Fuzzy Clustering for Finding Fuzzy Partitions of Many-Valued Attribute Domains in a Concept Analysis Perspective

Yassine Djouadi[1] Basma Alouane[2] Henri Prade[1]

1. IRIT, University Paul Sabatier, Toulouse
118 Route de Narbonne, 31062 Toulouse, Cedex 9, France
2. LIFAB Laboratory, University of Boumerdes
Avenue de l'Independance, Boumerdes, Algeria
Email: djouadi@irit.fr, B_alouane@umbb.dz, prade@irit.fr

Abstract— *Although an overall knowledge discovery process consists of a distinct pre-processing stage followed by the data mining step, it seems that existing formal concept analysis (FCA) and association rules mining (ARM) approaches, dealing with many-valued contexts, mainly focus on the data mining stage. An "intelligent" pre-processing of input contexts is often absent in existing FCA/ARM approaches, leading to an unavoidable information loss. Usually, many-valued attribute domains need to be first fuzzily partitioned. However, it is unrealistic that the most appropriate fuzzy partitions can be provided by domain experts. In this paper, an unsupervised learning stage, based on Fuzzy C-Means algorithm, is proposed in order to get fuzzy partitions that are faithful to data for quantitative attribute domains, and consequently for avoiding the loss of valuable association rules due to the use of empirical fuzzy partitions. More precisely, the paper reports an experiment where it is shown that some rules are no longer found because their support or confidence is too low when using such empirical partitions. Experimental results show that the learned fuzzy partition outperforms human expert fuzzy partitions. More generally, the paper provide discussions about the handling of many-valued attributes in both fuzzy FCA and fuzzy ARM.*

Keywords— Many-valued formal contexts, fuzzy partitions, fuzzy C-means, association rules.

1 Introduction

Association rules mining (ARM for short) [1] is one of the most widely used data mining technique. This model of knowledge represents the patterns of co-occurrence of items in a collection of transactions. Typically, input data set is in the form of sets of items called transactions (e.g. an item may be things we buy in a market). Thus an example of an association rule is an expression of the form "*beer*" $\wedge$ *"sausage"* $\Rightarrow$ *"mustard"* which represents the fact that purchasing beer and sausage implies purchasing mustard with some degree. It is important to point out that, in the classical setting, considered input data sets (called *contexts* in the rest of this paper) are binary (i.e. are expressed in terms of binary attributes).

Formal concept analysis [2] (FCA for short) consists also of learning some knowledge representation model in an unsupervised way. In the classical setting of this theory, the considered input data set is called *formal context* (also called *context* in this paper). It consists of a binary relation between a set of objects and a set of properties. This relation is usually represented as a table with rows corresponding to objects, columns corresponding to properties (or conversely), and table entries containing 1's or 0's depending on whether an object has or not the corresponding property.

It appears that, in their classical setting, both FCA and ARM are concerned with crisp binary contexts. However real-world data usually contain heterogenous kinds of data (binary, categorical, quantitative, etc...). In order to generalize binary settings to many-valued contexts (i.e. with quantitative or categorical attributes) it is usually suggested to partition quantitative attribute domain into many fuzzy intervals.

Few approaches propose a complete pre-processing stage that is intended to get an optimal fuzzy partition. In [3], authors propose a genetic algorithm-based clustering method that adjusts the centroids of the clusters, which are to be handled as midpoints of the triangular fuzzy partitions. Another approach, also based on genetic algorithms, is proposed in [4]. Based on a linguistic representation model, authors propose to perform a genetic lateral tuning of the membership functions. Authors consider also that each attribute has a predefined number of linguistic terms associated to it. For example, *Low*, *Middle* and *High* are the fuzzy linguistic terms covering the domain of the attribute *Age*.

Concerning these approaches, one may notice that they consider fuzzy partitions restricted to triangular membership functions. On the other hand, the number of clusters has to be empirically fixed for each attribute. Then, the fuzzy partitioning may be considered as ad-hoc to some extent and is far to be information lossless. In other words, the distribution of the data into the context is not taken into account for the assessment of discovered patterns. For this purpose, we present an approach that consists of learning the fuzzy partitions by means of the well-known Fuzzy C-Means algorithm. We also consider cluster validity measures in order to get an optimal multi-prototype context representation.

The paper is organized as follows. Section 2 gives some basic notions of FCA and ARM frameworks, while both crisp and fuzzy partitioning methods are reviewed. The next section highlights the information loss problem with empirical or expert-based fuzzy partitions. In the fourth section, data faithful fuzzy partitions are learned using Fuzzy C-Means algorithm. Experimental results on the *adult* database (UCI Machine Learning Repository) show that the proposed approach avoids information loss and outperforms in terms of support and confidence measures, expert-based partitions. Finally, we conclude and give future perspectives.

ISBN: 978-989-95079-6-8

2 Basic Notions

This section introduces main notions and terminology for both FCA and ARM paradigms and then, presents quantitative context representation and its inherent sharp boundary problem.

2.1 Association Rules Mining

Formally, let $\mathcal{X}$ be a set of objects (called also transactions) $\left(\mathcal{X} = \{x_1, x_2, ..., x_{|\mathcal{X}|}\}\right)$ and let $\mathcal{Y}$ be a set of items $\left(\mathcal{Y} = \{y_1, y_2, ..., y_{|\mathcal{Y}|}\}\right)$. A context $\mathbb{K}$ is denoted as $\mathbb{K} = (\mathcal{X}, \mathcal{Y}, \mathcal{R})$ where $\mathcal{R}$ is a binary relation ($x\mathcal{R}y$ indicates that item y satisfies the transaction x). An association rule is of the form $A \Rightarrow B$ where $A, B \subseteq \mathcal{Y}$, $A \cap B = \emptyset$ and, $A, B \neq \emptyset$. The intended meaning of this rule is that the presence of all of the items of A in a transaction implies the presence of all of the items of B in the same transaction. Association rules are assigned support and confidence measures for a given context. The support defines the percentage of transactions that satisfy (contain) both A and B. Whereas the confidence expresses the conditional probability that B is satisfied given that A is satisfied. These measures denoted respectively `Sup` and `Conf` are defined as:

$$\texttt{Sup}\,(A \Rightarrow B) = \frac{|\{x \in \mathcal{X} \mid \forall y \in A \cup B,\, x\mathcal{R}y\}|}{|\mathcal{X}|} \quad (1)$$

$$\texttt{Conf}(A \Rightarrow B) = \frac{\texttt{Sup}(A \Rightarrow B)}{\texttt{Sup}(A)} \quad (2)$$

Hence, an itemset A is said frequent, if its support is greater than a fixed threshold ξ (i.e. $\texttt{Sup}(A) \geqslant \xi$). A large variety of representational and computational approaches for association rules mining have been published in the literature in the case of binary items. Among them, frequent-based [5] and closed-based [6] approaches. While the former generate the lattice of all frequent itemsets, the latter generate a more condensed representation through the so-called closed itemsets lattice. For instance, based on the notion of closed itemsets, and minimal implication rules base (e.g. Duquenne-Guigues base, Luxemburger base) [7], association rules may be easily inferred from a closed itemsets lattice structure without enumerating them [8]. This kind of approach is closely related to formal concept analysis since the closed itemsets lattice is isomorphic to the concepts lattice which is described below.

2.2 Formal Concept Analysis

Formal concept analysis theory, proposed by [2], deals with a particular kind of analysis of data based on a formal context. Let us consider a similar notation as used above. A set $\mathcal{X}$ of objects, a set $\mathcal{Y}$ of attributes and a formal context $\mathbb{K} = (\mathcal{X}, \mathcal{Y}, \mathcal{R})$ with $\mathcal{R}$ a binary relation ($x\mathcal{R}y$ means that object x verifies attribute y). Let A (resp. B) be a subset of $\mathcal{X}$ (resp. $\mathcal{Y}$), two mappings $\uparrow$ and $\downarrow$ are symmetrically defined:

$$A\uparrow = \{y \mid \forall x \in A,\, x\mathcal{R}y\}$$
$$B\downarrow = \{x \mid \forall y \in B,\, x\mathcal{R}y\}$$

From a formal context, one can construct pairs (A, B) such that $A\uparrow = B$ and $B\downarrow = A$ known as formal concepts, where A and B are called respectively the extent and the intent of the corresponding concept. The set $\mathfrak{B}(\mathbb{K})$ of all formal concepts in the formal context $\mathbb{K}$ is equipped with a partial order (denoted $\preceq$) defined as:

$$(A_1, B_1) \preceq (A_2, B_2) \text{ iff } A_1 \subseteq A_2 \text{ (or, equivalently, } B_1 \supseteq B_2)$$

$(\mathfrak{B}(\mathbb{K}), \preceq)$ forms a complete lattice, called the *concept lattice* of $\mathbb{K}$ (for more details on the lattice structure see [9]).

It appears that ARM and more generally FCA both require, in their classical setting, the same context representation which is a binary context.

2.3 Quantitative Context Representation

In order to generalize binary context settings to many-valued contexts, Srikant et al. [10] suggested to partition a quantitative attribute domain into intervals. In their approach, these authors proposed to map a quantitative or categorical attribute value into a Boolean attribute value in the following way. Let y be an attribute whose domain is quantitative (the principle remains the same with a categorical attribute). Let $Dom\,(y)$ denote the domain of y (according to the relational paradigm, $Dom\,(y)$ is necessarily a finite set). For a given object x we denote by $x[y]$ the value v of y for the object x (i.e. $x[y] = v$).

If the set $Dom\,(y)$ consists of few values, the mapping is straightforward. Conceptually, the authors propose to have instead of the attribute y, as many new attributes as the cardinality of the set $Dom\,(y)$. The new attributes are of the form "$y.v_k$" (where $v_k \in Dom\,(y)$). The values of the new binary formal context are given as:

$$x[y.v_k] = \begin{cases} 1 & \text{if } x[y] = v_k \\ 0 & \text{otherwise.} \end{cases}$$

If the set $Dom\,(y)$ is large, it is partitioned into different intervals $I_1, I_2, ..., I_p$. Attributes of the new binary context are of the form "$y.I_k$" and are given as:

$$x[y.I_k] = \begin{cases} 1 & \text{if } x[y] \in I_k \\ 0 & \text{otherwise.} \end{cases}$$

Note that in the two above cases only particular subsets (singletons or definite intervals) are allowed.

An example of a many-valued context is illustrated in Table 1, while Table 2 gives the corresponding binary mapping.

Table 1: An example of a many-valued context.

Obj\Attr	Age	Marital Status	Native Country	Work Class
Alice	23	single	USA	never-worked
Boris	24	married	USA	private
Cyril	31	married	USA	fed-gov
David	60	divorced	Haiti	self-emp

Table 2: An example of a many-valued context mapping.

Obj\Attr	Age.[20,29]	Age.[30,39]	Age.[40,49]	...
Alice	1	0	0	...
Boris	1	0	0	...
Cyril	0	1	0	...
David	0	0	0	...

ISBN: 978-989-95079-6-8

2.4 Crisp vs. Fuzzy Context Representation

In [10], the authors point out a dilemma between support and confidence measures caused by crisp partitions. On the one hand, if the number of intervals for a quantitative attribute is important, the support for any single interval may be low. Hence, without using larger intervals, some rules involving this attribute may not be found because they lack minimal support. On the other hand, if the intervals are too large, some rules may not have minimal confidence. Indeed, when interval becomes larger, the support of the antecedent A of a rule $A \Rightarrow B$ becomes higher. Consequently, the confidence of the rule i.e. the ratio $\texttt{Supp}\,(A \Rightarrow B)\,/\texttt{Supp}\,(A)$ decreases and the rule may not have minimal confidence.

Besides, in [11, 12] other authors point out some undesirable threshold effects caused by crisp partitions. Such effects are well-known, for instance, from histograms in statistics: A slight variation of the boundary points of the intervals can have a noticeable effect on the histogram induced by a number of observations. Likewise, the variation of a partition can strongly influence the evaluation of association rules.

Considering the above dilemma and the undesirable threshold effects, fuzzy partitions (in fuzzy intervals) have been widely used instead of crisp partitions (crisp intervals). This has led to fuzzy association rules [13, 14, 15] and to fuzzy concepts analysis [16, 17]. For example, Table 3 illustrates a fuzzy representation for the many-valued context given in Table 1, when choosing the fuzzy partition $\mathcal{F}_1$ (see Fig. 1).

Table 3: Fuzzy representation of a many-valued context.

Obj\Attr	Age.Very Young	Age.Young	Age. Middle Age	...
Alice	0.4	0.6	0	...
Boris	0.2	0.8	0	...
Cyril	0	1	0	...
David	0	0	0	...

3 Information Loss in Quantitative Context

There are two distinct tendencies to transform a many-valued context into a fuzzy partitioned one. The former is based on human expertise, or is even given empirically. The latter uses dedicated heuristics (e.g. [16]). These partitioning methods induce important information loss as it will be showed in this section. Let us before, recall the definition of a fuzzy partition.

Definition1.

A fuzzy partition of a set X on a universe U is a family $\mathcal{F}(X) = (F_1, ..., F_n)$ such that:

1. $\forall i = 1..n \quad F_i \neq \emptyset \wedge F_i \neq U$
2. $\forall u \in U, \quad \sum_i F_i(u) = 1$

3.1 Empirical Fuzzy Partitions

Classical fuzzy set theory has a powerful tool to manage granularity, namely linguistic variables, introduced by Zadeh [18]. Fuzzy partitions are generally based on those linguistic variables. However the complete specification of a linguistic variable may obviously differ from a user to another. For example, in a survey of the literature [11, 12, 14, 15] related to fuzzy association rules discovery, one may find different fuzzy representations of the attribute *Age*. These partitions $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$, $\mathcal{F}_4$ are respectively showed in Fig. 1, Fig. 2, Fig. 3, Fig. 4.

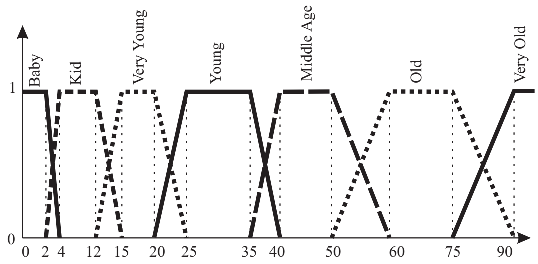

Figure 1: Fuzzy partition $\mathcal{F}_1$ given in [14].

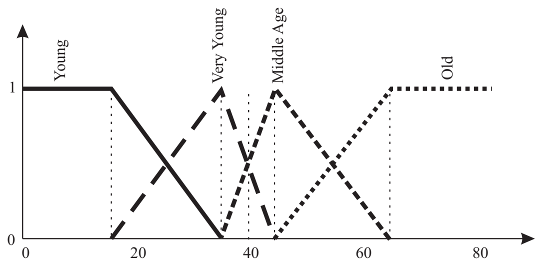

Figure 2: Fuzzy partition $\mathcal{F}_2$ given in [15].

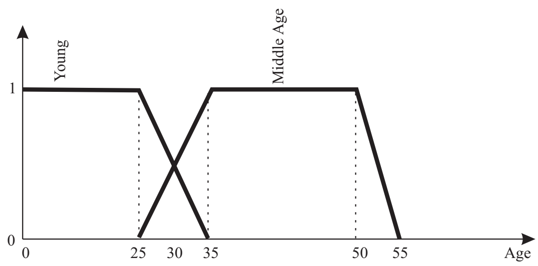

Figure 3: Fuzzy partition $\mathcal{F}_3$ given in [11].

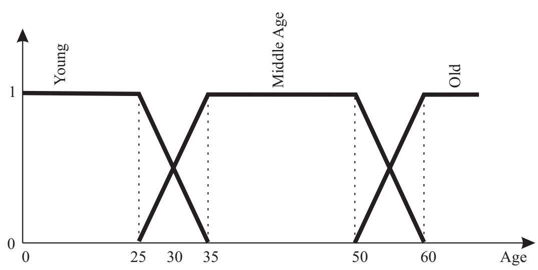

Figure 4: Fuzzy partition $\mathcal{F}_4$ given in [12].

In order to highlight the information loss that may arise with a given fuzzy context compared to another one, we induce all the fuzzy association rules w.r.t a fixed support threshold ξ. The *Adult* benchmark database (UCI Machine Learning Repository) available at http://www.ics.uciedu/mlearn/MLRepositery is taken as a target context for our experiments. Note that, we have restricted the *Adult* context up to six categorical attributes (i.e. *Marital_Status*, *Native_Country*, *Work_Class*, *Salary_Class*, *Education*, *Occupation*) and one quantitative attribute (i.e. *Age*). According to

ISBN: 978-989-95079-6-8

the fuzzy partitions represented respectively in Fig. 1, Fig. 2, Fig. 3, Fig. 4 (namely $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$, $\mathcal{F}_4$) we apply the fuzzy association rules mining algorithm that is fully described in [13] for each of the four fuzzy *Adult* contexts.

Note that, the algorithm in [13] uses a co-occurrence semantics [19] for fuzzy rules. However the information loss phenomenon could be observed as well with gradual or certainty semantics [19] .

Thus among the fuzzy rules induced by each partition $\mathcal{F}_1$-$\mathcal{F}_4$, we have chosen the following association rules to illustrate information loss:
$R1$:(Age.young)⇒(MaritalStatus.single)
$R2$:(Age.middle)⇒(WorkClass.private)
$R3$:(Age.young)⇒(WorkClass.private)
$R4$:(Age.middle)∧(WorkClass.private)⇒(NativeCountry.USA)

By analyzing the results, illustrated in Table 4, it may be remarked that:

1. Different support and confidence measures are found. This confers an uncertainty to the final decision process.

2. More important is a clear information loss materialized by the non discovery of many rules. For example, the rule $R1$ does not exist in the fuzzy context $\mathcal{F}_1$ even though this fuzzy context contains the fuzzy label *young*.

Besides, defining such fuzzy partitions may not be intuitive for humans. This is especially true for attributes where there does not exist a more or less standard linguistic vocabulary which fuzzily partition the attribute domain.

Table 4: Information loss for expert fuzzy partitioning.

	$R1$	$R2$	$R3$	$R4$
	Sup/Conf	Sup/Conf	Sup/Conf	Sup/Conf
$\mathcal{F}_1$	$\nvDash$	0.26/0.72	$\nvDash$	0.13/0.90
$\mathcal{F}_2$	0.08/0.77	0.17/0.68	0.31/0.75	0.08/0.88
$\mathcal{F}_3$	0.05/0.98	$\nvDash$	$\nvDash$	0.20/0.92
$\mathcal{F}_4$	0.13/0.59	0.47/0.77	$\nvDash$	0.30/0.94

where the mark $\nvDash$ indicates that the rule could not be discovered since it lacks support.

3.2 Heuristic-Based Contexts Mapping

The second family of approaches deals with a particular generalization of many-valued contexts based on some heuristics depending on the nature of the context. In [16] a "single-pivot" representation is proposed. The so-called pivot is taken as the maximum value in the attribute domain. The equivalent fuzzy context is obtained by dividing the quantitative values by the maximal value of the domain, thus achieving the estimation of a fuzzy proximity to the maximum value. For example, the equivalent fuzzy representation of the attribute *Age* given in Table 1 will be: {(Alice,0.38), (Boris,0.4), (Cyril,0.51), (David,1)}. One may remark that some non uniform data distribution (e.g. with few values near the pivot and almost all values far from it) will lead to an unavoidable information loss for any knowledge discovery process.

As a summary, it appears that empirical, or heuristic-based fuzzy partitioning for data mining applications may be not well-suited. Indeed, in [20] two important stages (i.e. *Processing* and *Transformation*) are intended to get a well-suited representation for the next data mining stage. This is why, we propose a fuzzy learning stage.

4 Learning Fuzzy Partitions

The Fuzzy C-Means algorithm [21, 22] is one of the strongest and widely used fuzzy clustering method. It is based on an iterative optimization of a fuzzy objective function. This function characterizes the partition and is given as:

$$J_m(X,U,V) = \sum_{j=1}^{n}\sum_{k=1}^{c} \mu_{jk}^{m} \left\|x_j - v_k\right\|^2 \tag{3}$$

where $X = \{x_1, ..., x_n\}$ represents the domain to partition, $V = \{v_1, ..., v_c\}$ represents the set of cluster centers and $U = (\mu_{ki})_{n\times c}$ is the fuzzy partition matrix composed of the membership degree of each item x_j in each cluster k. The exponent m is a parameter called *fuzzifier*.

The distance measure usually used is the Euclidian distance between a datum and a prototype (cluster center). Each cluster V_k is determined as:

$$V_k = \frac{\sum_{j=1}^{n} (\mu_{jk})^m x_j}{\sum_{j=1}^{n} (\mu_{jk})^m} \tag{4}$$

Whereas the elements μ_{jk} of the fuzzy partition matrix are given as:

$$\begin{cases} \mu_{jk} = 0 & \text{if } J_j \neq \emptyset \wedge k \notin J_j \\ \sum_{k\in J_j} \mu_{jk} = 1 & \text{if } J_j \neq \emptyset \wedge k \in J_j \\ \mu_{jk} = \frac{1}{\sum_{l=1}^{c}\left(\frac{\|x_j - v_k\|^2}{\|x_j - v_l\|^2}\right)^{\frac{1}{m-1}}} & \text{if } J_j = \emptyset \end{cases} \tag{5}$$

with $J_j = \{k \mid 1 \leqslant k \leqslant c, \|x_j - v_k\| = 0\}$

4.1 Algorithm

The proposed algorithm integrates cluster validity measures [23] in order to obtain an optimal fuzzy partition. It takes as input a minimal (resp. maximal) number of fuzzy partitions allowed, denoted $Cmin$ (resp. $Cmax$). It takes also a fixed threshold tolerance denoted θ. The vector V contains the cluster centers and it is randomly initialized. The algorithm is given as follows:

Begin.
1: $Cmin := 2$; $Cmax := 10$;
2: $\theta := 0.001$;
3: **For** l:=$Cmin$ To $Cmax$ **Do**
4: Initialize $V^{(l)}$ randomly;
5: Compute the partition matrix $U^{(l)}$ with expression (5);
6: Compute the objective function J with (3);
7: Update the cluster centers using (4);
8: Compute the new partition matrix with (5);
9: Compute the new objective function J_{new};
10: If $(J - J_{new}) < \theta$ goto 3 else goto 7;
11: **End For**
12: Determine l s.t. $Vpc(l) = \min(Vpc(k))_{k=Cmin,Cmax}$
13: Return $V^{(l)}$, $U^{(l)}$

ISBN: 978-989-95079-6-8

End.

Definition2.
The equivalent fuzzy representation of a quantitative attribute y corresponds to the pair $(\mathcal{F}_y, U_y)$ where:

$$\mathcal{F}_y = \{y.V_1, .., y.V_k, .., y.V_c\}$$
$$U_y\,(x_i, y.V_k) = \mu_{ik}$$

This definition means also that we have chosen to label each fuzzy cluster k by its prototype (cluster center) V_k. Considering that membership degrees to these fuzzy clusters are determined using a distance measure (i.e. Euclidian distance), the semantics attached to the discovered fuzzy association rules $Y.V_k(\overrightarrow{y}) \Rightarrow Z.V_l(\overrightarrow{z})$ is of the type: *the presence of* $\overrightarrow{y}$*'s near to the prototype* $Y.V_k$*, imply the presence of* $\overrightarrow{z}$*'s near to the prototype* $Z.V_l$.

4.2 Cluster validity measures

Our algorithm is implemented considering two options:
i) the end-user may specify the number c of partitions, and even the center of classes (i.e. $V_1, \ ..., V_c$),
ii) the learning algorithm determines itself the optimal number of classes using the *Vpc* validity measure as done in the above algorithm.

The partition coefficient (*Vpc*) [23] measure has been chosen since this measure is based on the subsethood degrees μ_{jk}. It is given as:

$$Vpc = \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{c}\mu_{jk}^2 \qquad (6)$$

5 Experimental Results

All experiments were performed on a 3.0 Ghz Pentium-IV PC with 512 MB main memory running on Microsoft Windows XP. Algorithms were implemented in C++.

The fuzzy association rules mining algorithm is an A-PRIORI like implementation. Cardinalities of fuzzy sets are computed using a $\Sigma-$count operator. While conjunction in expressions of support and confidence, resp. (1) and (2), is performed using the *min* operator. We have also considered weighted $\alpha-cuts$ for computing support and confidence measures which avoids accumulation of small (irrelevant) cardinalities. The reader is referred to [13] for complete details.

Experiments were performed on the *Adult* database (see section §3.1.). Many objects in the *Adult* context contain missing (null) values. Such objects have been deleted from the initial context which reduces the size of the considered context to 32561 objects. The probability distribution for the *Age* attribute is represented in Fig. 5.

Our fuzzy partition learning process for the *Age* attribute yields two fuzzy clusters and induces the partition:

$$\mathcal{F}_{age} = \{age.25, age.43\}$$

where 25 and 43 correspond to the prototypes (cluster centers). The complete partition is showed in Fig. 6.

According to the partition $\mathcal{F}_{age}$, we show below some new fuzzy association rules that are now discovered.

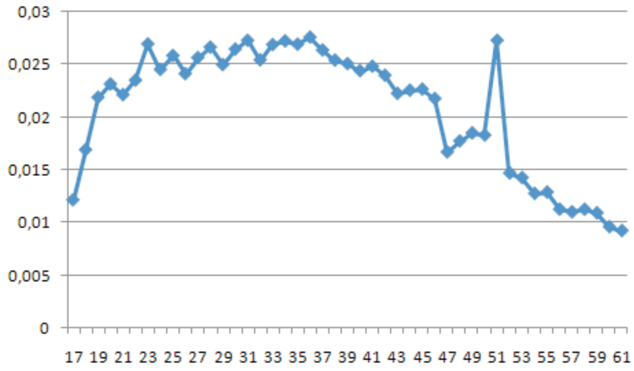

Figure 5: Probability distribution for *Age* attribute in *Adult* context.

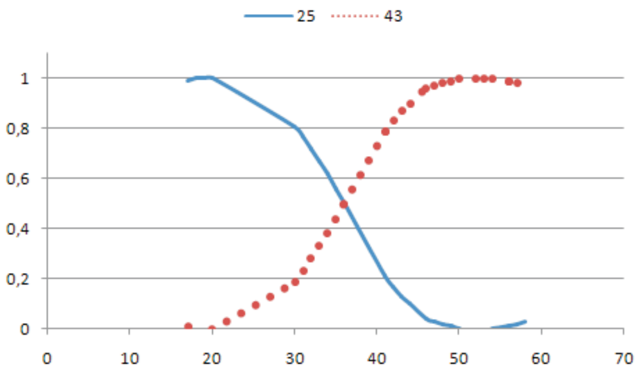

Figure 6: The learned fuzzy partition for *Age* attribute in *Adult* context.

(Age.25)∧(Workclass.*Private*)⇒(NativeCountry.*USA*)
[support=0.32;confidence=0.89]

(Age.25)∧(Workclass.*Private*)∧(NativeCountry.*USA*)⇒
(Salary_Class.*LeastThan50K$*) [support=0.31;confidence=1]

(Age.43)⇒(Workclass.*Private*)
[support=0.24;confidence=0.69]

(Age.25)⇒(Workclass.*Private*)
[support=0.35;confidence=0.94]

(Age.43)∧(Workclass.*Private*)⇒(NativeCountry.*USA*)
[support=0.32;confidence=0.89]

It has been checked that, related to the prototypes, almost all rules are now discovered. For instance, the fourth above rule is a counterpart of the rule $R3$ ((Age.young)⇒(WorkClass.private)), which was not found using the three partitions $\mathcal{F}_1$, $\mathcal{F}_3$, $\mathcal{F}_4$.

6 Conclusion and Perspectives

This paper deals with the transformation stage in a knowledge discovery process. An automatic approach is proposed for the mapping of a many-valued context into a fuzzy context. This mapping is achieved using a fuzzy clustering method. The difficulties of humans for determining fuzzy partitions and fuzzy grades are avoided, which makes things easier for end-users in real life applications. The Fuzzy C-Means algorithm has been

ISBN: 978-989-95079-6-8

chosen as a strong means to get optimized fuzzy partitions. Besides experimental results show that the proposed approach avoids information loss, which departs from the results of almost all existing approaches.

As a first perspective, we intend to enlarge the fuzzy partitions setting by allowing a hierarchical clusters representation. Similar ideas were already explored at least in two different ways. Wolff in [24] proposes a so-called conceptual scale theory which handles many-valued (quantitative) attributes in a fuzzy setting without being restricted to particular subsets (singletons or definite intervals). Bosc et al. in [25] have outlined a knowledge extraction method based on the idea on the fuzzy summarization of data, which may also be used in order to cluster data in a hierarchical-like way.

As pointed out in [26] fuzzy association rules may be also of interest for their increased expressivity. That's why, we intend secondly to investigate other semantics than co-occurrence, namely the one of gradual and certainty rules.

Lastly, it is worth of interest to enlarge fuzzy contexts to incomplete information fuzzy contexts (i.e. with unknown values) [27],[28].

Acknowledgment

The first author is grateful to the A.U.F. (Agence Universitaire de la Francophonie at Brussels-Belgium) who has supported this work under a mobility grant.

References

[1] R. Agrawal, T. Imielinski, and A. Swami. Mining associations rules between sets of items in large databases. *Proc. of the ACM SIGMOD Conf on Management of Data. Washington D.C.*, pages 207–216, 1993.

[2] Rudolf Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. *Rival. I. (Ed): Ordered Sets. Reidel, Dordrecht. Boston*, pages 445–470, 1982.

[3] M. Kaya and R. Alhajj. Genetic algorithm based framework for mining fuzzy association rules. *Fuzzy Sets and Systems*, 152(3):587–601, 2005.

[4] J. Alcala-Fdez, R. Alcala, M.J. Gacto, and F. Herrera. Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms. *Fuzzy Sets and Systems*, 160(7):905–921, 2009.

[5] R. Agrawal and R. Srikant. Fast algorithms for mining association rules. *Proc* 20^{th} *International Conf Very Large Databases. Santiago, Chile*, pages 478–499, 1994.

[6] M.J. Zaki and H.J. Hsiao. Efficient algorithms for mining closed itemsets and their lattice structure. *IEEE Transactions on Knowledge and Data Enginnering*, 17(4):462–478, 2005.

[7] V. Duquenne and J.L. Guigues. Famille minimale d'implications informatives resultant d'un tableau de donnes binaires. *Mathmatiques et sciences humaines*, 24(95):5–18, 1986.

[8] N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal. Efficient mining of association rules using closed itemsets lattices. *Information Systems*, 24:25–46, 1999.

[9] B. Ganter and R. Wille. *Formal Concept Analysis*. Springer Mathematical Foundations, edn, 1999.

[10] R. Srikant and R. Agrawal. Mining quantitative association rules in large relational tables. *Proc. of the ACM SIGMOD Conf on Management of Data. Montreal Canada*, pages 1–12.

[11] D. Dubois, H. Prade, and T. Sudkamp. On the representation, measurement and discovery of fuzzy associations. *IEEE Transactions on Fuzzy Systems*, 13(2):250–262, 2005.

[12] T. Sudkamp. Examples, counterexamples, and measuring fuzzy associations. *Fuzzy Sets and Systems*, 149:57–71, 2005.

[13] Y. Djouadi., S. Redaoui, and K. Amroun. Mining association rules under imprecision and vagueness: towards a possibilistic approach. *FUZZ'IEEE 07, International Conference on Fuzzy Systems. London, UK*, pages 722–727, July 2007.

[14] M. Delgado, N. Marin, D. Sanchez, and M.A. Vila. Fuzzy association rules : general model and applications. *IEEE Transactions on Fuzzy Systems*, 11(2):214–255, 2005.

[15] W.H. Au and K.C.C. Chan. Mining fuzzy association rules in a bank-account databases. *IEEE Transactions on Fuzzy Systems*, 11(2):238–248, 2003.

[16] N. Messai, M.D. Devignes, A. Napoli, and M.S. Tabbone. Many-valued concept lattices for conceptual clustering and information retrieval. *ECAI'08,* 18^{th} *European Conference on Artificial Intelligence. Patras, Greece*, pages 722–727, July.

[17] R. Belohlavek and V. Vychodil. What is a fuzzy concept lattice. *CLA 05, Proc. Olomounc. Czech Republic*, pages 34–45, 2005.

[18] L.A. Zadeh. The concept of a linguistic and its application to approximate reasoning. *Information Science*, 8:199–249, 1975.

[19] M. Serrurier, D. Dubois, H. Prade, and T. Sudkamp. Learning fuzzy rules with their implication operators. *Data and Knowledge Engineering*, 60:71–89, 2007.

[20] U. Fayyad, G.P. Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *Artificial Intelligence Magazine*, pages 37–54, 1996.

[21] J.C. Dunn. A fuzzy relative of the isodata process and its use in detecting compact, well-separated clusters. *Journal of Cybernetics*, 3:35–57, 1973.

[22] J.C. Bezdek, J. Keller, R. Krishnapuram, and N.R. Pal. *Fuzzy models and algorithms for pattern recognition and image processing*. Kluwer Academic Publisher, 1999.

[23] J.C. Bezdek and N.R. Pal. On cluster validity for the fuzzy c-means model. *IEEE Transactions on Fuzzy Systems*, 3(3):370–379, 1995.

[24] K.E. Wolff. Concepts in fuzzy scaling theory: order and granularity. *Fuzzy Sets and Systems*, 132:63–75, 2002.

[25] P. Bosc, D. Dubois, O. Pivert, H.Prade, and M. De Calmes. Fuzzy summarization of data using fuzzy cardinalities. *Proc. of the 9th Inter. Conf. on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU 2002), Annecy, France*, pages 1553–1559, 2002.

[26] M. Serrurier and H. Prade. Improving expressivity of inductive logic programming by learning different kinds of fuzzy rules. *Soft Computing*, 11(5):459–466, 2007.

[27] P. Burmeister and R. Holzer. Treating incomplete knowledge in formal concepts analysis. *Lecture Notes in Computer Science 3626, Bernhard Ganter (Eds.)*, pages 114–126, 2005.

[28] Didier Dubois, Florence Dupin de Saint Cyr Bannay, and Henri Prade. A possibilty-theoretic view of formal concept analysis. *Fundamenta Informaticae*, 75(1-4):195–213, 2007.

ISBN: 978-989-95079-6-8