Knowledge and Information Systems (2019) 61:821–846
https://doi.org/10.1007/s10115-018-1318-3

REGULAR PAPER

A novel fuzzy rule extraction approach using Gaussian kernel-based granular computing

Guangyao Dai[1] · Yi Hu[2] · Yu Yang[1] · Nanxun Zhang[1] · Ajith Abraham[3] · Hongbo Liu[1]

Received: 1 January 2018 / Revised: 8 September 2018 / Accepted: 29 November 2018 /
Published online: 1 January 2019
© Springer-Verlag London Ltd., part of Springer Nature 2019

Abstract

In this paper, we present a novel fuzzy rule extraction approach by employing the Gaussian kernels and fuzzy concept lattices. First we introduce the Gaussian kernel to interval type-2 fuzzy rough sets to model fuzzy similarity relations and introduce a few concepts and theorems to improve the classification performance with fewer attributes accordingly. Based on this idea, we propose a novel attribute reduction algorithm, which can achieve better classification performance of deducing reduction subset of fewer attributes, and this will be used in the subsequent decision rule extraction. Then we justify the necessary and sufficient conditions of our fuzzy rule extraction approach through three implicit rule theorems and present a novel fuzzy decision rule extraction algorithm using fuzzy concept lattices and introduce the concepts of frequent nodes and candidate 2-tuples to our pruning strategy. Also, comparative performance experiments are carried out on the UCI datasets, and the results of both reduction subset size and classification ability show the advantages of our algorithm.

Keywords Interval type-2 fuzzy rough sets · Gaussian kernel · Fuzzy formal concept · Fuzzy similarity relation · Granular computing

1 Introduction

Granular computing has a wide range of applications in artificial intelligence [46,52], clustering [24,34], and classification [2,25,27]. The basic idea of granular computing points out that mathematical universes of discourse must be partitioned in agreement with the limitations of human perception, and this is critical in knowledge representation and rule-based systems [12,60]. In general, the problems related to incompleteness, uncertainty, and vagueness are referred to as granular computing [29,43], in which the conventional tools include rough sets [55], fuzzy sets [48], and concept lattices [41].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10115-018-1318-3) contains supplementary material, which is available to authorized users.

✉ Hongbo Liu
lhb@dlmu.edu.cn

Extended author information available on the last page of the article

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

It is well known that rule extraction is a main task in granular computing-based information system [1,28,57], and the fuzzy logic can effectively extract fuzzy rules [26]. Typically, the fuzzy rule extraction methods are based on type-1 fuzzy logic system (FLS). It is noteworthy, however, that type-2 fuzzy sets can provide better performance than type-1 fuzzy sets, especially when there exists lots of uncertainties [22], so it is necessary to thoroughly study type-2 fuzzy sets-based granular computing models. Meanwhile, although association rules [30] can be extracted by known methods, there usually exists redundancy in its decision rules. This not only has tremendous restriction on rule extraction in time and space, but also affects the accuracy of decision classification [18,37]. Therefore, it is important to explore a new approach to extract decision rules without redundancy.

Interval type-2 fuzzy rough sets [11] combining with rough sets provide an effective means of overcoming the problem of discretization on rough sets. As we know, the attribute reduction and rule extraction are the main concerns in the fuzzy rough sets-based granular computing, in which attribute reduction can select useful features and can be used effectively in dimensionality reduction and rule extraction [56]. When addressing attribute reduction in the fuzzy rough sets-based granular computing, the fuzzy similarity relation [51] is used to measure the similarity between different samples and to formulate the fuzzy lower and upper approximations of a decision [53]. Thus, fuzzy similarity relations have great influence on the performance of interval type-2 fuzzy rough set-based intelligent data analysis. Fortunately, Gaussian kernel can be used to model fuzzy similarity relations on fuzzy rough sets and type-2 fuzzy sets [5,17], and it has significant advantages in nonlinear division of numerical and fuzzy data [40]; thus, it provides a new way to expand the theory of interval type-2 fuzzy rough sets.

Concept lattice is one of the most powerful mathematical tools to extract rules [44]. The relation between one object and one attribute is traditionally represented by 0 or 1, which means that the rules extracted are assumed to be consistent without uncertainty. However, when extracting knowledge in reality, deterministic rules are biased and less reliable, so it is reasonable to introduce fuzzy logic into concept lattice theory [42]. Although association rules [30] can be extracted with traditional concept lattices efficiently, much less is known about extracting non-redundant decision rules. It is therefore necessary to explore an approach to extract decision rules without redundancy, basing on fuzzy concept lattices.

In this paper, we propose a novel Gaussian kernel-based interval type-2 fuzzy rough sets attribute reduction algorithm and a fuzzy concept lattices-based rule extraction algorithm, respectively. The contributions of this paper are summarized as follows:

- We propose an attribute reduction algorithm on Gaussian kernel-based interval type-2 fuzzy rough sets. We also introduce the Gaussian kernel function into interval type-2 fuzzy rough sets to model fuzzy similarity relations. The relevant concepts and theorems, such as fuzzy upper and lower approximations, fuzzy equivalence relations, positive region and dependency, are defined. Then we present the attribute reduction algorithm on Gaussian kernel-based interval type-2 fuzzy rough sets.
- We fully justify the necessary and sufficient conditions of extracting non-redundant decision rules based on fuzzy concept lattices. Meanwhile, we characterize three intension cases of child node in fuzzy concept lattices, and we present three implicit rule theorems that guarantee the non-redundancy of the extracted decision rules.
- We propose a linear decision rule extraction algorithm based on fuzzy concept lattices (FCLRE), through introducing a novel pruning strategy in the search of frequent nodes and candidate 2-tuples in its concept lattices. Moreover, our implicit rule theorems ensure that the FCLRE can extract the non-redundant decision rules, which were concerned in few other methods before.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

The rest of the paper is organized as follows. In Sect. 2, we briefly review the related previous work. In Sect. 3, we present theoretical descriptions of Gaussian kernel-based interval type-2 fuzzy rough sets and propose a novel attribute reduction algorithm. In Sect. 4, we justify three implicit rule theorems that are used to extract non-redundant decision rules with respect to fuzzy concept lattices and further propose the corresponding rule extraction algorithm. In Sect. 5, we carry out comparative performance experiments on the UCI datasets, and the results show that our algorithm outperforms the FSRR, FRS, and IM-IT2FRS algorithms both in reduction subset size and classification ability. Lastly, in Sect. 6, we conclude the paper with some comments and propose some related future work.

2 Related work

In this section, we briefly introduce the notations and definitions of fuzzy rough sets, type-2 fuzzy sets, interval type-2 fuzzy sets, and concept lattices. Most of the details can be found in [13,35,39,62].

2.1 Fuzzy rough sets

Fuzzy rough set theory is a mathematical tool to deal with two distinct types of uncertainty in data, namely the vagueness and the incompleteness [50].

Definition 1 (*Fuzzy Rough Set*) A tuple (U, R) is called a fuzzy rough set, where U is the domain of objects and R is a set of a series of fuzzy equivalence relations. For each $R_i \in R$ and fuzzy set X on U, the fuzzy upper and lower approximations are defined as in Eqs. (1) and (2), respectively:

$$R_i X^*(x) = \sup_{y \in U} \min\{X(y), R(x, y)\}, \tag{1}$$

$$R_i X_*(x) = \inf_{y \in U} \max\{X(y), (1 - R(x, y))\}. \tag{2}$$

The fuzzy equivalence relations address the discretization of any real-valued features that rough set theory requires, and it should meet three properties, namely symmetry, reflexivity and max–min transitivity. We also know that kernel methods allow mapping data from low-dimension spaces to high-dimension spaces [39]. To model fuzzy similarity relations, Hu et al. proposed type-1 fuzzy rough sets using Gaussian kernel, and the fuzzy similarity matrix obtained satisfied symmetry, reflexivity and $T_{\cos}$-transitivity [17].

2.2 Type-2 fuzzy sets

The concept of type-2 fuzzy sets was proposed by Zadeh [62].

Definition 2 (*Type-2 Fuzzy Sets*) A type-2 fuzzy set, denoted by $\tilde{A}$ with domain U, is characterized by

$$\tilde{A} = \int_{x \in U} \frac{\mu_{\tilde{A}}(x)}{x} = \int_{x \in U} \frac{\int_{u \in J_x} \frac{f_x(u)}{u}}{x}. \tag{3}$$

Here $\mu_{\tilde{A}}(x) = \int_{u \in J_x} \frac{f_x(u)}{u}$ is the secondary membership function, where $J_x \subseteq [0, 1]$ is the primary membership of x, and $f_x(u)$ is the secondary membership of x.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

The union of J_x is called the footprint of uncertainty (FOU). The upper and lower limits of FOU are the upper and lower membership functions of $\tilde{A}$, respectively.

As an extension of type-1 fuzzy sets, the membership of type-2 fuzzy sets is not a precise set. In fact, it is a fuzzy set in the interval [0, 1]. Thus, the type-2 fuzzy sets possess more flexibility to present uncertainties than type-1 fuzzy sets [8]. Accordingly, the type-2 fuzzy sets and type-2 fuzzy systems have applications in various fields [15,54]. The computation cost of type-2 fuzzy sets is relative higher, as the operations of type-2 fuzzy sets are on three-dimensional spaces that generalized from two-dimensional spaces of type-1 fuzzy sets [62]. Recently, more attentions have been paid to the research of simple type-2 fuzzy sets, e.g., interval type-2 fuzzy sets [45], constrained type-2 fuzzy sets [14], and concavoconvex type-2 fuzzy sets [47].

The definition of interval type-2 fuzzy sets is as below.

Definition 3 (*Interval Type-2 Fuzzy Sets*) We call $\tilde{A}$ an interval type-2 fuzzy set if all the secondary membership grades of $\tilde{A}$ are l, namely, $\mu_{\tilde{A}}(x) = 1$, as shown in Eq. (4).

$$\tilde{A} = \int_{x \in U} \frac{\mu_{\tilde{A}}(x)}{x} = \int_{x \in U} \frac{\int_{u \in J_x} \frac{1}{u}}{x}, \quad J_x \subseteq [0, 1]. \tag{4}$$

Since interval type-2 fuzzy sets can simplify type-2 fuzzy sets from the spatial region to the band type region, the computational complexity of interval type-2 fuzzy sets is dramatically reduced [4,10].

Figure 1a illustrates two type-2 fuzzy sets $\tilde{N}$ and $\tilde{Z}$. The shadow parts in Fig. 1b are the FOU of $\tilde{N}$ and $\tilde{Z}$. Figure 1b illustrates an example of a secondary membership function $\mu_{\tilde{N}}(-0.4)$. For Fig. 1, it is easy to see that the primary membership of -0.4 is $J_{-0.4} = [0.4, 0.6]$, and the secondary membership of -0.4 is 1.

It is well known that fuzzy decision making [9] and rule extraction are two main tasks in interval type-2 fuzzy sets-based knowledge systems. By combining the Gaussian kernel with interval type-2 fuzzy sets, we can solve the fuzzy rules interpolation of sparse fuzzy rule-based systems and can achieve better performance than any other existing methods [5]. Nevertheless, the study of the Gaussian kernel-based interval type-2 fuzzy rough sets seems to have eluded our attention.

2.3 Concept lattices

The concept lattice is a core data structure in formal concept analysis [13,35]. Generally, the data analyzed by the concept lattice are represented by a formal context [3]. In formal concept analysis, a formal context is usually denoted as a triple (U, A, I), where U is the domain, A is the set of attributes, and I is a set of binary relations between U and A. In a formal context (U, A, I), there is a unique partially ordered set corresponding to I.

The definition of concept lattice is as below.

Definition 4 (*Concept Lattice*) A concept lattice L is a lattice structure generated by the partially ordered set. We also call each node in L a concept.

The concept lattice is complementary to other granular computing methods [59]. Because of fuzzy sets, we can use concept lattices to handle continuous data [42]. In fuzzy concept lattices-based fuzzy formal concept analysis, the uncertainties can be described directly by membership degrees and the similarity between any two different concepts can be computed [20]. Currently, concept lattices are widely used in rule extraction and information retrieval

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

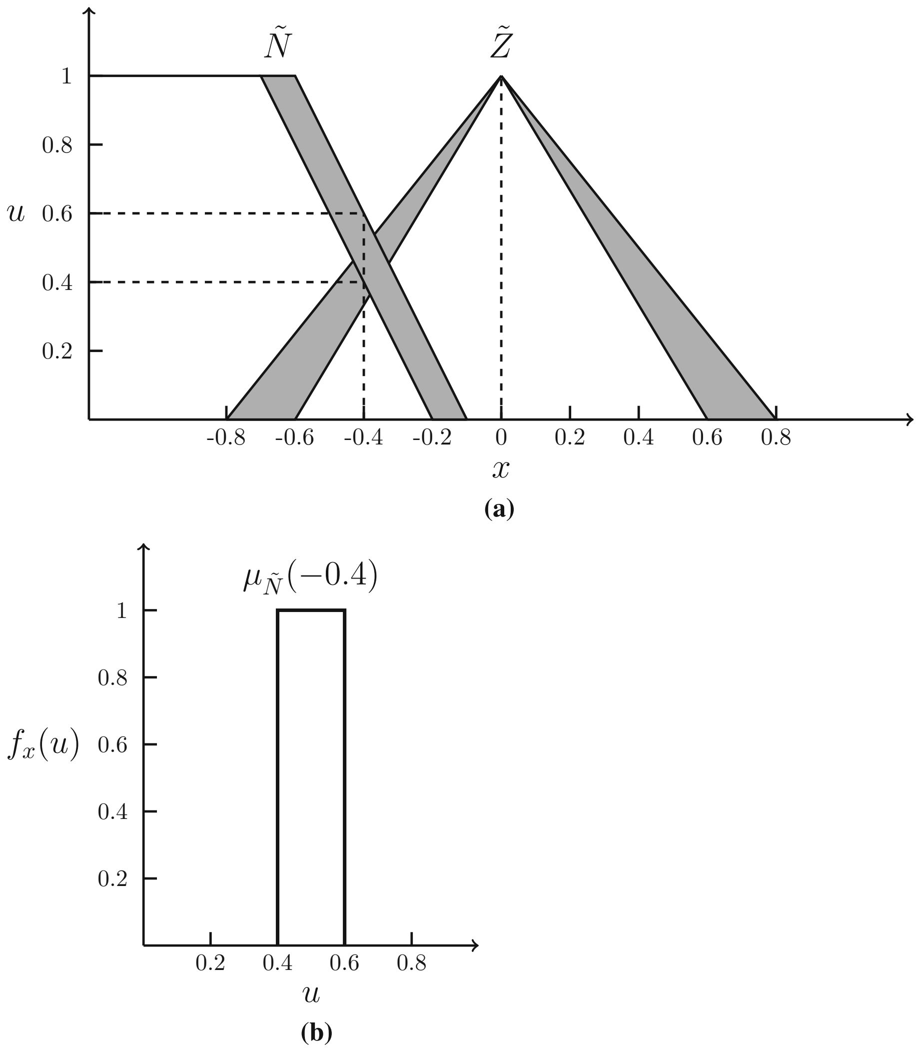


Fig. 1 Type-2 fuzzy sets

[21,23]. However, current rule extraction methods based on concept lattices generally focus on association rules [31], and little attention is paid to decision rules that can extract rules without redundancy.

3 Gaussian Kernel-based interval type-2 fuzzy rough set and its attribute reduction

In this section, we provide an overview of the relevant concepts and introduce the theoretical descriptions of Gaussian kernel-based interval type-2 fuzzy rough sets. Based on these, we propose a novel attribute reduction algorithm.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

3.1 Gaussian Kernel-based fuzzy rough sets

Let U be a non-empty finite set, and let an object $x_i \in U$ be represented as $x_i = \langle x_{i1}, x_{i2}, \ldots, x_{in} \rangle \in R^n$. Then, U can be considered as a subset of R^n. For convenience, let R represent R^n. For any two objects x_i and x_j, we define the similarity between them by Gaussian kernel function as

$$k(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\delta^2}\right), \tag{5}$$

where $\|\cdot\|$ is the Euclidean distance, δ is the kernel parameter, and $k : U \times U \to [0, 1]$ is a kernel function. Let $R_G(x_i, x_j) := k(x_i, x_j)$, and R_G is the similarity matrix of all objects. When $R_G(x_i, x_i) = 1$, the kernel function satisfies $T_{\cos}$-transitivity (see Eq. (6)):

$$T_{\cos}(r, s) = \max\{rs - \sqrt{(1 - r^2)(1 - s^2)}, 0\}, \tag{6}$$

where r and s represent the relationship values in R_G, and $T_{\cos}(r, s)$ is a triangular norm of r and s. It is easy to see that R_G is a fuzzy similarity matrix [36]. Since one can use the Gaussian kernel to model fuzzy relations between any two different objects, the fuzzy upper and lower approximations can be defined by the Gaussian kernel in fuzzy rough sets as follows.

Definition 5 (*Fuzzy Upper and Lower Approximations*) With the above notations, we define the fuzzy upper approximations of a fuzzy set X on domain U as in Eqs. (7) and (8):

$$R_{GT} X^*(x) = \sup_{y \in U} T_{\cos}(R_G(x, y), X(y)), \tag{7}$$

$$R_{G\sigma} X^*(x) = \sup_{y \in U} \sigma_{\cos}(N(R_G(x, y)), X(y)), \tag{8}$$

and the fuzzy lower approximations as in Eqs. (9) and (10):

$$R_{G\theta} X_*(x) = \inf_{y \in U} \theta_{\cos}(R_G(x, y), X(y)), \tag{9}$$

$$R_{GS} X_*(x) = \inf_{y \in U} S_{\cos}(N(R_G(x, y)), X(y)), \tag{10}$$

where N is an involutory negator, i.e., $N(N(x)) = x$ for each $x \in [0, 1]$, $S_{\cos}$ is the inverse triangular norm of $T_{\cos}$, $\theta_{\cos}$ is the residual implication of $T_{\cos}$ defined as

$$\theta_{\cos}(r, s) = \sup\{t | t \in [0, 1], T_{\cos}(r, t) \leq s\},$$

and $\sigma_{\cos}$ is a generalized triangular norm function of $S_{\cos}$ defined as

$$\sigma_{\cos}(r, s) = \inf\{t | t \in [0, 1], S_{\cos}(r, t) \geq s\}.$$

We say $T_{\cos}$ and $S_{\cos}$ are dual to N if $T_{\cos}(r, s) = N(S_{\cos}(N(r), N(s)))$.

3.2 Gaussian Kernel-based interval type-2 fuzzy rough sets

In order to combine interval type-2 fuzzy sets and rough sets, we propose the following dual theorem.

Theorem 1 *Suppose that $T_{\cos}$ and $S_{\cos}$ are dual to N. Then the residual implication $\theta_{\cos}$ of $T_{\cos}$ and the generalized triangular norm function $\sigma_{\cos}$ of $S_{\cos}$ are dual to N.*

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Proof Note that

$$\begin{aligned} N(\sigma_{\cos}(N(r), N(s))) &= N(\inf\{t|t \in [0,1], S_{\cos}(N(r), t) \geq s\}) \\ &= N(\inf\{t|t \in [0,1], N(T_{\cos}(r, N(t))) \geq s\}) \\ &= \sup\{N(t)|t \in [0,1], T_{\cos}(r, N(t)) \leq s\}. \end{aligned}$$

Let $d = N(t)$. Then,

$$N(\sigma_{\cos}(N(r), N(s))) = \sup\{d|d \in [0,1], T_{\cos}(r, d) \leq s\} = \theta_{\cos}(r, s).$$

Hence the theorem follows. □

The residual implication of $T_{\cos}$ can also be expressed as shown in Eq. (11):

$$\theta_{\cos}(r, s) = \begin{cases} rs + \sqrt{(1-r^2)(1-s^2)} & \text{if } r > s, \\ 1 & \text{otherwise.} \end{cases} \tag{11}$$

Combing Definitions 3, 5, Eq. (11) and Theorem 1, $T_{\cos}$ and $\theta_{\cos}$ which can be considered as the upper and lower approximation operator, respectively, are introduced to Gaussian kernel-based interval type-2 fuzzy rough sets (G-I2FRS) as in Definition 6.

Definition 6 (*G-I2FRS*) The Gaussian kernel-based interval type-2 fuzzy set is defined by the fuzzy $T_{\cos}$-upper and $\theta_{\cos}$-lower approximations. The $T_{\cos}$-upper approximation fuzzy operator, $R_{GT}X^*(x)$, is determined as shown in Eq. (12):

$$R_{GT}X^*(x) = [\overline{R_{GT}}X^*(x), \underline{R_{GT}}X^*(x)], \tag{12}$$

where

$$\overline{R_{GT}}X^*(x) = \sup_{y \in U} T_{\cos}(\overline{R_G(x,y)}, \overline{X(y)}),$$

and

$$\underline{R_{GT}}X^*(x) = \sup_{y \in U} T_{\cos}(\underline{R_G(x,y)}, \underline{X(y)}).$$

The $\theta_{\cos}$-lower approximation fuzzy operator, $R_{G\theta}X_*(x)$, is determined as in Eq. (13):

$$R_{G\theta}X_*(x) = [\overline{R_{G\theta}}X^*(x), \underline{R_{G\theta}}X^*(x)], \tag{13}$$

where

$$\overline{R_{G\theta}}X_*(x) = \inf_{y \in U} \theta_{\cos}(\overline{R_G(x,y)}, \overline{X(y)}),$$

and

$$\underline{R_{G\theta}}X_*(x) = \inf_{y \in U} \theta_{\cos}(\underline{R_G(x,y)}, \underline{X(y)}).$$

For a decision class d_i, its upper and lower approximations are defined as in Eqs. (14) and (15), respectively:

$$R_{GT}d_i^*(x) = \sup_{y \in d_i} R_G(x, y), \tag{14}$$

$$R_{G\theta}d_{i*}(x) = \inf_{y \notin d_i} \sqrt{1 - R_G^2(x, y)}. \tag{15}$$

The upper and lower bounds of the lower approximation are defined as in Eq. (16):

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

$$
\begin{aligned}
\overline{R_{G\theta}} d_{i_*}(x) &= \inf_{y \notin d_i} \sqrt{1 - \underline{R_G^2(x, y)}}, \\
\underline{R_{G\theta}} d_{i_*}(x) &= \inf_{y \notin d_i} \sqrt{1 - \overline{R_G^2(x, y)}}.
\end{aligned} \tag{16}
$$

In the rough set theory, the positive region of a set X is the maximal set of objects that fully belongs to X. Also, the dependency can measure the importance of attributes. The two important concepts in attribute reduction of rough sets, namely the positive region and dependency of our interval type-2 fuzzy rough sets with Gaussian kernel, are defined below.

Definition 7 (*Positive Region*) Let $U/P = \{P_1, P_2, \ldots, P_n\}$ and $U/D = \{D_1, D_2, \ldots, D_n\}$ be interval type-2 fuzzy partitions of U, in which P is a subset of condition attributes and D is the set of decision attributes. The positive region of D related to P is defined as

$$
POS_P(D) = \bigcup_{D_i \in U/D} R_{G\theta} D_{i_*}. \tag{17}
$$

The upper and lower membership functions of $POS_P(D)$, denoted by $\overline{POS_P(D)}$ and $\underline{POS_P(D)}$, are

$$
\begin{aligned}
\overline{POS_P(D)} &= \bigcup_{D_i \in U/D} \overline{R_{G\theta}} D_{i_*}, \\
\underline{POS_P(D)} &= \bigcup_{D_i \in U/D} \underline{R_{G\theta}} D_{i_*}.
\end{aligned} \tag{18}
$$

Definition 8 (*Upper and Lower Dependency*) The upper and lower dependency of D related to P, denoted by $\overline{\gamma}_P(D)$ and $\underline{\gamma}_P(D)$, are defined as follows:

$$
\overline{\gamma}_P(D) = \frac{\overline{POS_P(D)}}{|U|}, \tag{19}
$$

$$
\underline{\gamma}_P(D) = \frac{\underline{POS_P(D)}}{|U|}. \tag{20}
$$

3.3 Attribute reduction algorithm on Gaussian Kernel-based interval type-2 fuzzy rough sets

For an object x, we use $R_{G\theta} d_{i*}(x)$ (see Eq. (14)) to denote the degree that x certainly belongs to a decision class d_i, and use $R_{GT} d_i^*(x)$ (see Eq. (15)) to denote the degree that x possibly belongs to a decision class d_i. Note that, the objective of attribute reduction is to remove redundant attributes while keeping the classification performance. Based on the theoretical results in Sect. 3.2, we propose an attribute reduction algorithm on Gaussian kernel-based interval type-2 fuzzy rough sets (see Algorithm 1 for details).

First, we partition the objects against the condition attributes and decision attributes in decision table using the interval type-2 fuzzy according to the Gaussian kernel function (see Eq. (5)) as shown in line 1 of Algorithm 1. Then, the relationship matrix of D related to C and the positive region $POS_C(D)$ can be determined as in line 2 (see Eq. (17)). In line 3, we initialize the reduction set B, and set the threshold ε. Then we take an attribute subset $B - \{b_i\}$ of B one by one to determine the reduction subset (see lines 4 to 14). For the attribute subset $B - \{b_i\}$, we can also obtain the positive region $POS_{B-\{b_i\}}(D)$ (see line 7). Then we judge whether b_i is in an attribute that can be reduced (see line 8) by comparing $POS_C(D)$ and

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

$POS_{B-\{b_i\}}(D)$. If the condition (see line 9) is not satisfied, B will be updated (see line 10), and continue the cycle after updating variable i. Finally, the reduction set B will be outputted when the Algorithm 1 terminates (see line 15). It is easy to see that the time complexity of Algorithm 1 is $O(n^2 m \log m)$, where n is the number of attributes, and m is the number of objects.

Algorithm 1 Attribute reduction algorithm on Gaussian kernel-based interval type-2 fuzzy rough sets (ARGIRS)

Input:
The decision table $T = (U, C, D, V, f)$, threshold ε.
Output:
Reduction set B.
1: Conduct the interval type-2 fuzzy partition to each condition attribute $c_i \in C$ $(i = 1, 2 \ldots, |C|)$ and decision attribute $D = \{d\}$;
2: Compute the relation matrix R_G of D related to C, and get positive region $POS_C(D)$;
3: Let $B = C$ and set the value of ε in [0, 1];
4: **while** $|B| \neq 1$ and $E \neq \emptyset$ **do**
5: $\quad E = \emptyset$;
6: $\quad$ **for** $i = 1$ to $|B|$ **do**
7: $\quad\quad$ Compute the relationship matrix of D related to $B - \{b_i\}$, and get the positive region $POS_{B-\{b_i\}}(D)$, $b_i \in B(i = 1, 2 \ldots, |B|)$;
8: $\quad\quad$ Compute $E = E \cup \{b_i \in B \mid ||\overline{POS_{B-\{b_i\}}(D)}| - |\overline{POS_C(D)}||/|U| < \varepsilon \wedge ||\underline{POS_{B-\{b_i\}}(D)}| - |\underline{POS_C(D)}||/|U| < \varepsilon\}$
9: $\quad\quad$ **if** $E \neq \emptyset$ **then**
10: $\quad\quad\quad$ Take an attribute $b' \in E$, and let $B = B - \{b'\}$;
11: $\quad\quad$ **end if**
12: $\quad\quad$ $i = i + 1$.
13: $\quad$ **end for**
14: **end while**
15: Output reduction set B.

For the sake of better demonstrating the procedure of Algorithm 1, we provide an illustrative example as follows.

Example 1 Let $U = \{x_1, x_2, x_3, x_4, x_5\}$ be the five job application candidates. Four interviewers $C = \{a, b, c, e\}$ grade the candidates using centesimal system. The final result is shown in terms of decision attribute $D = \{d\}$, where '1' stands for passing while '0' represents failing. Example 1 is illustrated in Table 1.

We only consider the condition attributes $\{a, b, c, e\}$. After interval type-2 fuzzy operation, the Gaussian kernel similarity matrix R_G between any two objects can be computed as follows:

Table 1 Example 1

x	a	b	c	e	d
x_1	81.5	71	87	89.7	1
x_2	65	86.5	73.5	80.5	0
x_3	76.5	89	94	85	1
x_4	91	78.5	85.5	87	1
x_5	83	92	79	65.5	0

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

$$\begin{pmatrix} 1 & [0.848, 0.901] & [0.903, 0.952] & [0.901, 0.953] & [0.942, 0.960] \\ [0.848, 0.901] & 1 & [0.892, 0.959] & [0.900, 0.968] & [0.803, 0.883] \\ [0.903, 0.952] & [0.892, 0.959] & 1 & [0.917, 0.957] & [0.902, 0.927] \\ [0.901, 0.953] & [0.900, 0.968] & [0.917, 0.957] & 1 & [0.916, 0.933] \\ [0.942, 0.960] & [0.803, 0.883] & [0.902, 0.927] & [0.916, 0.933] & 1 \end{pmatrix}$$

From the above matrix R_G, we can see that R_G satisfies the symmetry, reflexivity, and the $T_{\cos}$-transitivity. Therefore, R_G is a $T_{\cos}$ fuzzy equivalence relation.

With Algorithm 1, we can divide the final score d into two decision classes, 1 (d_1) and 0 (d_2), respectively, where $d_1 = \{x_1, x_3, x_4\}$. The lower approximation sets of decision class d_1 can be obtained as follows:

$$\begin{aligned} \overline{R_{G\theta}} d_{1*}(x_1) &= \inf_{y \notin d_1} \sqrt{1 - \underline{R_G^2(x_1, y)}} = \inf_{y \in \{x_2, x_5\}} \sqrt{1 - \underline{R_G^2(x_1, y)}} \\ &= \min\{\sqrt{1 - 0.848^2}, \sqrt{1 - 0.942^2}\} = 0.3356, \\ \underline{R_{G\theta}} d_{1*}(x_1) &= \inf_{y \notin d_1} \sqrt{1 - \overline{R_G^2(x_1, y)}} = \inf_{y \in \{x_2, x_5\}} \sqrt{1 - \overline{R_G^2(x_1, y)}} \\ &= \min\{\sqrt{1 - 0.901^2}, \sqrt{1 - 0.960^2}\} = 0.2800. \end{aligned}$$

Similarly, we can obtain $\overline{R_{G\theta}} d_{1*}(x_i)$ and $\underline{R_{G\theta}} d_{1*}(x_i)$ for the case of $i = \{3, 4\}$. The lower approximation sets of decision class d_1 are

$$\begin{aligned} \overline{R_{G\theta}} d_{1*} &= \{0.3356/x_1, 0.4297/x_3, 0.4007/x_4\}, \\ \underline{R_{G\theta}} d_{1*} &= \{0.2800/x_1, 0.2811/x_3, 0.2490/x_4\}, \end{aligned}$$

and the lower approximation sets of decision class d_2 ($d = 0$) are

$$\begin{aligned} \overline{R_{G\theta}} d_{2*} &= \{0.4349/x_2, 0.3330/x_5\}, \\ \underline{R_{G\theta}} d_{2*} &= \{0.2490/x_2, 0.2779/x_5\}. \end{aligned}$$

Then the upper and lower membership degrees can be computed as follows:

$$\begin{aligned} \underline{POS_C(D)} &= \cup_{i \in \{1,2\}} \underline{R_{G\theta}} d_{i*} = \underline{R_{G\theta}} d_{1*} \cup \underline{R_{G\theta}} d_{2*} \\ &= \{0.2800/x_1, 0.2490/x_2, 0.2811/x_3, 0.2490/x_4, 0.2779/x_5\}, \\ \overline{POS_C(D)} &= \cup_{i \in \{1,2\}} \overline{R_{G\theta}} d_{i*} = \overline{R_{G\theta}} d_{1*} \cup \overline{R_{G\theta}} d_{2*} \\ &= \{0.3355/x_1, 0.4349/x_2, 0.4297/x_3, 0.4007/x_4, 0.3330/x_5\}. \end{aligned}$$

Let $B = C$ and $\varepsilon = 0.08$. We compute $\overline{POS_{B-\{i\}}(D)}$ and $\underline{POS_{B-\{i\}}(D)}$ of all elements in B ($i \in B$) in the first cycle similarly to computing $\underline{POS_C(D)}$ and $\overline{POS_C(D)}$. Then we compute $||\overline{POS_{B-\{i\}}(D)}| - |\overline{POS_C(D)}||/|U|$ and $||\underline{POS_{B-\{i\}}(D)}| - |\underline{POS_C(D)}||/|U|$ of all elements in B as follows:

$$\begin{aligned} ||\overline{POS_{B-\{a\}}(D)}| - |\overline{POS_C(D)}||/U &= 0.149967, \\ ||\underline{POS_{B-\{a\}}(D)}| - |\underline{POS_C(D)}||/U &= 0.0902363, \\ ||\overline{POS_{B-\{b\}}(D)}| - |\overline{POS_C(D)}||/U &= 0.0779337, \\ ||\underline{POS_{B-\{b\}}(D)}| - |\underline{POS_C(D)}||/U &= 0.0140327, \\ ||\overline{POS_{B-\{c\}}(D)}| - |\overline{POS_C(D)}||/U &= 0.0901133, \\ ||\underline{POS_{B-\{c\}}(D)}| - |\underline{POS_C(D)}||/U &= 0.057081, \end{aligned}$$

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

$$||\overline{POS_{B-\{e\}}(D)}| - |\overline{POS_C(D)}||/U = 0.0588185,$$
$$||\underline{POS_{B-\{e\}}(D)}| - |\underline{POS_C(D)}||/U = 0.0828353.$$

Similarly, we will have $E = \{b\}$ and $B = \{a, c, e\}$ when this cycle terminates, and we will obtain $E = \emptyset$ after the second running. Finally, the reduction subset $\{a, c, e\}$ will be outputted after the termination of Algorithm 1.

4 Fuzzy concept lattices and rule extraction

In this section, first we will introduce fuzzy concept lattice theory and a novel fuzzy pruning strategy for the sake of reducing the search space; then we present and prove three implicit rule theorems for extracting non-redundant decision rules. Lastly, we present the novel rule extraction algorithm.

4.1 Fuzzy concept lattices

In concept lattice theory, the relation between an object and an attribute in formal context is represented crisply by '1' or '0', where '1' indicates that the object possesses this attribute and '0' otherwise. In order to describe these non-crisp relations between objects and attributes, in what follows, we will introduce the fuzzy logic into concept lattices.

Definition 9 (*Fuzzy Formal Context*) A fuzzy formal context is a triple $(U, A^{'}, I^{'})$, where U and $A^{'}$ are finite sets of objects and attributes, respectively, and $I^{'}$ is a fuzzy set of $U \times A^{'}$.

Each $(x, a) \in I^{'}$, where $x \in U$ and $a \in A^{'}$, has a membership value $\mu(x, a) \in [0, 1]$. For a fuzzy formal context $(U, A^{'}, I^{'})$, given a confidence threshold T and $X \subseteq U$, $B \subseteq A^{'}$, we denote the fuzzy extension and fuzzy intension as

$$\begin{aligned} F_O(B) &= \{x \in U^{'} | \forall a (a \in B \wedge \mu(x, a) \geq T)\}, \\ F_A(X) &= \{a \in A^{'} | \forall x (x \in X \wedge \mu(x, a) \geq T)\}. \end{aligned} \tag{21}$$

If both $F_A(X) = B$ and $F_O(B) = X$ hold, we call (X, B) a fuzzy formal concept under T.

Let (X_1, B_1) and (X_2, B_2) be two fuzzy concepts in the fuzzy formal context $(U, A^{'}, I^{'})$. If $X_1 \subseteq X_2$, we call (X_1, B_1) a 'sub-concept' (or 'child node') of (X_2, B_2), and (X_2, B_2) a 'parent concept' (or 'parent node') of (X_1, B_1). If there is no node (X_3, B_3) satisfying $(X_1, B_1) \subseteq (X_3, B_3) \subseteq (X_2, B_2)$, we call (X_1, B_1) a 'direct sub-concept' (or 'direct child node') of (X_2, B_2), and (X_2, B_2) a 'direct parent concept' (or 'direct parent node') of (X_1, B_1).

For the subsets of objects X and attributes A, let $|X|$ and $|A|$ be the cardinalities of X and A, respectively. For any $a \in A$, we define the fuzzy pruning parameters σ and ψ as follows:

$$\sigma_a = \frac{\sum_{x \in X} \mu(x, a)}{|X|}, \tag{22}$$

$$\sigma = \frac{\sum_{a \in A} \sigma_a}{|A|}, \tag{23}$$

$$\psi_a = \sqrt{\frac{\sum_{x \in X} (\mu(x, a) - \sigma_a)^2}{|X|}}, \tag{24}$$

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

$$\psi = \frac{\sum_{a \in A} \psi_a}{|A|}. \tag{25}$$

Note that '$\sum$' in Eqs. (22)–(25) is not the summation symbol but an ordinary symbol in fuzzy sets.

The fuzzy pruning parameter σ in a fuzzy concept $K = (X, A)$ is in fact the average membership degree of all attributes that correspond to the whole objects, which can reflect the degree of each attribute in K. Also, the parameter ψ is actually the average value of all $\psi_a(K)$, which reflects the divergence degree of K. Then it is easy to see that when the fuzzy pruning parameters reach the thresholds, the pruning strategy will be triggered and the search space of rule extraction algorithm can be reduced.

4.2 Decision rule extraction algorithm

Definition 10 (*Association Rule and Decision Rule*) An implicit rule $B \Rightarrow B^{'}$ is called an association rule, if $B \in A^{'}$, $B^{'} \in A^{'}$ and $B \cap B^{'} = \emptyset$. Here B is the antecedent and $B^{'}$ is the consequent. If $B^{'}$ is a decision attribute, then we call $B \Rightarrow B^{'}$ a decision rule.

Definition 11 (*Support and Confidence*) Let $B \Rightarrow B^{'}$ be a decision rule in a fuzzy formal context $(U, A^{'}, I^{'})$, and let $S = (X, B)$ and $S^{'} = (X^{'}, B^{'})$ be the fuzzy concept lattices of B and $B^{'}$, respectively. Then the support of the rule $B \Rightarrow B^{'}$ is defined as follows:

$$Supp(B \Rightarrow B^{'}) = P(B \cup B^{'}). \tag{26}$$

The confidence of this rule is a conditional probability:

$$Conf(B \Rightarrow B^{'}) = P(B \cup B^{'})/P(B). \tag{27}$$

To reduce the search space of fuzzy concept lattices, we can introduce a pruning strategy while searching for the frequent nodes and candidate 2-tuples. We call the node S in a fuzzy concept lattice a frequent node, if its fuzzy pruning parameter σ is not smaller than the pruning threshold ϑ (i.e., $\sigma \geq \vartheta$), and its fuzzy pruning parameter ψ is not greater than the pruning threshold γ (i.e., $\psi \leq \gamma$ [61]). For two frequent nodes $S = (X, B)$ and $S^{'} = (X^{'}, B^{'})$ in a fuzzy concept lattice, we call $(S, S^{'})$ the 'frequent nodes pair'. Suppose that S is the parent node and $S^{'}$ is the child node. Then $(S, S^{'})$ is a 'candidate 2-tuple' if $Conf(B \Rightarrow B^{'}) \geq \eta$ and $Supp(B \Rightarrow B^{'}) \geq \zeta$, where η is the minimum confidence and ζ is the minimum support [33].

Next, we give a basic theorem of redundant rule.

Theorem 2 *For rules $B_1 \Rightarrow B_2$ and $B_3 \Rightarrow B_4$ with $B_1 \cup B_2 = C$ and $B_3 \cup B_4 = C^{'}$, if $C^{'} \subseteq C$, $B_1 \subseteq B_3$, $B_3 \Rightarrow B_4$ is redundant to $B_1 \Rightarrow B_2$, and $B_1 \Rightarrow B_2$ holds with a certain degree of support and confidence, then $B_3 \Rightarrow B_4$ must exist.*

Proof By Eq. (26), we have $Supp(B_1 \Rightarrow B_2) = P(B_1 \cup B_2) = P(C)$ and $Supp(B_1 \Rightarrow B_2) = P(B_3 \cup B_4) = P(C^{'})$. Since $C^{'} \subseteq C$, there is $P(C) \leq P(C^{'})$, and $Supp(B_1 \Rightarrow B_2) \leq Supp(B_3 \Rightarrow B_4)$. Meanwhile, since $B_1 \subseteq B_3$, we have $P(B_3) \leq P(B_1)$. By Eq. (27), there are

$$Conf(B_1 \Rightarrow B_2) = P(B_1 \cup B_2)/P(B_1) = Supp(B_1 \Rightarrow B_2)/P(B_1),$$

and

$$Conf(B_3 \Rightarrow B_4) = P(B_3 \cup B_4)/P(B_3) = Supp(B_3 \Rightarrow B_4)/P(B_3).$$

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Then we have $Supp(B_1 \Rightarrow B_2)/P(B_1) \leq Supp(B_3 \Rightarrow B_4)/P(B_3)$, namely, $Conf(B_1 \Rightarrow B_2) \leq Conf(B_3 \Rightarrow B_4)$. Therefore, if $B_1 \Rightarrow B_2$ exists with a certain degree of support and confidence, the $B_3 \Rightarrow B_4$ must exist, and $B_3 \Rightarrow B_4$ is redundant to $B_1 \Rightarrow B_2$. □

The following three implicit rule theorems (Theorems 3 to 5) follow from Theorem 2. They are used to extract non-redundant rules.

Theorem 3 *If a node $S = (X, B)$ has only one immediate parent $S^{'} = (X^{'}, B^{'})$, then:*

(i) *The antecedent of implicit rule generated by S consists of only one attribute;*
(ii) *The number of implicit rules generated by S is $|B| - |B^{'}|$;*
(iii) *For every attribute value $p_j \in B - B^{'}$, $j \in \{1, 2, \ldots, |B - B^{'}|\}$, there is $p \Rightarrow B - \{p_j\}$.*

Proof Suppose that the implicit rule generated by S is $p_1 p_2 \ldots p_n \Rightarrow B - \{p_1 p_2 \ldots p_n\}$ $(n > 1)$. Then we discuss it in two cases: (1) $p_1 p_2 \ldots p_n \in B^{'}$ and (2) $p_1 p_2 \ldots p_n \notin B^{'}$.

If $p_1 p_2 \ldots p_n \in B^{'}$, then we have the rule $p_1 p_2 \ldots p_n \Rightarrow B - \{p_1 p_2 \ldots p_n\}$ $(n > 1)$ for each $x \in X^{'}$. This rule should be extracted in parent node $S^{'}$ and thus S should be merged into $S^{'}$, contradicting with the fact that S is the child of $S^{'}$.

If $p_1 p_2 \ldots p_n \notin B^{'}$, then rules like $p_j \Rightarrow B - \{p_j\}$ can be extracted, where $p_j \in B - B'$, $j \in \{1, 2, \ldots, |B - B^{'}|\}$. By Theorem 2, we have that $p_1 p_2 \ldots p_n \Rightarrow B - \{p_1 p_2 \ldots p_n\}$ $(n > 1)$ is redundant to $p_j \Rightarrow B - \{p_j\}$. So the antecedent of implicit rules generated by S consists of only one attribute and the number of implicit rules generated by S is $|B| - |B^{'}|$. □

Theorem 4 *Suppose that a node $S = (X, B)$ has n direct parent nodes $S_1 = (X_1, B_1)$, $S_2 = (X_2, B_2), \ldots, S_n = (X_n, B_n)$, satisfying $B \supset B_1 \cup B_2 \cup \cdots \cup B_n$, and for each $p \in B - (B_1 \cup B_2 \cup \cdots \cup B_n)$, $p \Rightarrow B - \{p\}$. Then the number of implicit rules is $|B| - |B_1 \cup B_2 \cup \cdots \cup B_n|$.*

Proof Similar to Theorem 3, for each $p \in B - (B_1 \cup B_2 \cup \cdots \cup B_n)$, we have that p is in the sublattice and the rule $p \Rightarrow B - p$ can be extracted correspondingly. All rules containing p in its antecedent are redundant to $p \Rightarrow B - \{p\}$. The number of implicit rules for these types is thus $|B| - |B_1 \cup B_2 \cup \cdots \cup B_n|$. □

Theorem 5 *If a node $S = (X, B)$ has two direct parent nodes $S_1 = (X_1, B_1)$ and $S_2 = (X_2, B_2)$ with $B = B_1 \cup B_2$, and for each $p_i \in B_1 - (B_1 \cap B_2)$ and $q_j \in B_2 - (B_1 \cap B_2)$, $p_i q_j \Rightarrow B - \{p_i q_j\}$ $(i = 1, 2, \ldots, |B_1 - (B_1 \cap B_2)|$ and $j = 1, 2, \ldots, |B_2 - (B_1 \cap B_2)|)$. Then the number of implicit rules is $|B_1 - (B_1 \cap B_2)| \cdot |B_2 - (B_1 \cap B_2)|$.*

Proof If both p_i and q_j come from the same direct parent node, then the corresponding rules can be obtained when their direct parent node is visited.

If p_i and q_j come from different direct parent node, e.g., $p_i \in B_1 - (B_1 \cap B_2)$ and $q_j \in B_2 - (B_1 \cap B_2)$, then by Theorem 2, we have that all rules containing $p_i q_j$ in its antecedent are redundant to $p_i p_j \Rightarrow B - \{p_i p_j\}$. So the antecedent of implicit rules generated by S consists of only two attributes, and the number of implicit rules generated by S is $|B_1 - (B_1 \cap B_2)| \cdot |B_2 - (B_1 \cap B_2)|$. □

Theorem 5 can be naturally extended to the case that a node S has n direct parent nodes $S_1 = (X_1, B_1), S_2 = (X_2, B_2), \ldots, S_n = (X_n, B_n)$ with $B = B_1 \cup B_2 \cup \cdots \cup B_n$.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Based on the three implicit rule theorems, we propose the rule extraction Algorithm 2. Line 1 of Algorithm 2 initializes the set of decision rules that are extracted from the lattice L, and the rules can be extracted from all nodes in the lattice L one by one (see lines 2 to 26). Line 3 initializes the set of decision rules extracted from a node S_i in L, and line 4 interprets symbols used in the rest of Algorithm 2. The case that the intension of the child node properly includes the union of all intension of its immediate parent nodes is discussed in lines 5 to 13, and the corresponding decision rules are extracted according to Theorems 3 and 4. The other case is considered as well in lines 14 to 22, and the corresponding decision rules are extracted according to Theorem 5. The set I_s will be updated base on the union of I_s and the set of decision rules of node S_i. After extracting the decision rules of all nodes in L, the final rule set I_s can be obtained.

As the form of rules generated by a node S depends on its parent's form, and the number of rules depends on the number d of its parent nodes, we know that the maximum number of rules that can be extracted from S can be computed. Since the cycle for extracting rules executes C_d^k times when k ($k \leq d$) rules are extracted, the maximum number of generating all the rules of S is thus $1+C_d^2+\cdots+C_d^d = 2^d - d$. We can also see that the time complexity of Algorithm 2 is $O(2^n|L|)$, where n is the number of attributes and $|L|$ is the number of objects, and it is linear in terms of the number of objects.

Algorithm 2 Rule extraction algorithm based on fuzzy concept lattice (FCLRE)

Input:
A lattice L, set of decision attributes D.
Output:
Rule set I_s.
1: $I_s = \emptyset$;
2: **for** $i = 1$ to $|L|$ **do**
3: The rule of node $S_i = (X_i, B_i)$ is $Rules(S_i) = \emptyset$;
4: Let d be the number of candidate 2-tuples, where S_i is the child node and $S_{ik} = (X_{ik}, B_{ik})$ is the parent node ($k \geq 1$);
5: **if** $B_i \neq B_{i1} \cup B_{i2} \cup \cdots \cup B_{ik}$ **then**
6: $P = B_i - (B_{i1} \cup B_{i2} \cup \cdots \cup B_{ik})$;
7: **for** $j = 1$ to $|P|$ **do**
8: $p_j = P_j$;
9: **if** $p_j \notin D \wedge D \cap (B_i - p_j) \neq \emptyset$ **then**
10: $Rules(S_i) = Rules(S_i) \cup \{p_j \Rightarrow D \cap (B_i - p_j)\}$;
11: **end if**
12: $j = j + 1$;
13: **end for**
14: **else**
15: $Z = B_{i1} \cap B_{i2} \cap \cdots \cap B_{ik}, r = 2$;
16: **while** $r \leq d$ and $|B_i| > r$ **do**
17: Take all of the combination of r subsets from $\{B_{i1} - Z, B_{i2} - Z, \ldots, B_{ik} - Z\}$, b_{ir} is an attribute in $\{B_{ir} - Z\}$;
18: **if** $\{b_{i1}, b_{i2}, \ldots, b_{ir}\} \cap D = \emptyset \wedge B_i - \{b_{i1}, b_{i2}, \ldots, b_{ir}\} \cap D \neq \emptyset$ **then**
19: $Rules(S_i) = Rules(S_i) \cup \{b_{i1}, b_{i2}, \ldots, b_{ir}\} \Rightarrow B_i - \{b_{i1}, b_{i2}, \ldots, b_{ir}\}$;
20: **end if**
21: $r = r + 1$.
22: **end while**
23: **end if**
24: $I_s = I_s \cup Rules(S_i)$;
25: $i = i + 1$.
26: **end for**

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

5 Experimental results

In this section, we carry out the fuzzy rule extraction experiments on 14 datasets from the UCI Machine Learning Repository (as shown in Table 2). The maximum number of objects in these datasets is 690, and the maximum number of attributes is 60. First we analyze the impact of parameters involved in reduction algorithm ARGIRS, then we evaluate the ARGIRS by comparing it with other methods.

As we noted before, the goal of attribute reduction is to find a minimal attribute subset that maintains the classification performance. Furthermore, parameters δ and ε exhibit a significant impact on the effectiveness associated with the corresponding Gaussian kernel-based interval type-2 fuzzy rough set, where δ is the kernel parameter used to model fuzzy similarity relation, and ε is the precision parameter used to compare positive regions during attribute reduction. Although we know that the kernel parameter δ determines the granular space generated by Gaussian kernel function, and parameter ε determines the precision of final reduction results, there is no theoretical result on specifying these two parameters. In what follows, we will carry out several experiments to determine the range of 'optimal' values of δ and ε, and we illustrate the impact of δ and ε on two datasets, Sonar and Wine, respectively (see Figs. 2, 3).

From Figs. 2 and 3, we can see the trend that the number of attributes in reduction subsets decreases as ε increases and δ decreases. The effects of δ and ε on classification accuracy are the same. When the classification accuracy reaches 80%, we compare the number of attributes in reduction subsets obtained with different intervals of δ and ε and select the intervals that the length of their corresponding reduction subset is less than that of other intervals' on most of the datasets. We also find that when $\delta \in [0.3, 0.4]$ and $\varepsilon \in [0.02, 0.03]$, the higher the classification accuracy of reduced data, the less the number of attributes.

In order to verify the superiority of the proposed reduction algorithm, we carry out comparative experiments on datasets in Table 2 with ARGIRS, fuzzy swarm rough set reduction algorithm (FSRR) [58], fuzzy rough sets-based reduction algorithm (FRS) [32], and inclusion measures-based interval type-2 fuzzy rough sets reduction algorithm (IM-IT2FRS) [63]. The reduction subset sizes on different datasets with each algorithm are shown in Table 3 for comparison.

Table 2 UCI Datasets

ID	Datasets	Objects	Attributes	Classes
1	Credit	690	15	2
2	Heart	270	13	2
3	Hepatitis	155	19	2
4	Horse	368	27	2
5	Sonar	208	60	2
6	Wdbc	569	32	2
7	Wpbc	198	34	2
8	Wine	178	13	3
9	Iris	150	4	3
10	Zoo	101	17	7
11	Ionosphere	351	34	2
12	Climate	540	18	2
13	Qualitative Bankruptcy	250	7	2
14	Glass	214	10	7

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

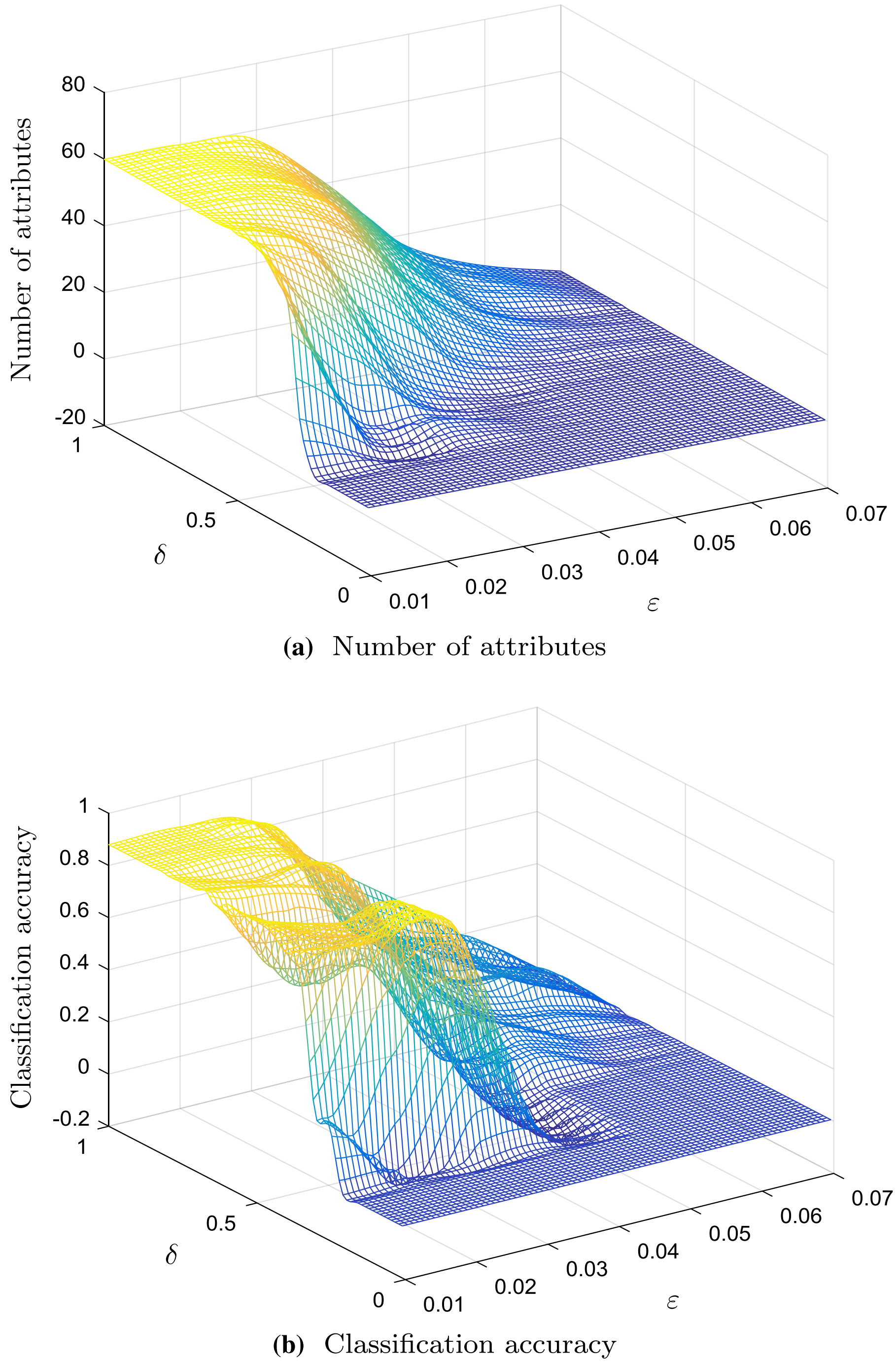
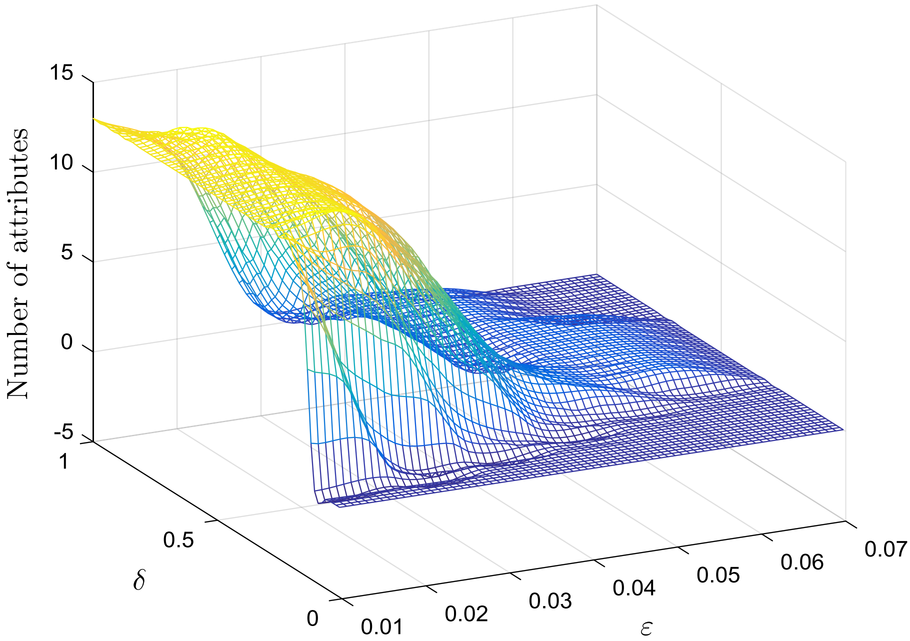
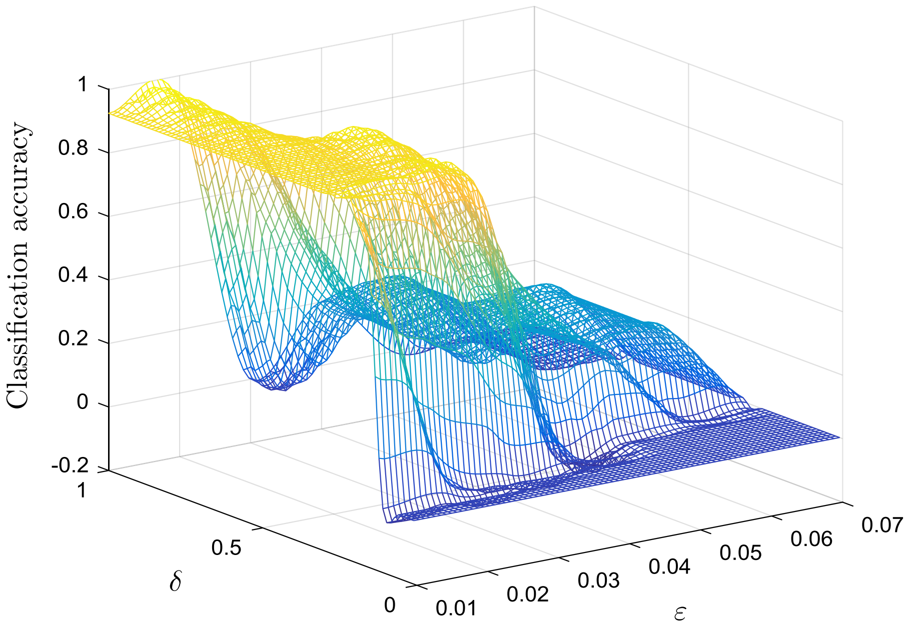


Fig. 2 Trends of the number of attributes (**a**) and classification accuracy (**b**) varying with ε and δ (Sonar)

Content courtesy of Springer Nature, terms of use apply. Rights reserved.



(a) Number of attributes



(b) Classification accuracy

Fig. 3 Trends of the number of attributes (**a**) and classification accuracy (**b**) varying with ε and δ (Wine)

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Table 3 Comparisons of reduction subset sizes on different datasets

ID	Datasets	FSRR	FRS	IM-IT2FRS	ARGIRS
1	Credit	8	11	7	5
2	Heart	10	9	8	8
3	Hepatitis	7	12	4	7
4	Horse	8	8	6	6
5	Sonar	10	7	7	4
6	Wdbc	12	12	9	14
7	Wpbc	14	7	8	7
8	Wine	7	12	6	6
9	Iris	3	3	2	3
10	Zoo	10	9	8	8
11	Ionosphere	12	9	8	7
12	Climate	7	6	6	6
13	Qualitative Bankruptcy	4	4	4	4
14	Glass	9	9	9	7

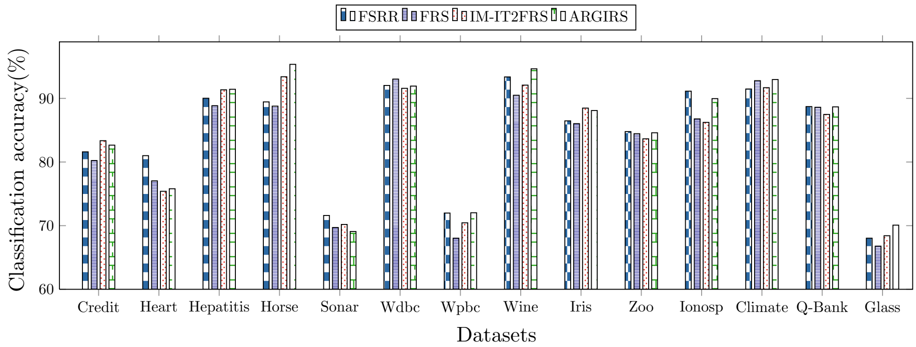


Fig. 4 Comparisons of classification accuracy using the kNN algorithm

As one can see, the performance of ARGIRS is better than FSRR, FRS, and IM-IT2FRS in terms of the number of attributes in reduction subsets on Credit, Sonar, Ionosphere and Glass datasets. On other datasets, the performance of ARGIRS is similar to other three algorithms. In general, we can obtain a minimal attribute subset with ARGIRS.

We also use different classifiers, such as k-Nearest Neighbor classifier (kNN) [16], Naive Bayes (BayesNet) [19] and decision tree algorithm (C4.5) [38], to compare the classification accuracy with our proposed reduction algorithm.

As shown in Figs. 4, 5 and 6, the classification performance of ARGIRS is better than the other algorithms on the majority of the datasets. On datasets such as Wine, Climate and Glass, ARGIRS is superior to other algorithms in terms of reduction subset size and classification accuracy. On other datasets, even if its performance is not optimal, the difference is subtle. The sizes of reduction subsets obtained by ARGIRS are small. Overall, with ARGIRS we can find a minimal reduction subset while maintain the classification performance of the original data, and the performance of ARGIRS is better than other methods.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

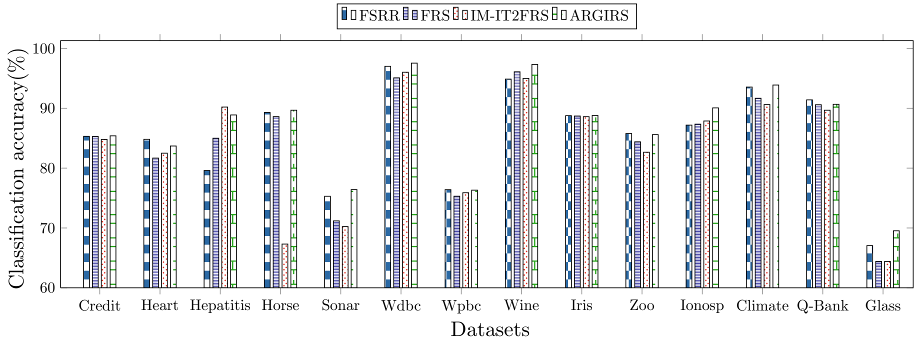


Fig. 5 Comparisons of classification accuracy using the BayesNet algorithm

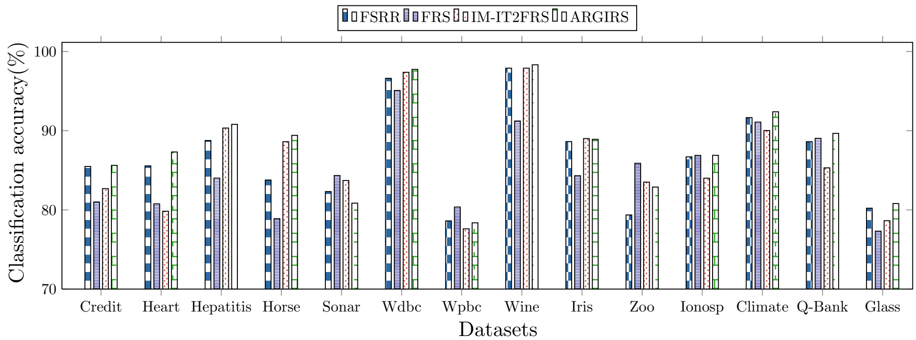


Fig. 6 Comparisons of classification accuracy using the C4.5 algorithm

Basing on the reduction results on these 14 datasets, we carry out the rule extraction experiments using the FCLRE algorithm, and the comparisons of the classification error rate between the FCLRE algorithm and the classical algorithms C4.5, kNN, and BayesNet are listed in Table 4.

As shown in Table 4, the error rate of proposed FCLRE algorithm is lower than the C4.5, kNN, and BayesNet algorithms on nine, eleven, and eight datasets, respectively. On datasets 'Credit', 'Wpbc', 'Iris', 'Zoo', 'Ionosphere', 'Qualitative Bankruptcy', and 'Glass', the error rate of FCLRE algorithm is the lowest. On the other datasets, although the FCLRE algorithm is not the best, it still has better classification performances than other compared algorithms. Also, the average error rate of these algorithms on these 14 datasets can be computed according to Table 4, and we can see that the average error rate of FCLRE is 11.02, which is lower than 11.97 of C4.5, 15.19 of kNN, and 14.13 of BayesNet. Taking all these into account in the analysis of their performances, we conclude that the proposed FCLRE algorithm is the best among all the compared algorithms.

As an illustration example, we list the decision rules extracted by the FCLRE algorithm on the Zoo dataset in Table 5. From Table 5, we see that there are 45 decision rules without redundancy. The letters A to G represent different reduction attributes, the number in the parenthesis after a letter represents the value of its corresponding conditional attribute, and d_i $(i = 1, 2, \ldots, 7)$ represents decision attribute. The corresponding concept lattices are illustrated in Fig. 7.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Table 4 Comparisons of classification error rate (%)

ID	Datasets	C4.5	kNN	BayesNet	FCLRE
1	Credit	14.38	17.35	14.60	**8.10**
2	Heart	**5.69**	24.2	16.30	16.4
3	Hepatitis	9.20	**8.57**	11.11	16.6
4	Horse	10.59	**4.65**	10.32	7.70
5	Sonar	**19.15**	30.92	23.59	25.4
6	Wdbc	**2.26**	8.07	2.45	7.10
7	Wpbc	21.63	27.97	23.68	**17.0**
8	Wine	**1.68**	5.35	2.67	8.40
9	Iris	11.09	11.91	11.21	**4.0**
10	Zoo	17.12	15.4	14.40	**7.60**
11	Ionosphere	13.11	10.04	9.93	**6.71**
12	Climate	7.61	7.04	**6.11**	6.32
13	Qualitative Bankruptcy	10.34	11.33	9.33	**4.56**
14	Glass	19.2	29.92	30.47	**18.4**

The value of the lowest error rate in each row is shown in bold

Table 5 Decision rules results from Zoo dataset

No.	Rule	No.	Rule
1	$A(0), E(1), H(4) \Rightarrow d_1$	24	$D(0), G(1) \Rightarrow d_4$
2	$A(0), D(0) \Rightarrow d_1$	25	$E(0), G(1) \Rightarrow d_4$
3	$A(0), F(0) \Rightarrow d_1$	26	$F(1), G(1) \Rightarrow d_4$
4	$A(0), H(2) \Rightarrow d_1$	27	$A(1), C(1), E(1), H(4) \Rightarrow d_5$
5	$A(0), G(1) \Rightarrow d_1$	28	$B(0), C(1), E(1), H(4) \Rightarrow d_5$
6	$B(0) \Rightarrow d_1$	29	$C(1), D(0), H(4) \Rightarrow d_5$
7	$E(1), G(1) \Rightarrow d_1$	30	$F(1), H(4) \Rightarrow d_5$
8	$G(1), H(2) \Rightarrow d_1$	31	$C(0), H(6) \Rightarrow d_6$
9	$A(1), H(2) \Rightarrow d_2$	32	$E(1), H(6) \Rightarrow d_6$
10	$B(0), H(2) \Rightarrow d_2$	33	$F(1), H(6) \Rightarrow d_6$
11	$C(1), D(0), G(0), H(2) \Rightarrow d_2$	34	$D(0), H(6) \Rightarrow d_6$
12	$A(0), B(0), C(1) \Rightarrow d_3$	35	$D(0), F(1) \Rightarrow d_6$
13	$A(0), B(0), E(0) \Rightarrow d_3$	36	$A(0), B(0), C(0) \Rightarrow d_7$
14	$A(0), B(0), G(0) \Rightarrow d_3$	37	$A(0), B(0), E(1) \Rightarrow d_7$
15	$C(0), D(1), E(1), F(1) \Rightarrow d_3$	38	$C(0), H(0) \Rightarrow d_7$
16	$C(0), D(1), E(1), H(0) \Rightarrow d_3$	39	$C(1), H(6) \Rightarrow d_7$
17	$A(1), B(0), E(0) \Rightarrow d_4$	40	$D(0), E(1), H(0) \Rightarrow d_7$
18	$A(1), C(1), H(0) \Rightarrow d_4$	41	$D(0), G(0), H(0) \Rightarrow d_7$
19	$A(1), G(1) \Rightarrow d_4$	42	$E(0), H(4) \Rightarrow d_7$
20	$B(0), G(1) \Rightarrow d_4$	43	$E(0), H(6) \Rightarrow d_7$
21	$C(1), F(0), H(0) \Rightarrow d_4$	44	$H(8) \Rightarrow d_7$
22	$C(1), D(0), H(0) \Rightarrow d_4$	45	$H(5) \Rightarrow d_7$
23	$D(0), E(0) \Rightarrow d_4$		

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

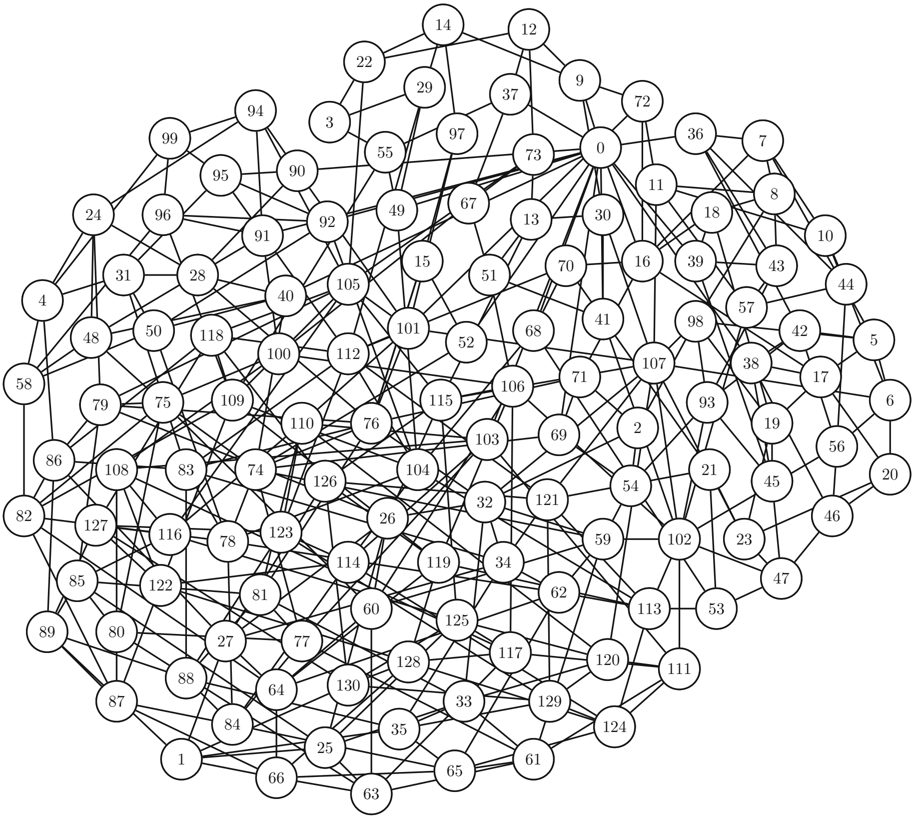


Fig. 7 The concept lattices of Zoo dataset

6 Conclusions

In this paper, we investigated the main concerns in interval type-2 fuzzy rough sets-based granular computing, namely the attribute reduction and rule extraction. We introduced the Gaussian kernel and fuzzy concept lattices to formulate fuzzy similarity relations in attribute reduction and extract fuzzy rules, respectively. We proposed a novel attribute algorithm with Gaussian kernel-based interval type-2 fuzzy rough sets, in which the kernel and precision parameters were comprehensively discussed. Then we generalized the attribute reduction algorithm, and the corresponding concepts and theorems to interval type-2 fuzzy rough sets. After obtaining reduction subsets, we proposed a rule extraction algorithm based on fuzzy concept lattices to extract the decision rules. Three implicit rule theorems were proved theoretically to ensure that the extracted decision rules were without redundancy. Besides, to reduce the search space of the proposed fuzzy rule extraction algorithm, we proposed a pruning strategy for searching the frequent nodes and candidate 2-tuples.

Experimental results illustrate that the proposed algorithms have significant advantages in terms of reduction subset size and classification accuracy. Hence, our rule extraction algorithm is a successful attempt to extract non-redundant decision rules. However, compared with classical fuzzy rough sets, our algorithm still has relatively high computational complexity and thus needs further improvement. For future work, we plan to study the parallelization

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

of the proposed algorithms. We may also explore how to tune the membership functions of interval type-2 fuzzy rough sets and increase the classification accuracy rates by employing swarm intelligence methods, such as ant colony optimization [6], particle swarm optimization [7], and cat swarm optimization [49].

Acknowledgements The authors would like to thank Zongmei Wang and Chao Yang for their scientific collaboration in this research. This work is partly supported by the Program for New Century Excellent Talents in University (NCET-11-0861) and the National Natural Science Foundation of China (Grant Nos. 61472058, 61602086, 61702291 and 61772102).

References

1. Ahmad SSS, Pedrycz W (2017) The development of granular rule-based systems: a study in structural model compression. Granul Comput 2(1):1–12
2. Antonelli M, Ducange P, Lazzerini B, Marcelloni F (2016) Multi-objective evolutionary design of granular rule-based classifiers. Granul Comput 1(1):37–58
3. Baixeries J, Sacarea C, Ojeda-Aciego M (2015) Formal concept analysis. Springer, Berlin
4. Celik E, Gul M, Aydin N, Gumus AT, Guneri AF (2015) A comprehensive review of multi criteria decision making approaches based on interval type-2 fuzzy sets. Knowl Based Syst 85:329–341
5. Chen SM, Chang YC, Pan JS (2013) Fuzzy rules interpolation for sparse fuzzy rule-based systems based on interval type-2 Gaussian fuzzy sets and genetic algorithms. IEEE Trans Fuzzy Syst 21(3):412–425
6. Chen SM, Chien CY (2011) Parallelized genetic ant colony systems for solving the traveling salesman problem. Expert Syst Appl 38(4):3873–3883
7. Chen SM, Kao PY (2013) TAIEX forecasting based on fuzzy time series, particle swarm optimization techniques and support vector machines. Inf Sci 247(15):62–71
8. Chen SM, Lee LW (2010) Fuzzy multiple criteria hierarchical group decision-making based on interval type-2 fuzzy sets. IEEE Trans Syst Man Cybern Part A Syst Hum 40(5):1120–1128
9. Chen SM, Wang CY (2013) Fuzzy decision making systems based on interval type-2 fuzzy sets. Inf Sci 242(Supplement2):1–21
10. Chen TY (2013) A signed-distance-based approach to importance assessment and multi-criteria group decision analysis based on interval type-2 fuzzy set. Knowl Inf Syst 35(1):193–231
11. Deng Z, Choi KS, Cao L, Wang S (2017) T2FELA: type-2 fuzzy extreme learning algorithm for fast training of interval type-2 TSK fuzzy logic system. IEEE Trans Neural Netw Learn Syst 25(4):664–676
12. Dubois D, Prade H (2016) Bridging gaps between several forms of granular computing. Granul Comput 1(2):115–126
13. Fkih F, Omri MN (2016) IRAFCA: an $O(n)$ information retrieval algorithm based on formal concept analysis. Knowl Inf Syst 48(2):465–491
14. Garibaldi JM, Guadarrama S (2011) Constrained type-2 fuzzy sets. In: Advances in type-2 fuzzy logic systems, pp 66–73
15. Gonzalez C, Castro JR, Melin P, Castillo O (2016) An edge detection method based on generalized type-2 fuzzy logic. Soft Comput 20(2):773–784
16. Hou J, Gao H, Xia Q, Qi N (2016) Feature combination and the kNN framework in object classification. IEEE Trans Neural Netw Learn Syst 27(6):1368–1378
17. Hu Q, Zhang L, Chen D, Pedrycz W, Yu D (2010) Gaussian kernel based fuzzy rough sets: model, uncertainty measures and applications. Int J Approx Reason 51(4):453–471
18. Huang CC, Tseng TL, Jiang F, Fan YN, Hsu CH (2014) Rough set theory: a novel approach for extraction of robust decision rules based on incremental attributes. Ann Oper Res 216(1):163–189
19. Janssen J, Laatz W (2017) Naive Bayes. Springer, Berlin
20. Kang X, Miao D (2016) A rough set model based on formal concept analysis in complex information systems. In: International conference on fuzzy systems and knowledge discovery, pp 206–214
21. Kaytoue M, Codocedo V, Buzmakov A, Baixeries J, Kuznetsov SO, Napoli A (2015) Pattern structures and concept lattices for data mining and knowledge processing. Springer, Berlin
22. Li H, Wu C, Wu L, Lam HK, Gao Y (2017) Filtering of interval type-2 fuzzy systems with intermittent measurements. IEEE Trans Cybern 46(3):668–678
23. Li J, Huang C, Mei C, Yin Y (2016) An intensive study on rule acquisition in formal decision contexts based on minimal closed label concept lattices. Intelligent automation and soft computing, pp 1–15

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

24. Lingras P, Haider F, Triff M (2016) Granular meta-clustering based on hierarchical, network, and temporal connections. Granul Comput 1(1):71–92
25. Liu H, Cocea M (2016) Granular computing-based approach for classification towards reduction of bias in ensemble learning. Granul Comput 2(3):131–139
26. Liu H, Cocea M (2017) Fuzzy information granulation towards interpretable sentiment analysis. Granul Comput 2(4):289–302
27. Liu H, Cocea M (2017) Semi-random partitioning of data into training and test sets in granular computing context. Granul Comput 2(4):357–387
28. Liu H, Gegov A, Cocea M (2016) Rule-based systems: a granular computing perspective. Granul Comput 1(4):259–274
29. Livi L, Sadeghian A (2016) Granular computing, computational intelligence, and the analysis of non-geometric input spaces. Granul Comput 1(1):13–20
30. Martinez-Ballesteros M, Troncoso A, Martinez-Alvarez F, Riquelme JC (2016) Improving a multi-objective evolutionary algorithm to discover quantitative association rules. Knowl Inf Syst 49(2):481–509
31. Mondal KC, Pasquier N, Maulik U, Maulik U, Bandhopadyay S (2016) A new approach for association rule mining and bi-clustering using formal concept analysis. In: International conference on machine learning and data mining in pattern recognition, pp 86–101
32. Parthal M, In N, Jensen R (2013) Unsupervised fuzzy-rough set-based dimensionality reduction. Inf Sci 229(6):106–121
33. Paul A (2016) Positive and negative association rule mining using correlation threshold and dual confidence approach. Springer India, New Delhi
34. Peters G, Weber R (2016) DCC: a framework for dynamic granular clustering. Granul Comput 1(1):1–11
35. Ren R, Wei L (2016) The attribute reductions of three-way concept lattices. Knowl Based Syst 99(C):92–102
36. Robniksikonja M, Kononenko I (2003) Theoretical and empirical analysis of relieff and rrelieff. Mach Learn 53(1):23–69
37. Sahoo J, Das AK, Goswami A (2015) An effective association rule mining scheme using a new generic basis. Knowl Inf Syst 43(1):127–156
38. Sathyadevan S, Nair RR (2015) Comparative analysis of decision tree algorithms: id3, c4.5 and random forest. Springer India, New Delhi
39. Shawe-Taylor J, Cristianini N (2004) Kernel methods for pattern analysis. J Am Stat Assoc 101(476):1730–1730
40. Sheikhpour R, Sarram MA, Chahooki MAZ, Sheikhpour R (2017) A kernelized non-parametric classifier based on feature ranking in anisotropic Gaussian kernel. Neurocomputing 267:545–555
41. Singh PK, Kumar CA (2017) Concept lattice reduction using different subset of attributes as information granules. Granul Comput 2(3):159–173
42. Singh PK, Kumar CA, Li J (2016) Knowledge representation using interval-valued fuzzy formal concept lattice. Soft Comput 20(4):1485–1502
43. Skowron A, Jankowski A, Dutta S (2016) Interactive granular computing. Granul Comput 1(2):95–113
44. Smatana M, Butka P, Covekova L (2017) Tree based reduction of concept lattices based on conceptual indexes. Springer, Berlin
45. Sola HB, Fernandez J, Hagras H, Herrera F, Pagola M, Barrenechea E (2015) Interval type-2 fuzzy sets are generalization of interval-valued fuzzy sets: toward a wider view on their relationship. IEEE Trans Fuzzy Syst 23(5):1876–1882
46. Song M, Wang Y (2016) A study of granular computing in the agenda of growth of artificial neural networks. Granul Comput 1(4):247–257
47. Tahayori H, Antoni GD (2008) Operations on concavoconvex type-2 fuzzy sets. Int J Fuzzy Syst 10(4):276–286
48. Tripathy BK, Sooraj TR, Mohanty RK (2017) A new approach to interval-valued fuzzy soft sets and its application in decision-making. Springer, Singapore
49. Tsai PW, Pan JS, Chen SM, Liao BY (2012) Enhanced parallel cat swarm optimization based on the Taguchi method. Expert Syst Appl 39(7):6309–6319
50. Vluymans S, Fernández A, Saeys Y, Cornelis C, Herrera F (2017) Dynamic affinity-based classification of multi-class imbalanced data with one-versus-one decomposition: a fuzzy rough set approach. Knowl Inf Syst 2017(1):1–30
51. Wang C, Qi Y, Shao M, Hu Q, Chen D, Qian Y, Lin Y (2017) A fitting model for feature selection with fuzzy rough sets. IEEE Trans Fuzzy Syst 25(4):741–753
52. Wang G (2017) DGCC: data-driven granular cognitive computing. Granul Comput 2(4):343–355
53. William-West TO, Singh D (2018) Information granulation for rough fuzzy hypergraphs. Granul Comput 3(1):75–92

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

54. Xie WX, Zhang QY, Sun ZM, Zhang F (2015) A clustering routing protocol for WSN based on type-2 fuzzy logic and ant colony optimization. Wirel Pers Commun 84(2):1165–1196
55. Xu W, Li W, Zhang X (2017) Generalized multigranulation rough sets and optimal granularity selection. Granul Comput 2(4):271–288
56. Xu WH, Zhang XY, Zhong JM, Zhang WX (2010) Attribute reduction in ordered information systems based on evidence theory. Knowl Inf Syst 25(1):169–184
57. Yang C, Liu H, McLoone S, Chen CLP, Wu X (2018) A novel variable precision reduction approach to comprehensive knowledge systems. IEEE Trans Cybern 48(2):661–674
58. Yang C, Liu H, Sun Y, Abraham A (2013) Multi-knowledge extraction from violent crime datasets using swarm rough algorithm. In: International conference on hybrid intelligent systems, pp 560–565
59. Yao Y (2016) Rough-set concept analysis: interpreting RS-definable concepts based on ideas from formal concept analysis. Inf Sci 346:442–462
60. Yao Y (2016) A triarchic theory of granular computing. Granul Comput 1(2):145–157
61. Yu Q, Liu Z, Lin W, Shi B, Li Y (2005) Research on fuzzy concept lattice in knowledge discovery and a construction algorithm. Acta Electron Sin 33(2):350–353
62. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning. Inf Sci 8(3):199–249
63. Zhao T, Xiao J (2013) Interval type-2 fuzzy rough sets based on inclusion measures. Acta Autom Sin 39(10):1714–1721

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Guangyao Dai is currently a Ph.D. candidate in the School of Information Science and Technology, Dalian Maritime University, China. His research interests include soft computing, granular computing, artificial intelligence and their applications.

Yi Hu received the Ph.D. degree in mathematics from the University of Illinois at Urbana-Champaign in 2012. He is with Department of Mathematical Sciences, Georgia Southern University. His research interests are pure and applied (Fourier) analysis and differential equations.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Yu Yang received the Ph.D. degree in computer science from the Dalian Maritime University, China, in 2016. He is with the Institute of Cognitive Intelligence Technology, Dalian Maritime University, Dalian, and also a Lecturer with the School of Computer Science, Pingdingshan University, Pingdingshan, China. His current research interests include graph theory and big data.

Nanxun Zhang is a Ph.D. candidate at Dalian University of Foreign Languages and Josai International University. Her main research interests are natural language processing and machine translation.

Ajith Abraham received Ph.D. degree in Computer Science from Monash University, Australia. Dr. Abraham is the Director of Machine Intelligence Research Labs, a Not-for-Profit Scientific Network for Innovation and Research Excellence connecting Industry and Academia. Dr. Abraham works in a multi-disciplinary environment involving machine intelligence, cyber-physical systems, Internet of things, network security, sensor networks, Web intelligence, Web services, data mining and applied to various real world problems. In these areas, he has authored/coauthored more than 1300+ research publications and has more than 30,000+ academic citations (h-index of 82 as per Google Scholar), and he has won seven best paper awards at prestigious International conferences. Currently, Dr. Abraham is the editor-in-chief of EAAI and serves/served the editorial board of over 15 International Journals indexed by Thomson ISI. He is actively involved in the organization of several academic conferences, and some of them are now annual events. More information at: http://www.softcomputing.net/.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Hongbo Liu received the Ph.D. degree in computer science from the Dalian University of Technology, China. He is with the Institute of Cognitive Information Technology (ICIT) and the School of Information Science and Technology, Dalian Maritime University. Professor Liu's research interests are in cognitive computing, machine learning, big data, etc. He participates and organizes actively international conference and workshop and international journals/publications. He received the New Century Excellent Talents Award in 2010.

Affiliations

Guangyao Dai[1] · **Yi Hu**[2] · **Yu Yang**[1] · **Nanxun Zhang**[1] · **Ajith Abraham**[3] · **Hongbo Liu**[1]

Guangyao Dai
delight.dai@foxmail.com

Yi Hu
yihu@georgiasouthern.edu

Yu Yang
yangyu@sjtu.edu.cn

Nanxun Zhang
zhangnanxun@foxmail.com

Ajith Abraham
ajith.abraham@ieee.org

1 School of Information Science and Technology, Dalian Maritime University, Dalian 116026, China

2 Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460, USA

3 Machine Intelligence Research Labs, Scientific Network for Innovation and Research Excellence, Auburn, WA 98071, USA

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH ("Springer Nature").

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users ("Users"), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use ("Terms"). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com