

Improving the Discovery of Association Rules with Intensity of Implication

Sylvie Guillaume¹, Fabrice Guillet¹ and Jacques Philippé^{1,2}

¹ IRIN, Equipe SIC, IRESTE - Rue Christian-Pauc - La Chantrerie,
44306 Nantes Cedex 3 – France

{sguillau, fguillet, jphilipp}@ireste.fr

² PERFORMANSE - 3 rue Racine,
44000 Nantes - France

Abstract. In this paper, we propose a way to improve the rule-learning step in a Knowledge Discovery in Databases (KDD) process. Our purpose is to make possible the discovery of relevant rules in a large database.

To achieve this goal, we merge :

- a quality index proposed by *R. Gras* : *intensity of implication*,
- together with a specific algorithm written by *Agrawal et al.*

The algorithm itself is efficient in a large database but delivers a prohibitively large set of knowledge.

Intensity of implication is a new measurement of the quality of association rules.

Hence, we analyze it in detail and compare it with conditional probability index.

We show that it is possible to significantly improve the relevance of association rules supplied by the algorithm proposed by *Agrawal et al.*, by using the quality index : *intensity of implication*.

An improved algorithm has been implemented, and has been tested both at the experimental level and on a real-life database.

1 Introduction

This study fits within the framework of KDD : how to extract knowledge from a database in the form of association rules [1], [2].

On large databases two main practical limits arise : algorithms are little efficient and they deliver a prohibitively large set of knowledge. So, relevant rules are not easy to highlight and the task of the end-user is heavy.

These two problems come from the fact that :

- the *complexity* of knowledge discovery grows exponentially with the number of attributes in the database; so algorithms may rapidly become inefficient.
- the *number of discovered rules* may be far larger than the number of examples, thus it may be impossible to exploit discovered knowledge in large databases.

To push back those two limits and work with large databases, we propose to improve an algorithm, published by *Agrawal et al.*, that is efficient on large databases. This

improvement is made by coupling it with a new index to measure the relevance of the rules : *intensity of implication*.

With this association we expect the following properties :

- With the algorithm proposed by Agrawal et al., discovery of association rules may run in reasonable time on large databases.
- Due to the properties of intensity of implication, the number of discovered rules may be reduced to the most relevant rules.

In this paper, we will show how intensity of implication can be used in order to find more relevant rules than rules discovered with conditional probability. Indeed, intensity of implication measures the relevance of a rule : the probability that a rule results from random phenomena. Therefore the efficiency of algorithm is improved, while the processing of large data volumes is more accessible.

2 Agrawal's Algorithm (A)

Agrawal [3], [4], [5] provides an efficient algorithm, called (A), for finding association rules between items in a database.

An association rule is an expression $X \rightarrow Y$ where X and Y are sets of items describing objects in the database. The intuitive meaning of such a rule is that objects in the database which are covered by Y tend to contain objects covered by X . Furthermore, rules are *statistical* : they can have negative examples.

Rules must meet the following requirements :

1. a sufficient number of objects must verify the rule (*support* also called *frequency*).
2. the number of objects which contradict the rule must be small (*confidence* also called *reliability*).

Formal Presentation. Let $D = \{d_0, d_1, \dots, d_p\}$ be a set of p literals, each literal d_i is a binary attribute called an *item*¹. Let E be a set of objects², and each element e of E is described by an itemset X , where X is a part of D .

To say that "object e is described by the itemset $X = \{d_0, d_1, d_3\}$ ", means "e shows all items of X and none of the others" or "e is described by the conjunction $d_0 \wedge d_1 \wedge \bar{d}_2 \wedge d_3 \wedge d_4 \wedge \dots \wedge d_p$ ". The description X may also be seen as a p -uple (x_0, x_1, \dots, x_p) where $x_i = 1$ if $d_i \in X$ and $x_i = 0$ otherwise.

An *association rule* is an implication of the form $X \rightarrow Y$, between disjointed itemsets ($X \cap Y = \emptyset$). $X \rightarrow Y$ means that "all objects that show the items of X show the items of Y , as well". This rule being of statistical nature, it is understood that some negative examples may exist.

The statistical feature of a rule $X \rightarrow Y$ is characterized by two indexes :

1. the *support* $T(X \rightarrow Y)$ (or coverage rate) and
2. the *confidence* $P(X \rightarrow Y)$ (or conditional probability).

¹ or variable in Data Analysis, feature in Machine Learning, attribute in Cognitive Science, field in Databases ...

² examples in Machine Learning, transactions in Databases, observations ...

Let (f,g) be two functions between parts of D and E . Let $f(e)=X$ be the itemset describing object e , and let $g(X)=\{e \in E / X \subseteq f(e)\}$ be the part of all objects covered by itemset X (subset of all objects showing at least all the items in X), then the two indexes are defined by :

$$T(X \rightarrow Y) = t(X \cup Y) = |g(X \cup Y)| / |g(X)| \tag{1}$$

$$P(X \rightarrow Y) = \text{Proba}(Y/X) = |g(X \cup Y)| / |E| \tag{2}$$

Given two thresholds α and β in $[0, 1]$, algorithm (A) uses the following criteria to select a convenient statistical rule $X \rightarrow Y$: $T(X \rightarrow Y) \geq \alpha$ and $P(X \rightarrow Y) \geq \beta$. When $T(X \rightarrow Y) \geq \alpha$, we say that itemset $X \cup Y$ is α -frequent, and when $P(X \rightarrow Y) \geq \beta$, we say that association rule $X \rightarrow Y$ is β -convenient.

Comments. The efficiency of algorithm (A) follows from the two previous indexes (1), (2) :

- The support limits the combinatorial explosion, because only α -frequent itemsets are retained: the itemsets that covers at least $\alpha\%$ of objects. Thus, algorithm (A) selects a restricted subset of 2^D , the set of all parts of D .
- The confidence improves the reliability of the discovered rules set, because only the β -convenient rules are kept : the rules with few negative examples, rules $X \rightarrow Y$ where at least $\beta\%$ of objects covered by X are also covered by Y .

3 Toward Intensity of Implication

3.1 Limits of Conditional Probability

The previous section shows that algorithm (A) is based upon the conditional probability (confidence) to select the convenient rules. But, the conditional probability of a rule $X \rightarrow Y$ is invariable when the size of $g(Y)$ or E varies and insensitive to cardinal dilatation [6]. Nevertheless, $X \rightarrow Y$ is more likely to happen when the size of $g(Y)$ increases or when the size of E decreases; and furthermore this implication will be more meaningful when the size of the all sets grows in the same proportion.

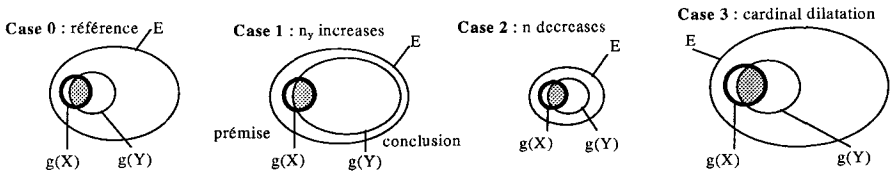


Fig. 1. Three cases with constant conditional probability

Fig. 1 shows that in case 1, when $g(Y)$, the number of objects covered by conclusion Y is close to the size of E , it is not surprising that most of the objects covered by premise

X, are included in $g(Y)$. In case 2, when the size of $g(X)$ and the size of $g(Y)$ are close to that of E , it is not surprising that $g(X)$ and $g(Y)$ share many objects. And to finish case 3, a more important sample will allow to be more confident in the statistical implication $X \rightarrow Y$.

It could be interesting for this algorithm to find a new index which would vary according to the size of $g(Y)$ and E . This would improve the quality of the discovered rules. *Intensity of implication* fulfils this goal.

3.2 Intensity of Implication

Let $A=g(X)$, $B=g(Y)$, $n=|E|$, $n_a=|A|$, $n_b=|B|$, $n_{a\bar{b}} = |A \cap \bar{B}|$; n is the number of objects, n_a and n_b are the number of objects covered by premise X and conclusion Y , and $n_{a\bar{b}}$ the number of negative examples (exceptions) of the rule $X \rightarrow Y$.

One property of intensity of implication proposed by R. Gras [7], [8], [9] is that it evolves with the sizes n , n_a , n_b , $n_{a\bar{b}}$. Its meaning is to measure the *statistical surprise* of having so few negative examples on a rule as compared with a random draw.

Definition. Let $X \rightarrow Y$ be the observed rule; premise X and conclusion Y in 2^D are associated with sets of objects $A=g(X)$ and $B=g(Y)$ in E . Let U and V be two sets randomly chosen with the same cardinality as A and B : $|U|=|A|=n_a$ and $|V|=|B|=n_b$ (see Fig. 2).

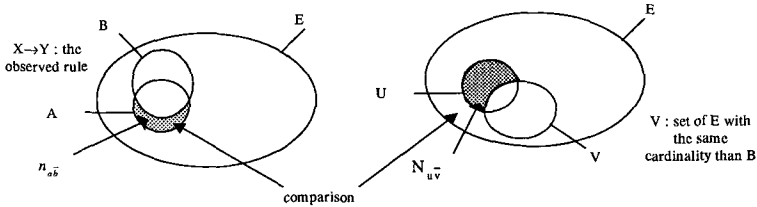


Fig. 2. U and V vary at random in E

Let $N_{u\bar{v}} = |U \cap \bar{V}|$ be the random variable measuring the number of random negative examples, and $n_{a\bar{b}} = |A \cap \bar{B}|$ the number of negative examples observed on the rule.

We shall compare $n_{a\bar{b}}$ with $N_{u\bar{v}}$, assuming that U and V are independent. If $n_{a\bar{b}}$ is unusually small as compared with $N_{u\bar{v}}$ the one we would expect at random, we will accept the statistical implication $X \rightarrow Y$.

The quality of implication is even better if the number of negative examples is smaller than the expected one, in other words, if the quantity $\text{Proba}(N_{u\bar{v}} \leq n_{a\bar{b}})$ is small. The observed "smallness" of $n_{a\bar{b}}$ is taken as a basic character of statistical implication $X \rightarrow Y$.

Intensity of implication is then defined by function φ : $\varphi(X \rightarrow Y) = 1 - \text{Proba}(N_{uv}^- \leq n_{ab}^-)$. The random variable $N_{uv}^- = |U \cap \bar{v}|$ follows the hypergeometric law [10], thus intensity of implication $\varphi(X \rightarrow Y)$ is defined by :

$$\varphi(X \rightarrow Y) = 1 - \sum_{k=\max(0, n_a - n_b)}^{n_{ab}^-} \frac{C_{n-a}^{n_b-k} \times C_n^k}{C_n^{n_b}} \tag{3}$$

Sensitivity to Cardinal

Fig. 3 shows the variation of intensity of implication $\varphi(X \rightarrow Y)$ with the number n_b of conclusion, with the number n of objects, and by dilatation of n when $\frac{n_a}{n}$, $\frac{n_b}{n}$ and $\frac{n_{ab}}{n}$ remain constant; meanwhile conditional probability (convenience) $\frac{n_{ab}}{n_a}$ stay constant.

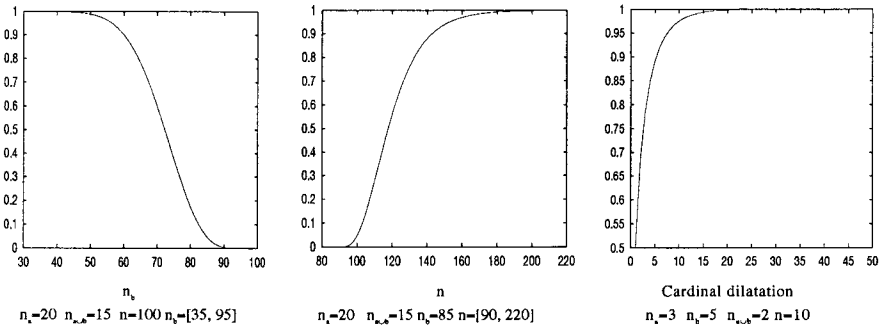


Fig. 3. Sensitivity of intensity of implication

Property of Symmetry. Intensity of implication is symmetric : $\varphi(Y \rightarrow X) = \varphi(X \rightarrow Y)$.
 Demonstration :

We have $\varphi(Y \rightarrow X) = 1 - \text{Proba}(N_{uv}^- \leq n_{ab}^-)$ and $N_{uv}^- = n_b - (n_a - N_{uv}^-)$ and $n_{ab}^- = n_b - (n_a - n_{ab}^-)$, thus $\text{Proba}(N_{uv}^- \leq n_{ab}^-) = \text{Proba}(n_b - n_a + N_{uv}^- \leq n_b - n_a + n_{ab}^-) = \text{Proba}(N_{uv}^- \leq n_{ab}^-)$, therefore : $\varphi(Y \rightarrow X) = \varphi(X \rightarrow Y)$

This symmetry of intensity of implication leads to use an other information to know the nature of the implication between X and Y : $Y \rightarrow X$ or $X \rightarrow Y$. In this paper we choose conditional probability. In another knowledge discovery system : FIABLE [11], [12], [13] intensity of implication is also associated with conditional probability.

Behavior on Logical Rules. When no negative example occurs on rule $X \rightarrow Y$, $n_{ab}^- = 0$, and

$$\varphi(X \rightarrow Y) = 1 - \frac{C_n^{n_b}}{C_n^{n_b}} = 1 - \frac{C_n^{n_b}}{C_n^{n_a}} \tag{4}$$

Hence, while conditional probability still equals 1, intensity of implication is more selective and may reject some logical rules.

In particular when $n_p = n$, $\phi(X \rightarrow Y) = 0$. So, according to intensity of implication it is not surprising to have the rule $X \rightarrow Y$ when $B = g(Y) = E$.

4 Algorithm (A) Improved with Intensity of Implication

Let us present two examples: the first essentially produces statistical rules and the second logical rules. We are going to observe the behavior of the new algorithm with these two examples.

First example. The example has 8 items ($D = \{a, b, c, d, e, f, g, h\}$) and 20 objects ($|E| = 20$).

Learning set									Computed rules					
Objets	a	b	c	d	e	f	g	h	Rules	$P(X \rightarrow Y)$	$\phi(X \rightarrow Y)$	Rules	$P(X \rightarrow Y)$	$\phi(X \rightarrow Y)$
1	1	0	1	0	1	1	1	0	$\{a\} \rightarrow \{c\}$	0.80	0.63	$\{e, g\} \rightarrow \{a\}$	0.87	0.69
2	1	1	0	1	1	0	0	1	$\{c\} \rightarrow \{a\}$	0.80	0.63	$\{f, g\} \rightarrow \{a\}$	1.00	0.87
3	0	1	1	1	1	0	0	0	$\{d\} \rightarrow \{a\}$	0.88	0.78	$\{f\} \rightarrow \{a, g\}$	1.00	0.96
4	1	0	1	0	1	1	1	0	$\{e\} \rightarrow \{a\}$	0.81	0.60	$\{a, f\} \rightarrow \{a\}$	1.00	0.87
5	1	0	1	1	0	0	1	0	$\{f\} \rightarrow \{a\}$	1.00	0.87	$\{d, g\} \rightarrow \{c\}$	0.83	0.48
6	1	1	1	0	1	0	1	1	$\{a\} \rightarrow \{g\}$	0.86	0.93	$\{f, g\} \rightarrow \{e\}$	0.83	0.88
7	1	0	1	0	0	0	1	0	$\{g\} \rightarrow \{a\}$	0.86	0.93	$\{e, f\} \rightarrow \{g\}$	1.00	0.80
8	0	1	0	0	0	0	1	0	$\{c\} \rightarrow \{g\}$	0.80	0.63	$\{f\} \rightarrow \{e, g\}$	0.83	0.98
9	1	0	1	1	0	1	1	0	$\{g\} \rightarrow \{c\}$	0.80	0.63	$\{cdg\} \rightarrow \{a\}$	1.00	0.81
10	1	1	1	1	1	0	0	0	$\{f\} \rightarrow \{e\}$	0.83	0.88	$\{adg\} \rightarrow \{c\}$	1.00	0.81
11	1	0	1	0	1	1	1	0	$\{f\} \rightarrow \{g\}$	1.00	0.87	$\{acd\} \rightarrow \{g\}$	0.83	0.48
12	1	0	1	1	1	0	1	0	$\{c, g\} \rightarrow \{a\}$	0.92	0.99	$\{dg\} \rightarrow \{ac\}$	0.83	0.81
13	0	1	1	0	1	0	1	0	$\{a, g\} \rightarrow \{c\}$	0.84	0.79	$\{ceg\} \rightarrow \{a\}$	0.83	0.48
14	1	0	1	0	0	0	1	0	$\{a, c\} \rightarrow \{g\}$	0.92	0.94	$\{ace\} \rightarrow \{g\}$	0.83	0.48
15	1	1	0	0	1	1	1	0	$\{c, d\} \rightarrow \{a\}$	0.86	0.59	$\{efg\} \rightarrow \{a\}$	1.00	0.81
16	1	0	1	1	0	0	1	0	$\{d, g\} \rightarrow \{a\}$	0.83	0.48	$\{afg\} \rightarrow \{e\}$	0.83	0.88
17	0	1	0	0	0	0	0	1	$\{e, f\} \rightarrow \{a\}$	1.00	0.81	$\{f\} \rightarrow \{aeg\}$	0.83	0.99
18	1	0	1	1	0	0	1	0	$\{a, f\} \rightarrow \{e\}$	0.83	0.88	$\{aef\} \rightarrow \{g\}$	1.00	0.81
19	0	1	1	0	0	0	0	1	$\{f\} \rightarrow \{a, e\}$	0.83	0.96	$\{fg\} \rightarrow \{ae\}$	0.83	0.96
20	1	0	0	1	1	1	1	0	$\{ef\} \rightarrow \{ag\}$	1.00	0.92	$\{af\} \rightarrow \{eg\}$	0.83	0.98

Fig. 4. Comparing discovered rules : confidence versus intensity of implication

Algorithm (A) is run on this example with a minimal support $\alpha = 0.25$ and a minimal confidence $\beta = 0.80$. It delivers 40 rules, that is twice the number of observations!

Intensity of implication with the same threshold (0.80) excludes 14 rules (in gray in Fig. 4). Therefore the number of rules decreases by 35%. This means that only 26 rules are not accidental and that 14 may happen quite by chance.

These discarded rules can be separated into two categories (see Sect. 3.1). Results are summarized in Fig. 5:

- the first category involves large conclusion ($g(Y)$ has a size close to E) and small premise ($g(X)$ is small) (see case 1 of Fig. 1), then having an important intersection is highly probable,

- the second one involves both large premise and conclusion ($g(X)$ and $g(Y)$ close to E) (see case 2 of Fig. 1); then sharing a lot of objects is also highly probable.

Case 1			Case 2		
Rules	$P(X \rightarrow Y)$	$\varphi(X \rightarrow Y)$	Rules	$P(X \rightarrow Y)$	$\varphi(X \rightarrow Y)$
$\{c,d\} \rightarrow \{a\}$	0.86	0.59	$\{a\} \rightarrow \{c\}$	0.80	0.63
$\{d,g\} \rightarrow \{a\}$	0.83	0.48	$\{c\} \rightarrow \{a\}$	0.80	0.63
$\{e,g\} \rightarrow \{a\}$	0.87	0.69	$\{d\} \rightarrow \{a\}$	0.88	0.78
$\{d,g\} \rightarrow \{c\}$	0.83	0.48	$\{e\} \rightarrow \{a\}$	0.81	0.60
$\{a,c,d\} \rightarrow \{g\}$	0.83	0.48	$\{c\} \rightarrow \{g\}$	0.80	0.63
$\{c,e,g\} \rightarrow \{a\}$	0.83	0.48	$\{g\} \rightarrow \{c\}$	0.80	0.63
$\{a,c,e\} \rightarrow \{g\}$	0.83	0.48	$\{a,g\} \rightarrow \{c\}$	0.84	0.79

Fig. 5. Discarded rules

With this example, we have seen that the intensity of implication has discarded the statistical rules which could be accidental.

Second Example. This example has 39 objects described by 21 items and studies the correlation between food and birds' shapes of beaks.

The following figure shows a part of learning set:

NAME	SHAPE	SIZE	THICKNESS	FOOD
Eagle	hooked	small	medium	animal
Snipe	straight	long	thin	worm
Crossbill	hooked	medium	medium	seed
Buzzard	hooked	small	thick	Animal and insect

Fig. 6. A part of learning set

The two algorithms produce the following number of rules, with a minimal support α varying from 0.01 to 0.20 and a minimal confidence $\beta=0.80$:

Support α	Number of statistical rules			Number of logical rules		
	$\beta=0.80$	$\beta=0.80, \varphi$	Rejected by φ	$\beta=1$	$\beta=1, \varphi=0.8$	Rejected by φ
0.20	52	52	0%	333	333	0%
0.10	53	53	0%	1586	1586	0%
0.01	53	53	0%	1848	1784	3.4%

Fig. 7. Number of discovered rules

First we can observe a prohibitive number of rules and a majority of logical rules. A deeper study shows that there are a lot of redundant rules, resulting from strong relationships between data and also because there are not many objects in the learning set.

After discarding redundant rules with the algorithm proposed by [14], only 43 rules are kept (initially, there were 1848 rules). Among these rules, 20 are considered as excellent rules (intensity of implication close to 1); 4 are considered as bad rules (intensity of implication below 0.8) and 19 are considered as good rules (intensity of implication between 0.8 and 1).

The non-redundant logical rules have values for intensity of implication varying from 0.63 to 1, while the conditional probability stays constant and equal to 1. Thus, this new index can measure the quality of this kind of rules where there is no counter-example.

The following figure shows examples of logical rules for each these three categories:

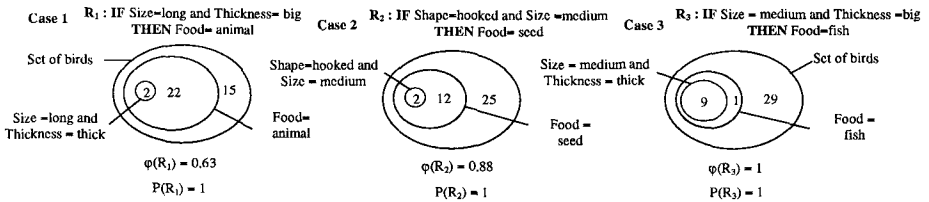


Fig. 8. Examples of logical rules

The first rule R_1 (see case 1 of Fig. 8) is regarded as a bad rule for intensity of implication while it is valuable for conditional probability. In fact, there are only two birds which own a long and large beak among the birds feeding on animals : this is not sufficient to infer rule R_1 because there are a lot of birds eating animals and there is an important probability of having such a small set included in the big one.

The second rule R_2 (see case 2 of Fig. 8) is regarded as a good rule for intensity of implication. In contrast of the previous rule, the small set of birds also composed of two objects is included in a smaller set. The chance of having this support is low and we can think that there is an implication between these two itemsets.

The last rule R_3 (see case 3 of Fig. 8) is regarded as an excellent rule. There is no doubt on the implication when we have two medium sets close to the learning set and when these two sets share a lot of objects.

Intensity of implication can measure the quality of logical rules while conditional probability cannot. This new index allows to discard rules which have a high probability of being accidental.

4.3 Study on a Real Case

Now, we are going to study the algorithm on two medical significant databases. One with 2178 transactions described by 75 items, and another with 1161 transactions and 103 items. Preliminary results, considering only the number of delivered rules, are presented in Fig. 9 :

DataBase	Support α	Number of statistical rules			Number of logical rules		
		$\beta=0.80$	$\beta=0.80,$ φ	Rejected by φ	$\beta=1$	$\beta=1,$ $\varphi=0.8$	Rejected by φ
Base 1	0.10	349	266	23.8%	4	4	0%
	0.01	21530	19812	8%	2758	2758	0%
Base 2	0.10	7760	6783	12.6%	2114	1560	26.2%

Fig. 9. Results in the two cases

Results are different for the two data sets. The first one (where there are essentially statistical rules) is close to the first experimental example. Meanwhile, the second one (where there is a lot of logical rules) is close to the second experimental example. These differences are due to the structure of the databases. The first has a strong logical structure of implication between items that explain the large set of delivered logical rules. In the second, some items are owned by nearly all the objects, and the intensity of implication is more selective.

5 Conclusion

From the previous comments and examples, the integration of intensity of implication into algorithm (A) proposed by Agrawal *et al.* allows :

- the possibility of working on large databases in a reasonable amount of time,
- to obtain *better* rules : intensity of implication measures finer and more *relevant* phenomena than conditional probability and is also *noise-resistant*. This point is very important for improving quality of discovered rules by knowledge discovery algorithms in databases

These real-case results are still to be complete by an in-depth analysis of the discovered rules.

Nevertheless, it is obvious that in those two cases the number of discovered rules is far too large to be understood by an end-user. As the rules are saved in the database, they can be manipulated by SQL requests. This is helpful for the end-user who wants to select and explore discovered knowledge; but this is clearly insufficient. An extension of our work is to add to the algorithm a browsing interface and to couple to the algorithm a redundancy elimination.

References

1. Matheus C.J., Chan P.K., Piatetsky-shapiro G.: Systems for Knowledge Discovery in Databases, IEEE Trans. Knowl. Data Eng., vol. 5, n°6, (1993)
2. Frawley W. J., Piatetsky-Shapiro G., Matheus J.: Knowledge Discovery in Databases : an overview, in Knowledge Discovery in Databases, Cambridge, MA :AAAI/MIT (1991) pp. 1-27

3. Agrawal R., Imielinski T., and Swami A.: Mining Association Rules between Sets of Items in Large Databases. In Proceedings, ACM SIG-MOD Conference on Management of Data, 207-216. Washington, D.C. (1993)
4. Agrawal R. And Srikant R.: Fast Algorithms for Mining Association Rules. IBM Research Report RJ9839, IBM Almaden Research Center, San Jose, Calif. (1994)
5. Agrawal R., Mannila H., Srikant R., Toivonen H. et Inkeri Verkamo A.: Fast Discovery of Association Rules, Advances in Knowledge Discovery and Data Mining, AAAI Press, The MIT Press (1996)
6. Guillaume S., Guillet F., Philippé J.: Contribution of the integration of intensity of implication into the algorithm proposed by Agrawal, EMCSR'98, Vienna, vol. 2, pp. 805-810, April (1998)
7. Gras R.: Contribution à l'Étude Expérimentale et à l'Analyse de Certaines Acquisitions Cognitives et de Certains Objectifs Didactiques en Mathématiques, Thèse d'État, Université de Rennes 1 (1979)
8. Gras R., Larher A.: L'Implication Statistique, une Nouvelle Méthode d'Analyse de Données, Mathématiques, Informatique et Sciences Humaines n°120 (1993)
9. Gras R., Ag. Almouloud S., Bailleuil M., Larher A., Polo M., Ratsimba-Rajohn H., Totohasina A.: L'Implication Statistique, Nouvelle Méthode Exploratoire de Données. Application à la Didactique, Travaux et Thèses, édition La Pensée Sauvage (1996)
10. Anderson T.W.: An Introduction to Multivariate Statistical Analysis, Series in Probability and Mathematical Statistics, John Wiley and Son (1984) reviewed edition
11. Fleury L., Masson Y., Gras R., Briand H., Philippe J.: A Statistical Measure of Rule Strength for Machine Learning, 2^d World Conference on the Fundamentals of Artificial Intelligence, WOCFAI, Paris (1995)
12. Fleury L., Briand H., Philippe J., Djeraba C.: Rules Evaluations for Knowledge Discovery in Database, 6th International Conference and Workshop on Database and Expert Systems Applications, DEXA, London (1995)
13. Fleury L., Masson Y.: The Intensity of Implication, a Measurement for Machine Learning, The 8th Int. Conf. On Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA/AEI, Melbourne (1995)
14. Lehn R., Guillet F., Briand H.: Eliminating redundant knowledge in an association rule-based system : An algorithm, EMCSR'98, Vienna, vol. 2, pp. 793-798, April (1998)