Soft Comput (2012) 16:883–901
DOI 10.1007/s00500-011-0775-3

FOCUS

Mining fuzzy association rules from low-quality data

A. M. Palacios · M. J. Gacto · J. Alcalá-Fdez

Published online: 13 October 2011

Abstract Data mining is most commonly used in attempts to induce association rules from databases which can help decision-makers easily analyze the data and make good decisions regarding the domains concerned. Different studies have proposed methods for mining association rules from databases with crisp values. However, the data in many real-world applications have a certain degree of imprecision. In this paper we address this problem, and propose a new data-mining algorithm for extracting interesting knowledge from databases with imprecise data. The proposed algorithm integrates imprecise data concepts and the fuzzy apriori mining algorithm to find interesting fuzzy association rules in given databases. Experiments for diagnosing dyslexia in early childhood were made to verify the performance of the proposed algorithm.

Keywords Data mining · Fuzzy association rules · Low-quality data

A. M. Palacios (✉)
Department of Computer Science, University of Oviedo,
33204 Gijón, Spain
e-mail: palaciosana@uniovi.es

M. J. Gacto
Department of Computer Science, University of Jaén,
23071 Jaén, Spain
e-mail: mgacto@ujaen.es

J. Alcalá-Fdez
Department of Computer Science and Artificial Intelligence,
CITIC-UGR, University of Granada, 18071 Granada, Spain
e-mail: jalcala@decsai.ugr.es

1 Introduction

Data mining (DM) is the process used for the automatic discovery of high-level knowledge from real-world, large and complex datasets. The use of DM to facilitate decision support can lead to improved performance in decision-making and can enable the tackling of new types of problems that have not been addressed before (Mladenic et al. 2002).

Discovering association rules is one of several data-mining techniques described in the literature (Han and Kamber 2006). Association rules are used to represent and identify dependencies between items in a database (Zhang and Zhang 2002). An association rule is an expression $X \rightarrow Y$, where X and Y are sets of items and $X \cap Y = \emptyset$. It means that if all the items in X exist in a transaction then all the items in Y are also in the transaction with a high probability, and X and Y should not have a common item (Agrawal et al. 1993; Agrawal and Srikant 1994).

Research in this field has mainly concentrated on boolean and quantitative association rules (Agrawal and Srikant 1994; Alatas and Akin 2006; Alcala-Fdez et al. 2010; Han et al. 2004; Sun and Fengshan 2008). However, in recent years many researchers have proposed methods to mine fuzzy association rules from quantitative data in order to solve some of the problems introduced by quantitative attributes (Chen et al. 2011; Hong et al. 2001; Hong and Lee 2008; Kaya 2006). The use of fuzzy sets to describe associations between data extends the type of relationships that may be represented, facilitates the interpretation of rules in linguistic terms, and avoids unnatural boundaries in the partitioning of attribute domains (Delgado et al. 2003; Dubois et al. 2005, 2006; Hullermeier and Yi 2007; Sudkamp 2005).

Various studies have proposed methods for mining association rules that have been focused on databases with crisp values, however the data in many real-world applications have a certain degree of imprecision (e.g., interval or fuzzy values). Sometimes, this imprecision is small enough to be safely ignored. On other occasions, the uncertainty of the data can be modeled by a probability distribution. However, there are other problems where the imprecision is significant and a probability distribution is not a natural model (Baudrit et al. 2008). Designing DM algorithms able to deal with the uncertainty of the data and exploit better the information contained in low-quality sets of data (LQD) presents a challenge to workers in this research field (Palacios et al. 2011; Villar et al. 2009).

Fuzzy statistic considers the use of fuzzy sets to represent imprecise knowledge about the data (Bertoluzza et al. 2003; Wu and Sun 2001). Recent works in fuzzy statistic suggest using a fuzzy representation when the data are known through a family of confidence intervals (Couso and Sanchez 2008), using a possibilistic representation to model these kinds of data (Sanchez et al. 2007, 2009). This representation assumes that a fuzzy set can be identified as a nested family of sets where each one of them contains the true value of the object with a probability greater than or equal to a certain bound (Couso and Sanchez 2008).

In this paper, we integrate LQD concepts with the fuzzy apriori mining algorithm proposed by Hong et al. (2001) in order to obtain high-quality fuzzy association rules from databases with LQD. We extend this algorithm considering a possibilistic representation to model the input data with inaccurate values, transforming each input value into a fuzzy set. Let us consider a set of linguistic terms L, $L = \{l_1, \ldots, l_n\}$, associated with a Ruspini (1969) fuzzy partition. According to Sanchez et al. (2008), an inaccurate input will be represented as a fuzzy subset of L, where the discrete probability distribution is generalized to an imprecise probability distribution and, where this imprecise probability distribution can also be interpreted as a possibility distribution (Dubois and Prade 1992). Let us illustrate this with a example. A crisp value, temperature of 45 grade, could be represented with the fuzzy subset $\{0.0/\text{Cold} + 0.2/\text{Warm} + 0.8/\text{Hot}\}$ where the sum of the memberships of a crisp measurement is 1. Nonetheless, an imprecise or vague measurement of the temperature could be represented by the fuzzy subset $\{0.1/\text{Cold} + [0.2, 0.4]/\text{Warm} + 0.9/\text{Hot}\}$ and a missing value by the set $\{1.0/\text{Cold} + 1.0/\text{Warm} + 1.0/\text{Hot}\}$, where the sum of the memberships of an inaccurate value can be greater or lower than 1. On the other hand, the confidence value of an association rule will be defined by an interval of probabilities, which will determine the probability that an association rule provides a high level of knowledge.

We will also present an experimental study to show the behavior of the proposed approach using a low-quality set of data for the diagnosis of dyslexia in early childhood, named Inexpert-57 (see Sect. 6.1). First, we will revise the fuzzy association rules obtained with our approach via support and confidence. Then, we will analyze the level of knowledge of the fuzzy association rule obtained by our proposal from the dataset Inexpert-57. Finally, a study of complexity and scalability of the proposal approach will be shown.

This paper is organized as follows. The next section describes the fuzzy mining algorithm proposed by Hong et al. to mine association rules from datasets with quantitative values. Section 3 introduces LQD, highlight their representation and interpretation. Section 4 details the fuzzy data-mining algorithm proposed to obtain fuzzy association rules from low-quality datasets. An example is given to illustrate the proposed algorithm in Sect. 5. Section 6 shows the results obtained by our proposal over a real-world dataset. Finally, in Sect. 7 some concluding remarks are made.

2 Fuzzy data-mining algorithm for quantitative values

The goal of the fuzzy data-mining algorithm presented by Hong et al. (2001) is to find interesting itemsets and fuzzy association rules in databases with quantitative values, discovering interesting patterns among them.

This method consists of transforming each quantitative value into a fuzzy set of linguistic terms using membership functions, which assumes that the membership functions are known in advance. The algorithm then calculates the scalar cardinality of each linguistic term in all the instances as the count value and checks whether the count of each linguistic term is larger than or equal to the minimum support value to put these items in the large itemsets L_r. The mining process, based on fuzzy counts, considers that the intersection between the membership value of each item is the minimum operator. Finally, this method obtains the fuzzy association rules by the criterion used in the apriori algorithm (Agrawal and Srikant 1994).

Hong et al. (1999) proposed a mining approach that integrated fuzzy-sets concepts with the apriori algorithm to find interesting itemsets and fuzzy association rules in the instances with quantitative values. Although this approach could quickly find interesting patterns, some patterns might be missed since only the linguistic term with the maximum cardinality in each item is used in the mining process.

In Hong et al. (2001), all the important linguistic terms in the mining process are considered, generating a more complete set of rules than the method proposed in Hong et al. (1999), although its computation time increases.

Hong et al. determine that there is a trade-off between the computation time and the completeness of rules. Choosing an appropriate learning method thus depends on the requirements of the application domains.

3 Low-quality data: representation and interpretation

The Ruspini (1969) fuzzy partition can be understood as a set of conditional probabilities; thus, the set of memberships of an object to a Ruspini's partition is a discrete probability distribution over the linguistic labels. For instance, let us assume an item with a finite domain of five linguistic labels $L = \{\text{Bad, Slow, Regulate, Normal, Good}\}$. In this case, a crisp value could be represented with the fuzzy subset $\{0.0/\text{Bad} + 0.2/\text{Slow} + 0.8/\text{Regulate} + 0.0/\text{Normal} + 0.0/\text{Good}\}$, where the sum of the memberships of a crisp measurement is 1. Nonetheless, the properties of one object or measurement cannot be accurately observed; the discrete probability distribution have to be generalized to an imprecise probability distribution (Sanchez et al. 2008), where the sum of the memberships of an inaccurate value can be greater or lower than 1. In certain cases this imprecise probability distribution can also be interpreted as a possibility distribution (Dubois and Prade 1992), where this possibility distribution can be defined by a fuzzy membership defined over the set of linguistic labels. For instance, the fuzzy set $\{0.0/\text{Cold} + [0.2, 0.4]/\text{Warn} + 0.9/\text{Hot}\}$ means that the probability of the temperature being 'Cold' is 0, the probability of 'Warm' is not greater than 0.4 and is not lower than 0.2 and the probability of 'Hot' is not greater than 0.9, and its corresponding lower bound of probability is: $p(\text{Hot}) \geq 1 - (p^*(\text{Cold}) + p^*(\text{Warm}) = 0.6$ (observe that with this interpretation the set $\{[0.0, 1.0]/\text{Cold} + [0.0, 1.0]/\text{Warm} + [0.0, 1.0]\text{Hot}\}$ represents the total absence of knowledge about the input value).

In order to model the imprecision case, in this paper we use a possibility representation which assumes that a fuzzy set can be identified or represented as a possibility distribution, i.e., a fuzzy set will be identified with the family of all the probability distributions, where each α-cut of a fuzzy feature is a random set that contains the unknown crisp value of the feature with a probability greater or equal than $1 - \alpha$-cut (Couso and Sanchez 2008) (see Fig. 1). Thus, this possibilistic representation consists of understanding a fuzzy membership function as a nested family of sets where each one contains the true value of the object with a probability greater than or equal to a certain bound. Notice that, this includes the interval and the crisp situations as particular cases.

These interpretations provide us a common framework for reasoning with numbers, words, interval, fuzzy values, missing values (one interval that spans the whole range of the variable), left and right censored data (the value is greater or lower than a cut-off value, or it is between a couple of bounds), compound measures (each item comprises a disperse list of values) or different values of the same attributes as described in Sanchez et al. (2009), where these different concepts are embodied in the term "low quality data".

There are several kinds of representations of LQD, with an interpretation based on a fuzzy statistic. In the next sections we will introduce the representation of the LQD used and the calculation of their membership value.

3.1 Representation of low-quality values

Real-world datasets are composed of groups of low-quality items where each item describes one property of an object but without observing the real value of the object in this item. The representation and dominion of each low-quality item can be defined by different kinds of inaccurate values:

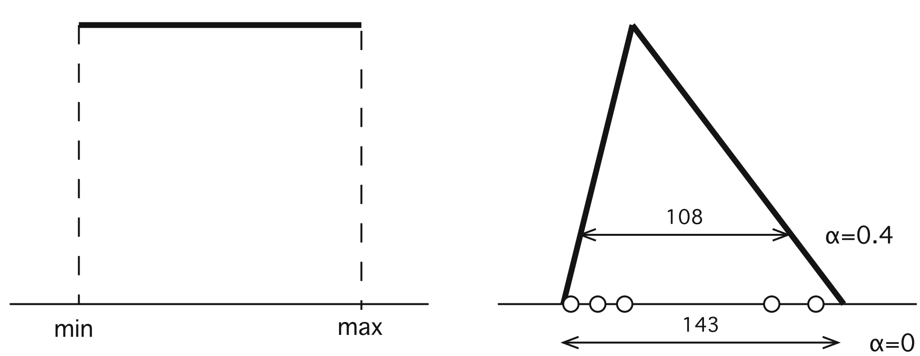

Fig. 1 Fuzzy representation of vague data. *Left* a missing value is codified with an interval that spans the whole range of the variable, or $P([\min, \max]) = 1$. *Right* a compound value (in this example, 5 different measurements of the variable) can be described by a fuzzy membership, that can also be understood as an upper probability. Each α-cut contains the true value of the variable with probability at least $1 - \alpha$

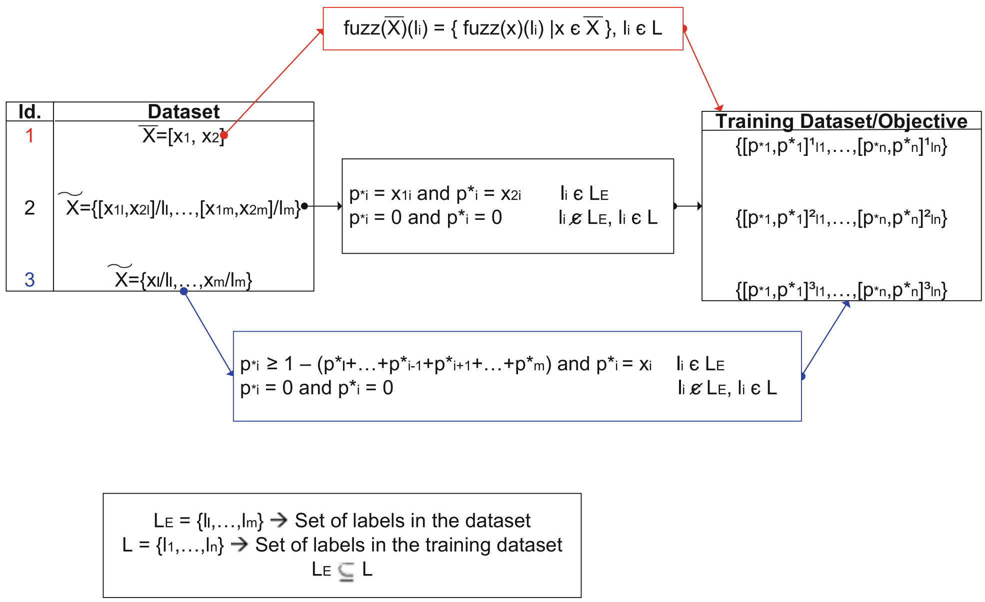

Fig. 2 Set of bounds of probabilities for each linguistic term $l_i \in L$ with $L = \{l_1, \ldots, l_n\}$

Table 1 Set of bounds of probabilities

Ins.	Item.Low	Item.Normal	Item.High
1	[0, 1]	[0, 1]	0
2	0	[0, 1]	[0, 1]
3	0	[0, 1]	[0, 1]
Count	[0, 1]	[0, 3]	[0, 2]
Count(%)	[0, 0.33]	[0, 1]	[0, 0.66]

- An interval $\bar{X} = [x_1, x_2]$ where x_1 and x_2 are included in the domain of the item. For instance, an item with a dominion between [0.0, 10.0] could be defined as $\bar{X} =$ [1.5, 3.6]; this implies that all our information is that "$x \in \bar{X}$" or as a missing value $\bar{X} =$ [0.0, 10.0].
- A fuzzy value $\widetilde{X} = (x_1, x_2, x_3)$. For instance, an item defined by three different values $\widetilde{X} =$ (1.0, 1.3, 2.0).
- A fuzzy subset of a finite set of linguistic labels associated with a Ruspini fuzzy partition (Sanchez et al. 2008). For instance, $\{0.1/\text{Bad} + 0.3/\text{Slow} + 0.9/\text{Regulate}\}$, where the sum of the memberships of an inaccurate value can be >1.
- A fuzzy subset $\widetilde{X}$ of a finite set of linguistic labels where the membership of each partition will be defined with the lower and upper membership. For the example of the previous representation, the fuzzy set could be defined as $\widetilde{X} = \{[0.0, 0.1]/\text{Bad} + [0.1, 0.3]/\text{Slow} + [0.7, 0.9]/\text{Regulate}\}$ where this fuzzy representation is interpreted as a possibility distribution.

3.2 Fuzzy membership with low-quality data

In this section we discuss how to compute the membership function given a vague input. Let us suppose that we have a crisp perception "x" of the properties of an object and a fuzzy set $\widetilde{A}$ with a finite set of linguistic labels, $L = \{l_1, \ldots, l_n\}$, where "n" is the number of labels. The membership function will be:

$$\text{fuzz}(x)(l_i) = P_x(l_i) \left| \sum_{i=1}^{n} P_x(l_i) = 1. \right. \quad (1)$$

In accordance with the representation of the imprecise inputs of the dataset and from the interpretation of LQD, which assumes that a fuzzy set can be identified as a possibility distribution, if the object is imprecise and all our information about the input is that it is in the set $\bar{X}$ ("$x \in \bar{X}$"), the membership function of this vague input in the finite set of linguistic labels of L are not implicit. The extension to set valued input is a set of membership function as follows:

$$\text{fuzz}(\bar{X})(l_i) = \{\text{fuzz}(x)(l_i) | x \in \bar{X}\}. \quad (2)$$

For instance, let us assume that the vague input is the interval $\bar{X} = [2.1, 3]$ and we want to calculate the membership function in the linguistic label "Low". We know that the interval [2.1, 3] provides incomplete and imprecise information about the real and unknown value (x_0) of the measurement object. All the information that we have is that $x_0 \in [2.1, 3]$ and each α-cut of this vague input is a random set that contains the unknown crisp value of the object with a probability greater or equal than 1 − α-cut). In this case, the interval [2.1, 3], for any α-cut, has the probability of 1 the contain the true real value x_0 ($P(x_0 \in [2.1, 3]) = 1$). Therefore, fuzzy([2.1, 3])(Low) will be a set of probabilities obtained from each value of x where $x \in$ [2.1, 3].

If the input is the imprecise value $\widetilde{X}$, the membership function will be a fuzzy set, according to the Extension Principle which is compatible with the possibilistic interpretation of fuzzy sets (Sanchez et al. 2009), computed as follows:

$$\mathrm{fuzz}(\widetilde{X})(l_i)(t) = \sup\{\alpha | t = \mathrm{fuzz(x)}(l_i) \quad \text{and} \quad x \in [\widetilde{X}]_\alpha\}. \tag{3}$$

Observe that this fuzzy set fuzz $(\widetilde{X})(l_i)$ is associated with the nested family of sets $\{\mathrm{fuzz}([\widetilde{X}]_\alpha)(l_i)\}_{\alpha\in[0,1]}$.

If the input is a fuzzy subset of a finite set of linguistic labels associated with a Ruspini fuzzy partition $\widetilde{X} = \{[x_{1l}, x_{2l}]/l_l, \ldots, [x_{1n}, x_{2n}]/l_n\}$, the upper ($p^*$) and lower ($p_*$) bounds of probabilities are directly obtained from the vague input. This means that, if we have a fuzzy set of linguistic labels $L = \{l_1, \ldots, l_n\}$, the item represented by the fuzzy subset $\widetilde{X} = \{[x_{1l}, x_{2l}]_{l_l}, \ldots, [x_{1m}, x_{2m}]_{l_m}\}$, where $L_E = \{l_l, \ldots, l_m\} \subseteq L$, is interpreted as a set of bounds of probabilities for each linguistic label l_i, where $l_i \in L_E$. Concretely: the lower (p_*) and upper (p^*) probabilities of l_i are $p_{*i} = x_{1i}$ and $p_i^* = x_{2i}$, respectively. If the linguistic label $l_i \in L$ is not contained in the subset of linguistic labels L_E then p_{*i} and p_i^* are zero.

From the fuzzy subset $\widetilde{X} = \{x_l/l_l, \ldots, x_m/l_m\}$ and according to Sanchez et al. (2008) the corresponding lower bound of each linguistic label is implicit from this fuzzy subset (4):

$$p_{*i} \geq 1 - (p_l^* + \cdots + p_{i-1}^* + p_{i+1}^* + \cdots + p_m^*), \quad l_i \in L_E. \tag{4}$$

For instance, from the fuzzy set $\{0.0/\mathrm{Cold} + 0.2/\mathrm{Warm} + 0.9/\mathrm{Hot}\}$, the lower bound of 'Warm' will be $p_{*\mathrm{Warm}} \geq 1 - (p^*_{\mathrm{Cold}} + p^*_{\mathrm{Hot}}) = 0.1$. As in the previous case, if the linguistic label $l_i \in L$ is not contained in the subset of linguistic labels L_E then p_{*i} and p_i^* are zero.

Figure 2 shows that the set of bounds of probabilities for each linguistic label l_i, which composes the fuzzy set, are obtained from the different representations of low-quality inputs ($[x_1, x_2], \{[x_1, x_2]/l_1, \ldots, [x_{1m}, x_{2m}]/l_m\}$ and $\{x_1/l_1, \ldots, x_m/l_m\}$) which will compose the dataset "Inexpert-57".

3.3 Fuzzy data-mining algorithm and low-quality data

In this section, we describe in detail our fuzzy data-mining algorithm to obtain fuzzy association rules from datasets with LQD. We take the following variables as the parameters or inputs of this new proposal:

- A low-quality dataset $\widetilde{D}$ that is composed of t instances, where each one contains m attributes, and where $\widetilde{X}_j^i$ represents the item j, $1 \leq j \leq m$, in the instance i, $1 \leq i \leq t$. This implies that the instance i of $\widetilde{D}$ will be formed by (5):

$$\widetilde{D}_i = \{\widetilde{X}_j^i\}_{j=1,\ldots,m} \tag{5}$$

- A set of membership functions $S = \{L_1, \ldots, L_m\}$, where m is the number of attributes and L_j represents the finite set of linguistic labels which, in turn, are associated with a Ruspini fuzzy partition $L_j = \{l_1, \ldots, l_n\}$, where n is the number of linguistic labels.
- A predefined minimum support value α.
- A predefined confidence value λ.
- The number of cuts to obtain the possible real values.
- The number of times γ that we sweep the probabilities of each possible value obtained with the cuts.

The objective is obtain a set of fuzzy association rules from LQD. To achieve this objective the steps are shown below:

Step 1 Transform each item $\widetilde{X}_j^i$, with $1 \leq j \leq m$, of each instance $\widetilde{D}_i$, $1 \leq i \leq t$, into a fuzzy set interpreted as a set of bounds of probabilities for each linguistic term of $L_j(\overline{P_j^i} = \{[p_{*1}, p_1^*]^i_{l_1}, \ldots, [p_{*n}, p_n^*]^i_{l_n}\})$.

Step 2 Calculate the frequency of occurrence of item j in each linguistic term "k" of L_j, that is to say, L_{j_k} where "k", $1 \leq k \leq n$, represents the partition k in the set of linguistic terms of the item j, therefore $L_j = \{l_1, \ldots, l_k, \ldots, l_n\}$ and $L_{j_k} = l_k$.

$$\overline{\mathrm{Count}}_{L_{j_k}} = \bigoplus_{i=1}^t \overline{P}^i_{j_k} = \bigoplus_{i=1}^t [p_{*k}, p_k^*]^i_{l_k} \tag{6}$$

where t represents the number of instances and $\bigoplus$ is the fuzzy arithmetic-based sum (Kaufmann and Gupta 1991). The percentage of $\overline{\mathrm{Count}}_{L_{j_k}}$ will be defined as:

$$\overline{\mathrm{Count}}_{L_{j_k}}(\%) = \frac{1}{t} \bigoplus_{i=1}^t [p_{*k}, p_k^*]^i_{l_k} \tag{7}$$

All L_{j_k} are collected to form the candidate set C_r of r-itemsets, where r represents the number of items kept in the candidate set, initially $r = 1$:

$$C_r = \{\cup\{L_{j_k}, \forall \mathrm{k}|1 \leq k \leq n\}, \quad \forall \mathrm{j}|1 \leq j \leq m\} \tag{8}$$

Step 3 Check whether $\overline{\mathrm{Count}}_{L_{j_k}}(\%)$ for all r-itemsets L_{j_k} of C_r ($k = 1$ to n for all j = 1 to m) is larger than or equal to the predefined minimum support α. If $\overline{\mathrm{Count}}_{L_{j_k}}(\%)$ satisfies this condition then the set r-itemsets (L_r) will contain L_{j_k}. As $\overline{\mathrm{Count}}_{L_{j_k}}(\%)$ is defined by a set of bounds of probabilities, this condition is satisfied if the upper bounds (p^*_k) is larger than or equal to $\alpha.L_r$, will be the next set:

$$L_r = \{L_{j_k} | \max\{\overline{\mathrm{Count}}_{L_{j_k}}(\%)\} \geq \alpha, \\ L_{j_k} \in C_r\} = \{L_{j_k} | p^*_k \geq \alpha, L_{j_k} \in C_r\} \tag{9}$$

Notice that, we have considered the upper bounds (p^*_k) in this condition because all possibly occurring itemsets are considered. For instance, if an item is imprecise in all the instances and we consider the option of the lower bounds (p_{*k}) then this item is not considered and never introduced in L_r. Let us assume that an item is imprecise in all instances of the dataset:

Item

1. $\{1/\mathrm{Low} + 1/\mathrm{Normal}\}$
2. $\{1/\mathrm{Normal} + 1/\mathrm{High}\}$
3. $\{1/\mathrm{Normal} + 1/\mathrm{High}\}$

The set of bounds of probabilities for each linguist term in this item is shown in Table 1. Therefore, if we consider the lower bounds (p_*) then the information about this item will be lost and the information provided will be that not exist information about this item in the dataset and as consequence this item would be determined as irrelevant.

Step 4 If L_r is not null, then do the next step; otherwise, exit the algorithm.

Step 5 Join the r-itemsets that compose L_r to generate the new candidate set C_{r+1}. This set C_{r+1} is obtained in a similar way to the apriori algorithm (Agrawal and Srikant 1994) except that two L_{j_k} with the same attribute j cannot simultaneously exist in an itemset in C_{r+1} (Hong et al. 2001).

Step 6 For each $(r + 1)$-itemset obtained in C_{r+1}, do the following substeps:

a. Calculate the fuzzy value of each $(r + 1)$-itemset (s), of C_{r+1}, for each instance D^i. The fuzzy value will be a set of bounds of probabilities obtained from each itemset that composes $(r + 1)$-itemset:

$$\overline{P}^i_s = \overline{P}^i_1 \bigwedge \cdots \bigwedge \overline{P}^i_{(r+1)}. \tag{10}$$

The product t-norm generalizes the aggregation or combination between the sets of probabilities of each itemset of $(r + 1)$-itemset.

$$\overline{P}^i_s = \overline{P}^i_1 \bigwedge \cdots \bigwedge \overline{P}^i_{(r+1)} = \bigotimes_{j=1}^{(r+1)} \overline{P}^i_j \tag{11}$$

b. Calculate the frequency of occurrence of each $(r + 1)$-itemset s as:

$$\overline{\mathrm{Count}}_s = \bigoplus_{i=1}^{t} \overline{P}^i_s \tag{12}$$

The percentage of $\overline{\mathrm{Count}}_s$ will be defined as:

$$\overline{\mathrm{Count}}_s(\%) = \frac{1}{t} \bigoplus_{i=1}^{t} \overline{P}^i_s \tag{13}$$

c. If $\max\{\overline{\mathrm{Count}}_s(\%)\}$, is larger than or equal to the minimum support α then, put the $(r + 1)$-itemset s in L_{r+1}.

Step 7 If $L_(r+1)$ is null then continue with the next steps; otherwise update the number of itemsets in the set of candidates ($r = r + 1$) and repeat the steps 5 and 6.

Step 8 Collect in R the itemsets of each L_i, where $2 \leq i \leq (r+1)$:

$$R = \{L_i, 2 \leq i \leq (r+1)\}. \tag{14}$$

Step 9 Construct all the possible association rules ($X \rightarrow Y$) from each large q-itemset s, $q \geq 2$, with items $(s_1, s_2, \ldots, s_q)$, of the set R (15):

$$s_1 \wedge \cdots \wedge s_{k-1} \wedge s_{k+1} \wedge \cdots \wedge s_q \\ \rightarrow s_k, \quad k = 1, \ldots, q. \tag{15}$$

Step 10 Determine whether the association rules obtained are relevant and provide interesting patterns or high-level knowledge from LQD. Two substeps are required to determine whether an association rule is relevant or not:

a. Calculate the confidence of the rule.
b. Compare the previous confidence with the predefined confidence threshold λ.

Let us suppose that we have a crisp dataset, D^i, the confidence of one association rule obtained from q-itemset s of the set R, Confidence $(X \rightarrow Y)_{(P_{s_1},\ldots,P_{s_q})}$, will be defined as:

$$\frac{\sum_{i=1}^{t} P^i_s}{\sum_{i=1}^{t}(P^i_{s_1} \wedge \cdots \wedge P^i_{s_{k-1}} \wedge P^i_{s_{k+1}} \wedge \cdots \wedge P^i_{s_q})} \\ = \frac{\mathrm{Count}_s}{\mathrm{Count}_{\mathrm{anteced}}}. \tag{16}$$

If the inputs are imprecise, $\widetilde{D}^i$, the confidence will be defined by an interval value between [0, 1] that represents the upper and lower bounds of this rule $X \rightarrow Y$ (17):

$$\overline{\mathrm{Confidence}}(X \rightarrow Y)_{(\bar{P}_{s_1},\ldots,\bar{P}_{s_q})} \\ = \{\mathrm{Confidence}(X \rightarrow Y)_{(x_{s_1},\ldots,x_{s_q})} | \\ \mathrm{Count}_{\mathrm{anteced.}} > 0, \forall x_{s_j} \in \bar{P}_{s_j}\}. \tag{17}$$

The computational cost of $\overline{\mathrm{Confidence}}(X \rightarrow Y)$ is very high and moreover, as the confidence value of a rule is defined by an interval-value that contains the real and unknown exact value of the confidence, depending on the values of x_{s_j}, the rule could be relevant or not. So, if the value of λ is contained in this interval-value we do not know whether or not the rule provides interesting information. An approximation of such an interval-value is defined in this proposal.

Let us consider a q-itemset s of the set R where the association rule is $s_1 \wedge \cdots \wedge s_{q-1} \rightarrow s_q$ and where $\overline{\text{Count}}_s = [x_1, x_2] = \overline{X}$ and $\overline{\text{Count}}_{\text{anteced.}} = [y_1, y_2] = \overline{Y}$. In order to determine the real value of $\overline{X}$ and $\overline{Y}$ some "cuts" are applied to obtain the possible real value of each interval, $\overline{X}_{\text{cut}} = [x_1, x_2]_{\text{cut}} = x_j$ where $x_j \in \overline{X}$, so that each cut (x_j) is assigned a random probability (P_{x_j}) of being the real value and, where the sum of all probabilities, of all cuts, have to be 1:

$$P_{\overline{X}} = \sum_{c=1}^{\text{cuts}} P_{\overline{X}_c} = \sum_{c=1}^{\text{cuts}} P_{[x_1,x_2]_c} = 1 \tag{18}$$

$$P_{\overline{Y}} = \sum_{c=1}^{\text{cuts}} P_{\overline{Y}_c} = \sum_{c=1}^{\text{cuts}} P_{[y_1,y_2]_c} = 1 \tag{19}$$

where cuts indicates the number of possible real values that are obtained from $\overline{X}$ and $\overline{Y}$. The possible values of each set of bounds of probabilities $\overline{X}$ and $\overline{Y}$ are:

$$V_X = \{\overline{X}_c, c = 1 \text{ to cuts}\} \tag{20}$$

$$V_Y = \{\overline{Y}_c, c = 1 \text{ to cuts}\}. \tag{21}$$

From a value of $V_Y(y_j \in V_Y)$ and a value of the set $V_X(x_j \in V_X)$, the rule would be deemed relevant if:

$$x_j \geq y_j * \lambda, \quad x_j \leq y_j, \quad y_j > 0 \tag{22}$$

The possible value x_j could have a low probability of being the real value. As a consequence, to determine if the rule is relevant besides satisfying (22) the $P_{x_j} \geq 0.5$. Notice that, for each value of the set $V_Y(y_j \in V_Y)$, all possible values of the set V_X have been considered and a new set $C = \{C_{y_j} | y_j \in V_Y\}$ is obtained from the probabilities that determine whether one rule is relevant or not for each value of the set $V_Y(y_j \in Y)$:

$$C_{y_j} = \sum_{i=1}^{\text{cuts}} P_{x_i} \geq 0.5 | x_i \geq y_j * \lambda, \quad x_i \leq y_j, \quad y_j > 0. \tag{23}$$

For instance, let us suppose $\overline{X} = [0.4, 0.8]$ and $\overline{Y} = [0.4, 1]$, where cuts $= 4$ and $\lambda = 0.6$. This implies that:

$$V_X = \{x_1, x_2, x_3, x_4\} = \{0.4, 0.53, 0.66, 0.8\}$$

$$V_Y = \{y_1, y_2, y_3, y_4\}) = \{0.4, 0.6, 0.8, 1\}$$

and the random probabilities obtained of being the possible values are:

$$P_{\overline{X}} = 0.15 + 0.58 + 0.2 + 0.07 = 1$$

$$P_{\overline{Y}} = 0.01 + 0.28 + 0.4 + 0.31 = 1$$

where, the set C will be:

$$C = \{C_{y_1}, C_{y_2}, C_{y_3}, C_{y_4}\}$$

where:

$$\begin{aligned} C_{y_1} = {} & (0.15) \leq 0.5 | 0.4 \geq 0.4 * 0.6, \\ & 0.4 \leq 0.4, 0.4 > 0 \end{aligned}$$

$$\begin{aligned} & C_{y_2} = (0.15 + 0.58) \geq 0.5 | \\ & (0.4 \geq 0.6 * 0.6, 0.4 \leq 0.6, 0.6 > 0) \\ & (0.53 \geq 0.6 * 0.6, 0.53 \leq 0.6, 0.6 > 0) \end{aligned}$$

$$\begin{aligned} & C_{y_3} = (0.58 + 0.2 + 0.07) \geq 0.5 | \\ & (0.4 \leq 0.8 * 0.6, 0.4 \leq 0.8, 0.8 > 0) \\ & (0.53 \geq 0.8 * 0.6, 0.53 \leq 0.8, 0.8 > 0) \\ & (0.66 \geq 0.8 * 0.6, 0.66 \leq 0.8, 0.8 > 0) \\ & (0.8 \geq 0.8 * 0.6, 0.8 \leq 0.8, 0.8 > 0) \end{aligned}$$

$$\begin{aligned} & C_{y_4} = (0.2 + 0.07) \leq 0.5 | \\ & (0.4 \leq 1 * 0.6, 0.4 \leq 1, 1 > 0) \\ & (0.53 \leq 1 * 0.6, 0.53 \leq 1, 1 > 0) \\ & (0.66 \geq 1 * 0.6, 0.66 \leq 1, 1 > 0) \\ & (0.8 \geq 1 * 0.6, 0.8 \leq 1, 1 > 0) \end{aligned}$$

$$C = \{0.15, 0.73, 0.85, 0.27\}.$$

This process is repeated γ times in order to sweep the possible probabilities of each possible value of V_X. Thus, for each value of $V_Y(y_j \in Y)$, it will not only have one probability that determines if the rule is relevant or not in this value y_j with respect to all the possible values of V_X with the corresponding probabilities, but will have a set of possible probabilities depending on the random probability assigned to each possible values of V_X.

In the previous example, if $\gamma = 2$ we have to assign new random probabilities to the possible values of the set V_X, obtaining another set C_2. The sets C_1 and C_2 are obtained with $\gamma = 1$ and $\gamma = 2$, respectively:

$$C_1 = \{0.15, 0.73, 0.85, 0.27\}$$

$$C_2 = \{0.25, 0.53, 0.45, 0.37\}.$$

For each possible value $y_j \in Y$, for example $y_1 = 0.4$, the possible probabilities that determine whether the rule is relevant or not in this value 0.4, will be the set:

$$\overline{C}_{y_j} = \{C_{y_j}, \forall C_i, 1 \leq i \leq \gamma\} = \{0.15, 0.25\}.$$

Finally, to determine whether one rule is relevant or not, considering all values of V_Y, we need to choose one sets of probabilities $\overline{C}_{y_j}$ and if the minimum of $\overline{C}_{y_j}$ is larger than or

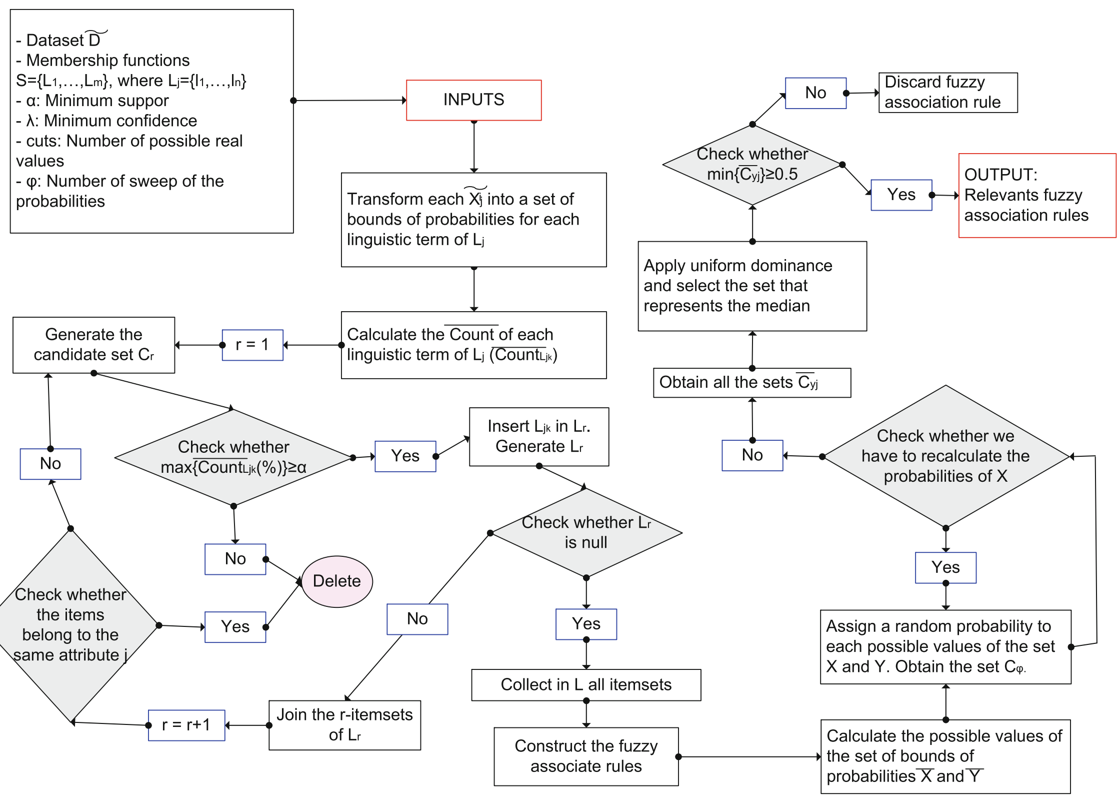

Fig. 3 Outline of the proposed algorithm

Table 2 Dataset that define the event of Jump as well as whether one athlete is relevance or not in such event

Ins.	DPE	MC	EI	VM	RA
1	[8.7, 9.7]	[47, 52]	[2.5, 2.62]	[4.77, 4.83]	{1}
2	[0.7, 1.3]	[62, 67]	[2.2, 2.24]	[5.18, 5.21]	$\{1/0 + 1/1\}$
3	[6.8, 7.7]	[33, 38]	[2.09, 2.16]	[5.87, 5.9]	{0}
4	[3.3, 4.1]	[44, 47]	[2.23, 2.27]	[4.92, 5]	{1}
5	[0, 0.8]	[46, 50]	[2.04, 2.14]	[5, 5.04]	{0}
6	[10.7, 11.6]	[53, 57]	[2.64, 2.72]	[4.34, 4.4]	{1}
7	[3.9, 4.7]	[47, 55]	[2.55, 2.6]	[4.25, 4.3]	{1}
8	[4.9, 5.6]	[36, 44]	[2.15, 2.18]	[5.01, 5.03]	$\{0.9/0 + 0.4/1\}$
9	[7.4, 8]	[45, 46]	[2.3, 2.37]	[4.96, 5]	{0}
10	[11.5, 12]	[36, 40]	[1.9, 1.94]	[5.37, 5.46]	$\{0.5/0 + 0.5/1\}$
11	[4.9, 5.7]	[45, 50]	[2.1, 2.14]	[4.87, 4.94]	$\{0.6/0 + 0.8/1\}$
12	[3, 3]	[47, 52]	[2.2, 2.29]	[4.92, 5.01]	{1}
13	[3.6, 4.3]	[47, 53]	[2.3, 2.36]	[4.86, 4.9]	{1}
14	[9, 9.3]	[34, 35]	[2.34, 2.35]	[4.99, 5.1]	{0}
15	[7.4, 8.3]	[34, 35]	[2.2, 2.26]	[5.77, 5.83]	{0}
16	[8.7, 10.1]	[45, 47]	[2, 2.15]	[5, 5.1]	{1}
17	[8.2, 9]	[36, 39]	[2.12, 2.24]	[5.06, 5.14]	{1}

equal to 0.5 then we can determine that the rule is relevant. For the reason, these sets are arranged according to the uniform dominance defined in Limbourg (2005), which induces a total order and the median set is chosen. In the previous example, $\overline{C}_{y_1} = [0.15, 0.25]$, $\overline{C}_{y_2} = [0.53, 0.73]$, $\overline{C}_{y_3} = [0.45, 0.85]$ and $\overline{C}_{y_4} = [0.27, 0.37]$. The total order

Table 3 Set of bounds of probabilities for each linguistic term

Ins.	DPE			MC			EI		
	DPE_L	DPE_M	DPE_H	MC_L	MC_M	MC_H	EI_L	EI_M	EI_H
1	0	[0.44, 0.44]	[0.55, 0.55]	[0, 0.11]	[0.88, 1]	[0, 0.11]	0	[0.33, 0.44]	[0.55, 0.66]
2	[0.78, 0.88]	[0.11, 0.21]	[0]	0	[0, 0.22]	[0.77, 1]	[0.22, 0.22]	[0.77, 0.77]	0
3	0	[0.77, 0.77]	[0.22, 0.22]	[0.77, 1]	[0, 0.22]	0	[0.44, 0.44]	[0.55, 0.55]	0
4	[0.33, 0.44]	[0.55, 0.66]	0	[0.22, 0.33]	[0.66, 0.77]	0	[0.11, 0.11]	[0.88, 0.88]	0
5	[0.88, 1]	[0, 0.11]	0	[0, 0.22]	[0.77, 1]	0	[0.44, 0.55]	[0.44, 0.55]	0
6	0	[0.11, 0.11]	[0.88, 0.88]	0	[0.66, 0.77]	[0.22, 0.33]	0	[0, 0.11]	[0.88, 1]
7	[0.22, 0.33]	[0.66, 0.77]	0	[0, 0.11]	[0.77, 1]	[0, 0.22]	0	[0.33, 0.33]	[0.66, 0.66]
8	[0.11, 0.11]	[0.88, 0.88]	0	[0.44, 0.77]	[0.22, 0.55]	0	[0.33, 0.33]	[0.66, 0.66]	0
9	0	[0.66, 0.66]	[0.33, 0.33]	[0.23, 0.29]	[0.70, 0.76]	0	0	[0.88, 1]	[0, 0.11]
10	0	0	[1, 1]	[0.66, 0.77]	[0.22, 0.33]	0	[1, 1]	0	0
11	[0.11, 0.11]	[0.88, 0.88]	0	[0, 0.22]	[0.77, 1]	0	[0.44, 0.44]	[0.55, 0.55]	0
12	0.5	0.5	0	[0, 0.11]	[0.88, 1]	[0, 0.11]	[0.11, 0.22]	[0.77, 0.88]	0
13	[0.33, 0.33]	[0.66, 0.66]	0	[0, 0.11]	[0.88, 1]	[0, 0.11]	0	[0, 0.11]	[0.88, 1]
14	0	[0.44, 0.5]	[0.5, 0.55]	[0.88, 0.88]	[0.11, 0.11]	0	0	[0.90, 0.92]	[0.07, 0.09]
15	0	[0.66, 0.66]	[0.33, 0.33]	[0.88, 0.88]	[0.11, 0.11]	0	[0.22, 0.22]	[0.77, 0.77]	0
16	0	[0.33, 0.44]	[0.55, 0.66]	[0.22, 0.22]	[0.77, 0.77]	0	[0.44, 0.66]	[0.33, 0.55]	0
17	0	[0.55, 0.55]	[0.44, 0.44]	[0.66, 0.77]	[0.22, 0.33]	0	[0.22, 0.44]	[0.55, 0.77]	0
Count	[3.28, 3.71]	[8.28, 8.83]	[4.83, 4.99]	[5.01, 6.84]	[8.70, 10.98]	[1, 1.88]	[3.99, 0.44]	[9.68, 10.81]	[2.18, 2.65]
Count (%)	[0.19, 0.21]	[0.48, 0.52]	[0.28, 0.29]	[0.29, 0.40]	[0.51, 0.64]	[0.05, 0.11]	[0.23, 0.27]	[0.56, 0.63]	[0.12, 0.15]

Table 4 Set of bounds of probabilities for each linguistic term

Ins.	VM			RP	
	VM_L	VM_M	VM_H	RP_{No}	RP_{Yes}
1	[0.33, 0.33]	[0.66, 0.66]	0	0	1
2	0	[0.83, 0.87]	[0.12, 0.16]	[0, 1]	[0, 1]
3	0	0	[1, 1]	1	0
4	[0.11, 0.11]	[0.88, 0.88]	0	0	1
5	[0.04, 0.09]	[0.90, 0.95]	0	1	0
6	[0.88, 0.88]	[0.11, 0.11]	0	0	1
7	[1, 1]	0	0	0	1
8	[0.05, 0.07]	[0.92, 0.94]	0	[0.6, 0.9]	[0.1, 0.4]
9	[0.11, 0.11]	[0.88, 0.88]	0	1	0
10	0	[0.55, 0.55]	[0.44, 0.44]	0.5	0.5
11	[0.22, 0.22]	[0.77, 0.77]	0	[0.2, 0.6]	[0.4, 0.8]
12	[0.11, 0.11]	[0.88, 0.88]	0	0	1
13	[0.22, 0.22]	[0.77, 0.77]	0	0	1
14	0	[1, 1]	0	1	0
15	0	[0.11, 0.11]	[0.88, 0.88]	1	0
16	0	[1, 1]	0	0	1
17	0	[1, 1]	0	0	1
Count	[3.09, 3.16]	[11.33, 11.44]	[2.46, 2.49]	[6.29, 8]	[9, 10.7]
Count (%)	[0.182, 0.186]	[0.66, 0.67]	[0.144, 0.146]	[0.37, 0.47]	[0.52, 0.62]

Table 5 Set of 1-itemsets that compose L_1

Itemset	$\max\{\overline{\text{Count}}_{jk}\}$
DPE_{Low}	0.21
$\text{DPE}_{\text{Middle}}$	0.52
DPE_{High}	0.29
MC_{Low}	0.4
$\text{MC}_{\text{Middle}}$	0.64
MC_{High}	0.11
EI_{Low}	0.27
$\text{EI}_{\text{Middle}}$	0.63
EI_{High}	0.15
VM_{Low}	0.18
$\text{VM}_{\text{Middle}}$	0.67
VM_{High}	0.14
RP_{No}	0.47
RP_{Yes}	0.62

Table 6 Subset of C_2. DPE_{L} with the rest of itemset of L_1

Itemset
$(\text{DPE}_{\text{Low}}, \text{MC}_{\text{Low}})$
$(\text{DPE}_{\text{Low}}, \text{MC}_{\text{Middle}})$
$(\text{DPE}_{\text{Low}}, \text{MC}_{\text{High}})$
$(\text{DPE}_{\text{Low}}, \text{EI}_{\text{Low}})$
$(\text{DPE}_{\text{Low}}, \text{EI}_{\text{Middle}})$
$(\text{DPE}_{\text{Low}}, \text{EI}_{\text{High}})$
$(\text{DPE}_{\text{Low}}, \text{VM}_{\text{Low}})$
$(\text{DPE}_{\text{Low}}, \text{VM}_{\text{Middle}})$
$(\text{DPE}_{\text{Low}}, \text{VM}_{\text{High}})$
$(\text{DPE}_{\text{Low}}, \text{RP}_{\text{No}})$
$(\text{DPE}_{\text{Low}}, \text{RP}_{\text{Si}})$

will be: $\overline{C}_{y_1}, \overline{C}_{y_4}, \overline{C}_{y_2}$ and $\overline{C}_{y_3}$. In this case, as the number of sets is par, the sets that represents the median are the sets $\overline{C}_{y_4}$ and $\overline{C}_{y_2}$ and, the minimum will be $(0.27 + 0.53)/2 = 0.4$. As the value obtained is <0.5, the rule is not relevant although for several possible values of V_Y the rule seems relevant.

In Fig. 3, we show a diagram outlining the steps that are needed to achieve the proposed algorithm.

4 Illustrative example

A small real-world dataset is given to illustrate the proposed data-mining algorithm.

This dataset is a study of the "Athletics event" in the University of Oviedo; in this case the Jump event, that includes 17 instances and 5 attributes. These attributes are (Vinuessa and Coll 1984; J.P. Martin, 2009, personal communication): (1) the ratio between the weight and the height (DPE); (2) tests of central (abdominal) muscles (MC); (3) test of lower extremities (EI); (4) the maximum speed in the 40-m race (VM) and (5) relevance of the athlete (RA). Table 2 shows this dataset where the low-quality items are represented by interval-values or fuzzy subsets (see Sect. 3.1).

Let us use three fuzzy regions for each attribute. Figure 4 shows the fuzzy membership functions for the different attributes: "R_{Low}", "R_{Middle}" and "R_{High}".

These partitions are only relevant for the attributes represented by interval-values (DPE, MC, EI and VM) due to RA being defined by a fuzzy set.

The values considered for the input parameters of this illustrative example are:

- Minimum support (α) = 0.05 (5%)
- Confidence (λ) = 0.9 (90%)
- Number of cuts = 10
- Number of times that we sweep the probabilities (γ) = 1.

The steps needed to obtain the fuzzy association rules from LQD are:

Step 1 Transform each LQD $\widetilde{X}_j^i$ to obtain the fuzzy membership value. For instance, the fuzzy membership value interpreted as a set of bounds of probabilities from the interval value [47, 52], of the first instance of the attribute MC, will be:

$$\{[0, 0.11]_{\text{Low}} + [0.88, 1]_{\text{Middle}} + [0, 0.11]_{\text{High}}\}$$

where the upper and lower bounds of probabilities of the linguistic label Middle, for example, are determined from the set of possible membership values of x, with $x \in \overline{X}$, which means with $x \in [47, 52]$. To obtain these possible values we apply α-cuts to obtain a random set that contains the unknown crisp value of the feature with a probability greater than or equal to $1 - \alpha$.

Step 2 Calculate the frequency of each fuzzy. For instance, the frequency of the attribute MC and the linguistic term Low will be: $\text{Count}_{\text{MC}_{\text{Low}}} = [0, 0.11] \oplus [0, 0] \oplus [0.77, 1] \quad \oplus \cdots \oplus \quad [0.22, 0.22] \oplus [0.66, 0.77] = [5.01, 6.84]$. Tables 3 and 4 show the frequency of each fuzzy item where each one will be a candidate (1-itemset).

$$C_{r=1} = \{\text{DPE}_{\text{L}}, \text{DPE}_{\text{M}}, \text{DPE}_{\text{H}}, \text{MC}_{\text{L}}, \text{MC}_{\text{M}}, \text{MC}_{\text{H}}, \text{EI}_{\text{L}}, \text{EI}_{\text{M}}, \text{EI}_{\text{H}}, \text{VM}_{\text{L}}, \text{VM}_{\text{M}}, \text{VM}_{\text{H}}, \text{RP}_{\text{No}}, \text{RP}_{\text{Yes}}\}$$

Step 3 Check whether $\overline{\text{Count}}_{L_{j_k}}(\%)$ for all r-items in C_r is larger than or equal to the predefined minimum support α. Table 5 shows all the 1-itemsets that compose L_1. The item $\text{DPE}_{\text{Middle}} = \text{DPE}_{\text{M}}$, with $\overline{\text{Count}}_{\text{DPE}_{\text{M}}}(\%) = [0.48, 0.52]$, will be part of L_1 due to $0.52 > 0.05$.

Step 4 Since L_1 is not null, the next step is then done. If L_1 is null the algorithm finishes.

Table 7 fuzzy membership value of $\mathrm{DPE_L} \bigwedge \mathrm{MC_M}$

Ins.	$\mathrm{DPE_L}$	$\mathrm{MC_M}$	$\mathrm{DPE_L} \bigwedge \mathrm{MC_M}$
1	0	[0.88, 1]	0
2	[0.78, 0.88]	[0, 0.22]	[0, 0.19]
3	0	[0, 0.22]	0
4	[0.33, 0.44]	[0.66, 0.77]	[0.22, 0.34]
5	[0.88, 1]	[0.77, 1]	[0.69, 1]
6	0	[0.66, 0.77]	0
7	[0.22, 0.33]	[0.77, 1]	[0.17, 0.33]
8	[0.11, 0.11]	[0.22, 0.55]	[0.02, 0.06]
9	0	[0.70, 0.76]	0
10	0	[0.22, 0.33]	0
11	[0.11, 0.11]	[0.77, 1]	[0.08, 0.11]
12	0.5	[0.88, 1]	[0.44, 0.5]
13	[0.33, 0.33]	[0.88, 1]	[0.29, 0.33]
14	0	[0.11, 0.11]	0
15	0	[0.11, 0.11]	0
16	0	[0.77, 0.77]	0
17	0	[0.22, 0.33]	0
Count	[3.28, 3.71]	[8.70, 10.98]	[1.93, 2.88]
Count (%)	[0.19, 0.21]	[0.51, 0.64]	[0.11, 0.16]

Table 8 A subset of r-itemset from L_2

Itemset	$\max\{\overline{\mathrm{Count}_s}\}$
$(\mathrm{DPE_L}, \mathrm{MC_M})$	0.16
$(\mathrm{DPE_L}, \mathrm{MC_H})$	0.06
$(\mathrm{DPE_L}, \mathrm{EI_L})$	0.05
$(\mathrm{DPE_L}, \mathrm{EI_M})$	0.15
$(\mathrm{DPE_L}, \mathrm{VM_M})$	0.17
$(\mathrm{DPE_L}, \mathrm{RP_{No}})$	0.12
$(\mathrm{DPE_L}, \mathrm{RP_{Yes}})$	0.15

Step 5 Join L_r ($r = 1$) to generate the candidate C_{r+1}. C_2 is generated as follows: $(\mathrm{DPE_L}, \mathrm{MC_L}), (\mathrm{DPE_L}, \mathrm{MC_M}), \ldots, (\mathrm{VM_H}, \mathrm{RP_{Yes}})$. Table 6 shows a subset of the candidate C_2 from the combination between $\mathrm{DPE_L}$ and the rest of the itemset of L_1. Notice that the itemsets $(\mathrm{DPE_L}, \mathrm{DPE_M})$, $(\mathrm{DPE_L}, \mathrm{DPE_H})$ and $(\mathrm{DPE_M}, \mathrm{DPE_H})$ are not in C_2 since the items belong to the same item DPE.
Step 6 For each r-itemset of C_r make the following substeps:

a. Transform each r-itemset to obtain the fuzzy membership value. Table 7 shows the results of all the instances with respect to $(\mathrm{DPE_L}, \mathrm{MC_M})$. For instance, the value [0, 0.19], of the 2-itemset $(\mathrm{DPE_L}, \mathrm{MC_M})$, in the instance $i = 2$, is calculated from the product of the sets of probabilities of each itemset of r-itemset ([0.78, 0.88] $\otimes$ [0, 0.22] = [0, 0.19]).
b. Calculate the frequency of each r-itemset. In Table 7 the results of the r-itemset $(\mathrm{DPE_L}, \mathrm{MC_M})$ in all the instances are shown.
c. Check whether these sets of bounds are larger than or equal to the minimum support to insert the r-itemset in L_2. From the subset of candidate C_2 (Table 6), the r-itemsets s that compose $L_{r+1}(L_2)$ are shown in Table 8.

Step 7 If $L_r + 1$ is null, then do the next step, otherwise update the number of itemsets in the set of candidates ($r = r + 1$) and repeat the steps 5 and 6.
Step 8 Collect in R the itemsets of each L_i, where $i \geq 2$.
Step 9 Construct all the possibles fuzzy association rules from the itemsets of R. From the subset of L_2 (Table 8), the association rules possible are shown in Table 9.
Step 10 Determine whether the fuzzy association rules obtained are relevant or must be deleted. Table 10 shows the results of $\overline{C}_{yj}$ with $\gamma = 1$ in the rule "If $\mathrm{DPE_M}$ and $\mathrm{MC_L}$ and $\mathrm{VM_H}$ then $\mathrm{RP_{No}}$", with $\overline{X}$ = [0.0665, 0.0767] and $\overline{Y}$ = [0.0665, 0.0767]. The values of $\overline{C}_{yj}$ are arranged according to the uniform dominance, in this example a strict dominance due to the value of γ being 1, and in this way can choose the set that represents the median. This fuzzy association rule will be relevant because its median takes the value 0.632.

Table 9 Fuzzy association rules obtained from L_2

If DPE.Low then MC.Middle
If MC.Middle then DPE.Low
If DPE.Low then MC.High
If MC.High then DPE.Low
If DPE.Low then EI.Low
If EI.Low then DPE.Low
If DPE.Low then EI.Middle
If EI.Middle then DPE.Low
If DPE.Low then VM.Middle
If VM.Middle then DPE.Low
If DPE.Low then RP.No
If RP.No then DPE.Low
If DPE.Low then RP.Si
If RP.Si then DPE.Low

The fuzzy association rules obtained in this illustrative example are shown in Table 11.

5 Experimental study

Several experiments have been carried on a real-world dataset Inexpert-57 to evaluate the good behavior of this proposed algorithm, which is available in the repository KEEL-dataset (http://www.keel.es/dataset.php) (Alcala-Fdez et al. 2011). In the following subsections, we describe the real-world dataset as well as the experiments carried out. Then, we analyze the fuzzy association rules according to the value of the minimum support and confidence and will show the high level of knowledge obtained in these rules. Finally, we study the complexity and scalability of the proposed algorithm.

5.1 Description of the dataset

Dyslexia can be defined as a learning disability in people with normal intellectual coefficient, and without further physical or psychological problems that can explain such a disability. According to Thomson and Gilchrist (1996), dyslexia is a neurologically based, often familial, disorder which interferes with the acquisition and processing of language […]. Although dyslexia is lifelong, individuals with dyslexia frequently respond successfully to timely and appropriate intervention.

In this research we are interested in obtaining fuzzy association rules in the early diagnosis of dyslexia of schoolchildren in Asturias (Spain), where this disorder is not rare. It has been estimated that between 4 and 5% of these schoolchildren have dyslexia. The average number of children in a Spanish classroom is 25, therefore there are cases in most classrooms (Ajuriaguerra 1976). Notwithstanding the widespread presence of dyslexic children, detecting the problem at this stage is a complex process, that depends on many different indicators, mainly intended to detect whether reading, writing and calculus skills are acquired at the proper rate. Moreover, there are disorders different from dyslexia that share some of their symptoms and therefore the tests not only have to detect abnormal values of the mentioned indicators but in addition, must also separate those children that actually suffer from dyslexia from those where the problem can be related to other causes (inattention, hyperactivity, etc.).

We have considered a real-world dataset Inexpert-57 regarding this experimentation. For its elaboration, all schoolchildren in Asturias were examined by a psychologist in diagnose dyslexia from several tests.

With these tests, from the criterion and knowledge of the inexpert and expert, several LQD were obtained.

Table 10 Steps to determine whether the fuzzy association rule "If DPE_M and MC_L and VM_H then RP_{No}" is relevant or must be deleted

$\overline{Y}_{cut}$	$\overline{X}_{cut}$	$P_{\overline{X}_{cut}}$	Check whether $\overline{X}_{cut} \geq (\overline{Y}_{cut} * \lambda)$ and $\overline{X}_{cut} \leq \overline{Y}_{cut}$					C_{cut}
			$y_1 * \lambda = 0.069$	$y_2 * \lambda = 0.068$	$y_3 * \lambda = 0.067$	...	$y_{10} * \lambda = 0.059$	
$y_1 = 0.076$	$x_1 = 0.0767$	0.15	Yes	$(x_1 > y_2)$	$(x_1 > y_3)$	...	$(x_1 > y_{10})$	0.15
$y_2 = 0.075$	$x_2 = 0.0756$	0.07	Yes	Yes	$(x_2 > y_3)$	...	$(x_2 > y_{10})$	0.24
$y_3 = 0.744$	$x_3 = 0.0744$	0.087	Yes	Yes	Yes	...	$(x_3 > y_{10})$	0.307
$y_4 = 0.733$	$x_4 = 0.0733$	0.23	Yes	Yes	Yes	...	$(x_4 > y_{10})$	0.537
$y_5 = 0.0722$	$x_5 = 0.0722$	0.19	Yes	Yes	Yes	...	$(x_5 > y_{10})$	0.727
$y_6 = 0.0710$	$x_6 = 0.0710$	0.27	Yes	Yes	Yes	...	$(x_6 > y_{10})$	0.997
$y_7 = 0.0699$	$x_7 = 0.0699$	0.003	Yes	Yes	Yes	...	$(x_7 > y_{10})$	1
$y_8 = 0.0688$	$x_8 = 0.0688$	0	No	Yes	Yes	...	$(x_8 > y_{10})$	0.85
$y_9 = 0.0676$	$x_9 = 0.0676$	0	No	No	Yes	...	$(x_2 > y_{10})$	0.78
$y_{10} = 0.0665$	$x_{10} = 0.0665$	0	No	No	No	...	Yes	0

Table 11 Fuzzy association rules obtained from the illustrative example

Rule	Median
R0: If DPE is Middle and MC is Low and VM is High then RP is No	0.632
R1: IF DPE IS Middle AND EI IS Middle AND VM IS High THEN RP IS No	0.682
R2: IF DPE IS Middle AND VM IS High THEN RP IS No	0.921
R3: IF DPE IS High AND MC IS Middle AND EI IS High THEN RP IS Si	0.5
R4: IF DPE IS High AND EI IS High THEN RP IS Si	0.778
R5: IF DPE IS High AND EI IS High AND VM IS Low THEN RP IS Si	0.614
R6: IF DPE IS High AND VM IS Low THEN RP IS Si	0.581
R7: IF MC IS Low AND EI IS Middle AND VM IS High THEN RP IS No	1
R8: IF MC IS Low AND VM IS High THEN RP IS No	0.6
R9: IF MC IS Middle AND EI IS High THEN RP IS Si	0.5
R10: IF MC IS Middle AND EI IS High AND VM IS Low THEN RP IS Si	0.667
R11: IF EI IS Middle AND VM IS High THEN RP IS No	0.841
R12: IF EI IS High THEN RP IS Si	0.73
R13: IF EI IS High AND VM IS Low THEN RP IS Si	1

Table 12 Categories of the tests currently applied in Spanish schools for detecting dyslexia when an expert evaluates the children

Category	Test	Description
Verbal comprehension	BAPAE	Vocabulary
	BADIG	Verbal orders
	BOEHM	Basic concepts
Logic reasoning	RAVEN*	Color
	BADIG	Figures
	ABC*	Actions and details
Memory	Digit WISC-R*	Verbal-additive memory
	BADIG*	Visual memory
	ABC	Auditive memory
Level of maturation	ABC	Combination of different tests
Sensory–motor skills	BENDER*	Visual–motor coordination
	BADIG	Perception of shapes
	BAPAE*	Spatial relations, shapes, orientation
	STAMBACK	Auditive perception, rhythm
	HARRIS/HPL	Laterality, pronunciation
	ABC	Pronunciation
	GOODENOUGHT	Spatial orientation, body scheme
Attention	Toulose*	Attention and fatigability
	ABC*	Attention and fatigability
Reading–writing	TALE	Analysis of reading and writing

The tests marked with a "*" are not included in the version of the low-quality dataset obtained from inexpert in the field of dyslexia

The objective is to obtain relevant information through fuzzy association rules when an inexpert in the field of dyslexia (parents, tutors) evaluates the children. To this end, the inexpert expresses what he/she is observing from the tests obtained when one child is evaluated. The tests applied in Spanish schools for detecting this problem, when an expert evaluates the children, are shown in Table 12 (the tests marked with a "*" are not included in this version of the low-quality dataset obtained from inexpert in the field of dyslexia). Each test observed by the inexpert is described by several variables providing more than one item in the low-quality set. This implies that we will have sub-items for each test applied. For instance, the test T.A.L.E (Toro and Cervera 1980) could be defined directly by an item, expressed, for example with a linguistic terms "Medium"; however, this test is defined by several items.

Table 13 Analysis of reading of the test TALE defined by several items

Item	Domain
Analysis of reading. TALE	
1. Reading-comprehension	[0, 10]
2. Global-level	{Impossible, Just-read, Low, Regulate, Normal, Good}
3. Finger-tracking	{Yes, No, Little, A-lot}
4. Move-head-no-eyes	{Yes, No, Little, A-lot}
5. Heard	{Comprehensive, No-comprehensive}
6. Intonation	{Bad, Good, Regulate, Punctuation-no-respected}
7. Syllables	{Yes, No, Occasionally}
8. Investment	{Yes, No, Occasionally}
9. Nervous	{Yes, No}
10. Omission	{Yes, No, Occasionally}
11. Substitution	{Yes, No, Occasionally}
12. Rotation	{Yes, No, Occasionally}
13. Speed	{Bad, Good, Regulate, Normal, Slow}
14. Arrhythmic	{Yes, No, Occasionally}
15. Rectification	{A-lot, Never, Often, Normal, Just}
16. Silent	{Yes, No, Decreases-level}

In Table 13 the items that compose the analysis of reading of the test TALE are shown.

The real-world dataset of dyslexia used in this proposal, denominated "Inexpert-57", is composed of groups of low-quality items where each group of items describes the behavior of a child in one test. Besides these groups of items, this dataset will contain the level of dyslexia of this child when an expert in the field diagnoses the child from the same test that the inexpert has used. In this way, each case or child has been individually diagnosed by a psychologist into one or more of the values "no dyslexia", "control and revision", "dyslexic" and "other disorders" (inattention, hyperactivity, etc.).

This dataset contains vague data; we have collected these data from 52 schoolchildren of Asturias (Spain) during our research and where each case has been individually classified by a psychologist. Each schoolchild is composed of 57 items; these 57 items are obtained from different tests (Table 12). Figure 5 shows the major tests as well as the number of items that compose each test. Moreover, for each item we have indicated whether it is defined by an interval or by a finite set of linguistic terms associated with Ruspini's partitions.

5.2 Experiments settings

The linguistic partitions are composed of several linguistic terms with uniformly distributed triangulars. The number of partitions of each item depends on whether this item is represented by an interval value or fuzzy subset. In the first case the linguistic partitions are composed of five linguistic terms, in the second case the expert defines the linguistic terms and the membership functions in advance. For example, in Table 13, the linguistic terms are defined in advance of several items of the test TALE.Reading. Several experiments have been carried out with different minimum supports and confidences, where the number of cuts is 7 and γ is 2.

5.3 Analysis of the fuzzy association rules via supports and confidence

In this section several experiments have been carried out to analyze the number of fuzzy association rules obtained by the fuzzy data-mining algorithm from low-quality data. The relationship between the number of fuzzy association rules with respect to several values of the minimum support

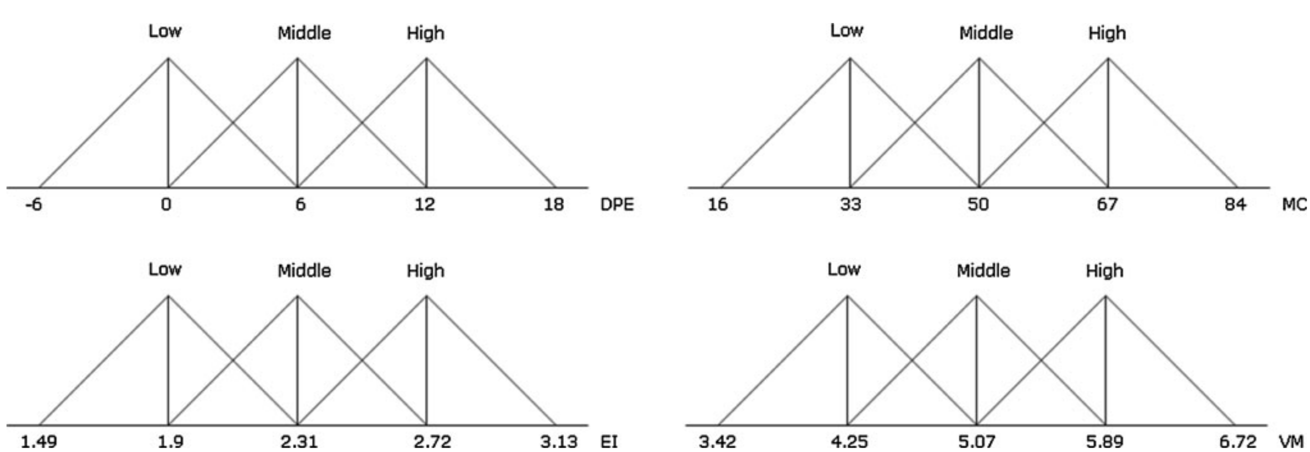

Fig. 4 The membership functions used in the items DPE, MC, EI and VM

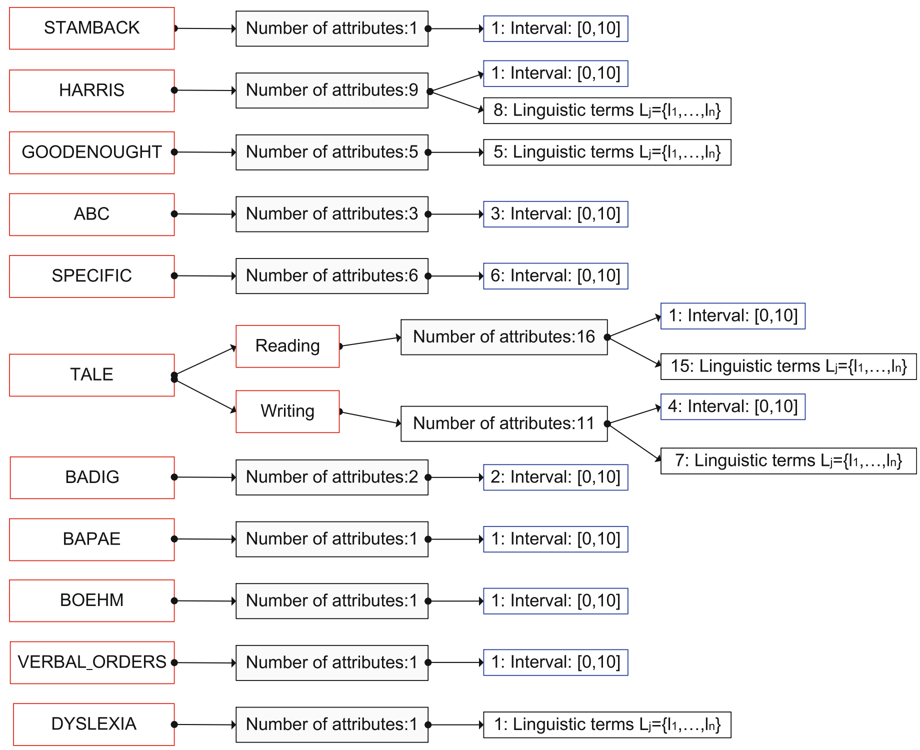

Fig. 5 Inexpert-57 with their main tests, indicating the number of items in each test as well as whether the items are intervals or are defined by linguistic terms

Table 14 Level of knowledge and interpretability in the fuzzy association rules of Inexpert-57

Rules	Support-rule	$\{\overline{C}_{yi}\}$	Median$\{\overline{C}_{yi}\}$
IF Harris IS Right AND Dyslexia IS No-Dyslexia THEN THEN Goodenought-Proportions IS Normal	[0.242, 0.248]	$\{[0.159, 0.209], [0.191, 0.715][0.191, 0.715], [1, 1], [1, 1], [1, 1], [1, 1]\}$	[1, 1]
IF Harris IS Right AND Dyslexia IS No-Dyslexia AND Goodenought-Proportions IS Normal THEN Reading-finger is No	[0.223, 0.248]	$\{[0.862, 0.996], [0.862, 0.996][0.873, 1], [0.873, 1], [0.873, 1], [0.873, 1], [1, 1]\}$	[0.873, 1]
IF Harris IS Right AND Dyslexia IS No-Dyslexia AND Goodenought-Proportions IS Normal THEN Reading is Comprehensive	[0.223, 0.248]	$\{[0.821, 0.924], [0.821, 0.924][0.963, 0.99], [0.963, 0.99], [0.963, 0.99], [0.963, 0.99], [1, 1]\}$	[0.963, 0.99]
IF Harris IS Right AND Dyslexia IS No-Dyslexia AND Goodenought-Proportions IS Normal THEN Reading-eyes-moved is No	[0.223, 0.248]	$\{[0.95, 0.996], [0.95, 0.996][0.957, 1], [0.957, 1], [0.957, 1], [0.957, 1], [1, 1]\}$	[0.957, 1]

along with different minimum confidences λ is shown in Fig. 6. We can observe that the number of rules decreases when the minimum support value increases. Moreover, we appreciate that the curves obtained have similar shapes and the distance between them is small with values of the minimum support larger than 0.2. With minimum support 0.2 and particularly with 0.1 the distances between the curves is more elevated, notably when the minimum confidence takes the value 0.5 or 0.6. This implies that there are number of rules that are uncommon or are special cases, highlighting the large distance between curves when the minimum support is 0.1.

Figure 7 shows the relations between the number of fuzzy association rules and several values of the minimum confidence along with different minimum support values. We can observe that the number of rules increases when

Table 15 Level of knowledge and interpretability in the fuzzy association rules of Inexpert-57 where the consequent is the level of dyslexia

Rules	Support-rule	$\{\overline{C}_{yi}\}$	Median$\{\overline{C}_{yi}\}$
IF Harris IS Right AND Reading-investments IS No AND Goodenought-Proportions IS Normal THEN Dyslexia IS No-Dyslexia	[0.204, 0.229]	$\{[0.115, 0.431], [0.22, 0.99][0.982, 0.991], [1, 1], [1, 1], [1, 1], [1, 1]\}$	[1, 1]
IF Harris IS Right AND Reading-substitutions IS No AND Goodenought-Proportions IS Normal THEN Dyslexia IS No-Dyslexia	[0.212, 0.231]	$\{[0.007, 0.018], [0.011, 0.027][0.014, 0.056], [1, 1], [1, 1], [1, 1], [1, 1]\}$	[1, 1]

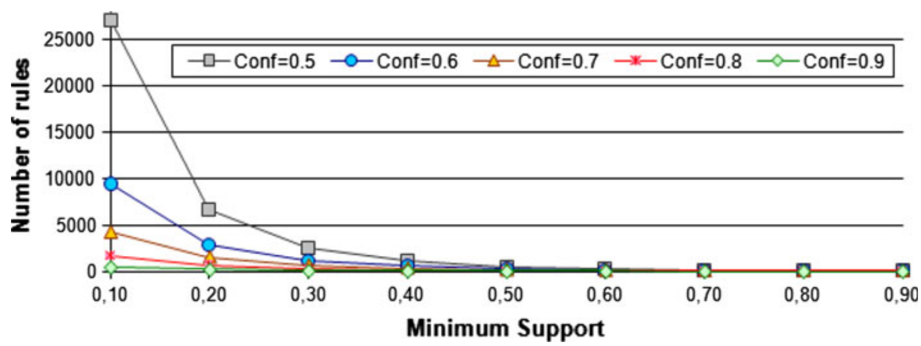

Fig. 6 Relationship between the number of fuzzy association rules and the minimum support along with different minimum confidences

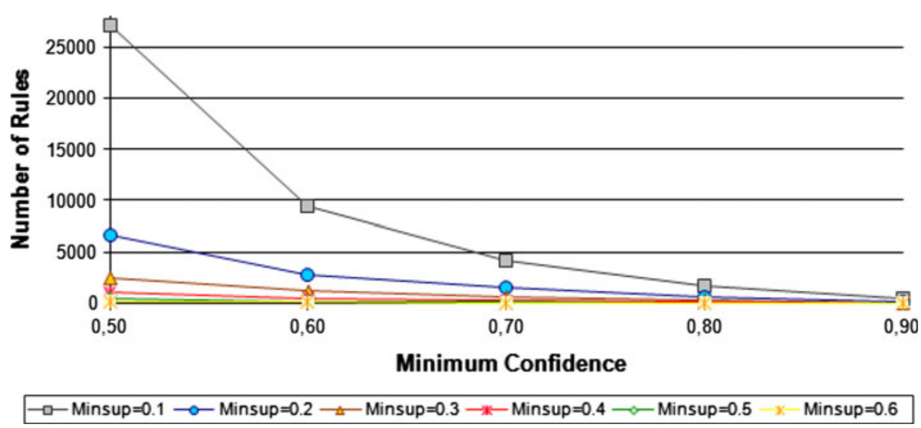

Fig. 7 Relationship between the number of fuzzy association rules and the minimum confidence along with different minimum supports

the minimum confidence decreases. Notice that the minimum confidence influences the number of fuzzy rules when the minimum support takes small values such as 0.1 and 0.2. On the other hand, we appreciate that with a minimum support larger than 0.2, there are many rules that satisfy the minimum confidence when this is increased.

5.4 High level of knowledge of Inexpert-57

To analyze the fuzzy rules obtained from this proposal, several experiments have been carried out with the dataset "Inexpert-57". The information obtained provides a high level of knowledge through the fuzzy association rules. The number of fuzzy association rules obtained with a minimum support of 0.2 and a minimum confidence of 0.8, in "Inexpert-57", was 669. An example of the level and interpretability of information obtained is shown in Table 14 with several relevant fuzzy association rules. These rules show that one child diagnosed as "No dyslexic" and, in the test of Goodenought, in the subitem denominated proportions, obtains a "normal" result, then the child is not going to have problems in reading due to

Table 16 Special cases of the dataset "Inexpert-57" with $\alpha = 0.1$ and $\lambda = 0.5$

Rules	Support-rule	$\{\overline{C}_{yi}\}$	Median$\{\overline{C}_{yi}\}$
IF Harris-dotted IS Right AND Reading-investments IS Yes AND Goodenought-Proportions IS Normal AND Dyslexia IS Dyslexic THEN Writing-Omissions IS Yes	[0.038, 0.106]	$\{[0.081, 0.358], [0.28, 0.285], [0.715, 0.72], [0.634, 0.899], [0.637, 0.899], [0.81, 0.978], [0.977, 0.98]\}$	[0.634, 0.899]
IF Harris-dotted IS Right AND Boehm-Concepts IS Medium AND Goodenought-Proportions IS Normal AND Dyslexia IS No-Dyslexic THEN Reading-nervous IS No	[0.039, 0.101]	$\{[0.046, 0.117], [0.191, 0.221], [0.77, 0.808], [0.776, 0.808], [0.883, 0.954], [0.964, 0.992], [0.989, 0.992]\}$	[0.776, 0.808]
IF Harris-dotted IS Right AND Reading IS Regular AND Goodenought-Global IS Regular AND Writing-unions IS Yes THEN Dyslexia IS Dyslexic	[0.079, 0.11]	$\{[0.008, 0.103], [0.045, 0.604], [0.067, 0.806], [0.897, 0.992], [1, 1], [1, 1], [1, 1]\}$	[0.897, 0.992]
IF Harris-dotted IS Right AND Dyslexia IS Other-disorders THEN THEN Goodenought-Global IS Regular	[0.095, 0.101]	$\{[1, 1], [1, 1][1, 1], [1, 1], [1, 1], [1, 1], [1, 1]\}$	[1, 1]

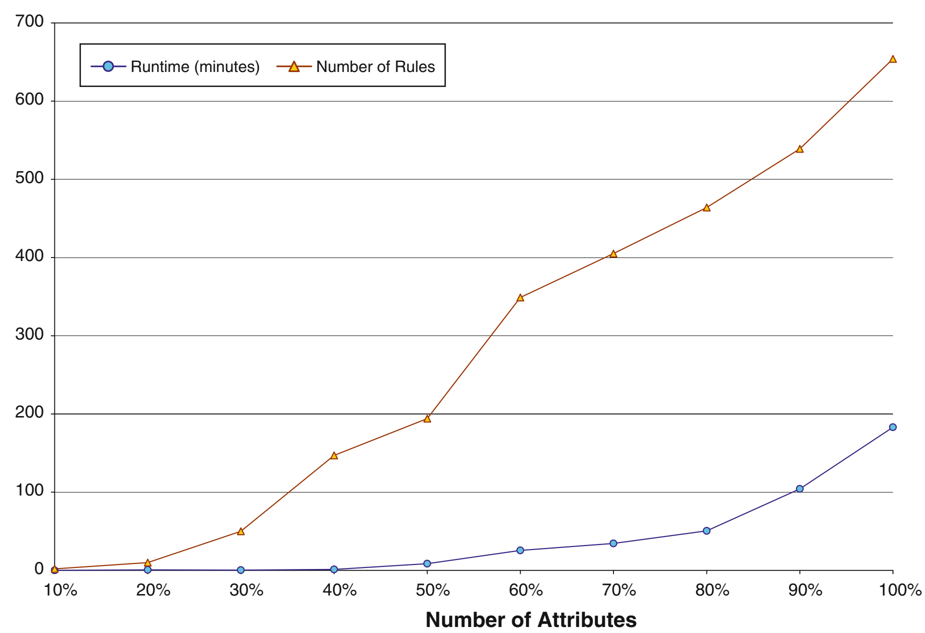

Fig. 8 Relationship between the runtime (minutes) and the number of attributes with the 100% of instances, $\alpha = 0.2$, $\lambda = 0.8$, cuts = 10 and $\gamma = 2$. The number of rules is also shown

the typical characteristics such as eyes move, read with finger, not understanding, etc. are not obtained. Also, this information provides us with the information that we will have to control the children that we see are "No dyslexic", have a "normal" result in the test of Goodenought but have problems in reading.

The rules shown in Table 15 provide different information to that found in the previous rules. In this case, these rules determine when one child has to be diagnosed as "No dyslexic". We can observe that the test of Goodenought and reading are very relevant tests to diagnose children and particularly two subitems of the test of reading: investments and substitutions. We can appreciate that whether the children are right-handed in the test of Harris is relevant or not due to this variable or test appearing in the most of rules obtained. This is the consequence of most children studied and diagnosed being right-handed so we can determine that this variable is irrelevant.

Others rules provide information in relation to children with dyslexia, for instance IF Reading-Syllables IS Yes AND Dyslexia IS Dyslexic THEN Writing-Unions IS Yes. In addition, it is important to highlight that when the values of the minimum confidences and support are small then the number of fuzzy rules increases and provides information about special cases of the dataset and new items appear in these rules (as BOEHM-Concepts). For instance, in Table 16 we show several rules obtained with $\alpha = 0.1$ and $\lambda = 0.5$ although, we can observe that the minimum of the median of the set $\{\overline{C}_{yi}\}$ in some rules is more elevated than 0.5, for instance in the third rule the minimum is 0.897 or in the last one it is 1.

5.5 Analysis of complexity and scalability

The complexity and scalability of this fuzzy data-mining algorithm from LQD has been analyzed from several experiments carried out with an HP EliteBook 8540w, processor Intel(R) Core(TM)i5, 2.4 GHz CPU, 4 Gb of RAM and running in Windows 7. All the experiments were performed with $\alpha = 0.2, \lambda = 0.8$, cuts = 10 and $\gamma = 2$.

To analyze the complexity and scalability, we compare the relationship between the runtime and the number of items. Figure 8 shows the relationship between the runtime and the number of items, observing that the time increases as well as the number of rules when the number of items also increases.

6 Conclusions

In this paper, we have proposed a new data-mining algorithm with the aim of getting high-quality fuzzy association rules from databases with interval and fuzzy values. This proposal is an extension of the algorithm proposed by Hong et al., which integrates fuzzy-set concepts with the apriori mining algorithm (Agrawal and Srikant 1994) from quantitative values. To do that, several important aspects have been considered due to the true value of one data being unknown and the fuzzy membership value interpreted as a set of bounds of probabilities. This affects the calculation of the frequency of occurrence of the items, due to it being defined by a set of bounds of probabilities, as well as the calculation of the confidence of a rule which is contained in a set of probabilities.

The behavior and performance of this new algorithm, able to obtain fuzzy association rules from low-quality data, is shown from one real-world dataset based on the Diagnosis of Dyslexia, obtaining as a result a high level of knowledge and interesting patterns. These fuzzy association rules also provide us with information about special cases, in the low-quality dataset of diagnosis of dyslexia, when the value of the minimum support and confidence decreases. Notice that these rules from LQD provide knowledge about the dependencies and relation between the items and, therefore, several items can be excluded or removed due to their being considered irrelevant.

Acknowledgments This study was supported by the Spanish Ministry of Education and Science under Grants no. TIN2008-06681-C06-{01 and 04}, TIN2011-28488 and by the Principado de Asturias under Grant PCTI 2006–2009.

References

Agrawal R, Imielinski T, Swami A (1993) Mining association rules between sets of items in large databases. In: SIGMOD, Washington, D.C., USA, pp 207–216

Agrawal R, Srikant R (1994) Fast algorithms for mining association rules. In: International conference on very large data bases, Santiago de Chile, pp 487–499

Ajuriaguerra J (1976) Manual de psiquiatría infantil. Barcelona, Toray-Masson

Alatas B, Akin E (2006) An efficient genetic algorithm for automated mining of both positive and negative quantitative association rules. Soft Comput Fusion Found Methodol Appl 10(3):230–237

Alcala-Fdez J, Fernandez A, Luego J, Derrac J, Garcia S, Sanchez L, Herrera F (2011) Keel data-mining software tool: data set repository, integration of algorithms and experimental analysis framework. J Multiple-Valued Log Soft Comput 17(2–3):255–287

Alcala-Fdez J, Flugy-Pape N, Bonarini A, Herrera F (2010) Analysis of the effectiveness of the genetic algorithms based on extraction of association rules. Fundamenta Informaticae 98(1):1–14

Baudrit C, Dubois D, Perror N (2008) Representing parametric probabilistic models tainted with imprecision. Fuzzy Sets Syst 15(1):1913–1928

Bertoluzza C, Gil M, Ralescu D (2003) Statistical modeling. Analysis and management of fuzzy data. Springer, Berlin

Chen C, Hong T, Tseng V (2011) Genetic-fuzzy mining with multiple minimum supports based on fuzzy clustering. Soft Comput Fusion Found Methodol Appl. doi:10.1007/s00500-010-0664-1

Couso I, Sanchez L (2008) Higher order models for fuzzy random variables. Fuzzy Sets Syst 159:237–258

Delgado M, Marín N, Sánchez D, Vila M (2003) Fuzzy association rules: general model and applications. IEEE Trans Fuzzy Syst 11(2):214–225

Dubois D, Hullermeier E, Prade H (2006) A systematic approach to the assessment of fuzzy association rules. Data Min Knowl Disc 13(2):167–192

Dubois D, Prade H (1992) When upper probabilities are possibility measures. Fuzzy Sets Syst 49:65–74

Dubois D, Prade H, Sudamp T (2005) On the representation, measurement, and discovery of fuzzy associations. IEEE Trans Fuzzy Syst 13(2):250–262

Han J, Kamber M (2006) Data mining: concepts and techniques, 2nd edn. Morgan Kaufmann, San Fransisco

Han J, Pei J, Yin Y (2004) Mining frequent patterns without candidate generation: a frequent-pattern tree approach. Data Min Knowl Dis 8(1):53–87

Hong T, Kuo C, Chi S (1999) Mining association rules from quantitative data. Intell Data Anal 3(5):363–376

Hong T, Kuo C, Chi S (2001) Trade-off between time complexity and number of rules for fuzzy mining from quantitative data. Int J Uncertain Fuzziness Knowl Based Syst 9(5):587–604

Hong T, Lee Y (2008) An overview of mining fuzzy association rules. In: Bustince H, Herrera F, Montero J (eds) Studies in fuzziness and soft computing, vol 220. Springer, Berlin, pp 397–410

Hullermeier E, Yi Y (2007) In defense of fuzzy association analysis. IEEE Trans Syst Man Cybern Part B Cybern 37(4):1039–1043

Kaufmann A, Gupta M (1991) Introduction to fuzzy arithmetic: theory and applications. Van Nostrand Reinhold, New York

Kaya M (2006) Multi-objective genetic algorithm based approaches for mining optimized fuzzy association rules. Soft Comput Fusion Found Methodol Appl 10(7):578–586

Limbourg P (2005) Multi-objective optimization of problems with epistemic uncertainty. In Proceedings of EMO, pp 413–427

Mladenic D, Lavrac N, Bohanec M, Moyle S (2002) Data mining and decision support: integration and collaboration. Kluwer, Norwell

Palacios A, Sanchez L, Couso I (2011) Future performance modelling in athletism with low quality data-based GFSs. J Multiple-Valued Log Soft Comput 17(2–3):207–228

Ruspini E (1969) A new approach to clustering. Inf Control 15:22–32

Sanchez L, Couso I, Casillas J (2007) Modelling vague data with genetic fuzzy systems under a combination of crisp and imprecise criteria. In: IEEE symposium on computational intelligence inmulticriteria decision making, pp 30–37

Sanchez L, Couso I, Casillas J (2009) Genetic learning of fuzzy rules on low quality data. Fuzzy Sets Syst 160(17):2524–2552

Sanchez L, Suarez M, Villar J, Couso I (2008) Mutual information-based feature selection and partition design in fuzzy rule-based classifiers from vague data. Int J Approx Reason 49:607–622

Sudkamp T (2005) Examples, counterexamples, and measuring fuzzy associations. Fuzzy Sets Syst 149(1):57–71

Sun K, Fengshan B (2008) Mining weighted association rules without preassigned weights. IEEE Trans Knowl Data Eng 20(4):489–495

Thomson P, Gilchrist P (1996) Dyslexia: a multidisciplinary approach. Chapman and Hall, London

Toro J, Cervera M (1980) TALE Test de Análisis de la lectoescritura. Pablo del Río, Madrid

Villar J, Otero A, Otero J, Sanchez L (2009) Taximeter verification using imprecise data from gps and multiobjective algorithms. Eng Appl Artif Intell 22:250–260

Vinuessa M, Coll J (1984) Tratado de atletismo. Servicio Geográfico del Ejército

Wu B, Sun C (2001) Interval-valued statistics, fuzzy logic, and their use in computational semantics. J Intell Fuzzy Syst 1–2(11):1–7

Zhang C, Zhang S (2002) Association rule mining: models and algorithms. Springer, Berlin