Linear and Nonlinear Analysis
Volume 9, Number 1, 2023, 1–23

ISSN 2188-8167

FUZZY ASSOCIATION RULE CLASSIFICATION MODEL ON FRFS AND CFFP-GROWTH ALGORITHMS

WEI-HSUAN LEE, CHIEN-HUA WANG, CHIA-HSUAN YEH, AND CHIN-TZONG PANG*

Abstract. Classification mining is to draw relative patterns from each class forming certain connections. In reality, data mining is all about a more proficient measurement and better storage capacity when it comes to processing enormous amount of data. This paper is targeted at introducing concepts of attribution reduction and later establishing new fuzzy data mining technologies. The attribution reduction through the fuzzy rough feature selection algorithm, as a preprocessing role, obtains reduced sets of attributes from a large amount of data by selecting relevant attributes in which human interventions are not necessary. Then, the proposed classified fuzzy FP-growth algorithm acts to manage classification data by means of providing a new tree structure for the facilitating construction of patterns and restricting the numbers of linguistic terms and combinations for efficiency. Results of this paper indicate that the reduced data sets have beneficial effects on the efficiency during mining processes. The new tree structure and the restriction involve lower computation without comprising the classification accuracy. These findings may have implications for the fields of attribute reduction and fuzzy data mining.

1. Introduction

In computerized period, what matters to the industrial field is the art of decision-making in the database management. Choosing rational and consequential information from extremely large data is what lies beneath data mining [17]. Bearing this in mind, it behooves one to combine association discovery and classification mining in which the former is to find frequent patterns among data whereas the latter is for these patterns to draw linkages with class labels [48].

When processing massive data, it appears that whole sizes of data are a matter of great importance. To operate effectively in this scenario is to utilize data reduction at a nascent stage so that the reduced representation of original data is served for data analysis and in turn similar results are gained [44]. Feature selection (FS) is on point in cases like these that one is able to use the selected features in reduced data with less computation time and less storage space. When compared with the outcome gained from the original data, FS generates identical results [31].

Rough set theory (RST), one approach of FS, benefits the possibility of maintaining the underlying semantics of reduced features. By collecting a minimal subset

---

2020 *Mathematics Subject Classification.* 68W01, 68T20, 68T35, 68U35.
Key words and phrases. Attribute reduction, fuzzy rough feature selection, data mining, fuzzy FP-growth algorithm, fuzzy associative classification rule.
*Corresponding author. This work was supported in part by the National Science Council of the Republic of China.

from the original features, RST analyzes data arbitrarily that human intervention is optional [41, 43, 10, 52]. It is this very reason that RST has been exploited in application domains such as classification, clustering, to name a few [13, 29, 20]. Enhanced in hybridization with fuzzy logic, RST is later perceived as fuzzy rough feature selection (FRFS) method. The FRFS better represents continuous data, a mixture of qualitative and quantitative data that we face in the mechanism of cognitive process. It is thus believed to be a promising alternative in the classification mining [32, 53].

Apriori algorithm, used in data mining, earns a reputation of obstructing association discovery and classification mining [17, 48]. Han et al. [19] came up with an associative mining method, called Frequent-Pattern growth (FP-growth), to tune this situation. Fuzzy FP growth (FFP-growth) method deals with cognitive uncertainty of blurriness and obscurity. In discerning users' perceived interpretation and subjective inception of ideas, FFP is brought into play and spotted in references [37, 38, 21, 49, 50, 22, 34, 35, 12]. Li et al. furthermore suggested classification on multiple association rules (CMAR) to deal with classification problems [36]. As we move along, FFP-growth does reach certain accuracy in the association-based classification yet with foretold weakness [11, 39, 7].

On one hand, during constructing the FP-tree structure, some methods [36, 7] simply attach class labels to the last nodes of the paths or within the nodes. If these nodes have many class labels, what pattern candidates a certain class label has is unsure of. On the other hand, it is difficult for computation time to bring about combinatorial explosive number of pattern candidates when the given support threshold is slight. While these pattern candidates become the antecedences of rules, it is massive for human users to absorb such a large number of rules, which concern conditions in information storage, retrieval, purging and connection with complicated fuzzy rules. Evaluating the options, we adjust the classified FFP-tree (CFFP-tree) structure and yield the suitable number of linguistic terms and combination quantities of pattern candidates.

In this paper, we present a new procedure considering both FRFS and classified FFP-growth algorithms for classification mining [33]. First, the FRFS algorithm can filter essential patterns (attributes) to reduce the attribute quantities of the original data set. We modified the QUICKREDUCT heuristic within the FRFS algorithm for efficiency. Second, a new classified FFP-growth (CFFP-growth) algorithm was posited to search for strong associations between frequent patterns and class labels. By adjusting the CFFP-tree structure, it is effortless to see the pattern candidates for each class from the top of the tree to the bottom and the paths belonged to different class labels from the bottom of the tree to the top. Then, we suggest different number of linguistic terms and pattern candidates of data sets by experimental results to improve on the performance and prevent too many meaningless rules before rule pruning. The details of our procedure are described in [33].

In order to access the performance of the proposed algorithms, 10 real-world data sets with the ranges of variables from 6 to 60 and the ranges of objects from 101 to 10992 are used for the following experiments. First, the experimental results are the parameters of the CFFP-growth algorithm. Second, we have shown the results that the reduction proportions of attributes in each data set. With regard to the CFFP-growth algorithm, we have compared the reduced data sets by the FRFS algorithm with the original ones to prove that the data sets only with relevant attributes have finer performance. In addition, from the comparison with an Apriori-based method, we have shown the results that the efficiency of the CFFP-growth algorithm is clearly superior to the Apriori algorithm. Finally, we have compared the performance of our study with that of another method.

2. Preliminaries

For a better understanding, this section introduces some necessary background information, including the fuzzy rough sets for attribute reduction and the fuzzy association rules for classification.

The rough set theory (RST) proposed by Pawlark [41, 42] serves as a mathematical approach for discovering relationships in data. The idea assumes that knowledge is based on the ability of classifying objects. Thus, in this approach, knowledge is necessarily connected with the variety of classification patterns related with specific parts of the real or abstract world. The benefits of RST is that it can analyze directly from data itself without additional artificial interventions such as thresholds or expert knowledge.

As most data sets contain quantitative attributes, it is natural to perform the standard fuzzification techniques. Fuzzification not only enables linguistic terms to be associated with attribute values but allows the membership degrees of attribute values to more than one fuzzy grid. To deal with linguistic variables, the grid partition method [26, 27, 25, 28] can divide quantitative variables into fuzzy sets with membership functions. Common membership functions are triangular or trapezoid membership functions. In this paper, symmetric triangle-shaped linguistic variables and the grid partition method are used for simplicity. We assume that a quantitative variable a is partitioned into K fuzzy sets $\{A_{a,l_1}, A_{a,l_2}, \ldots, A_{a,l_K}\}$, A_{a,l_i} is the ith fuzzy grid (linguistic term) and is defined by the triangular membership function [23]

$$\mu_{a,l_i}(t) = \max\{1 - |t - x_i^K|/y^K, 0\}, \tag{2.1}$$

where

$$\begin{aligned} x_i^K &= m_i + (m_a - m_i)(i-1)/(K-1), \qquad i = 1, 2, \ldots, K, \\ y^K &= (m_a - m_i)/(K-1), \end{aligned}$$

and m_a and m_i are the maximum and minimum values of the domain interval of A_{a,l_i}. x_i^K is the top where the membership degree is equal to 1 and y^K is the spread of the membership function of A_{a,l_i}.

In the data set, we utilize grid partition method to transform quantitative attributes into fuzzy grids. A quantitative attribute a is represented as:

$$(\frac{\mu_{a,l_1}(t_p)}{A_{a,l_1}} + \frac{\mu_{a,l_2}(t_p)}{A_{a,l_2}} + \cdots + \frac{\mu_{a,l_K}(t_p)}{A_{a,l_K}}) \tag{2.2}$$

using the triangular membership function defined by (2.1), where A_{a,l_i} is the ith fuzzy grid of K linguistic terms defined in the attribute a of pth transaction data t.

However, the complexity becomes excessively high while calculating the Cartesian product of fuzzy tolerance classes for large attribute data sets. And it is undesirable from a theoretical viewpoint while the fuzzy lower approximation might not be a subset of the fuzzy upper approximation in some situations. Instead of using the fuzzy partition method to determine fuzzy tolerance classes, another technique, proposed by Jensen and Shen [32], applies the fuzzy similarity relation as the fuzzy tolerance relation. The concepts would be illustrated in terms of the following notions [32].

A fuzzy similarity relation R is used to measure the appropriate equality of uncertain objects, instead of using a crisp equivalence relation to represent objects' indiscernibility. The definition [45, 14] assume that R is at least a fuzzy tolerance relation, which is reflexive and symmetric. For a nominal attribute a, we use the classical way of discerning objects, that is, $R_a(x, y) = 1$ if $a(x) = a(y)$. Or we can define, for any subset B of $\mathbb{A}$, the fuzzy B-indiscernibility relation by

$$R_B(x, y) = \mathcal{T}\{R_a(x, y)\}, \qquad {}^\forall a \in B \tag{2.3}$$

where $\mathcal{T}$ is a T-norm and $R_a(x, y)$ is the degree of similarity between id x and y for attribute a. If qualitative attributes are discrete only in the process, then we can recover the traditional concept of B-indiscernibility relation.

For the lower and upper approximations, a fuzzy tolerance relation R is used to approximate a fuzzy set A in X

$$\underline{R}A(y) = \inf_{x\in X} \mathfrak{I}(R(x, y), A(x)), \qquad {}^\forall y \in X \tag{2.4}$$

$$\overline{R}A(y) = \sup_{x\in X} \mathfrak{T}(R(x, y), A(x)), \qquad {}^\forall y \in X \tag{2.5}$$

where $\mathfrak{I}$ is a fuzzy implicator and $\mathfrak{T}$ is a t-norm.

Using fuzzy B-indiscernibility relations, the fuzzy B-positive region are defined by, for y in X,

$$POS_B(y) = \left(\bigcup_{x\in X} \underline{R_B}[x]_{R_d}\right)(y). \tag{2.6}$$

The resultant degree of dependency is calculated as

$$\gamma'_B = \frac{|POS_B|}{|X|}. \tag{2.7}$$

If $\gamma'_B = 1$, the subset B can be a reduct that preserves the same dependency degree as the entire data set.

Association mining is one of the most common data mining techniques that is used to explore, analyze knowledge, and discover meaningful patterns [8, 1, 2, 3]. In recent years, numerous studies were carried out on the integration of fuzzy set concept with the Apriori algorithm to find association rules in the data sets. In spite of its popularity, the Apriori algorithm is used to generate a huge number of candidate sets, repeatedly scan a database, and check a large set of candidates by pattern matching. On the other hand, FP-growth introduced by Han *et al.*[19] adopts a divide-and-conquer strategy to compress the database representing frequent items into a FP-tree and retain the itemset association information. Then it divides such a compressed database into a set of conditional database; each associated with one frequent item to mine the rules separately [18].

Most studies on the performance of the FP-growth method have shown that it is efficient and scalable for mining both long and short frequent patterns, and is about an order of magnitude faster than that of the Apriori algorithm. But there are still improvements to break through for the classification models. For the FP-tree structure, the class labels are attached to the last nodes of the paths or within the nodes. If these nodes have many class labels, it is not so intuitive to understand what pattern candidates a certain class label has. In addition, generating combinatorial explosive number of pattern candidates will be a challenge for computational efficiency when the given support threshold is small. A classifier thus may have to handle too many unnecessary rules through these pattern candidates, which concerns conditions in information storage, retrieval, and purging. Therefore, our new CFFP-tree structure puts all class labels as the second level nodes under the root nodes. The related fuzzy grids of tuples can be attached as nodes below their class labels. Besides, our new CFFP-growth method can choose the number of pattern candidates according to the experimental results of data sets to decrease the computational time (see Section 3.2).

3. FRFS and CFFP-growth algorithms

In this section, we describe a procedure combined with FRFS and CFFP-growth algorithms for generating fuzzy associative classification rules. First, a given data set adopts the FRFS algorithm to reduce attribute quantities. Then, the reduced data set will be utilized to a CFFP-growth algorithm with the GA process for the optimization of accuracy during classification. An example of the data set is given in Table 1 to illustrate the concepts of the proposed procedure. The data set contains six attributes from a_1 to a_6, two classes C_1 and C_2, and eight objects.

3.1. **Fuzzy-rough Attribute Reduction.** The FRFS algorithm, employs fuzzy similarity relations to construct approximations, allows several useful ways for attribute reduction, e.g., fuzzy lower approximation, fuzzy boundary region, and fuzzy discernibility matrix. Considering reduced size, runtime, and resulting classification accuracies in [32], we utilize the fuzzy lower approximation-based FS to compute degrees of data dependency for the minimal reducts.

Table 1. The data set

ID	a_1	a_2	a_3	a_4	a_5	a_6	Class
1	2	4	1	0	9	0	C_2
2	0	3	5	0	0	4	C_2
3	4	0	3	2	8	0	C_1
4	2	2	7	0	0	10	C_1
5	1	2	0	4	3	5	C_1
6	4	0	6	0	0	2	C_1
7	0	2	0	4	3	0	C_2
8	1	3	3	0	0	7	C_2

The first step is to compute fuzzy similarity relations among different values of a same attribute. This measure can be modeled as a fuzzy tolerance relation R [45, 14]. For quantitative values, a suitable measure is defined as

$$R_a(x,y) = \max\left\{\min\left\{\frac{(a(y)-(a(x)-\sigma_a))}{\sigma_a}, \frac{(a(x)+\sigma_a)-a(y)}{\sigma_a}\right\}, 0\right\} \tag{3.1}$$

where $R_a(x,y)$ is the degree of similarity between id x and y for attribute a, and σ_a is the standard deviation of attribute a.

Instead of the classical way of the equality metric for nominal attributes, the Value Difference Metric (VDM) is proposed to measure the closeness of two values if they have more similar classifications [46, 51, 15].

In distinguishing a subset of attributes, (2.3) gives a proper measure to evaluate. Fuzzy lower and upper approximations employ the fuzzy B-indiscernibility relation defined as (2.4) and (2.5), respectively. The fuzzy connectives chosen for this paper are the minimum t-norm $\mathfrak{T}(x,y) = min\{x,y\}$, for all $x,y \in [0,1]$, and the Lukasiewicz fuzzy implicator $\mathfrak{I}(x,y) = min\{1, 1-x+y\}$, for all $x,y \in [0,1]$.

Although (2.6) provides the most faithful way in defining the fuzzy positive region, the computational complexity is high. For easier computation, the definition of the positive region in [14] is replaced by

$$POS_B(y) = (\underline{R_B}R_d y)(y), \tag{3.2}$$

which results in smaller positive regions. Then, an increasing $[0,1]$-valued measure can be defined to implement a corresponding notion of fuzzy decision reducts. Also, a normalized extension of the resulting degree of dependency in [14] is calculated as

$$\gamma'_B = \frac{|POS_B|}{|POS_{\mathbb{A}}|}. \tag{3.3}$$

In the example of the data set in Table 1, the degrees of dependency of attributes in Table 1 are as the follows by (3.3).

$$\gamma'_{a_1} = 0.41, \qquad \gamma'_{a_2} = 0.55,$$
$$\gamma'_{a_3} = 0.23, \qquad \gamma'_{a_4} = 0.13,$$
$$\gamma'_{a_5} = 0.07, \qquad \gamma'_{a_6} = 0.30.$$

A fuzzy-rough attribute reduction can be seen that a subset of attributes with the maximum dependency degree have the same classified ability as the entire data set. Different from the original search algorithm proposed by Chouchoulas and Shen [13], a modified QUICKREDUCT heuristic [30] for a quick search of the dependency degrees of attributes in the descending order is presented in Fig. 1.

Input: C, a set of the attributes with the dependency degrees in the descending order; $\mathcal{D}$, a set of the classes.

Output: R, a subset of attributes.

$R \leftarrow$ the first attribute of C;

γ'_{best} = the dependency degree of R;

$\gamma'_{prev} = 0$;

foreach $x \in C - R$

  if $\gamma'_{R\cup\{x\}}(\mathcal{D}) > \gamma'_{best}(\mathcal{D})$

    $R \leftarrow R \cup \{x\}$;

    $\gamma'_{best} = \gamma'_R(\mathcal{D})$;

return R

FIG. 1. The modified fuzzy-rough QUICKREDUCT heuristic.

Following the example in Table 1, attribute a_2 is selected to be the reductive candidate because it has the largest degree of dependency, The algorithm shows that the remaining attributes are listed in the descending order by dependency degree and combined to the reductive candidate using (2.3). The following dependency degrees of combined attributes are

$$\gamma'_{\{a_2,a_1\}} = 0.80, \qquad \gamma'_{\{a_2,a_1,a_6\}} = 0.87,$$
$$\gamma'_{\{a_2,a_1,a_6,a_3\}} = 1.$$

Since the attribute $\{a_1, a_2, a_3, a_6\}$ has the maximum dependency degree, these attributes are considered to own the same ability to classify as the original one and the algorithm terminates. The data set can now be reduced to these attributes as Table 2.

Next, the reduced data set is to illustrate the proposed CFFP algorithm.

3.2. **Classified Fuzzy FP-growth Algorithm.** The purpose of CFFP-growth algorithm is to better the performance and reach efficiency by adjusting the CFFP-tree structure and restricting the combination number of patterns. The following describes the details of the algorithm.

TABLE 2. The reduced data set

ID	a_1	a_2	a_3	a_6	Class
1	2	4	1	0	C_2
2	0	3	5	4	C_2
3	4	0	3	0	C_1
4	2	2	7	10	C_1
5	1	2	0	5	C_1
6	4	0	6	2	C_1
7	0	2	0	0	C_2
8	1	3	3	7	C_2

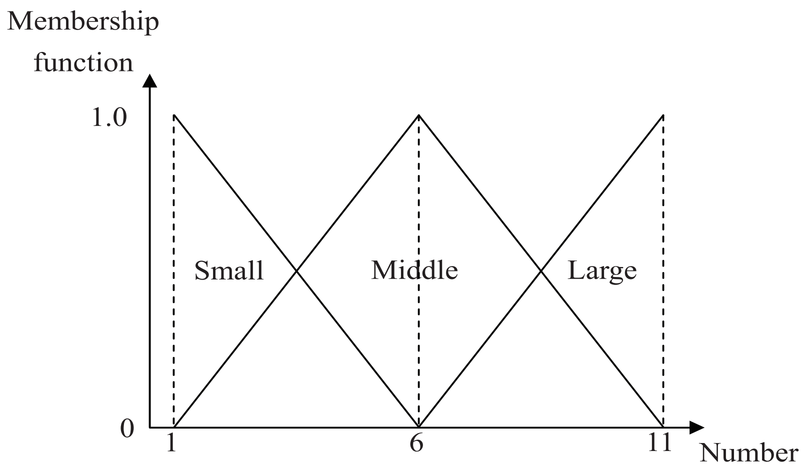

FIG. 2. The membership functions.

In the previous subsection, the reduced attributes have the same classification ability as the original ones after the reduction procedure. Then, we may ascertain the relationship between the concepts of reduced attributes and class labels to generate fuzzy associative classification rules. To lower computational complexity, we use grid partition method to transform quantitative values into fuzzy grids. For example, the grid partition method as (2.1) and the given membership functions in Fig. 2 would be used to transform values to fuzzy grids (as (2.2)). Then, the data set table in Table 2 can be replaced as a fuzzy grid table.

Then, we compute fuzzy supports of fuzzy grids to verify frequent fuzzy grids. Let A_{a,l_i} be a fuzzy grid for attribute a in the training data. A frequent fuzzy grid is a fuzzy grid if its fuzzy support is higher than or equal to the minimum support. If the fuzzy support $FS(A_{a,l_i})$ satisfies this condition, fuzzy grid A_{a,l_i} can be viewed as a frequent fuzzy grid to retain to the next process. By following the example, if the value of the minimum support is 0.31, the fuzzy grids that meet the minimum support are $a_1.Small$, $a_2.Small$, $a_3.Middle$, and $a_6.Middle$. These fuzzy grids can be regarded as the frequent fuzzy grids. Based on these frequent fuzzy grids, we rebuild a new fuzzy grid table as Table 3.

TABLE 3. The new fuzzy grid table

ID	Fuzzy Grid	Class
1	$\frac{0.80}{a_1.Small}, \frac{0.40}{a_2.Small}$	C_2
2	$\frac{0.60}{a_2.Small}, \frac{0.80}{a_3.Middle}, \frac{0.60}{a_6.Middle}$	C_2
3	$\frac{0.40}{a_1.Small}, \frac{0.40}{a_3.Middle}$	C_1
4	$\frac{0.80}{a_1.Small}, \frac{0.80}{a_2.Small}, \frac{0.80}{a_3.Middle}, \frac{0.20}{a_6.Middle}$	C_1
5	$\frac{1.00}{a_1.Small}, \frac{0.80}{a_2.Small}, \frac{0.80}{a_6.Middle}$	C_1
6	$\frac{0.40}{a_1.Small}, \frac{1.00}{a_3.Middle}, \frac{0.20}{a_6.Middle}$	C_1
7	$\frac{0.80}{a_2.Small}$	C_2
8	$\frac{1.00}{a_1.Small}, \frac{0.60}{a_2.Small}, \frac{0.40}{a_3.Middle}, \frac{0.80}{a_6.Middle}$	C_2

Next, we begin to proceed with the CFFP-growth method. First, we set a node ROOT of a CFFP-tree and put all the class labels as the second level nodes. Second, we scan the new fuzzy grids table to get the nodes of fuzzy grids and put them under the nodes of belonged class labels. Then, we link the nodes of fuzzy grids with one another based on the same class label and the same tuple. In this example, let us construct a CFFP-tree through the new fuzzy grid table in Table 3. We first set a node ROOT and link the class nodes C_1 and C_2 in the second level separately. Then we scan the tuples from the first one in Table 3. The first tuple has two fuzzy grids, $a_1.Small$, $a_2.Small$, and the belonged class "C_2". So, these two fuzzy grids can be the nodes of the CFFP-tree under the class node "C_2". Linking these nodes and the class node becomes a path of the first tuple. The other tuples repeat the step so the CFFP-tree are shown in Fig 3.

The difference of the CFFP-tree from other tree structures is that from top to bottom; it is easy to see the pattern candidates for each class. And from bottom to top, it explores the paths that belongs to different class labels. Then, we can scan the CFFP-tree to obtain the potential frequent patterns.

Based on the CFFP-tree, we construct a table with paths, a group of combinations of fuzzy grids as pattern candidates, and class labels belonged. And then we generate rules with pattern candidates as the antecedences of rules and class labels as the consequences. In terms of every path of CFFP-tree in Fig 3, we generate pattern candidates and class labels belonged listed in Table 4. Take the path $\{C_1 : a_1.Small, a_3.Middle, a_6.Middle\}$ as an example. $\{a_1.Small, a_3.Middle, a_6.Middle\}$ belonged to the class node C_1 means that we can use these fuzzy grids to generate pattern candidates of class C_1.

One can find it hard for users to grasp too many generated fuzzy rules and their relationship with antecedents. For this reason, we only generate the proper number of pattern candidates as antecedents of fuzzy rules according to the experiment results of accuracy. If the number of pattern combinations is 3, the pattern candidates

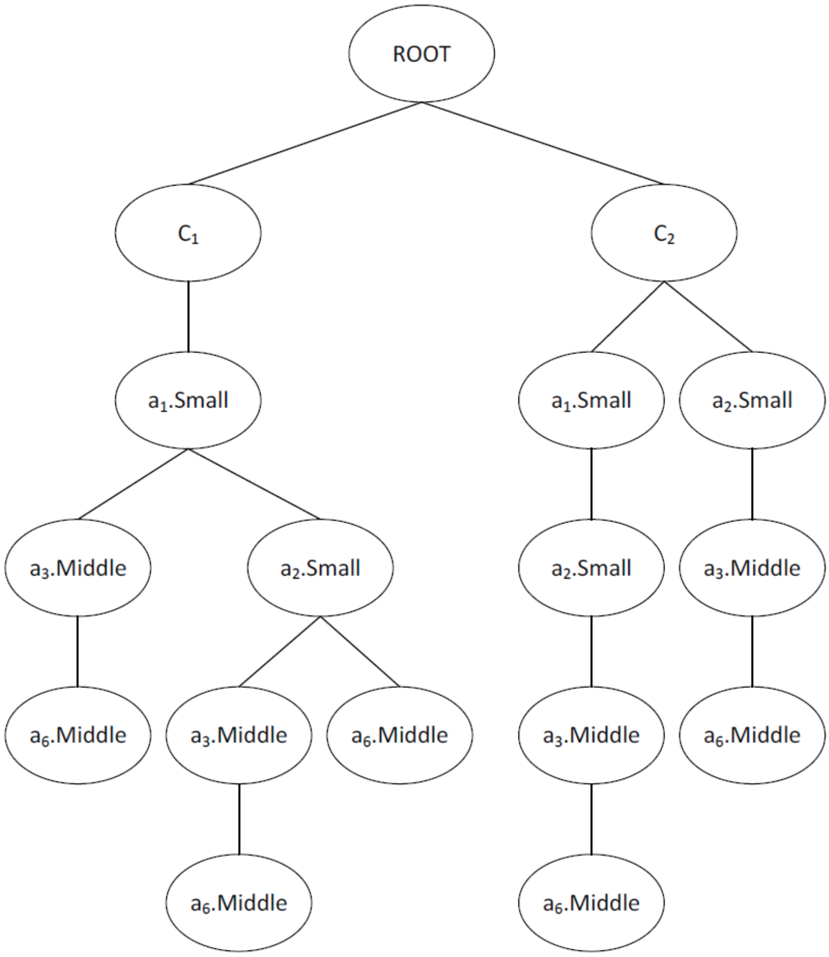

FIG. 3. The example of CFFP-tree.

of class C_1 are
$\{a_1.Small\}, \{a_3.Middle\}, \{a_6.Middle\}$, $\{a_1.Small, a_3.Middle\}$, $\{a_3.Middle, a_6.Middle\}, \{a_1.Small, a_6.Middle\}$, and $\{a_1.Small, a_3.Middle, a_6.Middle\}$.

Similar to fuzzy grids, the pattern candidates and class labels are used to calculate fuzzy supports. The fuzzy value $N_s^{C_j}(t_p)$ in pattern $A_s^{C_j} : (A_{s_1}^{C_j}, A_{s_2}^{C_j}, \ldots, A_{s_{r+1}}^{C_j})$ of class C_j is defined as

$$N_s^{C_j}(t_p) = \min\left\{\mu_{s_1}^{C_j}(t_p), \mu_{s_2}^{C_j}(t_p), \cdots, \mu_{s_{r+1}}^{C_j}(t_p)\right\}$$

TABLE 4. Pattern candidates

Path	Pattern Candidate	Class
$\{C_1 : a_1.Small, a_3.Middle, a_6.Middle\}$	$\{a_1.Small\}; \{a_3.Middle\}; \{a_6.Middle\};$ $\{a_1.Small, a_3.Middle\}; \{a_3.Middle, a_6.Middle\};$ $\{a_1.Small, a_6.Middle\};$ $\{a_1.Small, a_3.Middle, a_6.Middle\}.$	C_1
$\{C_1 : a_1.Small, a_2.Small, a_3.Middle, a_6.Middle\}$	$\{a_2.Small\}; \{a_1.Small, a_2.Small\};$ $\{a_2.Small, a_3.Middle\}; \{a_2.Small, a_6.Middle\};$ $\{a_1.Small, a_2.Small, a_3.Middle\};$ $\{a_1.Small, a_2.Small, a_6.Middle\};$ $\{a_2.Small, a_3.Middle, a_6.Middle\}.$	C_1
$\{C_1 : a_1.Small, a_2.Small, a_6.Middle\}$	*	C_1
$\{C_2 : a_1.Small, a_2.Small, a_3.Middle, a_6.Middle\}$	$\{a_1.Small\}; \{a_2.Small\}; \{a_3.Middle\}; \{a_6.Middle\};$ $\{a_1.Small, a_2.Small\}; \{a_1.Small, a_3.Middle\};$ $\{a_1.Small, a_6.Middle\}; \{a_2.Small, a_3.Middle\};$ $\{a_2.Small, a_6.Middle\}; \{a_3.Middle, a_6.Middle\};$ $\{a_1.Small, a_2.Small, a_3.Middle\};$ $\{a_1.Small, a_2.Small, a_6.Middle\};$ $\{a_1.Small, a_3.Middle, a_6.Middle\};$ $\{a_2.Small, a_3.Middle, a_6.Middle\}.$	C_2
$\{C_2 : a_2.Small, a_3.Middle, a_6.Middle\}$	*	C_2

where $A_{s_k}^{C_j}$ is the kth fuzzy grid in pattern $A_s^{C_j}$ of class C_j. The t-norm operator in the fuzzy intersection is the minimum operator. Then, the fuzzy support $FS(A_s^{C_j})$ is calculated as $FS(A_s^{C_j}) = \left[\sum_{t_p \in C_j} N_s^{C_j}(t_p) \right] /n$.

In considering the number of patterns for each class that can be different, we redefine the minimum support for class C_j [4] as

$$MinimumSupport_{C_j} = minSup * f_{C_j} \tag{3.4}$$

where minSup is minimum support, and f_{C_j} is the pattern ratio of the class C_j. The pattern s is a frequent pattern of the class C_j if $FS(A_s^{C_j})$ is larger than or equal to $MinimumSupport_{C_j}$.

Once all frequent patterns have been obtained, the proper fuzzy associative classification rule $R_s : A_s \to C_j$ can be generated by the frequent pattern $A_s^{C_j}$, setting the frequent pattern A_s in the antecedence of the rule and corresponding class C_j in the consequence.

The fuzzy confidence value assists to ensure that the rule is the effective fuzzy associative classification rule, calculated. If $FC(A_s \to C_j)$ is larger than or equal to the minimum confidence, the rule $R_s : A_s^{C_j} \to C_j$ is an effective fuzzy associative classification rule.

To improve the classification abilities of fuzzy rule-based systems and eliminate redundant rules, we adopt the method proposed by Nozaki et. al. [40] through incorporating the adaptive rules. To solve the shortcoming of user-specified minimum support and confidence, we introduce the learning process proposed by Hu et al. [24] incorporates genetic algorithm (GA) [16] to automatically determine these parameters for specific classification data set.

3.3. **Flowchart.** In the following, we propose a procedure of the proposed algorithms to form fuzzy associative classification rules.

INPUT: A data set with class labels.

OUTPUT: Fuzzy associative classification rules.

Phase 1. (Section 3.1)Fuzzy-rough Attribute Reduction.

Step 1. Calculate the fuzzy similarity degrees, the lower approximations, positive regions for every tuple of each attribute.

Step 2. Compute the resulting degrees of dependency for each attribute a.

Step 3. Arrange the attributes in the descending order by the degrees of dependency.

Step 4. Add the first attribute to the reduct candidate, and then sequentially evaluate the degree of dependency of this candidate with the addition of remaining attributes.

Step 5. If the reduct candidates produce the maximum dependency value for this data set, reduce the data set to these attributes only. Else, go to Step 4.

Phase 2. (Section 3.2)Classified Fuzzy FP-growth Algorithm

Step 6. Utilize the grid partition method to transform quantitative attributes into fuzzy grids.

Step 7. Scan the dataset and construct a table FGTTFS.

Step 8. Generate the initial population with P chromosomes.

Step 9. Select the frequent fuzzy grids to construct a CFFP-tree.

Step 10. Scan the CFFP-tree to generate pattern candidates.

Step 11. Select the frequent fuzzy patterns and generate fuzzy associative classification rules.

Step 12. Reduce redundant rules and employ adaptive rules to adjust fuzzy confidences.

Step 13. Evaluate the population.

Step 14. Generate the next population through the process of GA, selection, crossover, mutation, and elitist strategy.

Step 15. If the maximum generation is not reached, go to Step 9.

A structure is shown in Fig. 4.

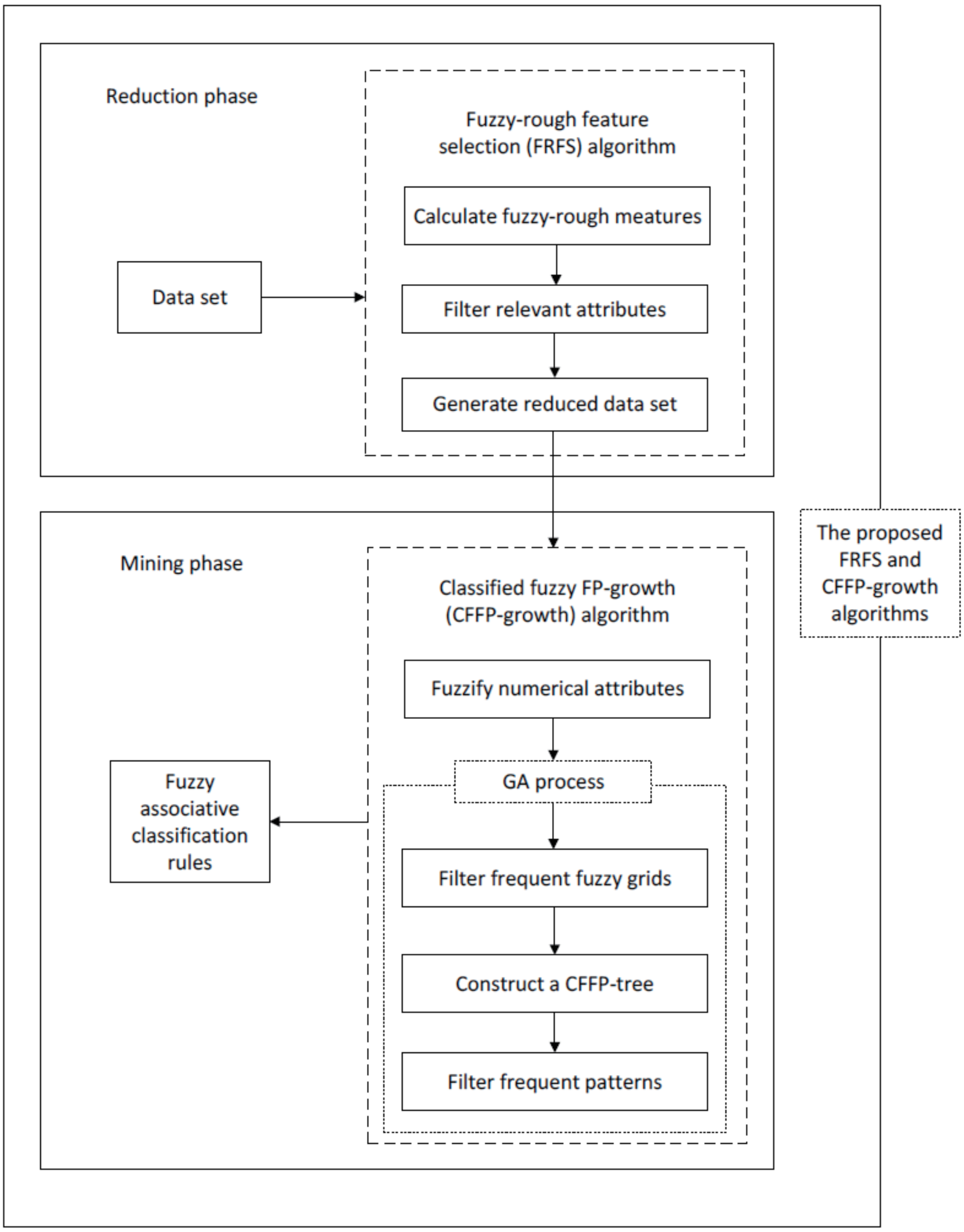

FIG. 4. The structure of the proposed algorithms.

4. Experimental framework and results

This section describes the classification data sets that are used in these experiments and presents several experiments to evaluate the utility of our proposal. First, an experiment shows the suitable parameters of each data set for the proposed CFFP-growth algorithm. Then, we introduce experiments to verify the reduction ability and reduction essentiality of the proposed FRFS algorithm for each data set. Finally, the experiments summarize the other two methods selected for comparison. The details of the experiments are described in [33].

To analyze the performance of the proposed approach, Table 5 summarizes the main properties of the 10 classification data sets from the Knowledge Extraction based on Evolutionary Learning(KEEL)-dataset repository [6, 5]. For each data set, "Attributes(R/I/N)" is the number of (Real/Integer/Nominal) attributes in data, "Examples" is the number of examples, and "Classes" is the number of classes. For continuous attributes, their values are fuzzified in the interval $[0, 1]$ by grid partition method (2.1) to equalize the influence of attributes with different range domains.

Table 5. Data sets for the experimental study

Name	Attributes(R/I/N)	Examples	Classes
Australian	14 (3/5/6)	690	2
Monks	6 (0/6/0)	432	2
Page-blocks	10 (4/6/0)	5472	5
Penbased	16 (0/16/0)	10992	10
Sonar	60 (60/0/0)	208	2
Vowel	13 (10/3/0)	990	11
Wdbc	30 (30/0/0)	569	2
Wine	13 (13/0/0)	178	3
Yeast	8 (8/0/0)	1484	10
Zoo	16 (0/0/16)	101	7

We use the ten fold cross-validation (10-fcv) procedure [47] to partition the data sets. Each data set is randomly partitioned into ten subsets, preserving the same number of examples and the same class distribution between partitions. In an iterative process, one partition is selected as the test set whereas the remaining is the training set. By averaging the results gained from the ten partitions, the final results are thus obtained.

4.1. **The experiment for parameters.**
In the parameters of the proposed CFFP-growth algorithm, the number of linguistic terms affects not only the range of fuzzy partition in quantitative variables but the formation of fuzzy rules. Also, a suitable number of combinations of fuzzy grids equips human users with the ability to handle pattern candidates to be appropriate amounts of rules. So this experiment is to explore these two parameters, records the average accuracy rates of the different number of linguistic terms within the three combinations of fuzzy grids, and the

different number of combinations based on the experiment results of the number of linguistic terms. "Tra" and "Tst" in Table 6 indicate the training data and testing data of data sets, respectively.

The results show that every data set has various average accuracy rates in the numbers of linguistic terms and combinations. In some cases, the highest accuracy rates fall in the different ranges of linguistic terms and combinations. So, the ranges are chosen according to the highest accuracy rates in the testing data of the data set. And we adopt these results as the parameters of the CFFP-growth algorithm for the later experiments. The best results for each one are highlighted in bold in each case.

The other given parameters of the FRFS and CFFP-growth algorithms are illustrated in Table 7.

4.2. **The experiment for attribute reduction.** The purpose of the attribute reduction is to reduce the representation of data and keep relevant attributes. The idea of the rough sets is introduced to prevent additional artificial interventions. So, the proposed FRFS algorithm combines these two concepts to make the operation of data mining simply. This experiment is to verify that the FRFS algorithm has the ability to reduce the representation of the given data sets.

For the later experiments, the minimum $t-$norm and the Lukasiewicz fuzzy implicator are used, with fuzzy similarity relation defined in (3.1).

Table 8 lists the average reduction size, the original size, and reduction rate for attributes in each data set. In considering the removed scales of attributes, the reduction rate is computed by

$$ReductionRate = \frac{|RemovedAttributes|}{|OriginAttributes|}. \tag{4.1}$$

4.3. **Experiments of reduction essentiality.** In data mining, most of studies only consider various methods for analyzing the contents of data. But there are two issues: one is whether original data sets containing redundant attributes would disarrange the rules and make the accuracy decrease, the other is whether the reduced data sets have impacts on the decrease of execution time. So, we want to explore the essentiality of attribute reduction by the proposed FRFS algorithm before the data mining process.

This experiment uses the proposed CFFP-growth algorithm with two kinds of data sets beforehand: the reduced ones treated by the FRFS algorithm (FRFS+ CFFP-growth) and the original ones (CFFP-growth) to compare the classification capability by the average accuracy rates and execution rates of two data sets. Table 9 shows the average accuracy rates in training and testing data for each method and data set. The execution rate is the reducing proportion of the execution time by the FRFS and CFFP-growth algorithms compared with the CFFP-growth algorithm. And the execution rate is calculated by

$$ExecutionRate = \frac{|\text{CFFP-growth}| - |\text{FRFS+CFFP-growth}|}{|\text{CFFP-growth}|} \tag{4.2}$$

TABLE 6. Average accuracy rates considering different number of linguistic terms and combinations

Data set	3 linguistic terms		4 linguistic terms		5 linguistic terms		3 combinations		4 combinations		5 combinations	
	Tra	Tst	Tra	Tst	Tra	Tst	Tra	Tst	Tra	Tst	Tra	Tst
Australian	86.39	85.07	**86.49**	**85.65**	86.34	85.65	**86.49**	**85.65**	85.93	84.78	86.49	85.51
Monks	97.22	97.27	**100.00**	**100.00**	100.00	100.00	99.72	99.77	**100.00**	**100.00**	100.00	100.00
Page-blocks	90.36	90.30	**92.51**	**92.31**	90.39	90.35	**92.51**	**92.31**	92.51	92.31	91.53	91.43
Penbased	64.81	64.53	**72.03**	**72.07**	18.66	18.60	**72.03**	**72.07**	67.58	67.00	70.08	69.91
Sonar	83.97	71.56	**87.51**	**74.22**	79.07	72.88	87.51	**74.22**	**89.85**	69.67	89.43	72.07
Vowel	**33.07**	**32.51**	0.00	0.00	0.00	0.00	32.15	32.23	**33.07**	**32.51**	0.00	0.00
Wdbc	96.20	93.76	96.52	**95.05**	**96.84**	93.60	96.52	**95.05**	96.41	94.09	**96.57**	94.25
Wine	**100.00**	**97.42**	100.00	93.88	99.77	93.35	**100.00**	**97.42**	100.00	97.42	99.94	95.90
Yeast	46.35	44.95	54.17	**52.83**	**54.23**	52.36	54.17	**52.83**	54.19	52.70	**54.41**	52.02
Zoo	95.17	92.17	94.94	92.56	**98.53**	**92.83**	**98.53**	**92.83**	98.53	92.83	95.17	92.17

Table 7. Parameter specification for the proposed algorithms

Algorithm	Parameters
FRFS	Gamma=1.0.
CFFP-growth	Number of generations=50; Population size=30; Length of support and confidence=10; Weight of the classification accuracy rate (WCAR)=10.0; Weight of the number of fuzzy rules (WV)=1.0; Crossover Probability=1.0; Mutation Probability (per gen)=0.01; Learning rate $\eta 1$=0.001; Learning rate $\eta 2$=0.1; Number of iterations Jmax=100.

Table 8. Results for reduction in data sets

Data set	After reduction / Origin (size)	Reduction rate (%)
Australian	11 / 14	21.43
Monks	5 / 6	16.67
Page-blocks	9 / 10	10.00
Penbased	14 / 16	12.50
Sonar	17 / 60	72.67
Vowel	11 / 13	15.38
Wdbc	21 / 30	30.00
Wine	6 / 13	53.85
Yeast	6 / 8	25.00
Zoo	15 / 16	6.25

where $|algorithm|$ is the execution time of the algorithm. The best accuracy results for each one are highlighted in bold in each data set.

The results in Table 9 reveal that most of the data sets have higher accuracy in FRFS and CFFP-growth algorithms. Although the reduced criterion of the FRFS algorithm may remove those attributes which have supports for the classification of data, the difference of the accuracy is not too far. In addition, the execution rates show that the data sets would have higher performance after the attribute reduction. So, the data sets with irrelevant attributes have larger chances to create surplus rules and decrease the accuracy. The reduced data sets treated by the FRFS algorithm not only increase the computation efficiency but retain essential attributes that can generate higher relevant rules by the CFFP-growth algorithm to better the ability for classification.

The experiment reported in this paper has demonstrated that the FRFS algorithm can be practically implemented and provide adequate results.

TABLE 9. Results for reduction essentiality

Data set	FRFS + CFFP-growth		CFFP-growth		Execution
	Tra	Tst	Tra	Tst	Rate (%)
Australian	**86.49**	**85.65**	86.31	85.51	37.89
Monks	**100.00**	**100.00**	99.55	98.97	35.66
Page-blocks	**92.51**	92.31	92.40	**92.44**	1.57
Penbased	**72.03**	**72.07**	54.56	54.39	93.90
Sonar	87.51	**74.22**	**89.74**	72.71	92.02
Vowel	**33.07**	**32.51**	0.00	0.00	47.06
Wdbc	96.52	**95.05**	**96.59**	94.89	80.46
Wine	**100.00**	**97.42**	99.83	96.94	43.43
Yeast	**54.17**	52.83	53.92	**53.17**	9.91
Zoo	**98.53**	92.83	95.24	**92.89**	72.83

4.4. **Comparisons with other methods.** Our aim is to solve the shortcoming of the Apriori algorithm which takes enormous time begetting high quantities fuzzy grids. For comparison, we introduce two methods based on the Apriori algorithm, which have given membership functions and are available in the KEEL software tool [6]. The following are brief descriptions of these two different methods, the fuzzy rules for classification problems based on the Apriori algorithm (FRCA) method and the classification with fuzzy association rules (CFAR) method. The other given parameters of these methods are illustrated in Table 10.

TABLE 10. Parameter specification for the compared methods

Algorithm	Parameters
FCRA	Number of generations=50; Population size=30; Length of support and confidence=10; Weight of the classification accuracy rate (WCAR)=10.0; Weight of the number of fuzzy rules (WV)=1.0; Crossover Probability=1.0; Mutation Probability (per gen)=0.01; Learning rate $\eta 1$=0.001; Learning rate $\eta 2$=0.1; Number of iterations Jmax=100; Number of linguistic values=5.
CFAR	Number of rules combining for every example (δ)=0.05; Minimum Gain=0.7; Weight decay factor (α)=0.66; Number of rules used in prediction=5.

The FRCA method, proposed by Hu et al. [24], uses the Apriori algorithm to discover fuzzy grids and the evolutionary process to learn the appropriate minimum support and confidence for the fuzzy classification rules. To differentiate the process for searching the antecedents of rules, our algorithm with the reduced attributes adopts the evolutionary process of the FRCA method but replaces the Apriori algorithm with the adjusted CFFP-growth structure.

The CFAR method, proposed by Chen and Chen [9], uses different measures of the minimum support and confidence to generate fuzzy classification rules based on the extend Apriori-type rule mining. However, to be more suitable for each data set, our proposed algorithm has the learning of the minimum support and confidence. Because the CFAR method has no iteration process, this experiment only utilizes the accuracy to examine the performance of our study.

Table 11 shows the accuracy results and computation rates between the FRCA method and our proposed algorithm. The computation rate is similar to (4.2) where the CFFP-growth algorithm is replaced by the compared method. Table 12 shows the accuracy results between the CFAR method and our proposed algorithm. The best accuracy results for each one are highlighted in bold in each data set.

TABLE 11. Results for out study and the FRCA method

Data set	FRFS + CFFP-growth		FRCA		Execution
	Tra	Tst	Tra	Tst	Rate(%)
Australian	86.49	**85.65**	**87.31**	86.65	99.99
Monks	**100.00**	**100.00**	100.00	100.00	99.16
Page-blocks	**92.51**	**92.31**	90.24	90.20	88.41
Penbased	**72.03**	**72.07**	55.78	55.67	94.48
Vowel	**33.07**	**32.51**	0.00	0.00	90.00
Wine	**100.00**	**97.42**	99.72	92.31	99.97
Yeast	54.17	52.83	**55.07**	**53.03**	99.52

TABLE 12. Results for out study and the CFAR method

Data set	FRFS + CFFP-growth		CFAR	
	Tra	Tst	Tra	Tst
Australian	86.49	85.65	**87.54**	**86.73**
Monks	**100.00**	**100.00**	47.22	47.34
Page-blocks	**92.51**	**92.31**	89.78	89.78
Penbased	**72.03**	**72.07**	10.40	10.40
Vowel	**33.07**	**32.51**	9.09	9.09
Wine	**100.00**	**97.42**	98.52	91.80
Yeast	**54.17**	**52.83**	16.44	16.47

Consequently, we can observe that the proposed FRFS and CFFP-growth algorithm achieves the highest accuracies of 6 data sets. The classification ability of our proposed FRFS and CFFP-growth algorithm is more accurate than that of the other two methods in most of data sets. Also, concerning time elapsed, our study takes lower execution time than the time takes on the FRCA method by an average over 90%. The results show that the performance of the CFFP-growth algorithm is better than that of the Apriori algorithm.

However, the comparison results cannot be revealed in other data sets. Except for the limitation of the computer hardware, we observe that too many number of attributes may increase the execution time of these methods even if they have small sizes of data. A possible reason is that while these attributes partition are fuzzy sets, the total amounts of fuzzy sets may enlarge the computing objects for the later process to generate fuzzy rules. So, the quantity of attributes has influences on the computational complexity of the fuzzy classification mining methods. The introduced concepts of attribute reduction can have some assistants.

5. Conclusions

This study concludes a research on the attribute reduction and fuzzy classification mining methods. With the application of FRFS algorithm and the appropriate parameters of the CFFP-growth algorithm, the attributes pertaining to the data sets beget fuzzy associative classification rules. The results unfold that the data sets with redundant attributes and indelicate number of linguistic terms and combinations have the tendency to issue excess rules and increase the likelihood of inaccuracy for classification. One can argue that the compositions of antecedents of rules mark the classification ability. It explains why the FRFS algorithm contributes much in fuzzy classification mining since it detects relevant attributes. Not only does it lessen the computational complexity but it brings forth fuzzy associative classification rules. Our findings in the end reinforce the efficiency of CFFP-growth algorithm and its surpassing quality than the Apriori algorithm.

References

[1] R. Agrawal, T. Imieliński, and A. Swami, *Mining association rules between sets of items in large databases*, ACM SIGMOD Record **22** (1993), 207–216.

[2] R. Agrawal, T. Imieliński, and A. Swami, *Database mining: A performance perspective*, Knowledge and Data Engineering, IEEE Transactions on **5** (1993), 914–925.

[3] R. Agrawal, R. Srikant, et al., *Fast algorithms for mining association rules*, Proc. 20th int. conf. very large data bases, VLDB **1215** (1994), 487–499.

[4] J. Alcalá-Fdez, R. Alcala, and F. Herrera, *A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning*, Fuzzy Systems, IEEE Transactions on **19** (2011), 857–872.

[5] J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S García, L. Sánchez, and F. Herrera, *Keel data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework*, Journal of Multiple-Valued Logic and Soft Computing **17** (2010), 255–287.

[6] J. Alcalá-Fdez, L. Sanchez, S. Garcia, M. J. del Jesus, S. Ventura, J. Garrell, J. Otero, C. Romero, J. Bacardit, V. M. Rivas, et al., *KEEL: a software tool to assess evolutionary algorithms for data mining problems*, Soft Computing **13** (2009), 307–318.

[7] M. Antonelli, P. Ducange, F. Marcelloni, and A. Segatori, *A novel associative classification model based on a fuzzy frequent pattern mining algorithm*, Expert Systems with Applications **42** (2015), 2806–2097.

[8] M. J. Berry and G. Linoff, *Data Mining Techniques: for Marketing, Sales, and Customer Support*, John Wiley & Sons, Inc, 1997.

[9] Z. Chen and G. Chen, *Building an associative classifier based on fuzzy association rules*, Building an associative classifier based on fuzzy association rules **1** (2008), 262–273.

[10] D. Chen, C.Wang, and Q. Hu, *A new approach to attribute reduction of consistent and inconsistent covering decision systems with covering rough sets*, Information Sciences **177** (2007), 3500–3518.

[11] H. Cheng, X. Yan, J. Han, and P. S. Yu, *Direct Discriminative Pattern Mining for Effective Classification*, Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference, 2008, pp. 169–178.

[12] M.-Y. Chou, W.-H. Lee, C.-H. Wang, and C.-T. Pang, *On the fuzzy genetic algorithm*, Journal of Nonlinear and Convex Analysis **17** (2016),921–929.

[13] A. Chouchoulas and Q. Shen, *Rough set-aided keyword reduction for text categorization*, Applied Artificial Intelligence **15** (2001), 843–873.

[14] C. Cornelis, R. Jensen, G. Hurtado, and D. Slezak, *Attribute selection with fuzzy decision reducts*, Information Sciences **180** (2010), 209–224.

[15] J. Derrac, C. Cornelis, S. Garcia, and F. Herrera, *Enhancing evolutionary instance selection algorithms by means of fuzzy rough set based feature selection*, Information Sciences **186** (2012), 73–92.

[16] D. E. Goldberg, *Genetic Algorithms in Search Optimization and Machine Learning*, Addison-wesley, 1989.

[17] J.W. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, San Francisco, 2011.

[18] J. Han and J. Pei, *Mining frequent patterns by pattern-growth: methodology and implications*, ACM SIGKDD explorations newsletter **2** (2000), 14–20.

[19] J. Han, J. Pei, and Y. Yin, *Mining Frequent Patterns Without Candidate Generation*, Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data (2000), pp. 1–12.

[20] T. B. Ho, S. Kawasaki, and N. B. Nguyen, *Documents clustering using tolerance rough set model and its application to information retrieval*, Intelligent exploration of the web, Springer, 2003, pp. 181–196.

[21] T.-P. Hong, T.-C. Lin, and T.-C. Lin, *Mining complete fuzzy frequent itemsets by tree structures*, Systems Man and Cybernetics (SMC), 2010 IEEE International Conference, 2010, pp. 563–567.

[22] T.-P. Hong, C.-W. Lin, and T.-C. Lin, *The MFFP-tree fuzzy mining algorithm to discover complete linguistic frequent itemsets*, Computational Intelligence **30** (2014), 145–166.

[23] Y.-C. Hu, R.-S. Chen, and G.-H. Tzeng, *Mining fuzzy association rules for classification problems*, Computers & Industrial Engineering **43** (2002), 735–750.

[24] Y.-C. Hu, R.-S. Chen, and G.-H. Tzeng, *Finding fuzzy classification rules using data mining techniques*, Pattern Recognition Letters **24** (2003), 509–519.

[25] H. Ishibuchi, T. Nakashima, and T. Murata, *Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems*, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on **29** (1999), 601–618.

[26] H. Ishibuchi, K. Nozaki, and H.Tanaka, *Distributed representation of fuzzy rules and its application to pattern classification*, Fuzzy Sets and Systems **52** (1992), 21–32.

[27] H. Ishibuchi, K. Nozaki, and H.Tanaka, *Selecting fuzzy if-then rules for classification problems using genetic algorithms*, IEEE Transactions on Fuzzy Systems **3** (1995), 260–270.

[28] J.-S. R. Jang and C.-T. Sun, *Neuro-fuzzy modeling and control*, Proceedings of the IEEE **83** (1995), 378–406.

[29] R. Jensen and Q. Shen, *Fuzzy-rough data reduction with ant colony optimization*, Fuzzy sets and systems **149** (2005), 5–20.

[30] R. Jensen and Q. Shen, *Fuzzy-rough sets assisted attribute selection*, Fuzzy Systems, IEEE Transactions on **15** (2007), 73–89.

[31] R. Jensen and Q. Shen, *Computational Intelligence and Feature Selection: Rough and Fuzzy Approaches*, vol. 8, John Wiley & Sons, 2008.

[32] R. Jensen and Q. Shen, *New approaches to fuzzy-rough feature selection*, IEEE Transactions on Fuzzy Systems **17** (2009), 824–838.

[33] W.-H. Lee, *A Study on Fuzzy Classification Mining Applying Fuzzy Rough Attribute Reduction*, doctoral dissertation, Yuan Ze University, Taiwan, 2016.

[34] W.-H. Lee, C.-H. Wang, J.-Y. Tsai, and C.-T. Pang, *Fuzzy weighted FP-growth algorithm with linguistic term thresholds*, Journal of Nonlinear and Convex Analysis **17** (2016),1175–1789.

[35] W.-H. Lee, C.-H. Wang, J.-Y. Tsai, and C.-T. Pang, *Association rules mining combined with FAHP*, International Journal of Computer Science and Engineering **5** (2016), 33–44.

[36] W. Li, J. Han, and J. Pei, *CMAR: Accurate and efficient classification based on multiple class-association rules*, Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference, 2001, pp. 369–376.

[37] C.-W. Lin, T.-P. Hong, and W.-H. Lu, *An efficient tree-based fuzzy data mining approach*, International Journal of Fuzzy Systems **12** (2010), 150–157.

[38] C.-W. Lin, T.-P. Hong, and W.-H. Lu, *Linguistic data mining with fuzzy FP-trees*, Expert Systems with Applications **37** (2010), 4560–4567.

[39] L. T. Nguyen, B. Vo, T.-P. Hong, and H. C. Thanh, *CAR-Miner: An efficient algorithm for mining class-association rules*, Expert Systems with Applications **40** (2013), 2305–2311.

[40] K. Nozaki, H. Ishibuchi, and H. Tanaka, *Adaptive fuzzy rule-based classification systems*, Fuzzy Systems, IEEE Transactions on **4** (1996), 238–250.

[41] Z. Pawlak, *Rough Sets*, International Journal of Information and Computer Science **11** (1982), 341–356.

[42] Z. Pawlak, *Rough Sets:Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Boston, 1991.

[43] Z. Pawlak and A. Skowron, *Rudiments of rough sets*, Information Sciences **177** (2007), 3–27.

[44] D. Pyle, *Data Preparation for Data Mining*, Morgan Kaufmann, San Francisco, 1999.

[45] A. M. Radzikowska and E. E. Kerre, *A comparative study of fuzzy rough sets*, Fuzzy Sets and Systems **126** (2002), 137–155.

[46] C. Stanfill and D. Waltz, *Toward memory-based reasoning*, Communications of the ACM **29** (1986), 1213–1228.

[47] M. Stone, *Cross-validatory choice and assessment of statistical predictions*, Journal of the royal statistical society. Series B (Methodological) (1974), 111–147.

[48] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*, Pearson Addison Wesley, Boston, 2006.

[49] C.-H. Wang, W.-H. Lee, and C.-T. Pang, *Applying fuzzy FP-Growth to mine fuzzy association rules*, World Academy of Science, Engineering and Technology **65** (2010),956–962.

[50] C.-H. Wang, W.-H. Lee, and C.-T. Pang, *Fuzzy data mining for quantitative transactions with FP-Growth*, Journal of Nonlinear and Convex Analysis **14** (2013),193–207.

[51] D. R. Wilson and T. R. Martinez, *Improved heterogeneous distance functions*, Journal of artificial intelligence research (1997), 1–34.

[52] X. Yang, J. Yang, C. Wu, and D. Yu, *Dominance-based rough set approach and knowledge reductions in incomplete ordered information system*, Information Sciences **178** (2008), 1219–1234.

[53] L. Zadeh, *Fuzzy sets*, Information and Control **8** (1965), 338–353.

Manuscript received 13 September 2023
revised 4 September 2023

Wei-Hsuan Lee
Dept. of Information of Management,Yuan Ze University, Taoyuan, 32003, Taiwan
E-mail address: sandymeta@gmail.com

Chien-Hua Wang
School of Business, Lingnan Normal University, Zhanjiang, 524048, China
E-mail address: wangthuck@gmail.com

Chia-Hsuan Yeh
Dept. of Information of Management,Yuan Ze University, Taoyuan, 32003, Taiwan
E-mail address: imcyeh@saturn.yzu.edu.tw

Chin-Tzong Pang
Department of Information Management, and Innovation Center for Big Data and Digital Convergence, Yuan Ze University, Taoyuan, Taiwan
E-mail address: imctpang@saturn.yzu.edu.tw