International Journal of Intelligent Information Systems
2013; 2(4): 64-69
Published online August 30, 2013 (http://www.sciencepublishinggroup.com/j/ijiis)
doi: 10.11648/j.ijiis.20130204.12

Some extensions of positive and negative rules for discovering basic interesting rules

Nguyen Duc Thuan

Dept. Information Systems of Nha Trang University, Khanh Hoa, Vietnam

Email address:
ngducthuan@ntu.edu.vn

To cite this article:
Nguyen Duc Thuan. Some Extensions of Positive and Negative Rules for Discovering Basic Interesting Rules. *International Journal of Intelligent Information Systems.* Vol. 2, No. 4, 2013, pp. 64-69. doi: 10.11648/j.ijiis.20130204.12

Abstract: Positive reasoning and negative reasoning have been applied to be very useful in practice as clear from the record of many real life applications, especially in medicine. These reasoning mechanisms play important role in cutting the search space, reflecting experts' decision, supporting decision by the cooperation of experts and computers. This paper proposes the concepts of extended negative rule, minimal rule and explores their properties. Furthermore, an algorithm for finding all minimal positive rule and minimal negative rule is given. This algorithm is effective to discover positive and negative rules which have not redundant formula. These rules support to deduce the other important positive and negative rules. Experiments are carried out on data sets of UCI machine learning repository to analyze the performance study.

Keywords: Positive Rule, Negative Rule, Minimal Rule, Minimal Positive Rule, Minimal Negative Rule

1. Introduction

Traditional classification rules take the deterministic or probabilistic form as if X then Y ($X\rightarrow Y$). The common feature of both deterministic and probabilistic rules is that they reduce their consequence positively if an example satisfies their conditional parts. The reasoning by these rules can be called positive reasoning. But this form as rules couldn't improve classification accuracy sharply because it is not easy to improve the matured algorithms.

In many applications, the most typical is in medicine, experts use not only positive reasoning but also negative reasoning of selection of candidates, which is represented as if –then rules whose consequences include negative term. New rules in form of $\rceil X\rightarrow \rceil Y$, $X\rightarrow \rceil Y$, $\rceil X\rightarrow Y$ are introduced into data mining.

In recent years negative rule mining got much focus. Many of the algorithms developed for mining positive and negative rule. The concept of negative relationships mentioned for the first time in the literature by Brin et.al (1997). To verify the independence between two variables, they use the statistical test. To verify the positive or negative relationship, a correlation metric was used. Their model is chi-squared based. The chi-squared test rests on the normal approximation to the binomial distribution (more precisely, to the hyper geometric distribution). This approximation breaks down when the expected values are small. Wu et al (2004) discussed how to use negative association rule and designed constraint to reduce the search space. Ji and Tan (2004) studied how to inducing negative and positive rules for gene expression, which also based on association rules. Tsumoto (2001) use positive and negative rules which based on rough sets to predict clinical case. This paper presents some results concerning extended positive and negative rules, which proposed by Tsumoto [7]. The main result is effective to discover positive and negative rules which have not redundant formula. These rules support to deduce the other important positive and negative rules. Moreover, this paper also proposes the concepts of extended negative rule for discovering interesting potential negative rules.

The structure of this paper is as follows:

Section 2 recalls some basic concepts: information system, formula, classification accuracy and coverage, positive and negative rules. Section 3 presents our results and experiments on UCI data sets. We summarize our research and discuss some future work directions in the section 4.

2. Basic Concepts

To illustrate and compare the results easily, in this section we present the basic concepts have been presented in [7].

2.1. Information System

Let U denote a nonempty finite set called the universe and A denote a nonempty, finite set of attributes, i.e., a: $U \to V_a$, $\forall a \in A$, where Va is called the domain of a, respectively. Then a decision table is defined as an information system $S=(U, A \cup \{d\})$.

For example, Table 1 is an information system with U={1,2,3,4,5,6} and A={Age, Location, Nature, Prodrome, Nausea, M1} and d = Class. For Location $\in$ A, $V_{Location}$ is defined as {Occular, Whole, Lateral}.

Table 1. *An example of a data set*

No	Age	Location	Nature	Prodrome	Nausea	M1	Class
1	50 – 59	Occular	Persistent	No	No	Yes	M.c.h
2	40 – 49	Whole	Persistent	No	No	Yes	M.c.h
3	40 – 49	Lateral	Throbbing	No	Yes	No	Migra
4	40 – 49	Whole	Throbbing	Yes	Yes	No	Migra
5	40 – 49	Whole	Radiating	No	No	Yes	M.c.h
6	50 – 59	Whole	Persistent	No	Yes	Yes	Pyscho

2.2. Formula

Atomic formulas over $B \subseteq A \cup \{d\}$ are expression V of the form [a=v], called descriptors over B, where $a \in B$ and $v \in V_a$. The set F(B, V) of formulas over B is the least set containing all atomic formulas over B and closed with respect to disjunction, and negation. For example, [location=occular] is a descriptor of B.

For each $f \in F(B, V)$, f_S is the set of all objects in S satisfy f, defined inductively as follows:

1. If $f = [a = v]$ then $f_S = \{s \in U \mid a(s) = v\}$.
2. $(f \wedge g)_S = f_S \cap g_S$; $(f \vee g)_S = f_S \cup g_S$; $(\neg f)_S = U - f_S$

Example 1: In the preceding example, with Table 1
$f = [Location = Whole] \Rightarrow f_S = \{2, 4, 5, 6\}$
$g = [Location = Whole] \wedge [Nausea = No]$
$\Rightarrow g_S = \{2, 4, 5, 6\} \cap \{1, 2, 5\} = \{2, 5\}$.

2.3. Classification Accuracy and Coverage

2.3.1. Definition 1

Let R and Q denote a formula in F(B,V), R is a formula over the conditional attribute set A, Q is a formula over the decision attribute set D={d}. Classification accuracy and coverage (true positive rate) for R→Q is is defined as:

$$\alpha_R(Q) = \frac{|R_A \cap Q_D|}{|R_A|} \quad (1)$$

$$\kappa_R(Q) = \frac{|R_A \cap Q_D|}{|Q_D|} \quad (2)$$

where |S|, $\alpha_R(Q)$, $\kappa_R(Q)$, denote the cardinaliry of a set S, a classification accuracy of R as to classification of Q, and coverage (a true positive rate of R to Q) respectively.

Example 2: As show in Table 1,
R = [Nausea=Yes] $\Rightarrow R_S = \{3, 4, 6\}$
D = [Class = Migra] $\Rightarrow Q_D = \{3, 4\}$
$\Rightarrow R_S \cap Q_D = \{3, 4, 6\} \cap \{3, 4\} = \{3, 4\}$
$\Rightarrow \alpha_R(Q) = 2/3, \kappa_R(Q) = 1$.

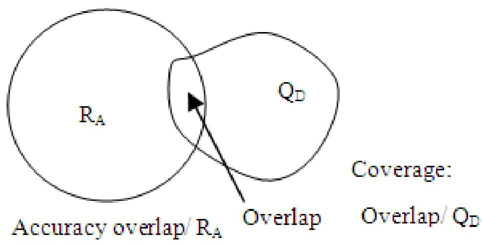

Figure 1. *Venn diagram of accuracy and coverage*

2.4. Atomic Rule

2.4.1. Definition 2

Rule R→Q is a atomic rule if R is a atomic formula.

2.5. Positive Rule

2.5.1. Definition 3

Rule R→Q is a positive rule if
$R = \wedge_j [a_j = v_k]$ and $\alpha_R(Q) = 1.0$.

Thus, Rule R→Q is a positive rule $\Leftrightarrow R_S \subseteq Q_D$
Example 3: As show in Table 1,
R= [Age=50-59]∧[Location=Whole]
$\Rightarrow R_S = \{1, 6\} \cap \{2, 4, 5, 6\} = \{6\}$ and Q=[Class=Psycho]$\Rightarrow D_S = \{6\} \Rightarrow \alpha_R(Q) = 1/1 = 1.0$.

Thus,[Age=50-59]∧[Location=Whole] → [Class=Psycho] is a positive rule.

2.5.2. Definition 4

Rule R→Q is an atomic positive rule if R is atomic formula and R→Q is a positive rule.

Example 4: As show in Table 1,
R= [Nausea=No] $\Rightarrow R_S = \{1, 2, 5\}$ and Q=[Class=M.c.h] $\Rightarrow D_S = \{1, 2, 5\} \Rightarrow \alpha_R(Q) = 3/3 = 1$.

Thus, [Nausea=No] → Q=[Class=M.c.h] is an atomic positive rule.

2.6. Exclusive Rule and Negative Rule

2.6.1. Definition 5 (Exclusive Rule)

Rule R→Q is an exclusive rule if

$R = \vee_j [a_j = v_k]$ and $\kappa_R(Q) = 1.0$

Thus, if rule R→Q is a exclusive rule then $D_S \subseteq R_S \Rightarrow$ Q→R is a positive rule. We also have Q→R ⇔ ¬R→¬Q.

2.6.2. Definition 6 (Negative Rule)

Rule R→Q is a negative rule if $R = \wedge_j \neg[a_j = v_k]$, $Q = \neg[d = v_d]$ and $\forall[a_j = v_k], \kappa_{[a_j=v_k]}([d = v_d]) = 1.0$.

If j=1 then negative rule is an atomic negative rule.

Example 5: As show in Table 1, we have

[Nature=Persistent]∨[Location=Whole]→[Class=M.c.h] is an exclusive rule.

¬[M1=Yes]∧¬[Nausea=no]→ ¬[Class=M.C.h] is a negative rule.

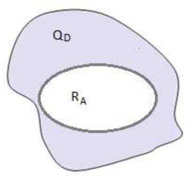

***Figure 2** Venn diag. for positive rule.*

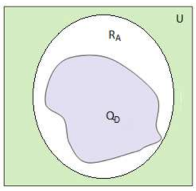

***Figure 3.** Venn diag. for neg. rule.*

3. Main Results

3.1. Extended Negative Rule

From definition 6, the discovery of the negative rules is very difficult, because having to check all atomic formulas involved these rules:

($\forall[a_j = v_k], \kappa_{[a_j=v_k]}(D) = 1.0$), here $D = [d = v_d]$

This condition also missed some sense as a negative rule in practice. Thus, we propose definition extended negative rule follow:

3.1.1. Definition 7 (Extended Negative Rule)

Rule R→Q be a extended negative rule if

$R = \wedge_j \neg[a_j = v_k]$, $Q = \neg[d = v_d]$ and $\kappa_{R'}([d = v_d]) = 1.0$. Here R'$= \vee_j [a_j = v_k]$.

Obviously, negative rule is a special case of the extended negative rule. The difference between negative rule and extended negative rule can be seen Venn diagram:

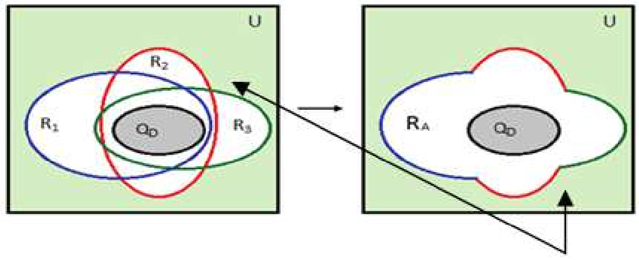

***Figure 4.** Venn diagram for negative rule*

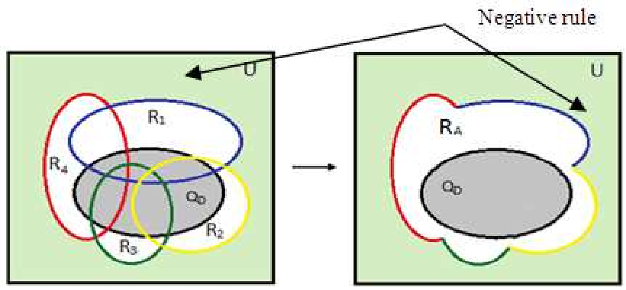

***Figure 5.** Venn diagram for extended negative rule*

∀i=1,2,3, R_i→Q is not a negative rule, but R→Q is a extended negative rule.

Example 6: As show in Table 1, we have

Q= ¬[Class=M.c.h], R_1= f=[Nature=Persistent],

R_2= g = [Location=Whole].

Let Q'= Q ⇒ $Q'_D = \{1,2,5\}$, $f_S = \{1,2,6\}$, $g_S = \{2,4,5,6\}$

$$\kappa_{R_1}(Q') = \frac{|f_S \cap Q'_D|}{|Q'_D|} = \frac{2}{3} \neq 1.0 \; ;$$

$$\kappa_{R_2}(Q') = \frac{|g_S \cap Q'_D|}{|Q'_D|} = \frac{2}{3} \neq 1.0$$

$$\kappa_{R'}(Q') = \frac{|(f_S \cup g_S) \cap Q'_D|}{|Q'_D|} = \frac{3}{3} = 1.0$$

Therefore,

R→Q⇔¬[Naure=Persistent]∧¬[Location=Whole]→¬[Class=M.c.h] is not negative rule, but it is extended negative rule.

3.2. Minimal Rule

3.2.1. Definition 8

Rule R→Q is a minimal rule if any formula R' be constructed from R by removing a component of the formula R then R→Q and R'→Q are not the same type (positive rule, extended negative rule, exclusive rule).

If R→Q is both a positive (negative/extende negative/ exclusive) rule and a minimal rule then R→Q is called minimal (negative/extende negative/ exclusive) rule.

Example 7: As show in Table 1

We have:

a. [*Age* = *50 – 59*] ∧ [*Location* = *Whole*] → [*Class* = *Psycho*] is positive rule.

If remove [*Age* = *50 – 59*] then [*Location* = *Whole*] → [*Class*= *Psycho*] is not a positive rule.

If remove [L*ocation* = *Whole*] then [*Age* = *50 – 59*] → [*Class*= *Psycho*] is not a positive rule.

Therefore, [*Age* = *50 – 59*] ∧ [*Location* = *Whole*] → [*Class* = *Psycho*] is a minimal positive rule.

b. Let R = [M1 = Yes] ∨ [Location = Whole]

⇒ R_S = {1, 2, 5, 6} ∨ {2, 4, 5, 6} = {1, 2, 4, 5, 6},

with D = [*Class* = *M.c.h*] ⇒ D_S = {1, 2, 5} ⊆ R_S

⇒ R → D is an exclusive rule

⇒ (¬[*M1* = *Yes*] ∧ ¬[*Location* = *Whole*]) → ¬[*Class* = *M.c.h*] is an extended negative rule.

But, we have: ¬[*M1* = *Yes*] → ¬[*Class* = *M.c.h*] is an extended negative rule.

Therefore, (¬[*M1* = *Yes*] ∧ ¬[*Location* = *Whole*]) → ¬[*Class* = *M.c.h*] is not minimal extended negative rule.

We have some properties of minimal rule:

3.2.2. Proposition 1

If [a=v]→[d=u] is an atomic positive rule then [a=v] is not appear in any minimal positive rule that the left hand side there is more than one formula.

3.2.3. Proposition 2

If [a=v]→[d=u] is an atomic positive rule then any formula:

$$R=\wedge_j [a_j = v_k]$$

we have [a=v]∧R→[d=u] is an positive rule.

3.2.3. Proposition 3

If ¬[a=v]→ ¬[d=u] is an atomic negative rule then [a=v]] is not appear in any minimal negative rule that the left hand side there is more than one formula.

Proposition 1, proposition 2 and proposition 3 are obtained from definitions 1, 3, 6, 8.

3.2.4. Proposition 4

If $f = [a = v], D = [d = u]$ and $f_S \cap D_S = 0$ then any R∈F(B,V), we have:

a. $f \wedge R \rightarrow D$ is not positive rule.

b. Let $R' = f \vee R$ then $\neg R' \rightarrow \neg D$ is not minimal negative rule.

Proof

a. From $f_S \cap D_S = 0$, it is obviously that $f \wedge R \rightarrow D$ is not positive rule.

b. Suppose $\neg R' \rightarrow \neg D$ is a minimal negative rule

⇒ $D_S \subseteq R'_S$, we have and $f_S \cap D_S = 0$ ⇒ $D_S \subseteq R_S$ ⇒$\neg R \rightarrow \neg D$ is a negative rule, a contradiction (*because* $\neg R' \rightarrow \neg D$ *is a minimal negative rule*). The proof is complete.

3.2.5. Algorithm for Finding all Minimal Positive and Minimal Negative Rules

Tsumoto (2005) gave an algorithm for deduce all positive and negative rules based on rough set theory. Xindong Wu, ChengQi Zhang and Shichao Zhang (2005), Alatas và Akin (2006), presented algorithms mining of both positive and negative associate rules. Tinghuai Ma, Jiazhao Leng, Mengmeng Cui and Wei Tan (2009) proposed RGCA algorithm to deduce negative and positive rules. The RGCA algorithm based on rough set, records are processed item-by-item. In RGCA algorithm, for the rule reduction that based on classification attribute, it does not need to deal with records one by one so that to improve calculation efficiently.

In this paper, the main result is effective to discover positive and negative rules which have not redundant formula. From proposition 1, 2, 3, 4, we proposed an algorithm for finding all minimal positive and negative rules as follow:

Input: S=(U,A∪{d}); L = {[a_j=v_k] | a_j∈ A, v_k∈ V_{aj}}; D = [d= d_i]

Output: RPos, RNeg: set of minimal positive rule and set of minimal negative rule respectively

Method

L_a = L

For each d_i ∈ Dom(d) do

{*//induct atomic positive and negative rules and candidate set can generate minimal positive and negative rules}*

L_p = L_a; L_n = L_a

For each R = [a_j = v_k] in L_a do

{ Caculate $\alpha_R(D)$ and $\kappa_R(D)$;

If $\alpha_R(D)$ = 1.0 then

{ Pos = Pos ∪ {R};

L_a = L_a – {R};

L_p = L_p – {R};

L_n = L_n – {R}; }

If $\kappa_R(D)$ = 1.0 then

{ Neg = Neg ∪ {R};

L_n = L_n – {R}; }

If $\alpha_R(D) = \kappa_R(D) = 0$ then

{ L_p = L_p – {R};

L_n = L_n – {R}; }

}

For i = 2 to n do // n =|A|

{LP_i = { $\overset{i}{\underset{j=1}{\wedge}} R_j \mid R_j \in L_p, R_k \neq R_m, when\ k\neq m$}

LN_i= { $\overset{i}{\underset{j=1}{\vee}} R_j \mid R_j \in L_p, R_k \neq R_m, when\ k\neq m$}}

For each R = $\wedge_i$[a_j = v_k] in LP_i do

{ if $\alpha_R(D)$ = 1.0 then

{ Pos = Pos ∪ {R}; LP_i = LP_i – {R};}

if $\alpha_R(D)$ = 0 then LP_i = LP_i – {R};}

For each R = $\vee_i$[a_j = v_k] in LN_i do

if $\kappa_R(D)$ = 1.0 then

{Neg = Neg ∪ {R}; LN_i = LN_i– {R};}

// Generate minimal positive and negative rules
LPos = {R→[d=d_i]| R∈Pos};
LNeg={¬R→¬[d=d_i]|R∈Neg};}
The worst-case complexity of inducing minimal positive and minimal negative rules is:

$$O(\sum_{i=1}^{|A|} |dom(a_i)| \times |dom(d)|)$$

3.3. Experiment Results

We use data sets from UCI repository of machine learning databases and domain theories to verify the extended negative rule concept and the presented algorithm above.

Table 2 summarizes our analysis results, which shows the number of negative rules and extended negative rules.

For Car Evaluation dataset, if we consider only two attributes Buying and Maint then the number of negative rule (definition 6) is zero, but there is one extended negative rule (definition 7):

¬[Buying=low] ^ ¬[Maint=low] → ¬[Class=good].

Similarly, if we consider three attributes Buying, Maint, Door, the number of negative rule (definition 6) is zero, but there are four extended negative rules (definition 7):

1. ¬[Buying=low] ^ ¬[Maint=low] ^ ¬[Doors=2] → ¬[Class=good]
2. ¬[Buying=low] ^ ¬[Maint=low] ^ ¬[Doors=3] → ¬[Class=good]
3. ¬[Buying=low] ^ ¬[Maint=low] ^ ¬[Doors=4] → ¬[Class=good]
4. ¬[Buying=low] ^ ¬[Maint=low] ^ ¬[Doors=5more] → ¬[Class=good]

We can see that the expansion of the negative rules to avoid omitting the rule mean in practice.

4. Conclusion and Future work

Our study is mainly focused on mathematical properties of negative rules and minimal rules. In this paper we introduced the concept of extended negative rules, which extracts hidden knowledgeable useful information from medical databases. Furthermore, an algorithm to generate all minimal positive and minimal negative rules is introduced.

Interestingly, very few have focused on negative association rules due to the difficulty in discovering these rules. Although some researchers pointed out the importance of negative associations, only few groups of researchers proposed an algorithm to mine these types of associations. This not only illustrates the novelty of negative association rules, but also the challenge in discovering them.

Experiment results with some data sets from UCI repository of machine learning databases showed the meaning of these extensions of positive and negative rules.

In future we wish to conduct experiments on some other real datasets and compare the performance of our algorithm with other related algorithms.

Table 2. *Comparision between the number of negative rules and the number of extended negative rules*

The number attributes combined	Attributes	Negative rules	Extended negative rules
Car Evaluation (1728 samples)			
2	Buying, maint	0	1
3	Buying, maint, door	0	4
4	Buying, maint, door, persons	0	12
5	Buying, maint, door, persons, lud_boot	0	36
6	Buying, maint, door, persons, lud_boot, Safety	0	699
Nursery (12960 samples)			
2	parents, has_nurs	1	7
3	parents, has_nurs, form	1	36
4	parents, has_nurs, form, children	1	168
5	parents, has_nurs, form, children, housing	1	576
6	parents, has_nurs, form, children, housing, finance	1	1296
7	parents, has_nurs, form, children, housing, finance, social	0	3888

Table 3. *Comparision between the number of negative rules and the number of extended negative rules*

Data set	The number of minimal positive rules	The number of minimal negative rules
Car Evaluation	913	10
Iris	272	0
Tictactoe	29570	14927

References

[1] B.Kavitha Rani, K.Srinivas, B.Ramasubba Reddy, Dr.A.Govardhan, "Mining Negative Association Rules", International Journal Of Engineering and Technology, Vol.3 (2), 100-105, 2011

[2] Brin S, Motwani R, Silverstein cristiano Beyond market: Generalizing association rules to correlations [A]. Processing of the ACM SIGMOD Conferrence 1997 [m]. New York: ACM Press,265-276, 1997

[3] Chris Cornelis, Peng Yan, Xing Zhang, Guoqing Chen, Mining Positive and Negative Association Rules from Large Databases, IEEE Conference Cybernetics and Intelligent Systems, BangKok, June 2006

[4] D.R. Thiruvady and G.I.Webb, Mining Negative Rules using GRD, Advances in Knowledge Discovery and Data Mining, Lecture Notes on Computer Science Vol 3056, pp 161-165. Proceedings of the 8th Pacific-Asia Conference, PAKDD (2004) Sydney, Australia, May 26-28, 2004

[5] E.Ramaraj, N.Venkatesan, "Positive and Negative Association Rule Analysis in Health care Databases", IJCSNS International Journal of Computer Science and Network Security, Vol 8 (10), October, 2008

[6] Maria-Luiza, Antonie Osmar R. Za¨ıane, Mining Positive and Negative Association Rules:An Approach for Confined Rules, 8th European Conference on Principles and Practice of Knowledge Discovery in Databases, Pisa, Italy, September 20-24, 2004

[7] Nikhil R.Pal, Lakhmi Jain (Eds), Advanced Techniques in Knowledge Discovery and Data Mining, Springer-Verlag London Limited , 233-252, 2005

[8] Tinghuai Ma, Jiazhao Leng, Mengmeng Cui and Wei Tian, "Inducing Positive and Negative Rules Based on Rough Set", Infomation Technology Journal 8(7): 1039-1043, 2009

[9] Xingdong Wu, ChengQi Zhang, Shichao Zhang, "Efficient Mining of Both Positive and Negative Association Rules", ACM Transactions on Information Systems, Vol. 22, No. 3, July, 2004

[10] Zdzislaw Pawlak, Rough sets, Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, London 1991