evolution of research in this field. We refer the reader to the syntheses proposed by Gras and al. (2004), Geng and Hamilton (2006), Lallich and al. (2007), Lenca and al. (2007), Geng and Hamilton (2007) and Suzuki (2008).

Several tools already implement some of these steps and their algorithms, but most of them rely on the support-trust or support-lift, or sometimes support-conviction pair. The pruning threshold is a value arbitrarily chosen as $minconf$ and $minlif$ for trust and lift respectively.

In ASI-MGK, we implemented the more famous algorithms and introduced some 40 measures of interest and optimized the chi-square threshold based on the M_{GK} interest measure.

2 ASI-MGK DESCRIPTION

Software allowing the extraction of such rules and propose to use the lift (IBM, XLSTAT, ORANGE, WEKA) in addition to support and trust. The tool Felix (Lehn, 2000) integrates the intensity of implication and its entropic version. The Herbs experimentation platform (Vaillant, 2002) integrates 20 quality measures *(support, confidence, linear correlation coefficient, centered confidence, conviction, Piatetsky-Shapiro, Loevinger, Sebag-Schoenauer information gain, lift, Laplace, less contradiction, odds multiplier, rate of examples and counterexamples, Cohen quality index, Zhang, implication index, intensity of involvement, intensity of entropic implication and discriminating probabilistic index).*



Donnée: Donne AGRO1csv.txt

Sortie: ıne AGRO1csv Apriori.txt

Algo: Apriori

minsup (%): 0.25 (e.g. 0.25 ou 25%)

□ Lecteur externe

```
4 6 8 9 13 17 18 19 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 41 42 43 52 56 59 60 64 65
1 5 7 13 17 19 21 24 25 26 27 28 30 31 32 33 34 35 36 37 54 55 56 57 58 59 62 63 65
3 6 11 13 17 18 19 20 23 25 26 27 28 29 30 34 35 36 37 38 41 42 43 44 48 52 54 55 56 59 60 62 65
1 6 8 16 17 19 21 23 24 27 34 35 36 37 44 45 51 54 55 56 59 60 64 65
3 5 8 10 13 17 18 19 21 25 26 27 29 30 31 32 33 35 36 37 44 45 51 54 56 59 60 64 65
1 5 7 12 13 17 18 20 23 27 29 30 34 35 36 37 44 52 54 55 56 58 60 64 65
3 5 7 9 14 17 18 19 21 24 27 29 31 34 35 36 37 44 45 52 56 59 60 63 65
3 6 7 12 13 17 18 19 21 23 25 26 28 29 30 32 33 35 37 38 40 43 44 45 47 51 52 54 55 56 65
3 5 7 12 13 17 19 21 23 24 25 27 29 30 31 32 33 34 35 36 37 43 44 45 46 51 54 56 59 63
3 5 8 9 13 17 19 21 24 27 29 30 31 33 34 35 36 37 44 45 51 54 59 60 62 65
3 5 7 12 13 17 19 21 23 25 27 29 30 34 35 36 37 44 48 52 54 55 56 59 60 61 65
2 5 8 10 13 14 17 18 22 23 27 29 30 31 34 35 36 44 45 46 48 51 54 55 59 60 61 63 65
4 5 7 10 13 17 19 21 23 24 27 31 34 35 36 37 44 46 51 54 55 56 59 60 64 65
4 6 7 10 13 17 19 21 24 27 29 30 31 32 33 34 35 36 37 44 45 51 54 56 59 60 64 65
3 5 7 10 13 17 18 19 21 23 24 27 29 30 31 32 33 34 35 36 37 38 44 45 51 54 55 56 59 60 64 65
4 5 7 12 13 17 18 19 21 24 27 29 30 31 32 33 34 35 36 37 44 45 46 48 51 54 55 56 59 60 64 65
4 6 8 10 13 19 21 23 25 26 27 28 29 30 31 32 33 34 35 36 37 40 41 42 43 44 45 51 54 55 56 59 60 63
2 6 8 9 13 18 21 23 24 25 26 27 28 29 30 31 32 33 34 37 44 45 51 54 55 56 59 60 64 65
1 6 8 10 14 18 21 25 27 29 34 37 56 57 58 60 61 65
1 6 7 10 13 17 18 19 21 25 26 29 30 31 32 34 35 36 37 38 40 44 45 51 54 55 56 57 58 59 60 63 65
```

```
Nombre de transaction : 73
Nombre de motif : 65
```

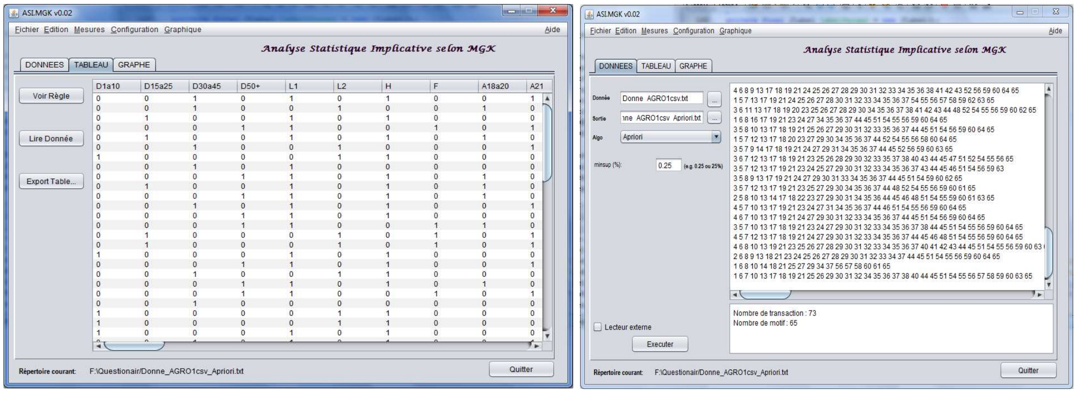
Répertoire courant: F:\Questionair/Donne_AGRO1csv_Apriori.txt

Figure 1: ASI-MGK window

The ASI-MGK software is a Java application (developed in Java programming language) designed to help young researchers in data analysis. More specifically, this software constitutes three tabs; The first two menus of the ASI-MGK guide the user through three stages of the association rule extraction process, using a graphical user interface: reading or data conversion, generating candidate items, or directly extracting rules. The last tab concerns the operations of analysis and visualization of the results.

2.1. Data reading and transformation

This is the first phase that was generally called the preprocessing phase. It consists of selecting from the database the data (attributes and objects) that are useful for extracting association rules and transforming them into data. an extraction context.



D1a10	D15a25	D30a45	D50+	L1	L2	H	F	A18a20	A21
0	0	1	0	1	0	1	0	0	1
0	0	1	0	0	1	0	0	1	0
0	1	0	0	1	0	1	0	0	0
0	0	0	1	1	0	0	1	0	1
0	1	0	0	1	0	1	0	1	0
0	0	1	0	0	1	0	0	0	1
1	0	0	0	0	1	1	0	0	0
0	0	1	0	1	0	0	0	0	0
0	0	0	1	1	0	1	0	1	0
0	1	0	0	1	0	1	0	1	0
0	0	0	1	1	0	1	0	1	0
0	0	1	0	1	0	1	0	0	1
0	0	0	1	1	0	1	0	0	0
0	0	0	1	1	0	0	1	1	0
0	1	0	0	0	1	0	1	0	1
0	1	0	0	0	1	0	1	0	1
1	0	0	0	1	0	1	0	0	0
0	0	0	1	1	0	1	0	0	1
0	0	1	0	0	1	1	0	0	0
0	0	0	1	1	0	1	0	1	0
0	0	0	1	1	0	0	1	0	1
0	0	1	0	0	0	0	0	0	0
1	0	0	0	0	1	1	0	0	0
1	0	0	0	0	1	1	0	0	0
1	0	0	0	1	0	1	0	0	0

Répertoire courant: F:\Questionair/Donne_AGRO1csv_Apriori.txt

Donnée: Donne AGRO1csv.txt
Sortie: ıne AGRO1csv Apriori.txt
Algo: Apriori
minsup (%): 0.25

```
4 6 8 9 13 17 18 19 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 41 42 43 52 56 59 60 64 65
1 5 7 13 17 19 21 24 25 26 27 28 30 31 32 33 34 35 36 37 54 55 56 57 58 59 62 63 65
3 6 11 13 17 18 19 20 23 25 26 27 28 29 30 34 35 36 37 38 41 42 43 44 48 52 54 55 56 59 60 62 65
1 6 8 16 17 19 21 23 24 27 34 35 36 37 44 45 51 54 55 56 59 60 64 65
3 5 8 10 13 17 18 19 21 25 26 27 29 30 31 32 33 35 36 37 44 45 51 54 56 59 60 64 65
1 5 7 12 13 17 18 20 23 27 29 30 34 35 36 37 44 52 54 55 56 58 60 64 65
3 5 7 9 14 17 18 19 21 24 27 29 31 34 35 36 37 44 45 52 56 59 60 63 65
3 6 7 12 13 17 18 19 21 23 25 26 28 29 30 32 33 35 37 38 40 43 44 45 47 51 52 54 55 56 65
3 5 7 12 13 17 19 21 23 24 25 27 29 30 31 32 33 34 35 36 37 43 44 45 46 51 54 56 59 63
3 5 8 9 13 17 19 21 24 27 29 30 31 33 34 35 36 37 44 45 51 54 59 60 62 65
3 5 7 12 13 17 19 21 23 25 27 29 30 34 35 36 37 44 48 52 54 55 56 59 60 61 65
2 5 8 10 13 14 17 18 22 23 27 29 30 31 34 35 36 44 45 46 48 51 54 55 59 60 61 63 65
4 5 7 10 13 17 19 21 23 24 27 31 34 35 36 37 44 46 51 54 55 56 59 60 64 65
4 6 7 10 13 17 19 21 24 27 29 30 31 32 33 34 35 36 37 44 45 51 54 56 59 60 64 65
3 5 7 10 13 17 18 19 21 23 24 27 29 30 31 32 33 34 35 36 37 38 44 45 51 54 55 56 59 60 64 65
4 5 7 12 13 17 18 19 21 24 27 29 30 31 32 33 34 35 36 37 44 45 46 48 51 54 55 56 59 60 64 65
4 6 8 10 13 19 21 23 25 26 27 28 29 30 31 32 33 34 35 36 37 40 41 42 43 44 45 51 54 55 56 59 60 63
2 6 8 9 13 18 21 23 24 25 26 27 28 29 30 31 32 33 34 37 44 45 51 54 55 56 59 60 64 65
1 6 8 10 14 18 21 25 27 29 34 37 56 57 58 60 61 65
1 6 7 10 13 17 18 19 21 25 26 29 30 31 32 34 35 36 37 38 40 44 45 51 54 55 56 57 58 59 60 63 65
```

```
Nombre de transaction : 73
Nombre de motif : 65
```

Répertoire courant: F:\Questionair/Donne_AGRO1csv_Apriori.txt

Figure 2: ASI-MGK data view windows: on the left it is a representation in binary table and on the right it is a representation in transaction.

This is a preliminary preparatory work to bring the data to the transactional format. In the preparation phase, the ASI-MGK offers us several algorithms; conversion algorithm: sequence in transaction, CSV in data. These possibilities will allow him to read multiple text data formats and CSV data.

2.2 Frequent pattern mining

An association rule extraction context is a $\mathcal{B} = (\mathcal{O}, I, \mathcal{R})$ triplet in which $\mathcal{O}$ is called set of objects (or transactions), I set of attributes (or items), and $\mathcal{R} \subseteq \mathcal{O} \times I$ is a binary relation. An association rule is a pair $(X, Y) \in 2^I \times 2^I$ of patterns, noted: $X \rightarrow Y$, where X and Y are of disjoint motives ($X, Y \subseteq I$ and $X \cap Y = \emptyset$), respectively called premise and consequent of the rule. In what follows, denote by $n = |\mathcal{O}|$ and P the uniform probability on $(\mathcal{O}, \mathcal{P}(\mathcal{O}))$, defined for all $X \subseteq I$, by: $P(X) = \frac{n_X}{n}$.

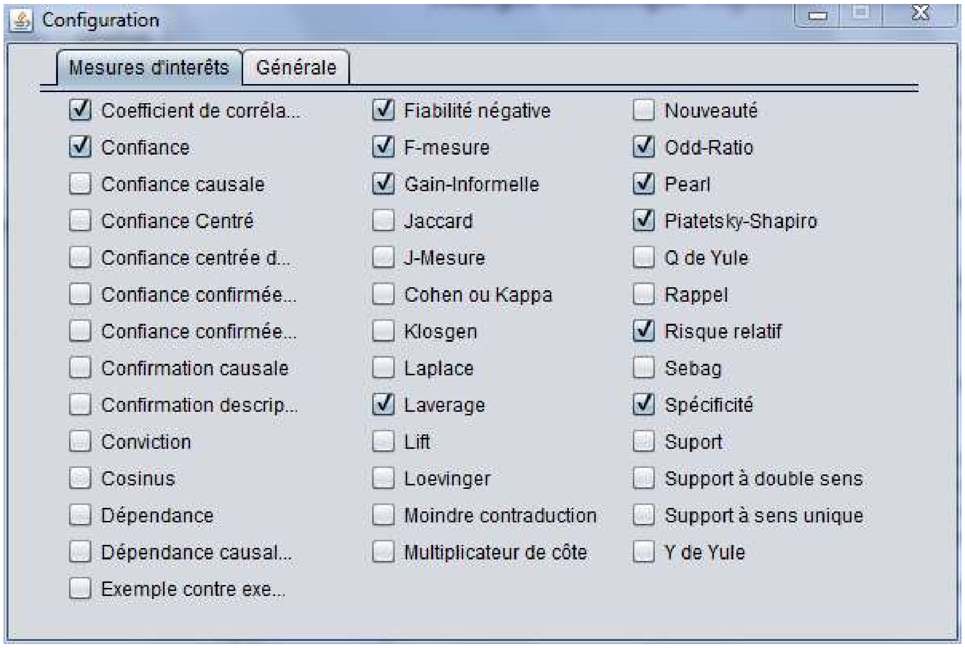
This phase consists of extracting from the context all sets of $X \subseteq I$ binary attributes, called sets of elements, which are frequent in the Y context. An itemset X is common if its support is the number of objects in the context containing Y, greater than or equal to the minimum support.

Knowledge extraction processes in databases vary depending on properties and different types of data. There are several efficient algorithms for extracting frequent sets in the literature, but we chose the following tool: Apiori, Close and A-Close, Agrawal and Srikant (1994).

2.3 Rule extraction

Much research on the association rules extraction in data mining define several measures of interest and their properties. We refer the reader to the syntheses proposed by Totohasina and Feno. (2008), Lenca and al., (2004).

Below, in this configuration window, we present the 42 interest measures of the association rules (A.R) used to reduce the number of valid A.R included in our ASI-MGK tools.



☑ Coefficient de corréla...
☑ Confiance
☐ Confiance causale
☐ Confiance Centré
☐ Confiance centrée d...
☐ Confiance confirmée...
☐ Confiance confirmée...
☐ Confirmation causale
☐ Confirmation descrip...
☐ Conviction
☐ Cosinus
☐ Dépendance
☐ Dépendance causal...
☐ Exemple contre exe...

☑ Fiabilité négative
☑ F-mesure
☑ Gain-Informelle
☐ Jaccard
☐ J-Mesure
☐ Cohen ou Kappa
☐ Klosgen
☐ Laplace
☑ Laverage
☐ Lift
☐ Loevinger
☐ Moindre contraduction
☐ Multiplicateur de côte

☐ Nouveauté
☑ Odd-Ratio
☑ Pearl
☑ Piatetsky-Shapiro
☐ Q de Yule
☐ Rappel
☑ Risque relatif
☐ Sebag
☑ Spécificité
☐ Suport
☐ Support à double sens
☐ Support à sens unique
☐ Y de Yule

Figure 3: ASI-MGK configuration window for the interest measures to be integrated

Definition 1. *(Support)* Support for a $X \rightarrow Y$ association rule is the proportion of transactions in the database that contain $X \wedge Y$. $supp(X \rightarrow Y) = P(X' \cap Y')$

A rule $r: X \rightarrow Y$ is valid according to support if $supp(X \rightarrow Y) \geq minsup$.

Studies on rule quality measures have been completed; Feno in his thesis in 2007 has characterized a basis for valid association rules in the sense of the M_{GK} quality measure. According to the eligibility criteria, Totohasina in its PhD thesis 2008 has confirmed that the use of this quality measure overcomes the problems of methods of the extraction of association rules using the quality measures Confidence.

Thus we chose the measures of interest denominated Confiance of Guillaume noted ConfG or measures called Guillaume-Kenchaff abbreviated M_{GK} , Implication oriented standardized abbreviated ION, conditional probability incremental ratio or $CPIR$.

Definition 2. *(M_{GK})* Let X and Y be two reasons for a data mining context. We define the measure M_{GK} by: (Guillaume, 2000)

$$M_{GK}(X \rightarrow Y) = \begin{cases} \frac{P(Y|X) - P(Y)}{1 - P(Y)}, & if \quad P(Y|X) \geq P(Y); \\ \frac{P(Y|X) - P(Y)}{P(Y)}, & if \quad P(Y|X) < P(Y). \end{cases} \tag{1}$$

The main mathematical properties characterizing this measure of quality are developed in (Feno and Totohasina 2007), we invite the interested reader to consult them in their works.

Definition 3. A rule $X \rightarrow Y$ is potentially interesting, if the support of its premise is inferior to that of its consequence.

Definition 4. Let X and Y be two patterns such that $P(Y|X) \geq P(Y)$ and $P(X) \leq P(Y)$, then $M_{GK}(Y \rightarrow X) \leq M_{GK}(X \rightarrow Y)$

Definition 5. Given a minimal threshold $minsupp$ for the measure of support and error risk α for M_{GKcr}, the set $\mathcal{R}$ of valid rules is:

$$\mathcal{R} = \{(X, Y) \subseteq I \times I | supp(X \rightarrow Y) \geq minsupp \;\; and \;\; M_{GK}(X \rightarrow Y) \geq M_{GKcr}(X \rightarrow Y, \alpha)\}$$

Given a $\mathcal{B}$ database of n transactions, where n_X and n_Y are supports of the patterns X and Y respectively. The critical value M_{GKcr} for the measure M_{GK}, proposed in (Totohasina and Feno),

is obtained in the following way. Consider a contingency table by crossing these two patterns X and Y, using critical value of Pearson's chi-square (χ^2_{cr}) statistic with one degree of freedom, such as:

$$M_{GKcr}(X \rightarrow Y, \alpha) = \sqrt{\frac{1}{n}\frac{n-n_X}{n_X}\frac{n_Y}{n-n_Y}\chi^2_{cr(\alpha)}}, \;\; if \;\; X \;\; favors \;\; Y. \tag{2}$$

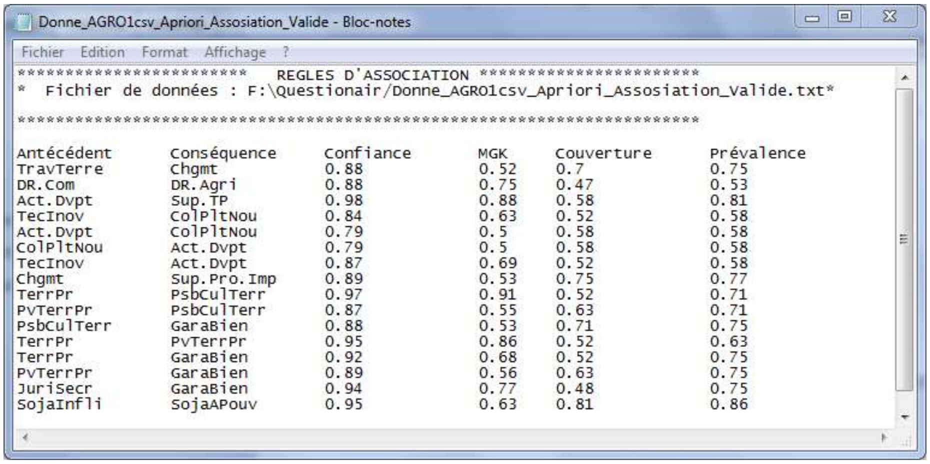
From the table of χ^2, this relation makes it possible to easily develop the abacuses of the critical values of M_{GK} to decide its meaning towards an association rule.

2.4 Results analysis and visualization

This phase consists in the visualization by the user of the association rules extracted from the context and their interpretation in order to deduce useful knowledge for the improvement of the activity concerned.

In this step our **ASI-MGK** tool offer three result presentation registers:

1. open text format represention, save in ASCI text format, the rules are printed in tabular data as tab-separated values (fig 4).
2. Interactive dynamic table, where the user can filter and sort the extracted association rules as needed.
3. Interactive dynamic implicative graph (fig 5), on which the user can reorganize the graph and move or even delete the vertices (items). The properties of the generated graph can be:
   - a directed digraph (between two items given $X \rightarrow Y$ and $Y \rightarrow X$ can all be validated.
   - a bi-valued graph (the arcs can carry two value of measure of interest: M_{GK} and a interest measures previously chosen to compare with M_{GK}).



Donne_AGRO1csv_Apriori_Assosiation_Valide - Bloc-notes
Fichier Edition Format Affichage ?

```
*************************   REGLES D'ASSOCIATION ************************
*  Fichier de données : F:\Questionair/Donne_AGRO1csv_Apriori_Assosiation_Valide.txt*

************************************************************************

Antécédent      Conséquence     Confiance       MGK     Couverture      Prévalence
TravTerre       Chgmt           0.88            0.52    0.7             0.75
DR.Com          DR.Agri         0.88            0.75    0.47            0.53
Act.Dvpt        Sup.TP          0.98            0.88    0.58            0.81
TecInov         ColPltNou       0.84            0.63    0.52            0.58
Act.Dvpt        ColPltNou       0.79            0.5     0.58            0.58
ColPltNou       Act.Dvpt        0.79            0.5     0.58            0.58
TecInov         Act.Dvpt        0.87            0.69    0.52            0.58
Chgmt           Sup.Pro.Imp     0.89            0.53    0.75            0.77
TerrPr          PsbCulTerr      0.97            0.91    0.52            0.71
PvTerrPr        PsbCulTerr      0.87            0.55    0.63            0.71
PsbCulTerr      GaraBien        0.88            0.53    0.71            0.75
TerrPr          PvTerrPr        0.95            0.86    0.52            0.63
TerrPr          GaraBien        0.92            0.68    0.52            0.75
PvTerrPr        GaraBien        0.89            0.56    0.63            0.75
JuriSecr        GaraBien        0.94            0.77    0.48            0.75
SojaInfli       SojaAPouv       0.95            0.63    0.81            0.86
```

Figure 4: Associations rules extracted in text format

Visualizations based on implicative graphs quickly show their limits when the number of rules is important: the arcs intersect very often and the abundance of nodes and arcs leads to occlusion.

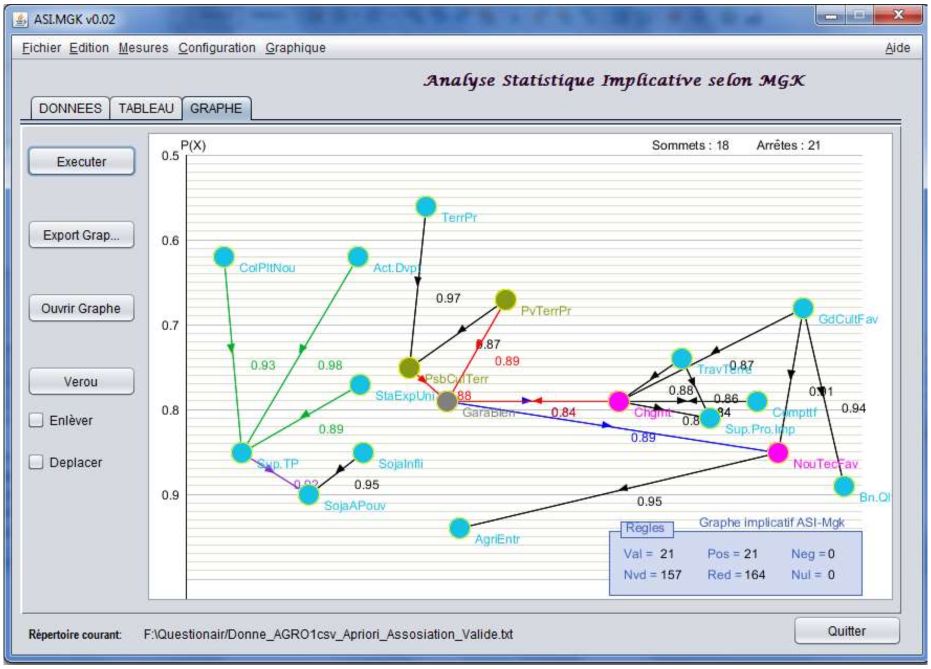


Figure 5: **ASI-MGK** interactive dynamic implicative graph viewer.

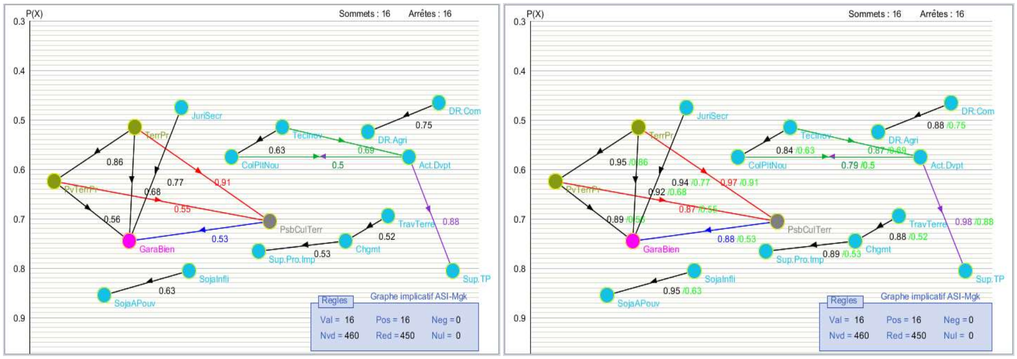


Figure 6: The implicative graphs displayed in the graphical window of the ASI-MGK: on the left a simple graph and on the right a bivalued digraph.

2.5 Results analysis

Here is the report of the processing of the **ASI-MGK** software, a data set of a survey conducted at L2 agronomy student, containing 74 transactions and 65 attibits.

```
 ============= STATISTIQUES APRIORI =============
Candidates number: 703 The algorithm is stopped at size 2
Frequent items number: 500
Maximum memory used : 13.183204650878906 mb
Total time ~ 0 ms
======== STATS DE GENERATION DES REGLES =========
 Valid association rules number generated : 16
Discarded : 910
Redundant rules number: 450
Invalid rules : 460 Rate : 0.03
Error risk : 0.001
Total time ~ 1326 ms
```

The main line of this result (fig 4) corresponds to this:

TravTerre Chgmt 0.88 0.52 0.7 0.75

It is interpreted as follows: The support of the premise is supp(TravTerre)=0.7 and that of consequence is supp(Chgmt) = 0.75 .

According to the interest measure confidence, the rule R1:(TravTerre -> Chgmt) is valid because conf(TravTerre-> Chgmt)= 0.88 .

According to our approach (Def. 5) Supp(TravTerre) < Supp(Chgmt) that is to say
R1:(TravTerre -> Chgmt) is potentially interesting,
we can evaluate MGK(TravTerre -> Chgmt) = 0.52
0.52 > MGKcr(TravTerre -> Chgmt, 0.001) = 0.44 .
So the rule R1:(TravTerre -> Chgmt) is also valid in the sense of M_{GK}.

3 Conclusion

In summary, we presented an implicit statistical analysis tool incorporating algorithms to extract knowledge from a binary database. In addition, we have also implemented and optimized the relevant algorithms. The specificity of this tool, among others, is that the proposed tool uses the quality measure of association rules M_{GK}. Presumably, the advantage of this approach lies in the use of M_{GK}, which is non-symmetric and implicative, and in the fact that the threshold is determined objectively from a chi-square independence test. We can therefore trust the results obtained if the risk defined by the user is very low. Is not it more motivating to use an application giving reliable results?

As future work, we will integrate the implicative cohesion and classification (Rakotomalala and al 2017) part in ASI-MGK and then we will present the procedure of the complete data processing following all the steps of the data mining with real data using this tool.

Acknowledgements

The authors thank the Department of Mathematics and Computer Science of the ENSET of the University of Antsiranana - Madagascar for the facilities provided.

References

[1] R. Agrawal and R. Srikant. (1994) “Fast algorithms for mining association rules.” *In Proc. of the 20th VLDB Conference*, pages 487-499, San Diego, Chile

[2] R. C. Agarwal, C. C. Aggarwal, and V. V. V. Prasad. (2000) “Depth first generation of long patterns. *In Proc. of SIGKDD’00*, pages 108-118. ACM Press.

[3] B. Liu, W. Hsu, and Y. Ma. (1999) “Mining association rules with multiple minimum supports”. *In Proc. of SIGKDD’99*, pages 337-341, New York, NY, USA. ACM Press.

[4] S. P. Bai and Kumar G. R. (2019). “Subset Significance Threshold: An Effective Constraint Variable for Mining Significant Closed Frequent Itemsets”. *In Emerging Technologies in Data Mining and Information Security* , pages 449-458. Springer, Singapore.

[5] D. Feno, (2007) Measure of quality the association rules: standardization and characterization of bases, University of the Reunion, France, PhD thesis.

[6] F. Guillet and H. Hamilton, (2007) Quality measures in data mining, Springer-Verlag,.

[7] S. Guillaume, (2000) Processing of large data. Measure and extraction algorithms and ordinal association rules, University of Nantes, France , PhD thesis.

[8] L. Geng and H.J. Hamilton, 2006. Interes-tingness Measures for Data Mining: A Survey. ACMComput. Surveys. DOI: 10.1145/1132960.1132963

[9] H. Yun, D. Ha, B. Hwang, and K. Ryu. (2003) “Mining association rules on significant rare data using relative support.” *Journal of Systems and Software*, 67(3):181-191.

[10] N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal. (1999) Discovering frequent closed itemsets for association rules. LNCS, 1540:398-416.

[11] P. Lenca, B. Vaillant, B. Meyer and S. Lallich. (2007) Association rule interestingness: experimental and theoretical studies. *Studies in Computational Intelligence*, Volume 43, pages 51–76.

[12] H. F. Rakotomalala, B. B. Ralahady, A. Totohasina. (2017) "A novel cohesitive implicative classification based on M_{GK} and application on diagnostic on informatics literacy of students of higher education in Madagascar”. *3rd Internationa Conference ICICT 2018-International Congress & Excellence Awards*.

[13] S. Lallich, O. Teytaud and E. Prudhomme, (2007) “Association Rule Interestingness: Measure and Statistical Validation”. *In Quality measures in Data Mining*, Guillet, Fabrice, Hamilton, Howard J. (Eds.), Springer, Berlin, ISBN-10: 3540449116, pp: 251- 275.

[14] A. Totohasina , (2008) Contribution to the study of measures of quality of association rules: normalization and constraints in five cases and MGK, properties, composite base and extension rules for applying statistical and physical sciences, University of Antsiranana, Madagascar, HDR thesis.

[15] A. Totohasina and D.Feno (2008) The quality of association rules: a comparative study of the MGK and *Confidence. In Proceedings of the 9th African conference on research in Computer Science and Applied Mathematics*, CARI-08,561-568 (in French)

[16] B.Vaillant, P. Picouet, and P. Lenca (2003). An extensible platform for rule quality measure benchmarking. HCP’03, Human Centered Processes, 187–191.

[17] Y. Bastide, R. Taouil, N. Pasquier, G. Stumme, and L. Lakhal. (2000) Mining frequent patterns with counting inference. SIGKDD Explor. Newsl., 2(2):66-75.

Authors

Bruno Bakys RALAHADY received his CAPEN (Certificate in Pedagogical Aptitude of the Normal School) degree *of Mathematics and Computer Science* in the year *2009.* He obtained his DEA (Degree of Profound Studies) in mechanics of fluids and energy systems 2013.

Since 2014, a PhD. Student in Problems of Education and Didactics of Disciplines (PE2Di) on *Education and Didactics of Mathematics and Computer Science.* Is currently a P Assc. Prof. In ENSET, University of Antsiranana, on Programming and Operational Calculus.

MADAGASCAR, Antsiranana.