Classification Methods based on Formal Concept Analysis

Nida Meddouri[1] and Mondher Meddouri[1,2]

[1] Research Unit on Programming, Algorithmics and Heuristics - URPAH
Faculty of Science of Tunis – FST
University of Tunis-El Manar
Campus Universitaire, EL Manar, 1060, Tunis, Tunisia
nmeddouri@gmail.com

[2] Department of Mathematic and Computer Sciences,
National Institute of Applied Sciences and Technology of Tunis – INSAT
University of 7th November at Carthage
Centre Urbain Nord, B.P. 676, 1080 TUNIS CEDEX, TUNISIE
mondher.maddouri@fst.rnu.tn

Abstract. Supervised classification is a spot/task of data mining which consists in building a classifier from a set of examples labeled by their class (learning step) and then predicting the class of new examples with a classifier (classification step). In supervised classification, several approaches were proposed [16] such as: Induction of Decision Trees [18], and Formal Concept Analysis [7]. The learning of formal concepts is based, generally, on the mathematical structure of Galois lattice (or concepts lattice). The complexity of generation of Galois lattice, limits the application fields of these systems [16]. In this paper, we present several methods of supervised classification based on Formal Concept Analysis. We present methods based on concept lattice, sub lattice and finally the cover of concepts.

Keywords: Formal Concept, Classification rules, Machine Learning, Data mining.

1 Introduction

Formal Concept Analysis is a formalization of the philosophical notion of concept defined as a couple of extension and comprehension [16]. The comprehension (called also intention) makes reference to the necessary and sufficient attributes which characterizes this concept. The extension is a set of examples which made it possible to find out the concept [16], [17].

The classification approach based on Formal Concept Analysis is a symbolic approach allowing the extraction of correlations, reasons and rules according to the concepts discovered from data. Classification is a process made up of two steps. In the learning step, we organize the information extracted from a group of objects in the

© Radim Belohlavek, Sergei O. Kuznetsov (Eds.): CLA 2008 (Posters), pp. 9–16,
Palacký University, Olomouc, 2008.

form of a lattice. In the classification step, we determine the class of new objects that are more or less deteriorated, based on the extracted concepts. Many learning methods based on Formal Concept Analysis were proposed, such as: GRAND [16], LEGAL [12], [16], GALOIS [3], [4], [16], RULEARNER [16], [19], CIBLe [6], [16], CLNN&CLNB [5], [16], [21], IPR [14], NAVIGALA [9], [10], [11] and more recently CITREC[5].

Unfortunately, systems based on Formal Concept Analysis encountered some problems such as an exponential complexity (in the worth case), a high error rate and an over-fitting. Fortunately, boosting algorithms are known by improving the error rate of any single learner.

In section 2, we present the basic notions of Formal Concept Analysis used for classification purposes. In section 3, we present several methods of supervised classification based on Formal Concept Analysis by evoking notions of concept lattices [10], [16], sub-lattice [8], [10], [16] and finally the cover of concept [14] , [15]. In section 4, a theoretical comparison of these methods is presented. Concluding remarks with future work directions are also given.

2 Basic notions of Formal Concept Analysis

A formal context is a triplet k = <*O*, *A*, *R*>, where $O = \{o_1, o_2, \dots, o_n\}$ is a finite set of elements called objects (instances, examples), $A = \{a_1, a_2, \dots, a_m\}$ a finite set of elements called properties (attributes) and *R* is a binary relation defined between *O* and *A*. The notation (g,m) , or R(g,m) = 1, means that object g verifies property m in relation *R* [2], [7]. The context is often represented by a cross-table or a binary-table as shown in Table 1 (taken from [16]).

Table 1: Binary formal Context describing the relation *R* [16]

O\A	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	***CLASS***
o_1	1	1	1	1	1	1	1	0	***1***
o_2	1	1	1	1	1	1	0	1	***1***
o_3	1	1	1	1	1	0	1	1	***1***
o_4	1	1	1	1	0	1	0	0	***1***
o_5	1	1	0	1	1	0	1	0	***2***
o_6	1	1	1	0	1	0	0	1	***2***
o_7	1	0	1	0	0	1	0	0	***2***

Let $B \subseteq O$ and $C \subseteq A$ two finite sets. For both sets B and C, operators φ (B) and δ (C) are defined as [4]:

- $\varphi(B) := \{m \mid \forall g, g \in B \rightarrow (g,m) \in R\}$.
- $\delta(C) := \{g \mid \forall m, m \in C \rightarrow (g,m) \in R\}$.

Operator φ defines the properties shared by all elements of B. Operator δ defines objects sharing the same properties included in set C. Operators φ and δ define a Galois connection between sets B and C [6]. The closure operators are $X'' = \delta \circ \varphi(X)$ and $Y'' = \varphi \circ \delta(Y)$. Finally, the closed sets (X, Y) are defined as if $X = \delta \circ \varphi(X)$ and $Y = \varphi \circ \delta(Y)$ [1], [2].

A formal concept of the context <O, A, R> is a pair (B, C), where $B \subseteq O$, $C \subseteq A$, and f (B) =C and h (C) =B. Sets B and C are called respectively the domain (extent) and range (intent) of the formal concept [2].

From a formal context <O, A, R>, we can extract all possible concepts. In [8], we prove that the set of all concepts may be organized as a complete lattice (called Galois lattice), when we define the following partial order relation << between two concepts, $(B_1, C_1) << (B_2, C_2)$ if and only if $(B_1 \subseteq B_2)$ and $(C_2 \subseteq C_1)$. The concepts (B_1, C_1) and (B_2, C_2) are called nodes in the lattice.

Figure 1 represents the concept lattice (Galois lattice) of the context presented in Table 1 taken from [16].

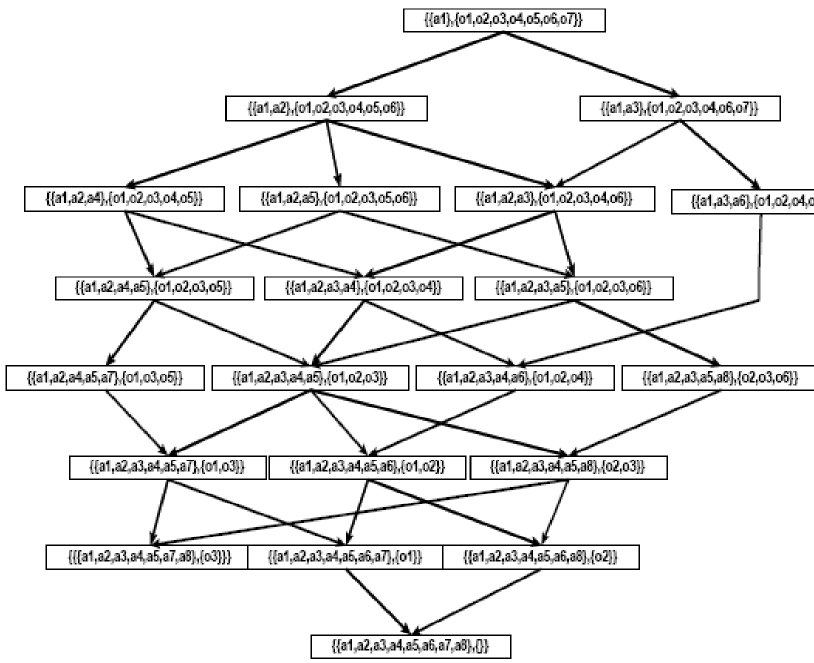

Fig. 1. The Galois lattice trained from the context of Table 1

3 FCA based methods for classification

In this section, we present several methods of supervised classification based on Formal Concept Analysis by evoking notions of concept lattices [8], [16], sub-lattice [6] and finally the cover of concept [14] , [15].

3.1 Concept lattice based classification

The classification has to determinate the class of new deteriorated objects. The Galois lattice can be seen as a space of search in which we evolve level to another, by validation of the characteristics associated to the concepts [8]. Navigation begins from the minimal concept where all the classes are candidates with the recognition and no attributes are validated. Then we have to progress concept by concept in the Galois lattice by validation of new attributes and consequently reducing the whole of remaining objects.

Many systems uses lattice concept based classification such as: GRAND [16], RULEARNER [16], [20], GALOIS [3], [4], [11], NAVIGALA [8], [9], [10] and CITREC [5]. For example, the authors in [16] have applied the system GRAND to the previous formal context. They obtained only one generated rule:

IF a1 $\wedge$ a2 $\wedge$ a3 $\wedge$ a4 THEN ***1***.

The common limit for the systems based on lattice concept, is the exponential complexity (temporally and spatially) of generating the lattice [16]. Then the navigation in huge research space becomes hard [13]. For these reasons, many researchers are oriented to the sub-lattice based classification.

3.2 Sub-lattice based classification

Systems like LEGAL [12], [16], CIBLe [6], [16] and CLNN&CLNB [5], [16], [21], have the characteristic to build sub-lattice, which reduces their theoretical complexity and their times of execution. A sub-lattice is a reflexive and transitive reduction of Galois lattice [9]. Classification based on sub-lattice is similar to that started from a lattice. The major difference between lattice based classification and sub-lattice based classification is the number of concepts generated.

For example, the authors of [16] have applied the system CIBLe to the previous formal context. They obtained the sub-lattice of figure 2. To extract rules from the sup lattice, the authors of [16] are looking for the pertinent concepts.

From the sub-lattice built by CIBLe, there are only 3 rules generated, characterized by a rectangular representation (means pertinent concept) in figure 2. The rules are obtained by associating each selected concept to a major class giving by a PPV function:

IF a1 $\wedge$ a2 $\wedge$ a4 THEN ***1***.
IF a1 $\wedge$ a2 $\wedge$ a3 THEN ***1***.
IF a1 $\wedge$ a3 $\wedge$ a6 THEN ***1***.

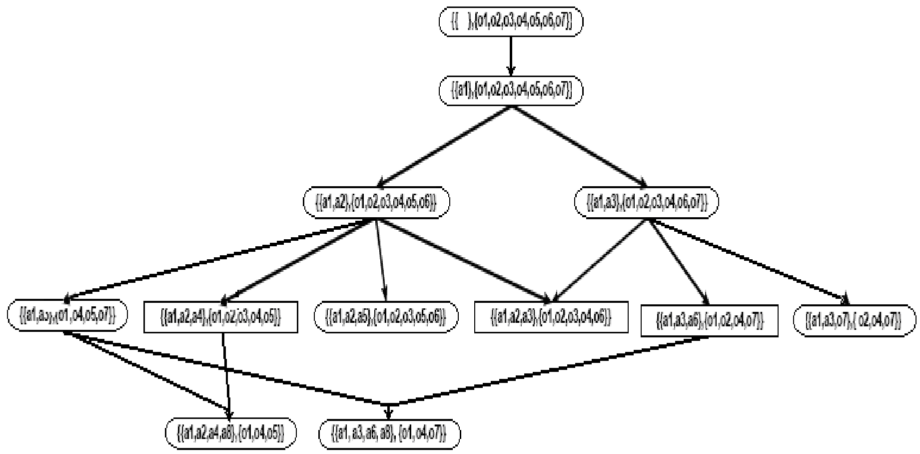

Fig. 2. The sub-lattice built by CIBLe on the previous context (h=3)[1] [16].

3.3 Cover based Classification

A concept cover is a part of the lattice containing only pertinent concepts [14], [15]. The construction of cover concept is based on heuristic algorithms which reduce the complexity of learning. The concepts are extracted one by one. Each concept is given by a local optimization of measure function (giving Pertinent Concept). However, rules are obtained from concepts. Each pertinent concept with associated major class constructs a rule.

The first method generating a concept cover was the so-called IPR (Induction of Product Rules [14], [15]). Each pertinent concept induced by IPR is given by a local optimization of entropy function. The sets of pertinent generated concepts are sorted from the more pertinent to the less pertinent and each pertinent concept induces a rule as described previously.

Fig. 3. The Cover Concept Built by IPR from the Initial Context

For example, applying the IPR method to the previous formal context; we obtain the concepts of figure 3.

---

[1] CIBLe is a parametrable system, which limits the construction of the sub-lattice concept by indicating the level 'h'. In the associated example, we have fixed h = 3.

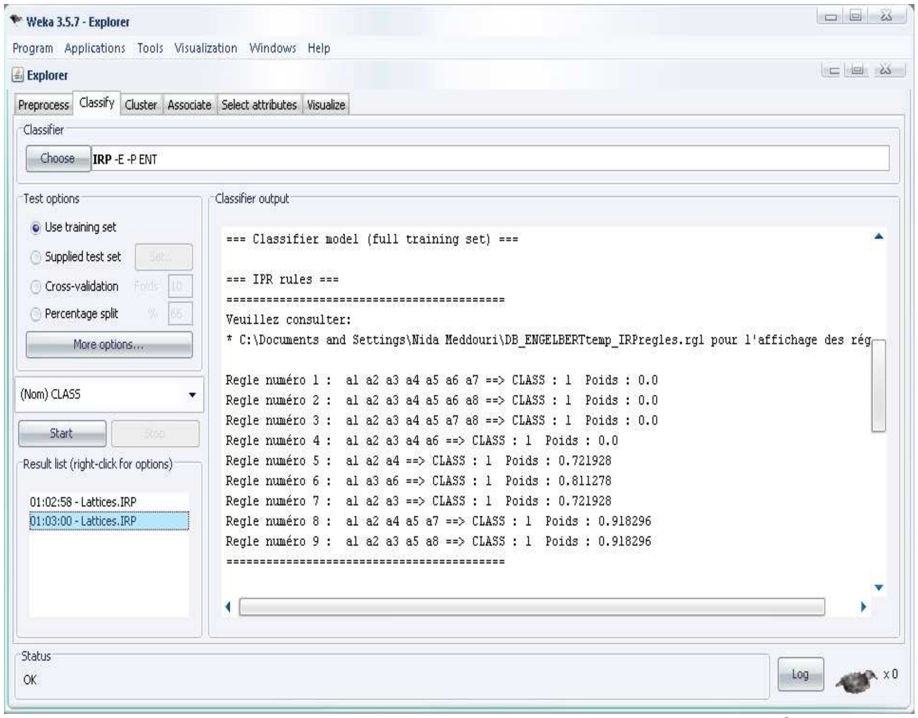

Fig. 4. Rules generated with IPR from the Initial Context under WEKA[2]

As shown in figure 4, IPR generates 9 rules more than GRAND (only one rule) or CIBLe (3 rules) from the initial context.

4. Discussion

In this paper, we have been interested by the classification approach based on Formal Concept Analysis. We have presented the methods GRAND (based on concept lattice), CIBLe (based on semi lattice of concepts) and IPR (based on cover of pertinent concepts). To compare the presented approaches, table 2 presents a theoretical comparison of these methods. Compared to the complexities of the other methods [16], we remark that the IPR method has the less temporally complexity. We remark also that the combination of methods is not largely used.

Known the disadvantages of these listed methods, especially their great complexity, we think that future works should focus on designing new FCA based methods that fix these problems. Certainly, such methods should be faster in order to compare it with well used classification methods (Decision trees, Nearest Neighbor, etc). Future work can focus also on the quality of the classification rules. In fact we plan to evaluate these methods on many machine learning datasets. Accordingly, we

[2] Available at http://www.cs.waikato.ac.nz/ml/Weka

think that we can improve the error rate of the FCA based methods by acting on the voting methods and the function allowing the selection of the best concepts.

Table 2. Theoretical comparison of the presented methods.

Systems	GRAND	CIBLe	IPR
Authors	OOSTHUIZEN	LIQUIRE M. MEPHUNGUIFO E.	MADDOURI M.
Kind of lattice	Complete	Sub-Lattice	Cover
Algorithms	Oosthuizen	Bordat	[Maddouri 2004]
Data	Binary	Numerical values Symbolic values	Binary
Number of classes	Multi-classes	Multi-classes	Multi-classes
Selection of concepts	Maximum Coherence	Height, function selection	Entropy
Combination of methods	No	K-PPV	No
Knowledge learned	Rules	Rules	Rules
Classification	Vote	K-PPV	More weighted rules
Theoretical complexity	$O(2^k \times k^4)$ K = min (m, n)[3]	O ($\|L\| \times m^3$) with $\|L\|$: sub lattice	$O(n^2 \times m^2 \times (m + n))$

References

1. I. Bouzouita, S. Elloumi. Integrated Generic Association Rule Based Classifier. 18th International Conference on Database and Expert Systems Applications: DEXA'07, Workshops 2007, p514-518, Regensburg, Germany (2007).
2. J.P Bordat. Calcul pratique du treillies de Galois d'une correspondence, Mathématiques & Sciences Humaines winter .96, pp31-47 (1986).
3. C. Caprineto, G. Romano. GALOIS An order-theoretic approach to conceptual clustering. In proceedings of ICML'93, pp33-40, Amherst, USA (July1993),.
4. C. Caprineto, G. Romano. GALOIS: A lattice conceptual clustering system and its application to browsing retrieval. Proceedings of the Tenth International Conference on Machine Learning, MA: Morgan Kaufmann, pp 33-40, Amherst, USA (1993).
5. B. Douar, C. Latiri, Y. Slimani. Approche hybride de classification supervisée à base de treillis de Galois: application à la reconnaissance de visages. 8èmes journées d'Extraction et de Gestion des Connaissances: EGC'08, pp309-320, Nice, France (2008),.
6. H. Fu et E. M. NGUIFO. Un algorithme de génération des itemsets fermés pour la fouille de données. 4èmes journées d'Extraction et de Gestion des Connaissances: EGC'04, Clermont Ferrand (2004).
7. B. Ganter and R. Wille. Formal analysis Concept: Mathematical Foundations, Springer Verlag (1997).
8. S. Guillas, K. Bertet, J-M. Ogier. Reconnaissance de symboles bruités à l'aide d'un treillis de Galois, Colloque International Francophone sur l'Ecrit et le Document : CIFED'06, pp85-90, Fribourg, Suisse (2006).
9. S. Guillas, K. Bertet, J-M. Ogier. Extension of Bordat's algorithm for attributes. Concept Lattices and Their Applications: CLA'07, Montpellier, France (2007).

[3] m' means number of examples and 'n' means number of attributes.

10. S. Guillas, K. Bertet, J-M. Ogier. Comment utiliser le treillis de Galois en reconnaissance d'images ? Atelier ECOI, 6èmes journées d'Extraction et de Gestion des Connaissances: EGC'06, pp31-36, Lille, France (2006).
11. M. J. Kearns, L. G. Valiant, A. Ehrenfeucht, D. Haussler. A General Lower Bound on the Number of Examples Needed for Learning. Proceedings of the 1988 Workshop on Computational Learning Theory: COLT'88, pp139-154, MIT, MA, USA (1988).
12. M. Liquiere, E. Mephu nguifo. « LEGAL »: learning with Galois Lattice. Actes des JFA, pp93-113, Lannion, France (Avril 1990).
13. M. Liquiere, J. Sallatin. Structural machine learning with Galois lattices and Graphs. In: Shavlik, J.W. (ed.), International Conference on Machine Learning: ICML'98, pp305-313, Madison, Wisconsin, USA (1998).
14. M. Maddouri. Contribution à l'apprentissage conceptuel: une approche incrémentale d'induction de règles à partir d'exemples. Thèse de doctorat en Informatique, soutenue le 29/05/2000. Faculté des Sciences de Tunis (2000).
15. M. Maddouri: Towards a machine learning approach based on incremental concept formation. Intelligent Data Analysis, Volume 8, Issue 3, pp267-280 (2004).
16. E.M. Nguifo, P. Njiwoua. Treillis de concepts et classification supervisé .Technique et Science Informatiques : TSI, Volume 24, Issue 4, pp449–488 (2005).
17. A. Napoli, Extraction de connaissances, gestion de connaissances et web sémantique. INFORSID'03, Nancy, France (2003).
18. K. Nehme, F. Douzidia. Analysis of the formal concepts applied to research of information. IFT6255 (April 2003).
19. J.R. Quinlan. Induction of decision Trees. Machine learning, Volume1, pp81-106 (1986).
20. M. Sahami. Learning classification Rules Using Lattices. N. Lavrac and S. Wrobel eds., pp343-346, Proc ECML, Heraclion, Crete, Greece (Avril 1995).
21. Z. Xie, W. Hsu, Z. Liu, M. Lee. Concept Lattice based Composite Classifiers for high Predictability. Artificial Intelligence, vol. 139, pp253-267, Wollongong, Australia (2002).