Iranian Journal of Fuzzy Systems

Volume 18, Number 6, (2021), pp. 67-82

Original paper

Different classes ratio and Laplace summation operator based intuitionistic fuzzy rough attribute selection

S. Shreevastava[1], S. Singh[2], A. K. Tiwari[3] and T. Som[4]

[1] *Division of Mathematics, SBAS, Galgotias University, Greater Noida, UP- 201310*

[2] *DST-Centre for Interdisciplinary Mathematical Sciences, Institute of Science, BHU, Varanasi, UP-221005*

[3] *Department of CSE, Koneru Lakshmaiah Education Foundation, Hyderabad, Telangana -500075*

[4] *Department of Mathematical Sciences, IIT (BHU), Varanasi, UP -221005*

shivam.rs.apm@itbhu.ac.in, shivanithakur030@gmail.com, anoop.phd2014@gmail.com, tsom.apm@itbhu.ac.in

Abstract

In real-world data deluge, due to insignificant information and high dimension, irrelevant and redundant attributes reduce the ability of experts both in predictive accuracy and speed, respectively. Attribute selection is the notion of selecting those attributes that are essential as well as enough to specify the target knowledge preferably. Fuzzy rough set-based approaches play a crucial role in selecting relevant and less redundant attributes from a high-dimensional dataset. Intuitionistic fuzzy set-based approaches can handle uncertainty as it gives an additional degree of freedom when compared to fuzzy approaches. So, it has a more flexible and practical ability to deal with vagueness and noise available in the information system. In this paper, we introduce two new robust approaches for attribute selection based on intuitionistic fuzzy rough set theory using the concepts of Different Classes ratio and Laplace Summation operator. Firstly, Different Classes ratio and Laplace Summation operator based lower and upper approximations are established based on intuitionistic fuzzy rough set concept. Moreover, we present algorithms and illustrative examples for a better understanding of our approaches. Finally, experimental analysis is performed on some real-valued datasets for attribute selection and classification accuracies.

Keywords: Attribute selection, rough set, intuitionistic fuzzy set, different classes ratio, Laplace summation operator.

1 Introduction

In many areas, data are usually interpreted by high-dimensional feature vectors, such as pattern recognition, bioinformatics, machine learning, signal processing, data mining, and computer vision. High-dimensionality significantly enhances space and time requirements for processing a large volume of data. Moreover, various machine learning and data mining problems, such as clustering, classification, regression and time series forecasting are computationally or analytically tractable in low dimensional spaces may become complex in spaces of several hundred or thousands of dimensions. To overcome this issue, attribute selection (also known as feature selection) methods are designed to select a reduced subset of attributes from the high-dimensional feature or attribute set for a compact and precise representation of data. By eliminating irrelevant and redundant attributes, attribute selection algorithm can degrade the dimensionality of the data, enhance the interpretability, speed up the learning process, increase the performance of different machine learning algorithms, and simplify the learned model.

Rough set theory (Pawlak [33, 34, 35]) can be applied as a tool to find dependencies on the data and to decrease the number of attributes or features held in a dataset by means of the information from the data alone. Rough set theory can be successfully applied only on symbolic data. Nevertheless, it is usually the situation that the values of various features may be both crisp as well as real-valued. Rough set theory can be applied on real-valued information system

Corresponding Author: S. Singh

Received: May 2020; Revised: April 2021; Accepted: June 2021.

https://doi.org/10.22111/IJFS.2021.6334

only after the conversion of real-valued data into discrete data, which may lead to information loss. As an outcome of this, numerous extensions to the original rough set concept have been proposed, such as neighborhood rough set [21, 48], fuzzy rough set [14, 15], tolerance rough set model [40], and variable precision rough set [52].

Dubois and Prade combined rough set with fuzzy set (Zadeh [49]) and developed the concept of fuzzy rough set [14, 15]. Fuzzy rough set theory gives a base for an adequate tool to feature selection or attribute reduction through the usage of a fuzzy similarity relation, which determines the similarity between pairs of data instances (i.e., samples). The fuzzy lower and upper approximations of a decision label are then established dependent on that fuzzy similarity. Numerical feature values are no longer required to be discretized before attribute reduction or feature selection. Rather, these values are transformed to the corresponding degrees of membership of samples (i.e., instances) to the lower and upper approximations. In the last few years, many techniques have been established to tackle this latter uncertainty. Intuitionistic fuzzy (IF) set (proposed by Atanassov [1, 2], which is an extension of Zadeh's fuzzy set (Zadeh [49]), based concepts are broadly used to handle uncertainty by using the concept of membership functions together with non-membership and hesitancy functions. Therefore, it can cope with uncertainty in a far better way in comparison with fuzzy set-based approaches. So, it provides a stronger facility to tackle with the information systems and carry a better sight of brittle ambiguities of the physical environment. Despite of the case that rough sets and IF sets both incarcerate specific characteristic of the similar idea- imprecision, the coalition of both IF set and rough set are hardly addressed by the researchers in their articles. Jena and Ghosh, Chakrabarthy et al. and Nanda and Majumdar explained that upper and lower approximations concept of IF rough sets are yet again IF sets [6, 26, 32]. Samanta and Mondal demonstrated a similar idea in their research article [38]. Coker discussed the relationship between rough set and IF set and discovered the information that fuzzy rough set is certainly an intuitionistic L-fuzzy set [12]. In present era, the intuitionistic fuzzy rough set model is emerging as a successful and effective tool to cope up with uncertainty and can be applied for decision making for solving many real-world problems. In the current study, two robust methods for attribute selection of an intuitionistic fuzzy information system using intuitionistic fuzzy rough set concept are discussed namely, Different Classes ratio (`DC_ratio`) based and Laplace Summation operator (LS operator) based approaches. `DC_ratio` is defined as a fraction of samples belonging to different classes in the neighborhood of a given sample. `DC_ratio` based fuzzy rough set approach can reduce the influence of noisy samples by identifying them [28]. Laplace distribution sets the lowest values to tail and assigns the maximum values nearby the distributed median when compared to the Normal and Uniform distributions. So, the Laplace distribution is more resistant to outliers. Hence, the LS operator is capable of enhancing the selection performance of attributes to attain better learning performance. A fuzzy rough approach for feature selection based on LS operator is proposed by Han et al. [19]. In this paper, we extend these ideas for intuitionistic fuzzy scenario due to its supremacy over fuzzy set. This paper is organized as follows. In Section 2, a literature survey related to the proposed approaches is presented. Some preliminaries required for understanding this paper are given in Section 3. In Section 4, `DC_ratio` based intuitionistic fuzzy rough set approach for attribute selection is presented. LS operator based intuitionistic fuzzy rough set approach for attribute selection is proposed in Section 5. Algorithms, along with illustrative examples, are also given for these approaches. We conclude our work in Section 6.

2 Literature survey

Fuzzy rough set theory is one of the widely used extensions of rough set theory, which can successfully tackle real-valued data. Jensen and Shen proposed a fuzzy rough set-based approach for attribute reduction using the degree of dependency concept [27]. A strong mathematical foundation, along with an algorithm, is set up for fuzzy rough set assisted attribute reduction via discernibility matrix approach is given by Tsang et al. [43]. Wu et al. [45, 46, 47] presented the concept of generalized fuzzy rough approximation operators by using different approaches. An incremental feature selection technique on a hybrid information system using fuzzy rough set-based approach is developed by Zeng et al. [50]. Due to the dependency of lower and upper approximations of a target sample on the nearest neighbor sample, classical fuzzy rough set theory is sensitive towards the noise. So, It is required to construct more robust fuzzy rough set models in order to perform various learning tasks by removing the noise successfully. Some researchers have presented robust fuzzy rough set models like soft fuzzy rough set [20], k-trimmed fuzzy rough set [22], -precision fuzzy rough set [37], vaguely quantified rough set [10], fuzzy variable precision rough set [51], ordered weighted average fuzzy rough set [11] to name a few. All such methods only consider the effect of the nearest sample of a target sample and totally ignore the influence of the neighbors of the nearest sample.

To cope with above problems, Li et al. established `DC_ratio` based fuzzy rough set model, which can minimize the effect of noise on the calculation of the lower and upper approximations by identifying noisy samples [28]. The `DC_ratio` based fuzzy rough set model deals with the similarity between samples as well as the influence of its neighbors. Han et al. [19] presented a laplace distribution-based fuzzy rough feature selection technique by using the LS operator. Laplace distribution is more impervious to outliers because of allocating the lowest values to tail and the maximum values nearby the distributed median when compared to the normal and uniform distributions. Moreover, the Laplace distribution is going to decline more rapidly than the normal distribution around the tails of the distribution. Therefore, LS operator is capable of enhancing the quality of features in order to attain more immeasurable classification performance.

Wang et al. introduced distance measure technique into fuzzy rough sets in order to avoid low discrimination of fuzzy decision values in real valued datasets which occurs due to intersection operation of fuzzy relations while defining dependency function [44]. Dai et. al presented a fuzzy rough set technique based on dominance concept for incomplete interval-valued decision system by defining a fuzzy dominance relation [13]. Singh et al. proposed a fuzzy similarity relation-based tolerance rough set approach for set-valued information systems using a threshold value. They applied their technique on real-valued datasets for attribute reduction and classification and got better results in terms of reduced attributes and higher classification accuracy [39]. In [3], authors introduced a novel approach based on fuzzy concept analysis considering the core ideas of rough set theory and studied the strength the relationship between these two concepts. Chen et al. constructed variable precision fuzzy neighbourhood rough set based multi-label feature reduction technique by using the parameterized fuzzy neighborhood granule to define the decision class of each samples [7]. Mohtashami and Eftekhari proposed a hybrid soft computing approach for feature selection by using the concepts of hesitant fuzzy set, rough set, and methods of discretization in microarray class imbalanced datasets [31]. However, all the above-proposed models are based on fuzzy set concept which is not capable of handling uncertainty completely. Intuitionistic fuzzy set is known as an intuitive generalization of fuzzy set, which simultaneously considers the membership and non-membership of objects belonging to objective sets into account. Along with this issue, much less research works have been injected in the area of feature selection of an intuitionistic fuzzy information system (feature values of objects are intuitionistic fuzzy values) based on intuitionistic fuzzy rough set concept. Lu et al. presented attribute reduction technique for intuitionistic fuzzy information system by using genetic algorithm [30]. Chen and Yang presented a new attribute reduction algorithm by combining intuitionistic fuzzy rough set with information entropy [8]. Esmail et al. considered about the structure of intuitionistic fuzzy rough set model and its characteristics and introduced an attribute reduction technique along with rule extraction [16]. Huang et al. proposed an attribute reduction approach based on intuitionistic fuzzy rough set model by using distance function [24]. Tiwari et al. presented a feature selection technique by establishing an intuitionistic fuzzy rough set model [42]. Huang et al. established dominance based intuitionistic fuzzy rough set model and presented its applications to rule extraction using attribute reduction approaches [25]. Liu and Lin proposed a novel intuitionistic fuzzy rough set model by means of distance method to solve many real-life conflict problems [29].

In this paper, we extend the concepts of two robust fuzzy rough set-based approaches to intuitionistic fuzzy rough scenario in order to tackle uncertainity in a much better way. The uniqueness of the proposed model is given below:

- It generalizes the two robust fuzzy rough set-based approaches, which are capable of handling the noise available in the datasets.
- It provides an attribute reduction algorithm for intuitionistic fuzzy information systems.
- The concept of dependency function is used for calculating the reduct set of an intuitionistic fuzzy information system.

3 Preliminaries

In this section, we propose `DC_ratio` based approach for attribute selection of an IF decision system.

Definition 3.1. [1] *Let U be a finite non-empty set, called universe of discourse. An intuitionistic fuzzy set (IFS) A on U is of the form $A = \{\langle x, \mu_A(x), \nu_A(x)\rangle | x \in U\}$, where $\mu_A : U \to [0,1]$ and $\nu_A : U \to [0,1]$ with the condition $0 \leq \mu_A(x) + \nu_A(x) \leq 1 (\forall x \in U)$ are known as membership degree and non-membership degree of the element x in A, respectively. $\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$ is the degree of hesitancy of the element x in IFS A.*

The cardinality of an IFS A is given by $|A| = \sum_{x\in A} \frac{1+\mu_A(x)-\nu_A(x)}{2}$, where 1 in numerator is a translation factor that guarantees the positivity of $|A|$ while 2 in denominator is a scaling factor which bounds the cardinality between 0 and 1. Any fuzzy set $A = \{\langle x, \mu_A(x)\rangle | x \in U\}$ is an IFS as it can be written in the form $\{\langle x, \mu_A(x), 1 - \mu_A(x)\rangle | x \in U\}$.

Property 3.2. [1] *For every two IFSs A and B the following relations and operations holds*

1. *$A \subset B$ iff $(\forall x \in U)(\mu_A(x) \leq \mu_B(x)$ and $\nu_A(x) \geq \nu_B(x))$;*
2. *$A = B$ iff $A \subset B$ and $B \supset A$;*

3. $N(A) = \{\langle x, \nu_A(x), \mu_A(x) > | x \in U\}$;

4. $A \cap B = \{\langle x, \min(\mu_A(x), \mu_B(x)), \max(\nu_A(x), \nu_B(x))\rangle | x \in U\}$;

5. $A \cup B = \{\langle x, \max(\mu_A(x), \mu_B(x)), \min(\nu_A(x), \nu_B(x))\rangle | x \in U\}$;

6. $A + B = \{\langle x, \mu_A(x) + \mu_B(x) - \mu_A(x).\mu_B(x), \nu_A(x).\nu_B(x)\rangle | x \in U\}$;

7. $A.B = \{\langle x, \mu_A(x).\mu_B(x), \nu_A(x) + \nu_B(x) - \nu_A(x).\nu_B(x)\rangle | x \in U\}$;

8. $\lambda.A = \{\langle x, 1 - (1 - \mu_A(x))^\lambda, \nu_A(x)^\lambda > | x \in U\}$.

where, N is negation operator.

Definition 3.3. [17] *Let U be a collection of finite objects and $C \subseteq A$, then Feng and Mi defined an intuitionistic fuzzy binary relation $R_c(x_i, x_j) = \langle \mu_{R_c}(x_i, x_j), \nu_{R_c}(x_i, x_j)\rangle, c \in C$ as*

$$\mu_{R_c}(x_i, x_j) = \begin{cases} 1 - 2K(x_i, x_j), & K(x_i, x_j) \leq \frac{1}{2} \\ 0, & otherwise \end{cases}$$

$$\nu_{R_c}(x_i, x_j) = \begin{cases} 2((\mu_c(x_i) - \mu_c(x_j))^2 + (\nu_c(x_i) - \nu_c(x_j))^2), & K(x_i, x_j) \leq \frac{1}{2} \\ 1, & otherwise \end{cases}$$

where, $K(x_i, x_j) = (|\mu_c(x_i) - \mu_c(x_j)| + |\nu_c(x_i) - \nu_c(x_j)|)^2, \forall x_i, x_j \in U$.

Definition 3.4. [23] *A quadruple $IS = (U, AT, V, h)$ is called an information system, where $U = \{u_1, u_2, \ldots, u_n\}$ is a non-empty finite set of objects, $AT = \{a_1, a_2, \ldots, a_m\}$ is a non-empty finite set of attributes. $V = \bigcup_{a \in AT} V_a$, where V_a is the set of attribute values with respect to each attribute $a \in AT$ and $h \colon U \times AT \to V$ is an information function that allocates particular values of the objects against to attribute set such that $\forall a \in AT, \forall u \in U, h(u, a) \in V_a$.*

An information system is called an IF information system if attribute values corresponding to objects are intuitionistic fuzzy values.

Definition 3.5. [23] *An IF information system $IS = (U, AT, V, h)$ is said to be an IF decision system if $AT = C \cup D$ with $C \cap D = \phi$, where C and D are non-empty finite sets of conditional and decision attributes, respectively. For example, Table 1 and Table 2 are IF decision systems. In Table 1, $U = \{x_1, x_2, x_3, x_4\}$ is a set of objects, $C = \{c_1, c_2, c_3, c_4\}$ is a set of conditional attributes and $D = \{d\}$ is a decision attribute. In Table 2, object set $U = \{x_1, x_2, \ldots, x_{10}\}$, set of conditional attributes $C = \{c_1, c_2, c_3, c_4, c_5\}$ and set of decision attributes $D = \{d\}$.*

Table 1: Intuitionistic fuzzy decision system with 4 objects and 5 attributes

U	c_1	c_2	c_3	c_4	D
x_1	$< 0.5, 0.3 >$	$< 0.5, 0.4 >$	$< 0.7, 0.2 >$	$< 0.3, 0.1 >$	2
x_2	$< 0.4, 0.3 >$	$< 0.3, 0.4 >$	$< 0.9, 0.1 >$	$< 0.5, 0.2 >$	1
x_3	$< 0.4, 0.1 >$	$< 0.5, 0.3 >$	$< 0.5, 0.4 >$	$< 0.6, 0.2 >$	1
x_4	$< 0.6, 0.2 >$	$< 0.2, 0.5 >$	$< 0.4, 0.2 >$	$< 0.7, 0.1 >$	2

4 Different classes ratio based approach for attribute selection of an IF decision system

In this section, we propose `DC_ratio` based approach for attribute selection of an IF decision system.

Definition 4.1. *Suppose $IS = (U, C \cup D, V, h)$ is an IF decision system with $U = \{x_1, x_2, \ldots, x_n\}$ and $C = \{c_1, c_2, \ldots, c_m\}$. Suppose D partitions the object set U into r crisp equivalence classes $U/D = \{D_1, D_2, \ldots, D_r\}$. Let $d_C(x_i, x_j)$ is*

Table 2: Intuitionistic fuzzy decision system with 10 objects and 6 attributes

U	c_1	c_2	c_3	c_4	c_4	D
x_1	$< 0.2, 0.4 >$	$< 0.1, 0.7 >$	$< 0.2, 0.6 >$	$< 0.6, 0.4 >$	$< 0.2, 0.8 >$	1
x_2	$< 0.1, 0.7 >$	$< 0.1, 0.8 >$	$< 0.3, 0.6 >$	$< 0.5, 0.2 >$	$< 0.2, 0.7 >$	2
x_3	$< 0.1, 0.8 >$	$< 0.1, 0.8 >$	$< 0.2, 0.8 >$	$< 0.5, 0.4 >$	$< 0.6, 0.4 >$	1
x_4	$< 0.1, 0.9 >$	$< 0.6, 0.3 >$	$< 0.2, 0.7 >$	$< 0.2, 0.8 >$	$< 0.6, 0.4 >$	1
x_5	$< 0.4, 0.6 >$	$< 0.2, 0.6 >$	$< 0.2, 0.8 >$	$< 0.2, 0.8 >$	$< 0.2, 0.8 >$	2
x_6	$< 0.1, 0.6 >$	$< 0.2, 0.6 >$	$< 0.2, 0.8 >$	$< 0.2, 0.4 >$	$< 0.2, 0.8 >$	1
x_7	$< 0.6, 0.4 >$	$< 0.6, 0.4 >$	$< 0.6, 0.4 >$	$< 0.7, 0.3 >$	$< 0.4, 0.6 >$	2
x_8	$< 0.6, 0.2 >$	$< 0.6, 0.2 >$	$< 0.8, 0.2 >$	$< 0.4, 0.6 >$	$< 0.4, 0.5 >$	2
x_9	$< 0.6, 0.2 >$	$< 0.6, 0.2 >$	$< 0.8, 0.2 >$	$< 0.1, 0.6 >$	$< 0.8, 0.2 >$	3
x_{10}	$< 0.6, 0.4 >$	$< 0.6, 0.4 >$	$< 0.8, 0.2 >$	$< 0.8, 0.2 >$	$< 0.6, 0.4 >$	3

the distance between objects x_i and x_j for attribute $c \in C$. For $x_i \in D_i$, define $S = \{x_j | d_C(x_i, x_j) \leq l\}$ and $S' = \{x_k | d_C(x_i, x_k) \leq l, x_k \notin D_i\}$. Then, Li et al. [28] defined different classes ratio of x_i as

$$\texttt{DC\_ratio}(x_i) = \frac{|S'|}{|S|}. \tag{1}$$

where $|.|$ is the cardinality of a set. Using this concept, depending on the ratio of neighbours with different classes with respect to x, one can easily investigate noise sample.

4.1 Different classes ratio based IF rough set

Let A is any intuitionistic fuzzy set in $U, R_c(x_j, x_i)$ is an IF relation then Cornellis et al.[9] defined the lower and upper approximations of A as follows:

$$\underline{(R_I)}A(x_i) = \inf_{x_j \in U} I(R_c(x_j, x_i), A(x_j)), \forall x_i, x_j \in U. \tag{2}$$

$$\overline{(R_T)}A(x_i) = \sup_{x_j \in U} T(R_c(x_j, x_i), A(x_j)), \forall x_i, x_j \in U, \tag{3}$$

where, I is an IF implicator and T is an IF t-norm. Let $x_1 = \langle \mu_1, \nu_1 \rangle$ and $x_2 = \langle \mu_2, \nu_2 \rangle$ are two IF numbers then we take following IF t-norm T_w and IF implicator I_w [9]

$$T_w(x_1, x_2) = \langle \max(0, \mu_1 + \mu_2 - 1), \min(1, \nu_1 + \nu_2) \rangle. \tag{4}$$

$$I_w(x_1, x_2) = \langle \min(1, 1 + \mu_2 - \mu_1, 1 + \nu_1 - \nu_2), \max(0, \nu_2 - \nu_1) \rangle. \tag{5}$$

Definition 4.2. *In an IF decision system to classify the instances, we define lower and upper approximations of decision labels $D_i \in U/D, i \in 1, 2, \ldots, r$ as follows:*

$$\underline{(R_I)}D_i(x_i) = \inf_{x_j \in U} I(R_c(x_j, x_i), D_i(x_j)), \forall x_i, x_j \in U.$$

$$\overline{(R_T)}D_i(x_i) = \sup_{x_j \in U} T(R_c(x_j, x_i), D_i(x_j)), \forall x_i, x_j \in U.$$

Theorem 4.3.

$$\underline{(R_I)}D_i(x_i) = \inf_{x_j \notin D_i} \langle \nu_{R_c}(x_i, x_j), 1 - \nu_{R_c}(x_i, x_j) \rangle.$$

$$\overline{(R_T)}D_i(x_i) = \sup_{x_j \in D_i} \langle \mu_{R_c}(x_i, x_j), \nu_{R_c}(x_i, x_j) \rangle.$$

Proof.

$$\begin{aligned}
\underline{(R_I)}D_i(x_i) &= \inf_{x_j\in U} I(R_c(x_i,x_j),D_i(x_j))\\
&= \inf_{x_j\in D_i} I(\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)\rangle,\langle 1,0\rangle)\wedge\inf_{x_j\notin D_i} I(\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)\rangle,\langle 0,1\rangle)\\
&= \inf_{x_j\in D_i}\langle\min(1,1+1-\mu_{R_c}(x_i,x_j),1+\nu_{R_c}(x_i,x_j)-0),\max(0,0-\nu_{R_c}(x_i,x_j))\rangle\\
&\quad\wedge\inf_{x_j\notin D_i}\langle\min(1,1+0-\mu_{R_c}(x_i,x_j),1+\nu_{R_c}(x_i,x_j)-1),\max(0,1-\nu_{R_c}(x_i,x_j))\rangle\\
&= \inf_{x_j\in D_i}\langle 1,0\rangle\wedge\inf_{x_j\notin D_i}\langle\nu_{R_c}(x_i,x_j),1-\nu_{R_c}(x_i,x_j)\rangle\\
&= \inf_{x_j\notin D_i}\langle\nu_{R_c}(x_i,x_j),1-\nu_{R_c}(x_i,x_j)\rangle,\\
\overline{(R_T)}D_i(x_i) &= \sup_{x_j\in U} T(R_c(x_i,x_j),D_i(x_j))\\
&= \sup_{x_j\in D_i} T(\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)\rangle,\langle 1,0\rangle)\vee\sup_{x_j\notin D_i} T(\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)\rangle,\langle 0,1\rangle)\\
&= \sup_{x_j\in D_i}\langle\max(0,\mu_{R_c}(x_i,x_j)+1-1),\min(1,\nu_{R_c}(x_i,x_j)+0)\rangle\\
&\quad\vee sup_{x_j\notin D_i}\langle\max(0,\mu_{R_c}(x_i,x_j)+0-1),\min(1,\nu_{R_c}(x_i,x_j)+1)\rangle\\
&= \sup_{x_j\in D_i}\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)\vee\sup_{x_j\notin D_i}\langle 0,1\rangle\\
&= \sup_{x_j\in D_i}\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)\rangle.
\end{aligned}$$

□

However, above classical IF rough set only considers nearest sample of a target sample. Therefore, attribute reduction of an IFIS based on this model will be sensitive to noise.

Definition 4.4. *We define* `DC_ratio` *based IF rough set, which first detect the noisy sample and then calculate lower and upper approximations by ignoring such noisy samples with taking into account the influence of neighbors of nearest samples as follows:*

$$\underline{(R_I)}^{DC}D_i(x_i) = \inf_{x_j\in U} I(R_c(x_i,x_j),N(DCratio(x_j))).$$

$$\overline{(R_T)}^{DC}D_i(x_i) = \sup_{x_j\in U} T(R_c(x_i,x_j),DCratio(x_j)).$$

Theorem 4.5.

$$\underline{(R_I)}^{DC}D_i(x_i) = \inf_{DCratio\le\lambda}\langle\nu_{R_c}(x_i,x_j),1-\nu_{R_c}(x_i,x_j)\rangle. \tag{6}$$

$$\overline{(R_T)}^{DC}D_i(x_i) = \sup_{DCratio\le\lambda}\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)\rangle. \tag{7}$$

Proof.

$$\begin{aligned}
\underline{(R_I)}^{DC}D_i(x_i) &= \inf_{x_j\in U} I(R_c(x_i,x_j),N(DCratio(x_j)))\\
&= \inf_{DCratio\le\lambda} I(\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)N(\langle 1,0\rangle))\wedge\inf_{DCratio>\lambda} I(\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)N(\langle 0,1\rangle))\\
&= \inf_{DCratio\le\lambda} I(\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)\rangle,\langle 0,1\rangle)\\
&\wedge\inf_{DCratio>\lambda} I(\langle\mu_{R_c}(x_i,x_j),\nu_{R_c}(x_i,x_j)\rangle,\langle 1,0\rangle)\\
&= \inf_{DCratio\le\lambda}\langle\min(1,1+0-\mu_{R_c}(x_i,x_j),1+\nu_{R_c}(x_i,x_j)-1),\max(0,1-\nu_{R_c}(x_i,x_j))\rangle\\
&\quad\wedge\inf_{DCratio>\lambda}\langle\min(1,1+1-\mu_{R_c}(x_i,x_j),1+\nu_{R_c}(x_i,x_j)-0),\max(0,0-\nu_{R_c}(x_i,x_j))\rangle\\
&= \inf_{DCratio\le\lambda}\langle\nu_{R_c}(x_i,x_j),1-\nu_{R_c}(x_i,x_j)\rangle\wedge\inf_{DCratio>\lambda}\langle 1,0\rangle\\
&= \inf_{DCratio\le\lambda}\langle\nu_{R_c}(x_i,x_j),1-\nu_{R_c}(x_i,x_j)\rangle.
\end{aligned}$$

$$\begin{aligned}\overline{(R_T)}^{DC} D_i(x_i) &= \sup_{x_j \in U} T(R_c(x_i, x_j), DCratio(x_j)) \\ &= \sup_{DCratio \le \lambda} T(\langle \mu_{R_c}(x_i, x_j), \nu_{R_c}(x_i, x_j)\rangle, \langle 1, 0\rangle) \vee \sup_{DCratio > \lambda} T(\langle \mu_{R_c}(x_i, x_j), \nu_{R_c}(x_i, x_j)\rangle, \langle 0, 1\rangle) \\ &= \sup_{DCratio \le \lambda} \langle \max(0, \mu_{R_c}(x_i, x_j) + 1 - 1), \min(1, \nu_{R_c}(x_i, x_j) + 0)\rangle \\ &\vee \sup_{DCratio > \lambda} \langle \max(0, \mu_{R_c}(x_i, x_j) + 0 - 1), \min(1, \nu_{R_c}(x_i, x_j) + 1)\rangle \\ &= \sup_{DCratio \le \lambda} \langle \mu_{R_c}(x_i, x_j), \nu_{R_c}(x_i, x_j)\rangle \vee \sup_{DCratio > \lambda} \langle 0, 1\rangle \\ &= \sup_{DCratio \le \lambda} \langle \mu_{R_c}(x_i, x_j), \nu_{R_c}(x_i, x_j)\rangle,\end{aligned}$$

□

where N is an IF negation operator as defined in Property 3.1. Tuple $\langle \underline{(R_I)}^{DC} D_i(x_i), \overline{(R_T)}^{DC} D_i(x_i)\rangle$ is called as IF rough set.

4.2 Different classes ratio based IF rough set assisted attribute selection

For attribute reduction of an IF decision system, we use degree of dependency approach [42]. Similar to usual fuzzy rough set approach, the intuitionistic fuzzy positive region can be formulated as

$$(POS_C{}^{DC}(D))(x) = \sup_{X \in U/D} \underline{(R_I)}^{DC} D(x), \tag{8}$$

where U/D is defined as a collection of sets of objects having same decision values. Dependency degree of decision attribute D over subset C of conditional attributes can be given by

$$\Gamma_C{}^{DC}(D) = \frac{|(POS_C{}^{DC}(D))(x)|}{|U|}. \tag{9}$$

A subset R of the conditional attribute set C is called reduct set of a decision system if

$$\Gamma_R{}^{DC}(D) = \Gamma_C{}^{DC}(D), \tag{10}$$

$$\Gamma_{R-\{c_i\}}{}^{DC}(D) < \Gamma_R{}^{DC}(D), \forall c_i \in C.$$

4.3 Algorithm for finding reduct by using different classes ratio based IFRS model

We start with a null set and add those attributes, which gives a higher value for the degree of dependency of decision attribute over a set of conditional attributes. This process will continue until it gains highest possible value for an IF decision system (degree of dependency will be 1 in case of consistent system). The main advantage of proposed algorithm is that it provides a close-to-minimal reduct of an IF decision system without exhaustively calculating degree of dependency values for all possible subsets of conditional attributes. The algorithm can be given as follows:

4.4 Illustrative examples

In this subsection, we provide two illustrative examples in order to demonstrate the proposed approach.

Example 4.6. *An IF decision system as given in Table 1* [23] *is taken and* `DC_ratio` *based approach is applied to calculate the reduced subset of conditional attributes.*

For Table 1, $U/D = \{D_1, D_2\}$, where,$D_1 = \{x_1, x_4\}, D_2 = \{x_2, x_3\}$.

Distance [41] between two objects with respect to conditional attribute c is computed using following formula and given in Table 3.

$$d_c(x_i, x_j) = \sqrt{(\mu_c(x_i) - \mu_c(x_j))^2 + (\nu_c(x_i) - \nu_c(x_j))^2 + (\pi_c(x_i) - \pi_c(x_j))^2}. \tag{11}$$

Algorithm 1: Attribute selection by `DC_ratio` based IFRS model

Input : $IS = (U, C \cup D, V, h)$: an IF decision system, $D_i \in U/D, \Gamma_{c_i}^{DC}(D)$: degree of dependency, $i \in \{1, 2, \ldots, r\}, j \in \{1, 2, \ldots, |U|\}$

Output: Reduct R of decision system

Initialize $R = \phi; \Gamma_{best}^{DC}(D) = 0$;
while $\Gamma_{best}^{DC}(D) \neq \Gamma_{prev}^{DC}(D)$ **do**
Compute `DC_ratio` of each target sample $x_j \in U$
Compute the IF lower approximations $\underline{(R_I)}^{DC} D_i(x_j), \forall x_j \in D_i$
Compute positive region $(POS_C{}^{DC}(D))(x_j), \forall x_j \in U$
Compute degree of dependency $\Gamma_{c_i}^{DC}(D)$
Find R with maximum value of $\Gamma_{c_i}^{DC}(D)$
End while
Return R
End

Table 3: Distance between objects x_i and $x_j, i, j \in \{1, 2, 3, 4\}$ with respect to conditional attribute c_1

$d_{c_1}(x_i, x_j)$				$d_{c_2}(x_i, x_j)$				$d_{c_3}(x_i, x_j)$				$d_{c_4}(x_i, x_j)$			
0	0.141	0.374	0.141	0	0.283	0.141	0.374	0	0.245	0.283	0.424	0	0.374	0.509	0.565
0.141	0	0.283	0.245	0.283	0	0.245	0.141	0.245	0	0.509	0.648	0.374	0	0.141	0.245
0.374	0.283	0	0.374	0.141	0.245	0	0.374	0.283	0.509	0	0.374	0.509	0.141	0	0.141
0.141	0.245	0.374	0	0.374	0.141	0.374	0	0.424	0.648	0.374	0	0.565	0.245	0.141	0

Considering x_1, x_2, x_3 and x_4 as target samples one by one, we calculate `DC_ratio` with respect to target class of each conditional attributes. An IF similarity relation $\langle \mu_R(x_i, x_j), \nu_R(x_i, x_j) \rangle$ is calculated using Def. 3.2.

Taking $l = 0.2$. For attribute c_1, we calculate `DC_ratio` of each object by Eq. (1),

`DC_ratio` $(x_1) = 1/3 = 0.33$, `DC_ratio`$(x_2) = 1/2 = 0.5$, `DC_ratio`$(x_3) = 0/1 = 0$, `DC_ratio`$(x_4) = 0/2 = 0$

Choosing $\lambda = 0.4$, we calculate lower approximations of decision classes with the help of `DC_ratio` of objects using Eq. (6),

$$\begin{aligned}
\underline{(R_I)}^{DC} D_1(x_1) &= \inf_{DCratio(u) \leq \lambda} \langle \nu_R(x_1, u), 1 - \nu_R(x_1, u) \rangle = \langle \nu_R(x_1, x_3), 1 - \nu_R(x_1, x_3) \rangle =< 0.1, 0.9 >, \\
\underline{(R_I)}^{DC} D_1(x_4) &= \langle \nu_R(x_4, x_3), 1 - \nu_R(x_4, x_3) \rangle =< 0.1, 0.9 >, \\
\underline{(R_I)}^{DC} D_2(x_2) &= \inf \left[\langle \nu_R(x_2, x_1), 1 - \nu_R(x_2, x_1) \rangle, \langle \nu_R(x_2, x_4), 1 - \nu_R(x_2, x_4) \rangle \right] \\
&= \inf \left[\langle 0.02, 0.98 \rangle, \langle < 0.1, 0.9 > \rangle \right] = \langle 0.02, 0.98 \rangle, \\
\underline{(R_I)}^{DC} D_2(x_3) &= \inf \left[\langle \nu_R(x_3, x_1), 1 - \nu_R(x_3, x_1) \rangle, \langle \nu_R(x_3, x_4), 1 - \nu_R(x_3, x_4) \rangle \right] =< 0.1, 0.9 > .
\end{aligned}$$

The positive region of the decision attribute with respect conditional attribute is calculated by Eq. (7),

$$\begin{aligned}
POS_{c_1}{}^{DC}(D)(x_1) &= \sup \underline{(R_I)}^{DC} D_1(x_1) =< 0.1, 0.9 >, \\
POS_{c_1}{}^{DC}(D)(x_2) &= \sup \underline{(R_I)}^{DC} D_2(x_2) =< 0.02, 0.98 >, \\
POS_{c_1}{}^{DC}(D)(x_3) &= \sup \underline{(R_I)}^{DC} D_2(x_3) =< 0.1, 0.9 >, \\
POS_{c_1}{}^{DC}(D)(x_4) &= \sup \underline{(R_I)}^{DC} D_1(x_4) =< 0.1, 0.9 > .
\end{aligned}$$

The degree of dependency of decision attribute with respect to conditional attribute is evaluated by Eq. (8),

$$\Gamma_{c_1}{}^{DC}(D) = \frac{|(POS_C{}^{DC}(D))(x)|}{|U|} = \frac{0.01 + 0.02 + 0.01 + 0.01}{4} = 0.08.$$

Similarly, ${\Gamma_{c_2}}^{DC}(D) = 0, {\Gamma_{c_3}}^{DC}(D) = 0.1, {\Gamma_{c_4}}^{DC}(D) = 0.07$.

The degree of dependency of conditional attribute $\{c_3\}$ is highest. Therefore, $\{c_3\}$ is the first candidate of the reduct set. We add other conditional attributes to $\{c_3\}$ one by one and calculate the degree of dependencies of paired attributes as

$${\Gamma_{\{c_1,c_3\}}}^{DC}(D) = 0.55, {\Gamma_{\{c_2,c_3\}}}^{DC}(D) = 0, {\Gamma_{\{c_3,c_4\}}}^{DC}(D) = 0.085.$$

This time $\{c_1, c_3\}$ has the highest degree of dependency among others. Continuing the process again, we obtain

$${\Gamma_{\{c_1,c_2,c_3\}}}^{DC}(D) = 0.51, {\Gamma_{\{c_1,c_3,c_4\}}}^{DC}(D) = 0.05.$$

Since, there is no increment in degree of dependency after adding attributes $\{c_2\}$ and $\{c_4\}$ to $\{c_1, c_3\}$. The process terminates here and $\{c_1, c_3\}$ is the required reduct set of the given dataset.

Example 4.7. *An IF decision system is given in Table 2* [23] *with 10 objects* $\{x_1, x_2, \ldots, x_{10}\}$*, 5 conditional attributes* $\{c_1, c_2, c_3, c_4, c_5\}$ *and one decision attribute* $\{d\}$.

Here, $U/D = \{D_1, D_2, D_3\}$,where,$D_1 = \{x_1, x_3, x_4, x_6\}, D_2 = \{x_2, x_5, x_7, x_8\}, D_3 = \{x_9, x_{10}\}$.

Take $\lambda = 0.4$. Using `DC_ratio` based approach as in Example 4.4.1, we calculate reduct set for this decision system. For attribute c_1, `DC_ratio` of each object is calculated,

`DC_ratio` $(x_1) = 0.33$, `DC_ratio`$(x_2) = 0.67$, `DC_ratio`$(x_3) = 0.5$, `DC_ratio`$(x_4) = 0.33$,
`DC_ratio` $(x_5) = 0.4$, `DC_ratio`$(x_6) = 0.25$, `DC_ratio`$(x_7) = 0.4$, `DC_ratio`$(x_8) = 0.5$,
`DC_ratio` $(x_9) = 0.5$, `DC_ratio`$(x_{10}) = 0.6$.

After taking $\lambda = 0.4$, we compute lower approximations of decision classes,

$\underline{(R_I)}^{DC} D_1(x_1) =< 0.16, 0.84 >, \underline{(R_I)}^{DC} D_2(x_2) =< 0.02, 0.98 >, \underline{(R_I)}^{DC} D_1(x_3) =< 0.26, 0.7 >, \underline{(R_I)}^{DC} D_1(x_4) =< 0.36, 0.64 >, \underline{(R_I)}^{DC} D_2(x_5) =< 0.16, 0.84 >, \underline{(R_I)}^{DC} D_1(x_6) =< 0.18, 0.82 >, \underline{(R_I)}^{DC} D_2(x_7) =< 0.32, 0.68 >, \underline{(R_I)}^{DC} D_2(x_8) =< 0.4, 0.6 >, \underline{(R_I)}^{DC} D_3(x_9) =< 0.08, 0.92 >, \underline{(R_I)}^{DC} D_3(x_10) =< 0, 1 >$.

With the help of lower approximation, we calculate the degree of dependency for attribute c_1,

$${\Gamma_{c_1}}^{DC}(D) = \frac{0.16+0.02+0.26+0.36+0.16+0.18+0.32+0.4+0.08+0}{10} = 0.19,$$

Similarly, ${\Gamma_{c_2}}^{DC}(D) = 0.28, {\Gamma_{c_3}}^{DC}(D) = 0.56, {\Gamma_{c_4}}^{DC}(D) = 0, {\Gamma_{c_5}}^{DC}(D) = 0.58$.

The degree of dependency of attribute $\{c_5\}$ is highest. We select $\{c_5\}$ as the first member of reduct set and calculate Γ of
${\Gamma_{\{c_1,c_5\}}}^{DC}(D) = 0.38, {\Gamma_{\{c_2,c_5\}}}^{DC}(D) = 0.61, {\Gamma_{\{c_3,c_5\}}}^{DC}(D) = 0.26, {\Gamma_{\{c_4,c_5\}}}^{DC}(D) = 0.33$.

Since $\{c_2, c_5\}$ has greater value than $\{c_5\}$ and other paired attributes, we select $\{c_2, c_5\}$ and continue the process,
${\Gamma_{\{c_1,c_2,c_5\}}}^{DC}(D) = 0.44, {\Gamma_{\{c_2,c_3,c_5\}}}^{DC}(D) = 0.32, {\Gamma_{\{c_2,c_4,c_5\}}}^{DC}(D) = 0.75$.

Now, we get $\{c_2, c_4, c_5\}$ has greater value of Γ among others,
${\Gamma_{\{c_1,c_2,c_4,c_5\}}}^{DC}(D) = 0.75, {\Gamma_{\{c_2,c_3,c_4,c_5\}}}^{DC}(D) = 0.74$.

Finally, above two pairs has value of degree of dependency less than or equal to $\{c_2, c_4, c_5\}$. Hence, $\{c_2, c_4, c_5\}$ is the reduct set of Table 2.

We performed the proposed approach on two example datasets (IF decision systems) for attribute selection in order to find close-to-minimal reduct set. In Example 4.4.1, conditional attribute set $\{c_1, c_2, c_3, c_4\}$ reduces to $\{c_1, c_3\}$ and in Example 4.4.2, conditional attribute set $\{c_1, c_2, c_3, c_4, c_5\}$ reduces to $\{c_2, c_4, c_5\}$.

5 Laplace summmation operator based approach for attribute selection of an IF decision system

In this section, we propose LS operator based approach for attribute selection of an IF decision system.

Definition 5.1. [19] *In the probability theory, the density function of random variable X is defined as*

$$f(x) = \frac{1}{2\lambda} e^{\frac{-|x-\mu|}{\lambda}}, \tag{12}$$

where, distribution function follows the Laplace distribution with the parameter λ, μ, or $X \sim L_a(\mu, \lambda)$, and λ, μ are defined as constant location parameter and constant scale parameter respectively along with a constraint $\lambda > 0$.

5.1 Laplace summation operator based IF rough set

In the current study, we use the concept of Laplace weighted summation operator W_i as proposed by Han et al.[19]. For n number of aggregated objects,

$$W_i = \frac{1}{2b_n} e^{\frac{-|i-\theta_n|}{b_n}}, i = 1, 2, 3, \ldots, n, \tag{13}$$

b_n is the scale of Laplace distribution and θ_n is the mean of the number of objects defined as

$$\theta_n = \frac{1}{n}.\frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Now, standard deviation σ for Laplace distribution is calculated as

$$\sigma = \sqrt{2}.b_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(i-\theta_n)^2}. \tag{14}$$

For simplicity, we can take normalized weighted operator as

$$\omega_i = \frac{\frac{1}{2b_n} e^{\frac{-|i-\theta_n|}{b_n}}}{\sum_{i=1}^{n} \frac{1}{2b_n} e^{\frac{-|i-\theta_n|}{b_n}}}, i = 1, 2, 3, \ldots, n, \tag{15}$$

Definition 5.2. *We define LS operator based IF lower and upper approximations as follows:*

$$\begin{aligned}\underline{(R_I)}^{LS} D_i(x_i) &= LS_{x_j \in U} I(R_c(x_i, x_j), \langle \mu_{D_i}(x_j), \nu_{D_i}(x_j)\rangle) \\ &= \sum_{j=1}^{n} \omega_i I(\langle \mu_{R_c}(x_i, x_j), \nu_{R_c}(x_i, x_j)\rangle, \langle \mu_{D_i}(x_j), \nu_{D_i}(x_j)\rangle).\end{aligned} \tag{16}$$

$$\begin{aligned}\overline{(R_T)}^{LS} D_i(x_i) &= LS_{x_j \in U} T(R_c(x_i, x_j), \langle \mu_{D_i}(x_j), \nu_{D_i}(x_j)\rangle) \\ &= \sum_{j=1}^{n} \omega_i T(\langle \mu_{R_c}(x_i, x_j), \nu_{R_c}(x_i, x_j)\rangle, \langle \mu_{D_i}(x_j), \nu_{D_i}(x_j)\rangle).\end{aligned} \tag{17}$$

Here I, T and LS are IF implicator, IF t-norm and an aggregate function called Laplace weighted summation respectively, and $R_c(x_i, x_j)$ is an IF similarity relation defined in Def. 3.2, ω_i is named weight vector of LS operator. The pair $\langle \underline{(R_I)}^{LS} D_i(x_i), \overline{(R_T)}^{LS} D_i(x_i)\rangle$ is called an IF rough set.

Now, using Eq. (4) & Eq. (5), Eq.(10) & Eq.(11) can be re-written as

$$\underline{(R_I)}^{LS} D_i(x_i) = \sum_{j=1}^{n} \omega_i \langle \min(1, 1 + \mu_{D_i}(x_j) - \mu_{R_c}(x_i, x_j), 1 + \nu_{R_c}(x_i, x_j) - \nu_{D_i}(x_j)), \max(0, \nu_{D_i}(x_j) - \nu_{R_c}(x_i, x_j))\rangle. \tag{18}$$

$$\overline{(R_T)}^{LS} D_i(x_i) = \sum_{j=1}^{n} \omega_i \langle \max(0, \mu_{R_c}(x_i, x_j) + \mu_{D_i}(x_j) - 1), \min(1, \nu_{R_c}(x_i, x_j) + \nu_{D_i}(x_j))\rangle. \tag{19}$$

5.2 Laplace summation operator based IF rough set assisted attribute reduction

Similar to usual fuzzy rough set approach, the intuitionistic fuzzy positive region [42] can be formulated as

$$({POS_C}^{LS}(D))(x) = \sup_{X \in U/D} \underline{(R_I)}^{LS} D(x). \tag{20}$$

LS intuitionistic fuzzy-rough dependency degree of decision attribute D over subset P of conditional attributes can be given by

$$ {\Gamma_C}^{LS}(D) = \frac{|({POS_C}^{LS}(D))(x)|}{|U|}, \tag{21}$$

where, $|.|$ in numerator is an IF cardinality whereas $|.|$ in denominator is crisp cardinality.

$$ {\Gamma_R}^{LS}(D) = {\Gamma_C}^{LS}(D). \tag{22}$$

$$ {\Gamma_{R-\{c_i\}}}^{LS}(D) < {\Gamma_R}^{LS}(D), \forall c_i \in C.$$

The process to find close-to-minimal reduct set is same as in Section 4.2.

5.3 Algorithm for finding reduct by using Laplace Summation operator based IFRS model

Now, we present a greedy algorithm in order to apply proposed approach on an IF information/decision system.

Algorithm 2: Attribute selection by Laplace Summation operator based IFRS model

Input : C: collection of conditional attributes; U: collection of samples, $R =< \mu_{R_C}(x_i, x_j), \nu_{R_C}(x_i, x_j) >$: intuitionistic fuzzy similarity relation, $\Gamma_C^{LS}(D)$: degree of dependency, $i \in \{1, 2, \ldots, r\}, j \in \{1, 2, \ldots, |U|\}$.

Output: R: selected a feature subset

Initialize $R = \phi; \Gamma_{best}^{LS}(D) = 0$;
while $\Gamma_{best}^{LS}(D) \neq \Gamma_{prev}^{LS}(D)$ **do**
Compute Laplace weights ω_i
Compute the IF lower approximations $\underline{(R_I)}^{LS} D_i(x_j)$ and $({POS_C}^{LS}(D))(x_j), \forall x_j \in U, D_i \in U/D$
Compute degree of dependency $\Gamma_{c_i}^{LS}(D)$
Find R with maximum value of $\Gamma_{c_i}^{LS}(D)$
End while
Return R
End

5.4 Illustrative examples

In this subsection, two demonstrative examples are presented for better understanding of LS operator based IF rough set model.

Example 5.3. *Taking same IF decision system tabulated in Table 1* [23]*, we apply LS operator based approach for finding reduct set as follows:*

First, we compute the scale of Laplace distribution using Eq. (14),
$b_4 = \sqrt{\frac{1}{2\times 4} \sum_{i=1}^{4} (i - \frac{4+1}{2})^2} = \frac{5}{8} = 0.79.$

Weights of LS operator is calculated by Eq. (13) and Eq. (16),
$W_1 = \frac{1}{2\times 0.79} e^{\frac{-|1-\frac{5}{2}|}{0.79}} = 0.094, W_2 = 0.433, W_3 = 0.433, W_3 = 0.094,$

and $\omega_1 = \frac{W_1}{\sum_{i=1}^{n} W_1} = \frac{0.094}{1.057} = 0.09, \omega_2 = 0.41, \omega_3 = 0.41, \omega_4 = 0.09.$

For conditional attribute c_1, the degree of dependency of decision attribute over c_1 is calculated as follows:
The lower approximations of decision class D_1 over each object are computed using Eq. (18)

$$\begin{aligned}\underline{(R_I)}^{LS}D_1(x_1) &= \sum_{i=1}^{n}\omega_i I(\langle\mu_{R_c}(x_1,x_i),\nu_{R_c}(x_1,x_i)\rangle\rangle,\langle\mu_{D_1}(x_1),\nu_{D_1}(x_1)\rangle)\\
&= 0.09I(<1,0>,<1,0>)+0.41I(<0.98,0.02>,<0,1>)+0.41I(0.82,0.1>,<0,1>)\\
&+0.09I(<0.92,0.04>,<1,0>)\\
&= 0.09<min(1,1,1),max(0,0)>+0.41<min(1,0.02,0.02),max(0,0.98)>\\
&+0.41<min(1,0.18,0.1),max(0,0.9)>+0.09<min(1,1.08,1.04),max(0,-0.04)>\\
&= 0.09<1,0>+0.41<0.02,0.98>+0.41<0.1,0.9>+0.09<1,0>\\
&=<1,0>+<0.0083,0.9917>+<0.0423,0.9577>+<1,0>=<0.502,0.949>.\end{aligned}$$

Similarly, $\underline{(R_I)}^{LS}D_1(x_2)=<0.033,0.966>, \underline{(R_I)}^{LS}D_1(x_3)=<0.033,0.966>, \underline{(R_I)}^{LS}D_1(x_4)=<0.084,0.915>$.
And lower approximations of decision class D_2 over each object are calculated,

$$\begin{aligned}\underline{(R_I)}^{LS}D_2(x_1) &= \sum_{i=1}^{n}\omega_i I(\langle\mu_{R_c}(x_1,x_i),\nu_{R_c}(x_1,x_i)\rangle\rangle,\langle\mu_{D_2}(x_1),\nu_{D_2}(x_1)\rangle)\\
&= 0.09I(<1,0>,<0,1>)+0.41I(<0.98,0.02>,<1,0>)+0.41I(0.82,0.1>,<1,0>)\\
&+0.09I(<0.92,0.04>,<0,1>)\\
&= 0.09<min(1,0,0),max(0,1)>+0.41<min(1,1.02,1.02),max(0,0.02)>\\
&+0.41<min(1,1.18,1.18),max(0,-0.1)>+0.09<min(1,0.08,0.04),max(0,0.96)>\\
&= 0.09<0,1>+0.41<1,0>+0.41<1,0>+0.09<0.04,0.96>\\
&=<0,1>+<1,0>+<1,0>+<0.0064,0.9936>=<0.006,0.993>.\end{aligned}$$

Similarly, $\underline{(R_I)}^{LS}D_2(x_2)=<0.033,0.966>, \underline{(R_I)}^{LS}D_2(x_3)=<0.033,0.966>, \underline{(R_I)}^{LS}D_2(x_4)=<0.084,0.915>$.
Taking supremum of lower approximations, we get the positive region as in Eq. (20),
$POS_{c_1}{}^{LS}(D)(x_1)=\sup[\underline{(R_I)}^{LS}D_1(x_1),\underline{(R_I)}^{LS}D_2(x_1)]=<max(0.502,0.006),min(0.949,0.993)>=<0.502,0.949>$.
and $POS_{c_1}{}^{LS}(D)(x_2)==<0.033,0.966>, POS_{c_1}{}^{LS}(D)(x_3)=<0.033,0.966>, POS_{c_1}{}^{LS}(D)(x_4)=<0.084,0.915>$

The degree of dependency of decision attribute D with respect to conditional attribute c_1 is calculated by (21),

$$\Gamma_{c_1}{}^{LS}(D)=\frac{|(POS_C{}^{LS}(D))(x)|}{|U|}=\frac{0.05+0.033+0.033+0.84}{4}=0.05.$$

Similarly, $\Gamma_{c_2}{}^{LS}(D)=0.06, \Gamma_{c_3}{}^{LS}(D)=0.22, \Gamma_{c_4}{}^{LS}(D)=0.05$.

Attribute $\{c_3\}$ consists the highest value of degree of dependency among others. We coupled other conditional attributes to $\{c_3\}$ and calculate their degree of dependencies. So, we obtain

$\Gamma_{\{c_1,c_3\}}{}^{LS}(D)=0.22, \Gamma_{\{c_2,c_3\}}{}^{LS}(D)=0.24, \Gamma_{\{c_3,c_4\}}{}^{LS}(D)=0.23$.
Now, $\{c_2,c_3\}$ obtains the highest value of degree of dependency. Continuing the procedure as above, we get,
$\Gamma_{\{c_1,c_2,c_3\}}{}^{LS}(D)=0.24, \Gamma_{\{c_2,c_3,c_4\}}{}^{DC}(D)=0.24$.

The process of getting reduct set stops as there is no increament in Γ after adding attributes c_1 and c_4 to $\{c_2,c_3\}$. Hence, $\{c_2,c_3\}$ is the required reduct set of the IS presented in Table 1.

Example 5.4. *The IF decision system is given in Table 2* [23]. *We use LS operator based approach for finding the reduct set of the dataset.*

The scale of Laplace distribution is calculated
$b_{10}=\sqrt{\frac{1}{2\times10}\sum_{i=1}^{10}(i-\frac{11}{2}^2}=\frac{51}{20}=1.6$.
And weights are computed
$\omega_1=0.02=\omega_{10}, \omega_2=0.04=\omega_9, \omega_3=0.07=\omega_8, \omega_4=0.04=\omega_7, \omega_5=0.04=\omega_6$.
$U/D=\{D_1,D_2,D_3\}$,where,$D_1=\{x_1,x_3,x_4,x_6\}, D_2=\{x_2,x_5,x_7,x_8\}, D_3=\{x_9,x_{10}\}$.

For object x_1, conditional attribute c_1, lower approximations $\underline{(R_I)}^{LS}D_i(x_j)$, positive region and degree of dependency is given by,

$$\underline{(R_I)}^{LS}D_1(x_1) =< 0.110, 0.895 >, \underline{(R_I)}^{LS}D_2(x_1) =< 0.131, 0.872 >, \underline{(R_I)}^{DC}D_3(x_1) =< 0.227, 0.781 >$$

$$POS_{c_1}{}^{LS}(D)(x_1) = \sup [\underline{(R_I)}^{LS}D_1(x_1), \underline{(R_I)}^{LS}D_2(x_1), \underline{(R_I)}^{LS}D_3(x_1)]$$
$$=< max(0.110, 0.131, 0.227), min(0.895, 0.872, 0.781) >=< 0.227, 0.781 > .$$

$$\Gamma_{c_1}{}^{LS}(D) = \frac{|(POS_C{}^{LS}(D))(x)|}{|U|} = 0.15.$$

Similarly, $\Gamma_{c_2}{}^{LS}(D) = 0.11, \Gamma_{c_3}{}^{LS}(D) = 0.05, \Gamma_{c_4}{}^{LS}(D) = 0.16, \Gamma_{c_5}{}^{LS}(D) = 0.09$.

Conditional attribute $\{c_4\}$ has the highest value of degree of dependency. Adding other attributes to $\{c_4\}$,

$\Gamma_{\{c_1,c_4\}}{}^{LS}(D) = 0.17, \Gamma_{\{c_2,c_4\}}{}^{LS}(D) = 0.14, \Gamma_{\{c_3,c_4\}}{}^{LS}(D) = 0.12, \Gamma_{\{c_4,c_5\}}{}^{LS}(D) = 0.14$.
We get $\Gamma_{\{c_1,c_4\}}(D)$ has greater value of Γ, So we choose $\{c_1, c_4\}$ for further processing,

$\Gamma_{\{c_1,c_2,c_4\}}{}^{LS}(D) = 0.16, \Gamma_{\{c_1,c_3,c_4\}}{}^{LS}(D) = 0.12, \Gamma_{\{c_1,c_4,c_5\}}{}^{LS}(D) = 0.12$.
Above pairs have value of Γ less than $\{c_1, c_4\}$. Therefore, $\{c_1, c_4\}$ is the reduct set of Table 2.

On applying LS operator based approach, the reduct sets $\{c_2, c_3\}$ of Table 1, and $\{c_1, c_4\}$ of Table 2 are obtained. Therefore, proposed approach is capable in finding close-to-minimal reduct set for any IF information/decision system.

Table 4: Reduct sets with `DC_ratio` and LS operator based approaches

	Conditional attribute set	**Reduct set with `DC_ratio`**	**Reduct set with LS operator**
Table 1	$\{c_1, c_2, c_3, c_4\}$	$\{c_1, c_3\}$	$\{c_2, c_3\}$
Table 2	$\{c_1, c_2, c_3, c_4, c_5\}$	$\{c_2, c_4, c_5\}$	$\{c_1, c_4\}$

6 Experimental analysis

In this section, we have implemented our proposed feature selection techniques in Jupyter Notebook 6.1.4 using Python code on hardware platform having configuration of Intel(R) Core(TM) i3-5005U CPU @ 2.00 GHz with 8.00 GB RAM. Performances of the learning algorithms are evaluated in WEKA 3.8 using 10-fold cross validation. We have taken datasets from UCI repository [4]. Experiments have been conducted and the characteristics of the datasets along with size of the reduct set produced by `DC_ratio` based attribute selection as well as LS operator based attribute selection are recorded in Table 5. Further, we evaluated the performances of the learning algorithms based on classification accuracies with standard deviation in terms of percentage. We have utilized three well known learning algorithms namely: SMO [36], PART [5], and Random Forest [18] to compute the evaluation metrics using 10- fold cross-validation over original datasets and reduced datasets produced by `DC_ratio` based attribute selection as well as LS operator based attribute selection. These results are recorded in Table 6. From Table 5, it can be observed that our proposed feature selection techniques are producing reduced datasets by eliminating irrelevant and redundant features. From Table 6, we can easily conclude that reduced datasets provided by our proposed approaches are producing better results when compared to original datasets in form of classification accuracies along with standard deviation for all the learning algorithms.

Table 5: Reduct sets with `DC_ratio` and LS operator based approaches

Dataset	**Instances**	**Attributes**	**Reduct size**	
			`DC_Ratio` based Attribute Selection	**LS Operator based Attribute Selection**
Wine	178	13	3	7
Glass	214	9	3	4
Sonar	208	60	13	17
Tae	151	5	3	2
Wpbc	198	33	25	21
Heart	267	13	10	8

Table 6: Comparison of Classification Accuracies for original datasets and reduced datasets by DC_ratio based attribute selection and LS operator based attribute selection approach using 10-fold cross validation

Dataset/ Learning Algorithm	Original			DC ratio based Attribute Selection			LS Operator based Attribute Selection		
	SMO	PART	Random forest	SMO	PART	Random forest	SMO	PART	Random forest
Wine	98.76±2.73	92.08±6.28	97.86±3.17	99.47±1.97	95.26±4.05	98.45±1.21	96.78±4.13	92.13±5.88	98.34±1.85
Glass	78.11±21.68	80.41±14.59	88.47±11.56	86.99±8.03	83.20±9.09	95.60±1.17	79.37±7.80	85.64±8.40	88.41±9.08
Sonar	76.60±8.27	77.40±9.43	84.00±8.56	77.74±7.33	79.65±10.81	89.54±6.09	79.61±9.33	83.04±9.85	87.57±8.73
Tae	80.79±2.00	82.13±6.64	86.74±7.00	80.79±2.00	83.22±6.12	84.94±7.33	80.79±2.00	84.68±6.39	83.89±7.47
Wbpe	76.75±2.96	73.64±8.62	80.58±4.83	76.29±2.00	74.96±5.46	76.79±4.92	76.29±2.00	75.12±4.21	77.29±5.38
Heart	83.43±6.47	77.59±7.48	82.09±7.04	85.76±7.33	83.47±8.48	86.09±7.28	88.12±1.39	82.93±10.23	85.88±9.27

7 Conclusions

In this paper, we have proposed two robust approaches for attribute selection of an IF decision system: `DC_ratio` based approach and Laplace Distribution based approach. First approach is based on the ratio of samples belonging to different classes in the neighbors of a given sample and second approach is based on Laplace weighted summation operator. We propose different lower and upper approximations for both approaches and applied degree of dependency based method for attribute selection. Algorithms for both approaches have been presented for better understanding. Proposed approaches have been applied on two illustrative examples for attribute reduction of IF decision systems and results have been recorded in Table 4. Moreover, we have implemented both proposed techniques on some real-valued datasets. From the experimental results, it can be observed that both the techniques are capable of removing irrelevant and redundant features as the performances of learning algorithms are improving for the reduced datasets produced by both the techniques.

In future, we intend to generalize type-2 IF rough set model and want to apply it for feature selection. Moreover, we wish to investigate a generalized method for conversion of fuzzy information system into intuitionistic fuzzy information system so that these approaches can be implemented for real valued data sets also.

Acknowledgement

Corresponding author would like to thank the Research Foundation-CSIR for funding her research and all authors are grateful to Ms. Chhavi Jain, Department of Mathematical Sciences, IIT (BHU), Varanasi, UP -221005 for her contribution in the review process of the manuscript.

References

[1] K. T. Atanassov, *Intuitionistic fuzzy sets*, Fuzzy Sets and Systems, **20**(1) (1986), 87-96.

[2] K. T. Atanassov, *Intuitionistic fuzzy sets*, In Intuitionistic Fuzzy Sets, Physica, Heidelberg, (1999), 1-137.

[3] M. J. Benıtez-Caballero, J. Medina, E. Ramırez-Poussa, D. Šiĕzak, *Rough-set-driven approach for attribute reduction in fuzzy formal concept analysis*, Fuzzy Sets and Systems, **391** (2020), 117-138.

[4] C. L. Blake, C. J. Merz, *UCI repository of machine learning databases*, 1998.

[5] L. Breiman, *Bagging predictors*, Machine Learning, **24**(2) (1996), 123-140.

[6] K. Chakrabarty, T. Gedeon, L. Koczy, *Intuitionistic fuzzy rough set*, In Proceedings of 4th Joint Conference on Information Sciences (JCIS), Durham, NC (1998), 211-214.

[7] P. Chen, M. Lin, J. Liu, *Multi-label attribute reduction based on variable precision fuzzy neighborhood rough set*, IEEE Access, **8** (2020), 133565-133576.

[8] H. Chen, H. C. Yang, *One new algorithm for intuitiontistic fuzzy-rough attribute reduction*, Journal of Chinese Computer Systems, **32**(3) (2011), 506-510.

[9] C. Cornelis, M. De Cock, E. E. Kerre, *Intuitionistic fuzzy rough sets: At the crossroads of imperfect knowledge*, Expert Systems, **20**(5) (2003), 260-270.

[10] C. Cornelis, M. De Cock, A. M. Radzikowska, *Vaguely quantified rough sets*, In International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing (pp. 87-94). Springer, Berlin, Heidelberg, (2007), 87-94.

[11] C. Cornelis, N. Verbiest, R. Jensen, *Ordered weighted average based fuzzy rough sets*, In International Conference on Rough Sets and Knowledge Technology, Springer, Berlin, Heidelberg, (2010), 78-85.

[12] D. Čoker, *Fuzzy rough sets are intuitionistic L-fuzzy sets*, Fuzzy Sets and Systems, **96**(3) (1998), 381-383.

[13] J. Dai, Y. Yan, Z. Li, B. Liao, *Dominance-based fuzzy rough set approach for incomplete interval-valued data*, Journal of Intelligent and Fuzzy Systems, **34**(1) (2018), 423-436.

[14] D. Dubois, H. Prade, *Rough fuzzy sets and fuzzy rough sets*, International Journal of General System, **17** (1990), 191-209.

[15] D. Dubois, H. Prade, *Putting rough sets and fuzzy sets together*, In Intelligent Decision Support, **177**(1) (1992), 203-232.

[16] H. Esmail, J. Maryam, L. Habibolla, *Rough set theory for the intuitionistic fuzzy information*, Systems International Journal of Modern Mathematical Sciences, **6**(3) (2013), 132-143.

[17] T. Feng, J. Mi, *Reductions of intuitionistic fuzzy covering systems based on discernibility matrices*, In International Conference on Rough Sets and Knowledge Technology, **34**(5) (2014), 111-120.

[18] E. Frank, I. H. Witten, *Generating accurate rule sets without global optimization*, 1998, Doi: 10.1.1.143.8073.

[19] X. Han, Y. Qu, A. Deng, *A Laplace distribution-based fuzzy-rough feature selection algorithm*, In 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI), IEEE, **96**(3) (2018), 776-781.

[20] Q. Hu, S. An, D. Yu, *Soft fuzzy rough sets for robust feature evaluation and selection*, Information Sciences, **180**(22) (2010), 4384-4400.

[21] Q. Hu, D. Yu, J. Liu, C. Wu, *Neighborhood rough set based heterogeneous feature subset selection*, Information Sciences, **178**(18) (2008), 3577-3594.

[22] Q. Hu, L. Zhang, S. An, D. Zhang, D. Yu, *On robust fuzzy rough set models*, IEEE Transactions on Fuzzy Systems, **20**(4) (2012), 636-651.

[23] S. Y. Huang, *Intelligent decision support: Handbook of applications and advances of the rough sets theory*, Springer Science and Business Media, **11** (1992).

[24] B. Huang, H. X. Li, D. K. Wei, *Dominance-based rough set model in intuitionistic fuzzy information systems*, Knowledge-Based Systems, **28** (2012), 115-123.

[25] B. Huang, Y. L. Zhuang, H. X. Li, D. K. Wei, *A dominance intuitionistic fuzzy-rough set approach and its applications*, Applied Mathematical Modelling, **37**(12-13) (2013), 7128-7141.

[26] S. P. Jena, S. K. Ghosh, B. K. Tripathy, *Intuitionistic fuzzy rough sets*, Notes on Intuitionistic Fuzzy Sets, **8**(1) (2002), 1-18.

[27] R. Jensen, Q. Shen, *Fuzzy-rough attribute reduction with application to web categorization*, Fuzzy Sets and Systems, **141**(3) (2004), 469-485.

[28] Y. Li, S. Wu, Y. Lin, J. Liu, *Different classes ratio fuzzy rough set based robust feature selection*, Knowledge-Based Systems, **120** (2017), 74-86.

[29] Y. Liu, Y. Lin, *Intuitionistic fuzzy rough set model based on conflict distance and applications*, Applied Soft Computing, **31** (2015), 266-273.

[30] Y. L. Lu, Y. J. Lei, J. X. Hua, *Attribute reduction based on intuitionistic fuzzy rough set*, Control and Decision, **3** (2009), 003.

[31] M. Mohtashami, M. Eftekhari, *A hybrid filter-based feature selection method via hesitant fuzzy and rough sets concepts*, Iranian Journal of Fuzzy Systems, **16**(2) (2019), 165-182.

[32] S. Nanda, S. Majumdar, *Fuzzy rough sets*, Fuzzy Sets and Systems, **45**(2) (1992), 157-160.

[33] Z. Pawlak, *Rough sets*, International Journal of Computer and Information Sciences, **11**(5) (1982), 341-356.

[34] Z. Pawlak, *Rough sets: Theoretical aspects of reasoning about data*, Springer Science and Business Media, **9** (2012).

[35] Z. Pawlak, A. Skowron, *Rudiments of rough sets*, Information Sciences, **177**(1) (2007), 3-27.

[36] J. Platt, *Sequential minimal optimization: A fast algorithm for training support vector machines*, 1998.

[37] J. F. Salido, S. Murakami, *Rough set analysis of a general type of fuzzy data using transitive aggregations of fuzzy similarity relations*, Fuzzy Sets and Systems, **139**(3) (2003), 635-660.

[38] S. K. Samanta, T. K. Mondal, *Intuitionistic fuzzy rough sets and rough intuitionistic fuzzy sets*, Journal of Fuzzy Mathematics, **9**(3) (2001), 561-582.

[39] S. Singh, S. Shreevastava, T. Som, G. Somani, *A fuzzy similarity-based rough set approach for attribute selection in set-valued information systems*, Soft Computing, **24**(6) (2020), 4675-4691.

[40] A. Skowron, J. Stepaniuk, *Tolerance approximation spaces*, Fundamenta Informaticae, **27**(2-3) (1996), 245-253.

[41] E. Szmidt, J. Kacprzyk, *Entropy for intuitionistic fuzzy sets*, Fuzzy Sets and Systems, **118**(3) (2001), 467-477.

[42] A. K. Tiwari, S. Shreevastava, K. K. Shukla, K. Subbiah, *New approaches to intuitionistic fuzzy-rough attribute reduction*, Journal of Intelligent and Fuzzy Systems, **34**(5) (2018), 3385-3394.

[43] E. C. Tsang, D. Chen, D. S. Yeung, X. Z. Wang, J. W. Lee, *Attributes reduction using fuzzy rough sets*, IEEE Transactions on Fuzzy Systems, **16**(5) (2008), 1130-1141.

[44] C. Wang, Y. Huang, M. Shao, X. Fan, *Fuzzy rough set-based attribute reduction using distance measures*, Knowledge-Based Systems, **164** (2019), 205-212.

[45] W. Z. Wu, Y. Leung, M. W. Shao, *Generalized fuzzy rough approximation operators determined by fuzzy implicators*, International Journal of Approximate Reasoning, **54**(9) (2013), 1388-1409.

[46] W. Z. Wu, J. S. Mi, W. X. Zhang, *Generalized fuzzy rough sets*, Information Sciences, **151** (2003), 263-282.

[47] W. Z. Wu, W. X. Zhang, *Constructive and axiomatic approaches of fuzzy approximation operators*, Information Sciences, **159**(3-4) (2004), 233-254.

[48] L. Yong, H. Wenliang, J. Yunliang, Z. Zhiyong, *Quick attribute reduct algorithm for neighborhood rough set model*, Information Sciences, **271** (2014), 65-81.

[49] L. A. Zadeh, *Fuzzy sets*, Information and Control, **8** (1965), 338-353.

[50] A. Zeng, T. Li, D. Liu, J. Zhang, H. Chen, *A fuzzy rough set approach for incremental feature selection on hybrid information systems*, Fuzzy Sets and Systems, **258** (2015), 39-60.

[51] S. Zhao, E. C. Tsang, D. Chen, *The model of fuzzy variable precision rough sets*, IEEE Transactions on Fuzzy Systems, **17**(2) (2009), 451-467.

[52] W. Ziarko, *Variable precision rough set model*, Journal of Computer and System Sciences, **46**(1) (1993), 39-59.