MINISTRY OF EDUCATION AND TRAINING

VIETNAM ACADEMY OF SCIENCE AND TECHNOLOGY

GRADUATE UNIVERSITY OF SCIENCE AND TECHNOLOGY

Pham Viet Anh

ADVANCED MODELS FOR INCREMENTAL ATTRIBUTE REDUCTION BASED ON INTUITIONISTIC FUZZY SETS AND WEIGHTED NEIGHBORHOOD ROUGH SETS

SUMMARY OF DISSERTATION ON COMPUTER

Major: Information system

Code: 9 48 01 04

Ha Noi - 2026

The dissertation is completed at: Graduate University of Science and Technology, Vietnam Academy Science and Technology

Supervisors:

1. Supervisor 1: Associate Professor . Dr. Nguyen Long Giang
2. Supervisor 2: PhD Nguyen Ngoc Thuy

Referee 1:...

Referee 2:...

Referee 3:...

The dissertation is examined by Examination Board of Graduate University of Science and Technology, Vietnam Academy of Science and Technology at……………………….. (time, date……)

The dissertation can be found at:

1. Graduate University of Science and Technology Library
2. National Library of Vietnam

Introduction

1. Significance of the dissertation

As an effective tool, the rough set model has established a solid foundation for the development of attribute reduction algorithms on decision tables [4]. However, when dealing with datasets containing attributes with continuous numerical domains, methods based on this model must undergo a data discretization step. Consequently, this process can significantly affect information preservation and reduce the effectiveness of the obtained reducts. In response to these challenges, several extensions of the rough set model have been developed to directly process original data, among which the most popular and effective extended models have evolved along two main approaches: the neighborhood rough set model and the fuzzy rough set model.

The neighborhood rough set model is one of the research branches extended from rough set theory. Methods based on the neighborhood rough set framework are highly effective in handling numerical or mixed decision tables, as they provide a more comprehensive characterization of objects compared to classical rough set theory. However, these methods mainly focus on the number of objects within an information granule. This implies that all objects in a granule are assigned the same level of importance with respect to a given decision. In practice, however, data are usually distributed heterogeneously, meaning that each object within a neighborhood plays a different role. Based on the combination of classical rough set theory and fuzzy set theory, fuzzy rough set theory was proposed by Dübois and Prade [34] and is regarded as the second major research branch for handling continuous data. However, several studies have indicated that attribute reduction methods based on fuzzy rough sets are less effective when dealing with noisy datasets with low classification accuracy.

In recent years, attribute reduction methods based on the intuitionistic fuzzy rough set model have attracted increasing attention. The main advantage of this model lies in the incorporation of the non-membership function, which effectively adjusts information contributed by noisy objects toward correct classification [52]. As a result, the intuitionistic fuzzy rough set model demonstrates superior classification capability compared to classical fuzzy sets, especially on noisy or low-consistency datasets. Numerous experimental results have confirmed the superior performance of algorithms based on intuitionistic fuzzy rough sets over those based on fuzzy rough sets. Nevertheless, the intuitionistic fuzzy rough set model still suffers from several limitations. First, the inclusion of the non-membership function increases storage requirements and computational time compared to traditional fuzzy rough set approaches. Second, objects with distributions that significantly deviate from the majority of the universe may create noise in the computation.

Moreover, modern data continuously increase and update over time, leading to decision tables of extremely large scale. To address this issue, incremental attribute reduction approaches have emerged as an extended and promising research direction. This trend has provided strong motivation for the investigation and development of incremental algorithms based on the intuitionistic fuzzy rough set model.

2. Research objectives

1) *Proposing several models extended from the intuitionistic fuzzy rough set model*: From the preceding discussion, the first objective of this dissertation is to construct several models that inherit and exploit the advantages of the intuitionistic fuzzy rough set model in order to overcome the limitations of the two main research branches extended from classical rough set theory. In addition, the proposed models aim to improve computational efficiency and reduce the influence of noisy objects in the data compared to the intuitionistic fuzzy rough set model.

2) *Designing incremental algorithms based on the proposed models*: Based on several key properties of the proposed models, the second objective of this dissertation is to design incremental algorithms to handle practical data scenarios involving the addition and removal of object sets.

3. Research subjects

This dissertation focuses on fundamental concepts such as decision tables, reducts, and attribute reduction methods on decision tables through several extended branches of rough set theory:
1) Investigating several extensions based on the neighborhood rough set model and the fuzzy rough set model and their corresponding attribute reduction methods.
2) Investigating the intuitionistic fuzzy rough set model, attribute significance measures, and heuristic-based attribute reduction methods.

4. Scope of the study

The scope of this dissertation focuses on the limitations and improvement directions of several extensions of rough set models and their corresponding attribute reduction methods applied to both static decision tables and dynamic decision tables with changing object sets, specifically:
1) Investigating several models extended from the intuitionistic fuzzy rough set framework to develop attribute reduction algorithms for static decision tables.
2)Investigating incremental algorithms based on the proposed models to identify reducts in decision tables with changing object sets.

5. Research methods

1) *Theoretical aspect:* Research and proving several important properties of the proposed models, and investigating heuristic algorithms for identifying reducts in static decision tables as well as in decision tables with object addition and removal, based on significance measures defined in the space of the proposed models.
2) *Experimental aspect:* Conducting experiments to compare and evaluate the proposed algorithms against existing published algorithms using benchmark datasets collected from the UCI data repository[1] and OpenML[2] to evaluate the effectiveness of the proposed algorithms.

6. Dissertation structure

In addition to the introduction and conclusion, this dissertation consists of three main research chapters.

Chapter 1. This chapter first introduces the attribute reduction problem through several major approaches and presents fundamental concepts related to decision tables. Accordingly, this chapter provides an overview of intuitionistic fuzzy rough set theory, which serves as the theoretical foundation for the development of extended models and the proposal of attribute reduction algorithms. The main contributions of the dissertation are presented in detail in Chapters 2 and 3.

Chapter 2. This chapter presents attribute reduction algorithms for decision tables with object addition and removal based on the α, β-level intuitionistic fuzzy rough set model.

Chapter 3. This chapter presents attribute reduction algorithms for decision tables with object addition and removal based on the weighted intuitionistic fuzzy neighborhood rough set model.

Through the proposed methods, the dissertation also presents some experimental studies to demonstrate the advantages of the proposed approaches in overcoming the limitations of several representative methods based on fuzzy rough sets, weighted neighborhood rough sets, and intuitionistic fuzzy rough sets. Finally, the conclusion summarizes the results achieved in this dissertation, outlines future research directions, and discusses issues of interest to the author.

[1] `https://archive.ics.uci.edu/datasets`
[2] `https://openml.org/search?type=data&status=active&sort=runs`.

Chapter 1: Overview of the attribute reduction problem on decision tables

Chapter 1 presents the main contributions as follows:
(1) Providing an overview of the attribute reduction problem and several representative approaches.
(2) Presenting an overview of several models applied to the attribute reduction problem based on the two main branches extended from rough set theory. Accordingly, the advantages and disadvantages of each model are analyzed to identify research motivations.
(3) Presenting fundamental concepts related to decision tables and the intuitionistic fuzzy rough set model, which serve as the basis for proposing several extended models with high effectiveness in applying attribute reduction algorithms.

The research results of this chapter have been published in work [CT3] listed in the section List of Publications of the dissertation.

1.1. Overview of attribute reduction

1.1.1 Basic concepts

Attribute reduction is an vital data preprocessing problem whose primary objective is to retain essential attributes while eliminating redundant ones, without compromising classification and prediction accuracy. At present, attribute reduction problems are commonly addressed on decision tables.

A *decision table* is represented by a pair $IS = (U, C \cup D)$, where U is a non-empty finite set of objects, C and D are non-empty finite sets of attributes satisfying $C \cap D = \emptyset$. Each attribute $c \in C \cup D$ determines a mapping $c : U \to V_c$, where V_c denotes the value domain of attribute c. For any $u \in U$ and $c \in C \cup D$, the value of attribute c for object u is denoted by $c\,(u)$. C is referred to as the set of condition attributes, while D denotes the set of decision attributes. In the case where D contains multiple decision attributes, it can be transformed into a single decision attribute through an appropriate transformation [88].

1.1.2 Some models in attribute reduction

Neighborhood rough sets (NRSs) and k-nearest neighbor rough sets (KNNRSs) were first introduced by Hu *et al.* in 2008 [13]. Based on this foundation, many variants of the neighborhood rough set model have been developed to enhance the effectiveness of attribute reduction. The main advantage of neighborhood rough set models lies in their ability to directly select attributes from numerical decision tables, eliminating the need for data discretization while still ensuring classification performance. In addition, neighborhood relations focus only on objects that belong to the neighborhood of a given object. Consequently, neighborhood rough set models help reduce the computational scope and improve the processing efficiency of attribute reduction algorithms.

However, neighborhood rough set models do not consider the influence of individual attributes on the decisions of objects. In other words, these models assume that all condition attributes have equal weights. This assumption may lead to an inaccurate description of the relationship between condition attributes and decision attributes. As a result, some attributes that are strongly related to the decision may not be sufficiently represented to reflect their true importance. This may cause meaningful attributes to be overlooked during the attribute reduction process.

To address this issue, Hu *et al.* [29] proposed the weighted neighborhood rough set model (WNRSs), which employs distance measures based on attribute weights:

$$\Delta_A^{\omega}\,(u, v) = \sqrt{\sum_{a \in A} \omega^2\,(a)\,.\big(a\,(u) - a\,(v)\big)^2} \tag{1.1}$$

where, $\omega(a)$ is the weight of attribute $a \in A$, which is determined as follows. Given a decision table $IS = (U, C \cup D)$, $U = \{u_1, u_2, \ldots, u_n\}$ is the set of objects, $C = \{a_1, a_2, \ldots, a_m\}$ is the set of condition attributes, and D is the set of decision attributes. The partition coefficient of attributes is represented by a vector $\lambda = \left(\lambda(a_1), \lambda(a_2), \ldots, \lambda(a_m)\right)^{\mathrm{T}}$ is determined based on the solution $\lambda = (\mathbf{A}^{\mathrm{T}}\mathbf{A} + \mathbf{E})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{Y}$, where $\mathbf{A}$ is the coefficient matrix defined as

$$\mathbf{A} = \begin{bmatrix} a_1(u_1) & a_2(u_1) & \cdots & a_m(u_1) \\ a_1(u_2) & a_2(u_2) & \cdots & a_m(u_2) \\ \vdots & \vdots & \ddots & \vdots \\ a_1(u_n) & a_2(u_n) & \cdots & a_m(u_n) \end{bmatrix},$$

and $\mathbf{Y} = (D(a_1), D(a_2), \ldots, D(a_m))^{\mathrm{T}}$ is a *decision vector.*

Based on the partition coefficients of the attributes, the weight of each attribute $a \in C$ is determined by the following formula:

$$\omega(a) = \frac{|C| \cdot |\lambda(a)|}{\sum_{a \in C} |\lambda(a)|} \tag{1.2}$$

where $|C|$ denotes the number of condition attributes in the decision table, and $|\lambda(a)|$ represents the absolute value of $\lambda(a)$.

In this study, to ensure generality, the weighted distance formula is redefined as follows:

$$\Delta_A^{\omega}(u, v) = \sqrt[p]{\sum_{a \in A} \left(\omega(a) \,.\, \left|a(u) - a(v)\right|\right)^p} \tag{1.3}$$

The *weighted neighborhood relation*, denoted by $\mathcal{R}_A^{\delta,\omega}$, is defined as follows:

$$\mathcal{R}_A^{\delta,\omega} = \{(u, v) \in U \times U : \Delta_A^{\omega}(u, v) \leq \delta\} \tag{1.4}$$

Accordingly, for an object $u \in U$, a neighborhood information granule of u generated by attribute weights is referred to as an *attribute-weighted neighborhood information granule*, denoted by $[u]_A^{\delta,\omega}$, and is defined as follows:

$$[u]_A^{\delta,\omega} = \{v \in U : (u, v) \in \mathcal{R}_A^{\delta,\omega}\} \tag{1.5}$$

Obviously, the family of all these information granules forms a covering of U, denoted by $U/\mathcal{R}_A^{\delta,\omega} = \{[u]_A^{\delta,\omega} : u \in U\}$. Accordingly, for an object set $X \subseteq U$, the *lower approximation* and *upper approximation* of X based on attribute-weighted neighborhood information granules with respect to A are defined as follows:

$$\underline{\mathrm{W}N_A}(X) = \left\{u \in U : [u]_A^{\delta,\omega} \subseteq X\right\} \tag{1.6}$$

$$\overline{\mathrm{W}N_A}(X) = \left\{u \in U : [u]_A^{\omega} \cap X \neq \emptyset\right\} \tag{1.7}$$

From this model, the studies in [32] and [33] proposed models incorporating object weights and demonstrated the effectiveness of the corresponding attribute reduction methods. However, several limitations still exist in these models, as follows:

- The structure of information granules in these models is often relatively simple, failing to fully reflect the role of each object within a granule, while also neglecting the uncertainty and hesitation that inherently exist in real-world data.
- The models have not yet integrated both attribute weights and object weights, although such a combination is expected to significantly improve the effectiveness of optimal reduct selection.

- Most models employ neighborhood dependency as the criterion for defining a reduct. Consequently, when the decision table is inconsistent, this measure considers only objects in the positive region of the decision table, ignoring a large number of other objects, including those in the boundary region.

The research branch on fuzzy rough set models (FRSs) has been developed in parallel and has alleviated several difficulties of neighborhood rough set models in representing the characteristics of information granules. Specifically, each condition attribute $a \in C$ sdetermines a fuzzy binary relation $\widetilde{\mathcal{R}}_{\{a\}}$ on $U \times U$ with $\widetilde{\mathcal{R}}_{\{a\}}(u,v) \in [0,1]$. Then, $\widetilde{\mathcal{R}}_{\{a\}}$ is referred to as a fuzzy equivalence relation if it satisfies the following properties for all objects $u, v \in U$.
1. Reflexivity: $\widetilde{\mathcal{R}}_{\{a\}}(u,u) = 1$,
2. Symmetry: $\widetilde{\mathcal{R}}_{\{a\}}(u,v) = \widetilde{\mathcal{R}}_{\{a\}}(v,u)$,
3. Transitivity: $\widetilde{\mathcal{R}}_{\{a\}}(u,v) \geq \sup_{t \in U} \min \left\{ \widetilde{R}_{\{a\}}(u,t), \widetilde{R}_{\{a\}}(t,v) \right\}$.

A fuzzy equivalence relation generates a fuzzy partition $U/\widetilde{\mathcal{R}}_{\{a\}} = \left\{ \widetilde{[u]}_{\{a\}} : u \in U \right\}$ on U, where $\widetilde{[u]}_{\{a\}}$ is referred to as the fuzzy information granule of object u. Note that $\widetilde{[u]}_{\{a\}}$ is also a fuzzy set, and the membership degree of each object $v \in U$ denoted by $\widetilde{[u]}_{\{a\}}(v)$, is exactly the value of $\widetilde{\mathcal{R}}_{\{a\}}(u,v)$. Accordingly, the cardinality of the information granule $\widetilde{[u]}_{\{a\}}$, denoted by $|\widetilde{[u]}_{\{a\}}|$ is defined as $\left|\widetilde{[u]}_{\{a\}}\right| = \sum_{v \in U} \widetilde{[u]}_{\{a\}}(v)$.

It can be observed that the structure of fuzzy information granules is represented in greater detail than that of neighborhood information granules. Consequently, data characteristics are more comprehensively captured in the FRSs model. However, the FRSs model still has several limitations, as follows:

- Many objects belonging to a fuzzy information granule with very small membership degrees may introduce noise during the computation process.
- Attribute evaluation measures constructed under this model are mainly based on the computation of the cardinalities of fuzzy information granules. As a result, when dealing with datasets with low initial classification accuracy, the obtained reducts are often ineffective.
- FRSs-based methods have not yet incorporated attribute weighting for condition attributes.

Overall, both extended branches of rough set theory face considerable challenges. These limitations provide strong motivation for the author to propose new extensions along both research directions. To address the aforementioned issues, the intuitionistic fuzzy rough set model is considered a promising solution for the development of new models.

1.2. Intuitionistic fuzzy rough set model

Definition 1.1. (Intuitionistic fuzzy relation [91]). Given a decision table $IS = (U, C \cup D)$, each attribute $a \in C$ determines an intuitionistic fuzzy binary relation $\ddot{\mathcal{R}}_{\{a\}}$ on $U \times U$ as follows:

$$\ddot{\mathcal{R}}_{\{a\}} = \left\{ \left((u,v), \gamma_{\ddot{\mathcal{R}}_{\{a\}}}(u,v), \eta_{\ddot{\mathcal{R}}_{\{a\}}}(u,v) \right) : (u,v) \in U \times U \right\} \tag{1.8}$$

where $\gamma_{\ddot{\mathcal{R}}_{\{a\}}}(u,v) \in [0,1]$ and $\eta_{\ddot{\mathcal{R}}_{\{a\}}}(u,v) \in [0,1]$ denote the degree of similarity and the degree of diversity between objects u and v with respect to $\ddot{\mathcal{R}}_{\{a\}}$, respectively, such that $0 \leq \gamma_{\ddot{\mathcal{R}}_{\{a\}}}(u,v) + \eta_{\ddot{\mathcal{R}}_{\{a\}}}(u,v) \leq 1$.

From Definition 1.1, a relation $\ddot{\mathcal{R}}_{\{a\}}$is called an intuitionistic fuzzy equivalence relation if it satisfies the following properties for all $u, v \in U$

1. Reflexivity: $\gamma_{\ddot{\mathcal{R}}_{\{a\}}}(u,u) = 1$ và $\eta_{\ddot{\mathcal{R}}_{\{a\}}}(u,u) = 0$,

2. Symmetry: $\gamma_{\ddot{\mathcal{R}}_{\{a\}}}(u,v) = \gamma_{\ddot{\mathcal{R}}_{\{a\}}}(v,u)$ và $\eta_{\ddot{\mathcal{R}}_{\{a\}}}(u,v) = \eta_{\ddot{\mathcal{R}}_{\{a\}}}(v,u)$,

3. Transitivity: $\begin{cases} \gamma_{\ddot{\mathcal{R}}_{\{a\}}}(u,v) \geq \sup\limits_{t \in U}\left\{\min\left(\gamma_{\ddot{\mathcal{R}}_{\{a\}}}(u,t), \gamma_{\ddot{\mathcal{R}}_{\{a\}}}(t,v)\right)\right\} \\ \eta_{\ddot{\mathcal{R}}_{\{a\}}}(u,v) \leq \inf\limits_{t \in U}\left\{\max\left(\eta_{\ddot{\mathcal{R}}_{\{a\}}}(u,t), \eta_{\ddot{\mathcal{R}}_{\{a\}}}(t,v)\right)\right\} \end{cases}$

Each attribute $a \in C$ in a decision table determines an intuitionistic fuzzy equivalence relation $\ddot{\mathcal{R}}_{\{a\}}$. Consequently, an intuitionistic fuzzy equivalence relation $\ddot{\mathcal{R}}_{\{a\}}$ greates an intuitionistic fuzzy partition $U/\ddot{\mathcal{R}}_{\{a\}} = \left\{[\ddot{u}]_{\{a\}} \,|u \in U\right\}$ on U, where $[\ddot{u}]_{\{a\}}$ is referred to as an intuitionistic fuzzy information granule of object u with respect to the relation $\ddot{\mathcal{R}}_{\{a\}}$.

Based on this model, numerous attribute reduction methods have been developed using different measures, such as methods in [53, 96, 54, 56], intuitionistic fuzzy entropy [57, 58, 97, 98], and intuitionistic fuzzy distance measures [59, 91]. These methods have demonstrated the effectiveness of the IFRSs model in improving classification performance on datasets with low initial classification accuracy. The advantages of IFRSs provide a solid foundation for developing extended models aimed at overcoming the limitations of different branches of rough set models.

- Compared with the NRSs and FRSs models, the IFRSs model exhibits a more detailed granule structure, which more clearly reflects the inherent vagueness and hesitation present in real-world data.
- The incorporation of the non-membership function in intuitionistic fuzzy information granules effectively adjusts information affected by objects generated from noise.

However, through investigation, the IFRSs model still has several limitations, as follows:

- The inclusion of the non-membership component also makes the algorithms require more storage space and results in less efficient processing time.
- Many objects belonging to intuitionistic fuzzy information granules with very small membership degrees and large non-membership degrees may create noise and redundancy during the computation process.

Therefore, the research motivation of this dissertation is not limited to employing the IFRSs model as an inherited framework for developing new models, but also focuses on addressing the inherent limitations of the IFRSs model itself. One important issue that deserves attention is that the number of objects in data often changes over time, posing significant challenges not only for attribute reduction algorithms based on the IFRSs approach but also for methods based on other models. In response to these challenges, incremental attribute reduction approaches have emerged as an extended research direction and have attracted considerable attention.

Chapter 2: Proposing attribute reduction algorithms based on the α, β-level intuitionistic fuzzy set model

2.1. Introduction

In Chapter 2, we present the main contributions as follows: (1) Proposing the α, β-level intuitionistic fuzzy set model and presenting its important properties, (2) Proposing an attribute reduction algorithm for fixed decision tables, (3) Proposing two incremental attribute reduction algorithms for dynamic decision tables, (4) Demonstrating the efficiency of the proposed algorithms in comparison with some algorithms based on fuzzy rough set and intuitionistic fuzzy rough set models. The research results of this chapter have been published in works [CT1], [CT2], [CT4], and [CT5] listed in the List of Publications.

2.1.1 Concept of the α, β-level intuitionistic fuzzy set

Definition 2.1. Let $[\ddot{u}]_{\{a\}}$ be an intuitionistic fuzzy information granule of object u with respect to attribute a, and let α and β be two real numbers in the interval $[0, 1]$ satisfying $\alpha + \beta \leq 1$. The α, β-cut set of $[\ddot{u}]_{\{a\}}$ is a crisp set defined as follows:

$$[u]^{\alpha,\beta}_{\{a\}} = \left\{ v \in U : \gamma_{[\ddot{u}]_{\{a\}}}(v) \geq \alpha \wedge \eta_{[\ddot{u}]_{\{a\}}}(v) \leq \beta \right\} \tag{2.1}$$

An α-β-level intuitionistic fuzzy set, denoted by $[\ddot{u}]^{\alpha,\beta}_{\{a\}}$, is defined based on each element of the α, β-cut set, where the similarity and dissimilarity degrees of each object are determined as follows:

$$[\ddot{u}]^{\alpha,\beta}_{\{a\}}(v) = \left(\gamma_{[\ddot{u}]^{\alpha,\beta}_{\{a\}}}(v), \eta_{[\ddot{u}]^{\alpha,\beta}_{\{a\}}}(v) \right) = \begin{cases} \left(\gamma_{[\ddot{u}]_{\{a\}}}(v), \eta_{[\ddot{u}]_{\{a\}}}(v) \right), v \in [u]^{\alpha,\beta}_{\{a\}} \\ (0, 1), v \notin [u]^{\alpha,\beta}_{\{a\}} \end{cases} \tag{2.2}$$

If a is a numerical attribute, the similarity and diversity values of v in the intuitionistic fuzzy information granule of u are determined according to the formulas in [91]. In this study, $[\ddot{u}]^{\alpha,\beta}_{\{a\}}$ is called an α, β-level intuitionistic fuzzy information granule of u. Accordingly, the family $U/\ddot{\mathcal{R}}^{\alpha,\beta}_{\{a\}} = \left\{ [\ddot{u}]^{\alpha,\beta}_{\{a\}} : u \in U \right\}$ is called an α, β-level intuitionistic fuzzy partition.

2.1.2 Properties of α, β-level intuitionistic fuzzy information granules

Proposition 2.1. Given a decision table $IS = (U, C \cup D)$, then:
1. If $A \subseteq B \subseteq C$, then $[u]^{\alpha,\beta}_{B} \subseteq [u]^{\alpha,\beta}_{A} \subseteq U$,
2. $\forall A, B \subseteq C$, $[\ddot{u}]^{\alpha,\beta}_{A} \subseteq [\ddot{u}]_{A}$ and $[u]^{\alpha,\beta}_{A \cup B} = [u]^{\alpha,\beta}_{A} \cap [u]^{\alpha,\beta}_{B}$.

Proposition 2.2. Given $IS = (U, C \cup D)$, $A \subseteq C$, if $\alpha_1 \leq \alpha_2$ and $\beta_1 \geq \beta_2$, then $[\ddot{u}]^{\alpha_2,\beta_2}_{A} \subseteq [\ddot{u}]^{\alpha_1,\beta_1}_{A}$.

2.2. Proposed attribute reduction algorithm based on α, β-level intuitionistic fuzzy sets

2.2.1 Proposed attribute reduction algorithm for fixed decision tables

Proposition 2.3. Given a decision table $IS = (U, C \cup D)$, the distance between two α, β-level intuitionistic fuzzy partitions induced by C and $C \cup D$ on U is computed as follows:

$$\ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}^{\alpha,\beta}_{C}, U/\ddot{\mathcal{R}}^{\alpha,\beta}_{C \cup D} \right) = \sum_{u \in U} \frac{\left(\left| [\ddot{u}]^{\alpha,\beta}_{C} \right| - \left| [\ddot{u}]^{\alpha,\beta}_{C} \cap [\ddot{u}]^{\alpha,\beta}_{D} \right| \right)}{|U|^2} \tag{2.3}$$

Proposition 2.4. Let $A, B \subseteq C$. If $A \subseteq B$, then $\ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_A^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{A\cup D}^{\alpha,\beta}\right) \geq \ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_B^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{B\cup D}^{\alpha,\beta}\right)$.

Definition 2.2. Given a decision table $IS = (U, C \cup D)$, a subset of condition attributes $A \subseteq C$ is called a reduct of C if:
1. $\ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_A^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{A\cup D}^{\alpha,\beta}\right) = \ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)$,
2. $\forall A' \subseteq A$, $\ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_{A'}^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{A'\cup D}^{\alpha,\beta}\right) \geq \ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_A^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{A\cup D}^{\alpha,\beta}\right)$.

Definition 2.3. Given a decision table $IS = (U, C \cup D)$, a subset of condition attributes $A \subseteq C$, and an attribute $a \in C \backslash A$, the significance of a with respect to A, denoted by $Sig_P(a, A)$, is determined by the formula:

$$Sig_P(a, A) = \ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_A^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{A\cup D}^{\alpha,\beta}\right) - \ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_{A\cup\{a\}}^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{A\cup\{a\}\cup D}^{\alpha,\beta}\right) \tag{2.4}$$

Algorithm 2.1 Attribute Reduction based on the α,β-level intuitionistic fuzzy Partition Distance (ARPD).

Input: A decision table $IS = (U, C \cup D)$, levels α, β.
Output: One reduct $\mathcal{A}$

compute $U/\ddot{\mathcal{R}}_D$.
for $a \in C$ **do**
  if a is a continuous numeric value domain attribute **then**
    compute $U/\ddot{\mathcal{R}}_{\{a\}}^{\alpha,\beta}$
  else
    $U/\ddot{\mathcal{R}}_{\{a\}}^{\alpha,\beta} = U/\ddot{\mathcal{R}}_{\{a\}}$
  end if
end for
compute $U/\ddot{\mathcal{R}}_C^{\alpha,\beta} = \bigcap_{a\in C} U/\ddot{\mathcal{R}}_{\{a\}}^{\alpha,\beta}$
$\mathcal{A} = \{a_0\}$ which satisfies: $\ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_{\{a_0\}}^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{\{a_0\}\cup D}^{\alpha,\beta}\right) = \min_{a\in C} \ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_{\{a\}}^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{\{a\}\cup D}^{\alpha,\beta}\right)$
while $\ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_{\mathcal{A}}^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{\mathcal{A}\cup D}^{\alpha,\beta}\right) \geq \ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)$ **do**
  compute $Sig(a, \mathcal{A})$, for all $a \in C \backslash \mathcal{A}$
  select a_0 which satisfies: $Sig_P(a_0, \mathcal{A}) = \max_{a\in C\backslash\mathcal{A}} \{Sig_P(a, \mathcal{A})\}$
  $\mathcal{A} \leftarrow \mathcal{A} \cup \{a_0\}$
end while
return $\mathcal{A}$

The computational complexity of Algorithm 2.1 is $O\left(|C|^2|U|^2\right)$, where $|C|$ and $|U|$ denote the numbers of condition attributes and objects in the decision table, respectively.

2.2.2 Attribute reduction on the decision table when adding object set

Proposition 2.5. Given $IS = (U, C \cup D)$ with $U = \{u_1, u_2, \ldots, u_n\}$, intuitionistic fuzzy equivalence relations $\ddot{\mathcal{R}}_C$, $\ddot{\mathcal{R}}_D$ and a new object set $\Delta U = \{U_{n+1}, U_{n+2}, \ldots, U_{n+z}\}$, where $z \geq 1$, the intuitionistic fuzzy partition distance between $U^+/\ddot{\mathcal{R}}_C^{\alpha,\beta}$ and $U^+/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}$ in the object set $U^+ = U \cup \Delta U$, denoted $\ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)$, is computed as follows:

$$\ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right) = \frac{n^2 \ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)}{(n+z)^2} + 2\sum_{i=1}^{z}\left(\left|[\ddot{u}_{n+i}]_C^{\alpha,\beta}\right| - \left|[\ddot{u}_{n+i}]_C^{\alpha,\beta} \cap [\ddot{u}_{n+i}]_D^{\alpha,\beta}\right| - \vartheta_i\right) \tag{2.5}$$

where $\vartheta_1 = 0$ and for any $i \geq 2$,

$$\vartheta_i = \sum_{j=1}^{z-1}\left(\gamma_{[\ddot{u}_{n+i}]_C^{\alpha,\beta}}\left(u_{n+j+1}\right) - \gamma_{[\ddot{u}_{n+i}]_{C\cup D}^{\alpha,\beta}}\left(u_{n+j+1}\right) - \eta_{[\ddot{u}_{n+i}]_C^{\alpha,\beta}}\left(u_{n+j+1}\right) + \eta_{[\ddot{u}_{n+i}]_{C\cup D}^{\alpha,\beta}}\left(u_{n+j+1}\right)\right).$$

Proposition 2.6. Assume that $A \subseteq C$ is a reduct based on the intuitionistic fuzzy partition distance on U and $\Delta U = \{U_{n+1}, U_{n+2}, \ldots, U_{n+z}\}$, where $z \geq 1$ is an incremental object set, then we have:
1. If all objects in ΔU have the same value of decision attribute then:

$$\ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right) = \frac{n^2\ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right) + 2\sum_{i=1}^{z}\left(\left|[\ddot{u}_{n+i}]_C^{\alpha,\beta}\right| - \left|[\ddot{u}_{n+i}]_C^{\alpha,\beta} \cap [\ddot{u}_{n+i}]_D^{\alpha,\beta}\right|\right)}{(n+z)^2} \tag{2.6}$$

2. $\ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_A^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{A\cup D}^{\alpha,\beta}\right) = \ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)$ if $[\ddot{u}_{n+i}]_A^{\alpha,\beta} \subseteq [\ddot{u}_{n+i}]_D^{\alpha,\beta}$ with $i = 1, ..., z$.

Algorithm 2.2 Incremental Attribute Reduction based on the α,β-level Intuitionistic Fuzzy Partition Distance when Adding Object set (IARPD-AO)

Input: Decision table $IS = (U, C \cup D)$ with $U = \{u_1, u_2, \ldots, u_n\}$, the reduct $\mathcal{A} \subseteq C$, partitions $U/\ddot{\mathcal{R}}_{\mathcal{A}}^{\alpha,\beta}$, $U/\ddot{\mathcal{R}}_C^{\alpha,\beta}$ and added set of objects $\Delta U = \{u_{n+1}, u_{n+2}, \ldots, u_{n+z}\}$.
Output: The approximation reduct $\mathcal{A}^+$ on $U^+ = U \cup \Delta U$

// Check the added set of objects
$\mathcal{A}^+ \leftarrow \mathcal{A}$, $Z \leftarrow \Delta U$
for $i = 1$ to z **do**
  if $[\ddot{u}_{n+i}]_{\mathcal{A}}^{\alpha,\beta} \subseteq [\ddot{u}_{n+i}]_D^{\alpha,\beta}$ **then**
    $Z \leftarrow Z \setminus \{u_{n+i}\}$
  end if
end for
if $Z = \emptyset$ **then**
  return $\mathcal{A}^+$
end if
$\Delta U \leftarrow Z$, $z \leftarrow |\Delta U|$
update $U^+/\ddot{\mathcal{R}}_{\mathcal{A}}^{\alpha,\beta}$, $U^+/\ddot{\mathcal{R}}_C^{\alpha,\beta}$ and $U^+/\ddot{\mathcal{R}}_{\{a\}}^{\alpha,\beta}$, for all $a \in C \setminus \mathcal{A}$
// Finding the reduct
compute $\ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+}^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+\cup D}^{\alpha,\beta}\right)$ and $\ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)$ from Equation 2.5.
while $\ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+}^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+\cup D}^{\alpha,\beta}\right) \geq \ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)$ **do**
  compute $Sig\left(a, \mathcal{A}^+\right)$, for all $a \in C \setminus \mathcal{A}^+$
  select a_0 which satisfies: $Sig_P\left(a_0, \mathcal{A}^+\right) = \max_{a \in C \setminus \mathcal{A}^+}\left\{Sig\left(a, \mathcal{A}^+\right)\right\}$
  $\mathcal{A}^+ \leftarrow \mathcal{A}^+ \cup \{a_0\}$
end while
// Remove redundant attributes
for $a \in \mathcal{A}^+$ **do**
  compute $\ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+\setminus\{a\}}^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+\setminus\{a\}\cup D}^{\alpha,\beta}\right)$
  if $\ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+\setminus\{a\}}^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+\setminus\{a\}\cup D}^{\alpha,\beta}\right) = \ddot{\mathcal{D}}\left(U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+}^{\alpha,\beta}, U^+/\ddot{\mathcal{R}}_{\mathcal{A}^+\cup D}^{\alpha,\beta}\right)$ **then**
    $\mathcal{A}^+ \leftarrow \mathcal{A}^+ \setminus \{a_0\}$
  end if
end for
return $\mathcal{A}^+$

The complexity of IARPD-AO is $\max\left\{O\left(\left|\mathcal{A}^+\right|\left|U^+\right||\Delta U|\right), O\left(\left(|C| - \left|\mathcal{A}^+\right|\right)^2\left|U^+\right||\Delta U|\right)\right\}$, which is carried out through three main stages. In the first stage, the algorithm checks the newly added objects. In the second stage, the algorithm considers the remaining attributes in the decision table that are not included in the reduct obtained from the previous stage. The attribute with the highest significance is then added to the reduct. In the third stage, the algorithm removes redundant attributes from the resulting reduct.

2.2.3 Attribute reduction on the decision table when deleting object set

Proposition 2.7. The intuitionistic fuzzy partition distance between $U^-/\ddot{\mathcal{R}}_C^{\alpha,\beta}$ and $U^-/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}$ on $U^- = U \setminus \Delta U$, with $\Delta U = \{u_n, u_{n-1}, \ldots, u_{n-z+1}\}$ is an object set is removed from U, where $z \geq 1$, is computed as follows:

$$\ddot{\mathcal{D}}\left(U^-/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U^-/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right) = \frac{n^2\ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)}{(n-z)^2} - 2\sum_{i=1}^{z}\frac{\left(\left|[\ddot{u}_{n-i+1}]_C^{\alpha,\beta}\right| - \left|[\ddot{u}_{n-i+1}]_C^{\alpha,\beta} \cap [\ddot{u}_{n-i+1}]_D^{\alpha,\beta}\right| - \xi_i\right)}{(n-z)^2} \tag{2.7}$$

where $\xi_1 = 0$ and for all $i \geq 2$,

$$\xi_i = \sum_{j=i}^{z}\left(\gamma_{[\ddot{u}_{n-i+1}]_C^{\alpha,\beta}}(u_{n-j+1}) - \gamma_{[\ddot{u}_{n-i+1}]_{C\cup D}^{\alpha,\beta}}(u_{n-j+1}) - \eta_{[\ddot{u}_{n-i+1}]_C^{\alpha,\beta}}(u_{n-j+1}) + \eta_{[\ddot{u}_{n-i+1}]_{C\cup D}^{\alpha,\beta}}(u_{n-j+1})\right)$$

Proposition 2.8. Assume that $A \subseteq C$ is a reduct based on the intuitionistic fuzzy partition distance on U and a object set $\Delta U = \{u_n, u_{n-1}, \ldots, u_{n-z+1}\}$ is removed from, where $z \geq 1$ and $U^- = U\backslash\Delta U$, then we have:
1. If all objects in ΔU have the same value of decision attribute then:

$$\ddot{\mathcal{D}}\left(U^-/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U^-/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right) = \frac{n^2\ddot{\mathcal{D}}\left(U/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)}{(n-z)^2} - 2\sum_{i=1}^{z}\frac{\left(\left|[\ddot{u}_{n-i+1}]_C^{\alpha,\beta}\right| - \left|[\ddot{u}_{n-i+1}]_C^{\alpha,\beta} \cap [\ddot{u}_{n-i+1}]_D^{\alpha,\beta}\right|\right)}{(n-z)^2} \tag{2.8}$$

2. $\ddot{\mathcal{D}}\left(U^-/\ddot{\mathcal{R}}_A^{\alpha,\beta}, U^-/\ddot{\mathcal{R}}_{A\cup D}^{\alpha,\beta}\right) = \ddot{\mathcal{D}}\left(U^-/\ddot{\mathcal{R}}_C^{\alpha,\beta}, U^-/\ddot{\mathcal{R}}_{C\cup D}^{\alpha,\beta}\right)$

Based on the presented foundations, this dissertation proposes the IARPD-RO algorithm with a computational complexity of $O\left(\left|\mathcal{A}^-\right|\left|U^-\right||\Delta U|\right)$,which consists of two main stages. In the first stage, the algorithm updates the α, β-level intuitionistic fuzzy partitions of the attribute set C together with the condition attributes in the new decision table. In the second stage, the algorithm examines the reduct obtained from the previous stage and removes each attribute that does not satisfy the properties of a reduct.

2.3. Experimental and evaluation of the proposed algorithms

2.3.1 Performance of the IARPD-AO Algorithm

The experimental process is conducted on several datasets described in Table 2.1. These datasets are divided into two approximately equal parts, denoted by U_{ori} and U_{inc}. The dataset U_{ori} is applied for the ARPD, F-FDAR [81], ARIFPD [91], IFPR [55], NIFS [82], and FMIFRFS [56] algorithms to obtain a reduct.

The reduct size and computational time of the algorithms when applied to U_{ori} are presented in Table 2.2. It can be observed that algorithms based on the fuzzy rough set approach exhibit faster processing times than those based on the intuitionistic fuzzy rough set approach.

However, when considering the intuitionistic fuzzy set space, the proposed algorithm achieves the best execution time. The average processing time of the proposed algorithm is 13.590 seconds, whereas those of ARIFPD, FMIFRFS, and IFPR are 29.545 seconds, 15.172 seconds, and 18.371 seconds, respectively. Algorithms based on the intuitionistic fuzzy set model generally obtain reducts with smaller sizes than those based on the fuzzy rough set

Algorithm 2.3 Incremental Attribute Reduction based on the α,β-level Intuitionistic Fuzzy Partition Distance when Removing Object set (IARPD-RO).

Input: Decision table $IS = (U, C \cup D)$ with $U = \{u_1, u_2, \ldots, u_n\}$, the reduct $\mathcal{A} \subseteq C$, partitions $U/\ddot{\mathcal{R}}_{\mathcal{A}}^{\alpha,\beta}$, $U/\ddot{\mathcal{R}}_{C}^{\alpha,\beta}$ and removed set of objects $\Delta U = \{u_{n+1}, u_{n+2}, \ldots, u_{n+z}\}$.
Output: The approximation reduct $\mathcal{A}^-$ on $U^- = U \cup \Delta U$

// initialize and update
$\mathcal{A}^- \leftarrow \mathcal{A}$
update $U^-/\ddot{\mathcal{R}}_{\mathcal{A}}^{\alpha,\beta}$, $U^-/\ddot{\mathcal{R}}_{C}^{\alpha,\beta}$ and $U^-/\ddot{\mathcal{R}}_{\{a\}}^{\alpha,\beta}$, for all $a \in C \backslash \mathcal{A}$
compute $\ddot{\mathcal{D}}\left(U^-/\ddot{\mathcal{R}}_{\mathcal{A}^-}^{\alpha,\beta}, U^-/\ddot{\mathcal{R}}_{\mathcal{A}^- \cup D}^{\alpha,\beta}\right)$ và $\ddot{\mathcal{D}}\left(U^-/\ddot{\mathcal{R}}_{C}^{\alpha,\beta}, U^-/\ddot{\mathcal{R}}_{C \cup D}^{\alpha,\beta}\right)$ by Equation 2.7.
// finding the reduct
for $a \in \mathcal{A}^-$ **do**
  compute $\ddot{\mathcal{D}}\left(U^-/\ddot{\mathcal{R}}_{\mathcal{A}^- \backslash \{a\}}^{\alpha,\beta}, U^-/\ddot{\mathcal{R}}_{\mathcal{A}^- \backslash \{a\} \cup D}^{\alpha,\beta}\right)$ by Equation 2.5.
  if $\ddot{\mathcal{D}}\left(U^-/\ddot{\mathcal{R}}_{\mathcal{A}^- \backslash \{a\}}^{\alpha,\beta}, U^-/\ddot{\mathcal{R}}_{\mathcal{A}^- \backslash \{a\} \cup D}^{\alpha,\beta}\right) = \ddot{\mathcal{D}}\left(U^-/\ddot{\mathcal{R}}_{\mathcal{A}^-}^{\alpha,\beta}, U^-/\ddot{\mathcal{R}}_{\mathcal{A}^- \cup D}^{\alpha,\beta}\right)$ **then**
    $\mathcal{A}^- \leftarrow \mathcal{A}^- \backslash \{a_0\}$
  end if
end for
return $\mathcal{A}^-$

Table 2.1: Datasets used for evaluating IARPD-AO and several algorithms

ID	Datasets	Objects	$\lvert U_{ori} \rvert$	$\lvert U_{inc} \rvert$	Attributes	Classes	Source
1	Ionosphere	351	175	176	34	2	UCI
2	Leaf	340	170	170	15	30	UCI
3	Movement	360	180	180	90	15	UCI
4	Urban	675	337	338	147	9	OpenML
5	Hill-valley	1212	606	606	100	2	UCI
6	Mfeat	2000	1000	1000	76	10	OpenML
7	Wall	5456	2728	2728	24	4	OpenML
8	Waveform2	5000	2500	2500	40	3	UCI

model. The proposed algorithm produces the smallest average reduct size. According to Table 2.3, the classification performance of ARPD is also considered superior to that of the remaining algorithms.

Next, the dataset U_{inc} is divided into five equal parts, denoted by U_1 to U_5. These subsets are then incrementally added to U_{ori} to evaluate the reducts obtained by the incremental algorithms. Specifically, the IARPD-AO algorithm is compared with several incremental algorithms based on fuzzy rough sets (IF-FDAR-AO [81], AIFSA-FKD [82]) and intuitionistic fuzzy sets (ARIFPD-AO [91]). The results are analyzed based on the performance of the algorithms across five stages of object set addition.

Table 2.3: Comparison between ARPD and several algorithms in terms of classification accuracy on U_{ori}

Datasets	Raw data	ARPD	FMIFRFS	ARIFPD	IFPR	NIFS	F-FDAR
Ionosphere	0.766 ± 0.092	**0.841 ± 0.097**	0.840 ± 0.112	0.789 ± 0.121	0.817 ± 0.068	0.822 ± 0.089	0.812 ± 0.113
Leaf	0.718 ± 0.101	**0.741 ± 0.096**	0.724 ± 0.118	**0.741 ± 0.096**	**0.741 ± 0.106**	**0.741 ± 0.124**	0.735 ± 0.121
Movement	0.867 ± 0.090	**0.883 ± 0.091**	**0.883 ± 0.084**	0.856 ± 0.097	0.856 ± 0.114	0.867 ± 0.112	0.867 ± 0.097
Urban	0.801 ± 0.065	**0.828 ± 0.054**	0.813 ± 0.051	0.810 ± 0.059	0.778 ± 0.052	0.804 ± 0.048	0.810 ± 0.066
Hill-valley	0.510 ± 0.037	**0.544 ± 0.040**	0.525 ± 0.055	0.500 ± 0.052	0.517 ± 0.046	0.490 ± 0.049	0.497 ± 0.040
Mfeat	0.925 ± 0.018	**0.949 ± 0.016**	0.945 ± 0.019	0.947 ± 0.017	0.939 ± 0.020	0.910 ± 0.019	0.942 ± 0.013
Wall	0.638 ± 0.094	**0.750 ± 0.045**	0.746 ± 0.064	0.682 ± 0.070	0.724 ± 0.067	0.684 ± 0.089	0.650 ± 0.076
Waveform2	0.782 ± 0.020	0.800 ± 0.022	**0.817 ± 0.013**	0.800 ± 0.022	0.762 ± 0.020	0.783 ± 0.022	0.787 ± 0.017
Average	0.751 ± 0.065	**0.792 ± 0.058**	0.787 ± 0.065	0.766 ± 0.067	0.767 ± 0.062	0.763 ± 0.069	0.763 ± 0.068

- *Runtime*: Incremental algorithms exhibit significantly faster processing times than algorithms executed on static decision tables, even though the number of objects is larger. Similar to the results in Table 2.2, incremental algorithms based on the fuzzy rough set approach continue to demonstrate superior execution efficiency. When compared with intuitionistic fuzzy–based algorithms, the proposed algorithm shows faster execution times across incremental stages. This highlights the impact of the α,β-cut set in eliminating noisy objects from intuitionistic fuzzy information granules, thereby reducing the search space.

Table 2.2: Reduct size and runtime of ARPD and comparative algorithms on U_{ori}

Datasets	ARPD		FMIFRFS		ARIFPD		IFPR		NIFS		F-FDAR	
	time	$\|\mathcal{A}\|$	*time*	$\|\mathcal{A}\|$	*time*	$\|\mathcal{A}\|$	*time*	$\|\mathcal{A}\|$	*time*	$\|\mathcal{A}\|$	*time*	$\|\mathcal{A}\|$
Ionosphere	0.0192	5	0.1257	**4**	0.0846	14	0.0902	9	**0.0033**	9	0.0348	14
Leaf	0.0186	9	0.1242	10	0.0171	9	0.0174	12	**0.0028**	10	0.0049	**8**
Movement	0.1123	**9**	1.1194	21	0.4815	20	0.3383	19	**0.0162**	54	0.0908	21
Urban	3.4501	48	9.1160	47	4.2266	**35**	30.652	43	1.2784	74	**0.6702**	38
Hill-valley	1.2533	**2**	1.3471	3	17.322	5	1.3789	4	**0.0521**	4	3.4982	24
Mfeat	23.382	**22**	36.674	23	56.114	30	30.890	29	**6.8312**	56	6.3366	35
Wall	8.0352	3	15.774	4	38.941	8	5.2763	**2**	**1.0913**	8	5.5972	14
Waveform2	72.449	24	57.093	**15**	119.17	24	78.322	38	**1.6991**	26	14.378	28
Average	13.590	**15.3**	15.172	15.9	29.545	16.3	18.371	19.5	**1.3718**	30.1	3.8263	22.8

- *Reduct size*: The reduct sizes obtained by algorithms based on the intuitionistic fuzzy model are generally smaller than those produced by fuzzy rough set–based algorithms on most datasets. The proposed algorithm achieves the smallest average reduct size among the four algorithms. In particular, for datasets with low initial classification accuracy, such as Movement, Hill-valley, and Wall, the number of attributes selected by IARPD-AO is very small. This indicates that IARPD-AO is more effective in removing redundant attributes from noisy data than the other algorithms.

- *Classification accuracy*: Reducts obtained from intuitionistic fuzzy set–based algorithms achieve higher classification accuracy than those based on fuzzy rough set models in most incremental stages. This emphasizes the importance of the non-membership component in adjusting evaluation measures, enabling the algorithms to select important attributes and achieve higher classification performance. These results also reveal a clear competition between the proposed algorithm and ARIFPD-AO. Specifically, across all cases, the proposed algorithm achieves the highest accuracy in 31 cases, whereas ARIFPD-AO does so in only 13 cases.

2.3.2 Performance of the IARPD-RO Algorithm

In this section, several experiments are conducted to demonstrate the effectiveness of the incremental IARPD-RO algorithm. The experimental process is carried out on standard benchmark datasets described in Table 2.4. The subset U_{dec} consists of half of the original dataset and is further divided into five approximately equal parts, denoted by U_1 to U_5, which are sequentially removed from U. According to Table 2.5, the ARPD algorithm achieves the most notable results in four cases, extracting fewer attributes than the other methods.

In the intuitionistic fuzzy rough set space, the proposed algorithm exhibits faster processing speed than the other three algorithms. Specifically, the average execution time of ARPD on the datasets is 40.995 seconds, whereas those of FMIFRFS, ARIFPD, and IFPR are 52.252 seconds, 57.514 seconds, and 63.478 seconds, respectively.

Table 2.4: Datasets used for evaluating IARPD-RO and several algorithms

ID	Datasets	Objects	$\|U_{dec}\|$	Attributes	Classes	Source
1	Robot-failures	164	84	90	5	OpenML
2	Ionosphere	351	176	34	2	UCI
3	Movement	360	180	90	15	UCI
4	Pizza	1043	523	35	2	OpenML
5	PD	756	381	754	2	UCI
6	Seismic-bumps	2584	1294	18	2	UCI
7	Waveform1	5000	2500	21	3	UCI
8	Wall	5456	2728	24	4	OpenML

Table 2.5: Reduct size and runtime of ARPD and comparative algorithms on U.

Tập dữ liệu	ARPD		FMIFRFS		ARIFPD		IFPR		NIFS		F-FDAR	
	Time	$\|\mathcal{A}\|$	*Time*	$\|\mathcal{A}\|$	*Time*	$\|\mathcal{A}\|$	*Time*	$\|\mathcal{A}\|$	*Time*	$\|\mathcal{A}\|$	*Time*	$\|\mathcal{A}\|$
Robot-failures	0.1280	**15**	0.7683	22	0.1392	16	1.6414	17	**0.0151**	19	0.0711	19
Ionosphere	0.1619	12	0.2356	9	0.1523	15	0.8762	**8**	**0.0084**	14	0.0584	14
Movement	0.4845	**15**	5.5189	18	0.8367	23	2.0156	23	**0.1163**	15	0.2482	20
Pizza	3.3650	**7**	9.6722	19	5.9454	14	5.0071	13	1.2779	22	**1.0891**	15
PD	218.54	41	211.70	**38**	267.13	49	251.56	47	**6.4523**	46	79.811	47
Seismic-bumps	4.1578	3	11.474	9	4.5112	3	1.6542	**2**	**0.1371**	5	1.8836	6
Waveform1	68.391	15	76.427	**9**	72.667	16	107.64	19	**2.9914**	16	19.477	15
Wall	32.730	**4**	102.22	7	108.73	16	137.43	13	**2.8172**	13	23.320	15
Trung bình	40.995	**14.0**	52.252	16.4	57.514	19.0	63.478	17.8	**1.7270**	18.8	15.745	18.9

Table 2.6: Comparison between ARPD and several algorithms in terms of classification accuracy on U

Tập dữ liệu	Tập gốc	ARPD	FMIFRFS	ARIFPD	IFPR	NIFS	F-FDAR
Robot-failures	0.535 ± 0.099	**0.658 ± 0.113**	0.622 ± 0.106	**0.658 ± 0.113**	0.609 ± 0.121	0.583 ± 0.133	0.597 ± 0.097
Ionosphere	0.838 ± 0.064	0.880 ± 0.074	0.903 ± 0.053	0.858 ± 0.055	**0.906 ± 0.055**	0.849 ± 0.060	0.860 ± 0.049
Movement	0.758 ± 0.117	0.761 ± 0.130	**0.769 ± 0.119**	0.756 ± 0.114	**0.764 ± 0.115**	0.733 ± 0.116	0.747 ± 0.126
Pizza	0.863 ± 0.022	**0.869 ± 0.027**	**0.869 ± 0.020**	0.865 ± 0.024	0.867 ± 0.020	0.866 ± 0.023	0.866 ± 0.022
PD	0.792 ± 0.061	**0.798 ± 0.036**	0.796 ± 0.046	0.774 ± 0.047	0.780 ± 0.044	0.742 ± 0.070	0.784 ± 0.057
Seismic-bumps	0.908 ± 0.043	0.925 ± 0.028	0.914 ± 0.032	0.925 ± 0.028	**0.932 ± 0.007**	0.913 ± 0.046	0.909 ± 0.036
Waveform1	0.821 ± 0.021	**0.823 ± 0.022**	0.812 ± 0.010	0.816 ± 0.022	0.806 ± 0.023	0.817 ± 0.019	0.814 ± 0.016
Wall	0.773 ± 0.059	**0.831 ± 0.040**	0.816 ± 0.038	0.784 ± 0.062	0.782 ± 0.065	0.773 ± 0.064	0.779 ± 0.074
Trung bình	0.786 ± 0.061	**0.818 ± 0.059**	0.813 ± 0.053	0.805 ± 0.058	0.806 ± 0.056	0.785 ± 0.066	0.795 ± 0.060

According to Table 2.6, ARPD achieves the highest average classification accuracy, with an improvement of 4.2% compared to using all attributes. When considering only the intuitionistic fuzzy rough set space, ARPD attains the highest classification performance in five cases. Finally, further analyses of incremental algorithms on decision tables with object removal are presented through the evaluation of the performance of three algorithms: IARPD-DO, IF-FDAR-DO [81], and AIFSD-FKD [82]. Similar to IARPD-AO, the proposed algorithm demonstrates strong performance in terms of execution time, reduct size, and classification accuracy across object set removal stages.

2.4. Conclusion of Chapter 2

This study redefines an optimal reduct for decision tables and designs an attribute reduction algorithm for fixed decision tables (ARPD). Experimental results demonstrate that the proposed algorithm outperforms existing algorithms based on fuzzy rough set and intuitionistic fuzzy rough set models. In addition, to address practical data scenarios, two distance measures for α, β intuitionistic fuzzy partitions are developed, which serve as the foundation for proposing incremental algorithms to handle decision tables with object set addition or removal.

Chapter 3: Proposing attribute reduction algorithms based on the intuitionistic fuzzy weighted neighborhood rough set model

3.1. Introduction

Chapter 3 focuses on proposing some new solutions with the following main contributions: (1) Proposing a weighted intuitionistic fuzzy neighborhood rough set model, in which each object is characterized not only by membership degree but also by non-membership degree and hesitation degree, (2) Constructing a distance measure for weighted intuitionistic fuzzy neighborhood families and redefining an effective reduct, (3) Defining attribute significance as a criterion for selecting important attributes, (4) Proposing an attribute reduction algorithm for fixed decision tables based on a filter-based approach, (5) Developing two incremental computation formulas based on the distance measure of weighted intuitionistic fuzzy neighborhood families for dynamic decision tables, (6) Proposing two incremental attribute reduction algorithms for dynamic decision tables, (7) Demonstrating the efficiency of the proposed algorithms through comparisons with some algorithms based on fuzzy rough set and weighted neighborhood rough set models. The research results of this chapter have been published in works [CT6] and [CT7] listed in the List of Publications of the dissertation.

3.2. Intuitionistic fuzzy weighted neighborhood rough set model

3.2.1 Concept of intuitionistic fuzzy weighted neighborhood rough sets

The initial study aims to assign a weight to each object within an information granule based on two components characterizing an intuitionistic fuzzy set. Specifically, the weight of each object $u_i \in [u]_A^{\delta,\omega}$ is denoted by $[\ddot{u}]_A^{\delta,\omega}(u_i) = \left(\gamma_{[\ddot{u}]_A^{\delta,\omega}}(u_i), \eta_{[\ddot{u}]_A^{\delta,\omega}}(u_i)\right)$, where $\gamma_{[\ddot{u}]_A^{\delta,\omega}}(u_i)$ and $\eta_{[\ddot{u}]_A^{\delta,\omega}}(u_i)$ represent the membership and non-membership degrees of object u_i in the information granule $[u]_A^{\delta,\omega}$, respectively, which are defined as follows:

- $\gamma_{[\ddot{u}]_A^{\delta,\omega}}(u_i) = 1 - \Delta_A^{\omega}(u, u_i)$, where $\Delta_A^{\omega}(u, u_i)$ denotes the weighted distance between u and u_i defined in Equation 1.3 for the case $p = \infty$,

- $\eta_{[\ddot{u}]_A^{\delta,\omega}}(u_i) = \dfrac{1 - \gamma_{[\ddot{u}]_A^{\delta,\omega}}(u_i)}{1 + w^o \times \gamma_{[\ddot{u}]_A^{\delta,\omega}}(u_i)}$, where $w^o = \dfrac{1}{|U|}\sum\limits_{u \in U} \dfrac{\left|[u]_C^{\delta,\omega} \cap [u]_D\right|}{\left|[u]_C^{\delta,\omega}\right|}$ is the *average classification ratio* based on the weighted neighborhood information granules of C.

For objects that do not belong to the weighted neighborhood information granule, $[\ddot{u}]_A^{\delta,\omega}(u_i) = (0, 1)$. By combining all these evaluations, we obtain a new information granule $[\ddot{u}]_A^{\delta,\omega} = \left\{[\ddot{u}]_A^{\delta,\omega}(u_1), \ldots, [\ddot{u}]_A^{\delta,\omega}\left(u_{|U|}\right)\right\}$, is referred to as an intuitionistic fuzzy weighted neighborhood information granule. Accordingly, considering all objects in U, the family $\mathcal{F}_A^{\delta,\omega} = \left\{[\ddot{u}]_A^{\delta,\omega} : u \in U\right\}$ is formed and is called the intuitionistic fuzzy weighted neighborhood family induced by attribute set A.

The lower and upper approximations of a set $X \subseteq U$ based on the weighted intuitionistic fuzzy neighborhood information granules with respect to A, denoted by $\underline{IWA}(X)$ and $\overline{IWA}(X)$, respectively, are defined as follows.

$$\underline{IW_A}(X) = \left\{ u \in U : \frac{\left|[\ddot{u}]_A^{\delta,\omega} \cap X\right|}{\left|[\ddot{u}]_A^{\delta,\omega}\right|} \geq \ddot{\alpha} \right\} \tag{3.1}$$

$$\overline{IW_A}(X) = \left\{ u \in U : \frac{\left|[\ddot{u}]_A^{\delta,\omega} \cap X\right|}{\left|[\ddot{u}]_A^{\delta,\omega}\right|} \geq \ddot{\beta} \right\} \tag{3.2}$$

where $\ddot{\alpha}$ and $\ddot{\beta}$ are parameters satisfying $0 \leq \ddot{\beta} \leq \ddot{\alpha} \leq 1$.

Proposition 3.1. Let $IS = (U, C \cup D)$ and $A, B \subseteq C$. If $p = \infty$, then $[\ddot{u}]_{A\cup B}^{\delta,\omega} = [\ddot{u}]_A^{\delta,\omega} \cap [\ddot{u}]_B^{\delta,\omega}$, for all $u \in U$.

3.2.2 Some Properties of IFWNRs

Proposition 3.2. Given a decision table $IS = (U, C \cup D)$, $A \subseteq C$ and $X \in U/D$. If $p = 2$, $\ddot{\alpha} = 1$ and $\ddot{\beta} = 0$, then $\underline{IW_A}(X) = \underline{WN_A}(X)$ and $\overline{IW_A}(X) = \overline{WN_A}(X)$.

Proposition 3.2 indicates that WNRs is a special case of IFWNRs.

Proposition 3.3. Let $IS = (U, C \cup D)$ and $A, B \subseteq C$. If $A \subseteq B$, then
1. $\forall u \in U$, $[\ddot{u}]_B^{\delta,\omega} \subseteq [\ddot{u}]_A^{\delta,\omega}$,
2. If $\ddot{\beta} = 0$, then $\forall X \subseteq U$, $\overline{IW_B}(X) \subseteq \overline{IW_A}(X)$,
3. If $\ddot{\alpha} = 1$, then $\forall X \subseteq U$, $\underline{IW_A}(X) \subseteq \underline{IW_B}(X)$.

3.3. Proposed attribute reduction algorithm based on IFWNRs

This section defines two types of reducts based on the IFWNRs approach and analyzes their effectiveness, providing a foundation for designing attribute reduction methods for fixed decision tables.

3.3.1 Proposed attribute reduction algorithm for fixed decision tables

Proposition 3.4. Given a decision table $IS = (U, C \cup D)$, the distance between two intuitionistic fuzzy weighted neighborhood families $\mathcal{F}_C^{\delta,\omega}$ and $\mathcal{F}_{C\cup D}^{\delta,\omega}$ is defined as follows:

$$\ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega}, \mathcal{F}_{C\cup D}^{\delta,\omega}\right) = \sum_{u\in U} \frac{\left(\left|[\ddot{u}]_C^{\delta,\omega}\right| - \left|[\ddot{u}]_C^{\delta,\omega} \cap [\ddot{u}]_D^{\delta,\omega}\right|\right)}{|U|^2} \tag{3.3}$$

Proposition 3.5. Given a decision table $IS = (U, C \cup D)$ and a subset of condition attributes $A \subseteq C$, if $\delta_1 \leq \delta_2$, then $\ddot{\mathcal{D}}\left(\mathcal{F}_A^{\delta_1,\omega}, \mathcal{F}_{A\cup D}^{\delta_1,\omega}\right) \leq \ddot{\mathcal{D}}\left(\mathcal{F}_A^{\delta_2,\omega}, \mathcal{F}_{A\cup D}^{\delta_2,\omega}\right)$.

Proposition 3.6. Given a decision table $IS = (U, C \cup D)$ and two subsets of attributes $A, B \subseteq C$. If $A \subseteq B$, then $\ddot{\mathcal{D}}\left(\mathcal{F}_A^{\delta,\omega}, \mathcal{F}_{A\cup D}^{\delta,\omega}\right) \geq \ddot{\mathcal{D}}\left(\mathcal{F}_B^{\delta,\omega}, \mathcal{F}_{B\cup D}^{\delta,\omega}\right)$.

Definition 3.1. Given a decision table $IS = (U, C \cup D)$, a subset of attributes $A \subseteq C$ is called a $\ddot{\mathcal{D}}$-reduct of C if the following conditions hold:
1. $\ddot{\mathcal{D}}\left(\mathcal{F}_A^{\delta,\omega}, \mathcal{F}_{A\cup D}^{\delta,\omega}\right) = \ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega}, \mathcal{F}_{C\cup D}^{\delta,\omega}\right)$,
2. $\ddot{\mathcal{D}}\left(\mathcal{F}_A^{\delta,\omega}, \mathcal{F}_{A\cup D}^{\delta,\omega}\right) > \ddot{\mathcal{D}}\left(\mathcal{F}_{A\backslash\{a\}}^{\delta,\omega}, \mathcal{F}_{A\backslash\{a\}\cup D}^{\delta,\omega}\right)$.

Definition 3.2. Given a decision table $IS = (U, C \cup D)$, a subset of attributes $A \subseteq C$ and an attribute $a \in C\backslash A$, the significance of a with respect to A based on the intuitionistic fuzzy weighted neighborhood family distance, denoted by $Sig_F(a, A)$, is defined as follows:

$$Sig_F(a, A) = \ddot{\mathcal{D}}\left(\mathcal{F}_{A\backslash\{a\}}^{\delta,\omega}, \mathcal{F}_{A\backslash\{a\}\cup D}^{\delta,\omega}\right) - \ddot{\mathcal{D}}\left(\mathcal{F}_A^{\delta,\omega}, \mathcal{F}_{A\cup D}^{\delta,\omega}\right) \tag{3.4}$$

Based on the above analysis, the ARIFW algorithm is proposed to find a reduct in decision tables, with a computational complexity of $O\left(|C|^2|U|^2\right)$.

Algorithm 3.1 Attribute reduction based on IFWNRs (ARIFW)

Input: A decision table $IS = (U, C \cup D)$ and a neighborhood radius δ.
Output: One reduct $\mathcal{B}$

compute the weight of each attribute by Equation 1.2
compute $\ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega}, \mathcal{F}_{C\cup D}^{\delta,\omega}\right)$
for $a \in C$ **do**
  compute $\mathcal{F}_{\{a\}}^{\delta,\omega}$
  compute $\ddot{\mathcal{D}}\left(\mathcal{F}_{\{a\}}^{\delta,\omega}, \mathcal{F}_{\{a\}\cup D}^{\delta,\omega}\right)$
end for
$\mathcal{B} = \{a_0\}$ which satisfies: $\ddot{\mathcal{D}}\left(\mathcal{F}_{\{a_0\}}^{\delta,\omega}, \mathcal{F}_{\{a_0\}\cup D}^{\delta,\omega}\right) = \min_{a\in C} \ddot{\mathcal{D}}\left(\mathcal{F}_{\{a\}}^{\delta,\omega}, \mathcal{F}_{\{a\}\cup D}^{\delta,\omega}\right)$
while $\ddot{\mathcal{D}}\left(\mathcal{F}_{\mathcal{B}}^{\delta,\omega}, \mathcal{F}_{\mathcal{B}\cup D}^{\delta,\omega}\right) > \ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega}, \mathcal{F}_{C\cup D}^{\delta,\omega}\right)$ **do**
  compute $Sig_F(a, \mathcal{B})$, for all $a \in C \backslash \mathcal{B}$
  select a_0 which satisfies: $Sig_F(a_0, \mathcal{B}) = \max_{a\in C\backslash\mathcal{B}} \{Sig_F(a, \mathcal{B})\}$
  $\mathcal{B} \leftarrow \mathcal{B} \cup \{a_0\}$
end while
return $\mathcal{B}$

3.3.2 Proposed attribute reduction algorithm for dynamic decision tables with object set addition

Suppose that a new object set $\Delta U = (u_{n+1}, u_{n+2}, \ldots, u_{n+z})$ is added to U, then, the process of updating the weight of an attribute $a \in C$ for objects $u_j \in U^+ = U \cup \Delta U$ within a weighted neighborhood information granule $[u_i]_{\{a\}}^{\delta,\omega}$ is defined as follows:

$$\omega^+ = \begin{cases} \omega & \text{if } 1 \le i, j \le n \\ \omega^\Delta & \text{otherwise} \end{cases} \tag{3.5}$$

where ω denotes the original weight generated from the initial object set, while ω^Δ represents the new weight applied to the added object set ΔU. The update mechanism of the intuitionistic fuzzy weighted neighborhood information granule is given by

$$[\ddot{u}_i]_{\{a\}}^{\delta,\omega^+}\left(u_j\right) = \begin{cases} [\ddot{u}_i]_a^{\delta,\omega}\left(u_j\right) & \text{if } 1 \le i, j \le n \\ [\ddot{u}_i]_a^{\delta,\omega^\Delta}\left(u_j\right) & \text{if } u_j \in [u_i]_{\{a}^{\delta,\omega^\Delta} \; (n+1 \le i \le n+z) \vee (n+1 \le j \le n+z) \\ (0,1) & \text{otherwise} \end{cases} \tag{3.6}$$

Based on the updated intuitionistic fuzzy weighted neighborhood information granules, a new intuitionistic fuzzy weighted neighborhood family induced by attribute a is formed, denoted by $\mathcal{F}_{\{a\}}^{\delta,\omega^+}$.

Proposition 3.7. Given a decision table $IS = (U, C \cup D)$ with $U = (u_1, u_2, \ldots, u_n)$ and a new object set $\Delta U = (u_{n+1}, u_{n+2}, \ldots, u_{n+z})$, where $z \ge 1$, the distance between two intuitionistic fuzzy weighted neighborhood families $\mathcal{F}_C^{\delta,\omega^+}$ and $\mathcal{F}_{C\cup D}^{\delta,\omega^+}$ in the object set $U^+ = U \cup \Delta U$, denoted $\ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega^+}, \mathcal{F}_{C\cup D}^{\delta,\omega^+}\right)$, is computed as follows:

$$\ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega^+}, \mathcal{F}_{C\cup D}^{\delta,\omega^+}\right) = \frac{n^2 \ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega}, \mathcal{F}_{C\cup D}^{\delta,\omega}\right) + 2\sum_{i=1}^{z}\left(\left|[\ddot{u}_{n+i}]_C^{\delta,\omega^+}\right| - \left|[\ddot{u}_{n+i}]_C^{\delta,\omega^+} \cap [\ddot{u}_{n+i}]_D^{\delta,\omega^+}\right| - \theta_i\right)}{(n+z)^2} \tag{3.7}$$

where $\theta_1 = 0$ and for all $i \ge 2$,

$$\theta_i = \sum_{j=1}^{z-1} \left(\gamma_{[\ddot{u}_{n+i}]_C^{\delta,\omega^+}} \left(u_{n+j+1}\right) - \gamma_{[\ddot{u}_{n+i}]_{C\cup D}^{\delta,\omega^+}} \left(u_{n+j+1}\right) - \eta_{[\ddot{u}_{n+i}]_C^{\delta,\omega^+}} \left(u_{n+j+1}\right) + \eta_{[\ddot{u}_{n+i}]_{C\cup D}^{\delta,\omega^+}} \left(u_{n+j+1}\right) \right).$$

Proposition 3.8. Given a decision table $IS = (U, C \cup D)$ with $U = (u_1, u_2, \ldots, u_n)$, $A \subseteq C$ is a reduct based on intuitionistic fuzzy weighted neighborhood family distance on U, and an incremental object set $\Delta U = (u_{n+1}, u_{n+2}, \ldots, u_{n+z})$ is added to U, where $z \geq 1$, then we have:
1. If all objects in ΔU have the same value of decision attribute then:

$$\ddot{\mathcal{D}} \left(\mathcal{F}_C^{\delta,\omega^+}, \mathcal{F}_{C\cup D}^{\delta,\omega^+} \right) = \frac{n^2 \ddot{\mathcal{D}} \left(\mathcal{F}_C^{\delta,\omega}, \mathcal{F}_{C\cup D}^{\delta,\omega} \right) + 2 \sum\limits_{i=1}^{z} \left(\left| [\ddot{u}_{n+i}]_C^{\delta,\omega^+} \right| - \left| [\ddot{u}_{n+i}]_C^{\delta,\omega^+} \cap [\ddot{u}_{n+i}]_D^{\delta,\omega^+} \right| \right)}{(n+z)^2} \tag{3.8}$$

2. If $[\ddot{u}_{n+i}]_B^{\delta,\omega^+} \subseteq [\ddot{u}_{n+i}]_D^{\delta,\omega^+}$ with $i = 1, 2, \ldots, z$, then $\ddot{\mathcal{D}} \left(\mathcal{F}_A^{\delta,\omega^+}, \mathcal{F}_{A\cup D}^{\delta,\omega^+} \right) = \ddot{\mathcal{D}} \left(\mathcal{F}_C^{\delta,\omega^+}, \mathcal{F}_{C\cup D}^{\delta,\omega^+} \right)$.

Algorithm 3.2 Incremental Attribute Reduction with IFWNR when Adding Object set (IARIF-AO)

Input: A decision $IS = (U, C \cup D)$ with $U = \{u_1, u_2, \ldots, u_n\}$, the reduct $\mathcal{B} \subseteq C$, $\mathcal{F}_{\mathcal{B}}^{\delta,\omega}$, $\mathcal{F}_C^{\delta,\omega}$, and added set of objects $\Delta U = \{u_{n+1}, u_{n+2}, \ldots, u_{n+z}\}$.
Output: the approximation reduct $\mathcal{B}^+$ on $U^+ = U \cup \Delta U$

// initialize and update
$\mathcal{B}^+ \leftarrow \mathcal{B}$
update ω^+ from Equation 3.5 and $\mathcal{F}_{\{a\}}^{\delta,\omega^+}$ on U^+ for each $a \in C \backslash \mathcal{B}^+$ from Equation 3.6.
// check the set of added objects
$Z \leftarrow \Delta U$
for $i = 1$ to z **do**
  if $[\ddot{u}_{n+i}]_{\mathcal{B}^+}^{\delta,\omega^+} \subseteq [u_{n+i}]_D$ **then** $Z \leftarrow Z \backslash \{u_{n+i}\}$
end for
if $Z = \emptyset$ **then** return $\mathcal{B}^+$
set $\Delta U \leftarrow Z$, $z \leftarrow |\Delta U|$
// finding the reduct
compute $\ddot{\mathcal{D}} \left(\mathcal{F}_{\mathcal{B}^+}^{\delta,\omega^+}, \mathcal{F}_{\mathcal{B}^+\cup D}^{\delta,\omega^+} \right)$ và $\ddot{\mathcal{D}} \left(\mathcal{F}_C^{\delta,\omega^+}, \mathcal{F}_{C\cup D}^{\delta,\omega^+} \right)$ by Equation 3.2.
while $\ddot{\mathcal{D}} \left(\mathcal{F}_{\mathcal{B}^+}^{\delta,\omega^+}, \mathcal{F}_{\mathcal{B}^+\cup D}^{\delta,\omega^+} \right) > \ddot{\mathcal{D}} \left(\mathcal{F}_C^{\delta,\omega^+}, \mathcal{F}_{C\cup D}^{\delta,\omega^+} \right)$ **do**
  compute $Sig_F \left(a, \mathcal{B}^+\right)$, for all $a \in C \backslash \mathcal{B}^+$ by Equation 3.4.
  select attribute a_0 which satisfies: $Sig_F \left(a_0, \mathcal{B}^+\right) = \max\limits_{a \in C \backslash \mathcal{B}^+} \left\{ Sig_F \left(a, \mathcal{B}^+\right) \right\}$
  $\mathcal{B}^+ \leftarrow \mathcal{B}^+ \cup \{a_0\}$
end while
// remove redundant attributes
for $a \in \mathcal{B}^+$ **do**
  compute $\ddot{\mathcal{D}} \left(\mathcal{F}_{\mathcal{B}^+\backslash\{a\}}^{\delta,\omega^+}, \mathcal{F}_{\mathcal{B}^+\backslash\{a\}\cup D}^{\delta,\omega^+} \right)$ by Equation 3.3.
  if $\ddot{\mathcal{D}} \left(\mathcal{F}_{\mathcal{B}^+\backslash\{a\}}^{\delta,\omega^+}, \mathcal{F}_{\mathcal{B}^+\backslash\{a\}\cup D}^{\delta,\omega^+} \right) = \ddot{\mathcal{D}} \left(\mathcal{F}_{\mathcal{B}^+}^{\delta,\omega^+}, \mathcal{F}_{\mathcal{B}^+\cup D}^{\delta,\omega^+} \right)$ **then** $\mathcal{B}^+ \leftarrow \mathcal{B}^+ \backslash \{a\}$
end for
return $\mathcal{B}^+$

The computational complexity of IARIF-AO is
$\max \left\{ O\left(\left|\mathcal{B}^+\right| \left|U^+\right| |\Delta U| \right), O\left(\left(|C| - \left|\mathcal{B}^+\right| \right)^2 \left|U^+\right| |\Delta U| \right) \right\}$.

The algorithm consists of four main stages: initialization and updating, checking the added object set, filter-based reduct search, and removal of redundant attributes.

3.3.3 Proposed attribute reduction algorithm for dynamic decision tables with object set deletion

Proposition 3.9. Given a decision table $IS = (U, C \cup D)$ with $U = \{u_1, u_2, \ldots, u_n\}$ and an object set $\Delta U = \{u_n, u_{n-1}, \ldots, u_{n-z+1}\}$ is removed from U, where $z \geq 1$, the distance between two intuitionistic fuzzy weighted neighborhood families on $U^- = U \backslash \Delta U$, denoted $\ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega^-}, \mathcal{F}_{C\cup D}^{\delta,\omega^-}\right)$, is computed as follows:

$$\mathcal{D}\left(\mathcal{F}_C^{\delta,\omega^-}, \mathcal{F}_{C\cup D}^{\delta,\omega^-}\right) = \frac{n^2\mathcal{D}\left(\mathcal{F}_C^{\delta,\omega}, \mathcal{F}_{C\cup D}^{\delta,\omega}\right) - 2\sum_{i=1}^{z}\left(\left|[\ddot{u}_{n-i+1}]_C^{\delta,\omega}\right| - \left|[\ddot{u}_{n-i+1}]_C^{\delta,\omega} \cap [\ddot{u}_{n-i+1}]_D^{\delta,\omega}\right| - \zeta_i\right)}{(n-z)^2} \tag{3.9}$$

where $\zeta_1 = 0$ and for all $i \geq 2$,

$$\zeta_i = \sum_{j=i}^{z}\left(\gamma_{[\ddot{u}_{n-i+1}]_C^{\delta,\omega}}\left(u_{n-j+1}\right) - \gamma_{[\ddot{u}_{n-i+1}]_{C\cup D}^{\delta,\omega}}\left(u_{n-j+1}\right) - \eta_{[\ddot{u}_{n-i+1}]_C^{\delta,\omega}}\left(u_{n-j+1}\right) + \eta_{[\ddot{u}_{n-i+1}]_{C\cup D}^{\delta,\omega}}\left(u_{n-j+1}\right)\right)$$

Proposition 3.10. Given a decision table $IS = (U, C \cup D)$ with $U = \{u_1, u_2, \ldots, u_n\}$, $A \subseteq C$ is a reduct based on intuitionistic fuzzy weighted neighborhood family distance on U, a set of objects $\Delta U = \{u_n, u_{n-1}, \ldots, u_{n-z+1}\}$ is removed from U, where $z \geq 1$ and $U^- = U \backslash \Delta U$ is the new object set, then we have:
1. If all objects in ΔU have the same value of decision attribute then:

$$\ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega^-}, \mathcal{F}_{C\cup D}^{\delta,\omega^-}\right) = \frac{n^2\mathcal{D}\left(\mathcal{F}_C^{\delta,\omega}, \mathcal{F}_{C\cup D}^{\delta,\omega}\right) - 2\sum_{i=1}^{z}\left(\left|[\ddot{u}_{n-i+1}]_C^{\delta,\omega}\right| - \left|[\ddot{u}_{n-i+1}]_C^{\delta,\omega} \cap [\ddot{u}_{n-i+1}]_D^{\delta,\omega}\right|\right)}{(n-z)^2} \tag{3.10}$$

2. $\ddot{\mathcal{D}}\left(\mathcal{F}_A^{\delta,\omega^-}, \mathcal{F}_{A\cup D}^{\delta,\omega^-}\right) = \ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega^-}, \mathcal{F}_{C\cup D}^{\delta,\omega^-}\right)$

Algorithm 3.3 Incremental Attribute Reduction with IFWNR when Deleting Object set (IARIF-DO)

Input: A decision $IS = (U, C \cup D)$ with $U = \{u_1, u_2, \ldots, u_n\}$, the reduct $\mathcal{B} \subseteq C$, $\mathcal{F}_\mathcal{B}^{\delta,\omega}$ and $\mathcal{F}_C^{\delta,\omega}$, a deleted set of objects $\Delta U = \{u_n, u_{n-1}, \ldots, u_{n-z+1}\}$
Output: One reduct red^- on $U^- = U \backslash \Delta U$

initialize $\mathcal{B}^- \leftarrow \mathcal{B}$
update $\mathcal{F}_C^{\delta,\omega^-}$
compute $\ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega^-}, \mathcal{F}_{C\cup D}^{\delta,\omega^-}\right)$ by Equation 3.9.
$\ddot{\mathcal{D}}\left(\mathcal{F}_{\mathcal{B}^-}^{\delta,\omega^-}, \mathcal{F}_{\mathcal{B}^-\cup D}^{\delta,\omega^-}\right) \leftarrow \ddot{\mathcal{D}}\left(\mathcal{F}_C^{\delta,\omega^-}, \mathcal{F}_{C\cup D}^{\delta,\omega^-}\right)$
for $a \in \mathcal{B}^-$ **do**
    compute $\ddot{\mathcal{D}}\left(\mathcal{F}_{\mathcal{B}^-\backslash\{a\}}^{\delta,\omega^-}, \mathcal{F}_{\mathcal{B}^-\backslash\{a\}\cup D}^{\delta,\omega^-}\right)$ by Equation 3.3.
    if $\ddot{\mathcal{D}}\left(\mathcal{F}_{\mathcal{B}^-\backslash\{a\}}^{\delta,\omega^-}, \mathcal{F}_{\mathcal{B}^-\backslash\{a\}\cup D}^{\delta,\omega^-}\right) = \ddot{\mathcal{D}}\left(\mathcal{F}_{\mathcal{B}^-}^{\delta,\omega^-}, \mathcal{F}_{\mathcal{B}^-\cup D}^{\delta,\omega^-}\right)$ **then** $\mathcal{B}^- \leftarrow \mathcal{B}^- \backslash \{a\}$
end for
return $\mathcal{B}^-$

The computational complexity of Algorithm IARIF-DO is $O\left(\left|\mathcal{B}^-\right| \left|U^-\right| |\Delta U|\right)$. With this complexity, the execution time of the proposed incremental algorithm IARIF-DO is significantly reduced compared to ARIFW.

3.4. Experimental and evaluation of the proposed algorithms

3.4.1 Performance of the IARIF-AO Algorithm

To evaluate the effectiveness of the proposed algorithm, the datasets listed in Table 3.1 are first divided into two approximately equal subsets, denoted as U_{ori} and U_{inc}. Specifically, U_{ori} is used for attribute reduction algorithms on fixed decision tables, including NIFS [82], WNRS [29], W-MGMN [86], FMIFRFS [56], ARPD, and ARIFW. Among these, NIFS is a fuzzy rough set–based algorithm, WNRS and W-MGMN are neighborhood rough set–based algorithms, while FMIFRFS and ARPD are based on intuitionistic fuzzy rough set models.

Next, the dataset U_{inc} is further divided into five approximately equal parts. Each part is then incrementally added to U_{ori} to evaluate the performance of the incremental algorithms. As described previously, for each dataset, the algorithms scan parameter values from 0 to 1 to identify the subset of attributes that yields the highest classification accuracy.

Table 3.1: Datasets used for evaluating IARIF-AO and several algorithms

ID	Datasets	Objects	$\|U_{ori}\|$	$\|U_{inc}\|$	Attributes	Classes	Source
1	Leaf	340	170	170	15	30	UCI
2	Ionosphere	351	175	176	34	2	UCI
3	Vehicle	846	423	423	18	4	OpenML
4	Vowel	990	495	495	12	11	OpenML
5	German	1000	500	500	20	2	UCI
6	LSVT	126	63	63	310	2	OpenML
7	Leukemia	72	36	36	7129	2	UCI
8	Mfeat	2000	1000	1000	76	10	OpenML
9	PD	756	378	378	754	2	UCI
10	Segmentation	2310	1155	1155	19	7	OpenML
11	Wall	5456	2728	2728	24	4	OpenML
12	Waveform2	5000	2500	2500	40	3	OpenML

Table 3.2 reports the sizes of the reducts and the corresponding parameter settings obtained by each algorithm. The FMIFRFS, ARPD, and ARIFW algorithms generally produce smaller reducts than the other methods. When considering all cases, ARPD and ARIFW most frequently achieve the smallest reducts across the majority of datasets. Specifically, ARPD yields the smallest reducts on four datasets, while ARIFW does so on six datasets. In addition, ARIFW achieves the smallest average reduct size among all compared algorithms. These results indicate that the proposed algorithm is highly effective in attribute reduction, particularly for high-dimensional datasets.

Table 3.3 presents the computational time of the attribute reduction algorithms when applied to the dataset U_{ori}. Clearly, algorithms based on weighted neighborhood rough sets and fuzzy rough sets exhibit faster execution times than those based on intuitionistic fuzzy rough set models. This is mainly because the WNRS algorithm performs computations on crisp sets, which involve simpler and more efficient processing. Although the NIFS algorithm is based on the fuzzy rough set framework, it does not require the construction of relation matrices and employs a relatively simple measure to select important attributes. In contrast, WNRS involves an additional process of computing attribute weights. As a result, NIFS demonstrates the fastest execution speed among all the compared algorithms.

Table 3.2: Reduct sizes and parameters of the algorithms using KNN on U_{ori}

Datasets	NIFS		WNRS		W-MGMN		FMIFRFS		ARPD		ARIFW	
	k	$\lvert\mathcal{B}\rvert$	*delta*	$\lvert\mathcal{B}\rvert$	*delta*	$\lvert\mathcal{B}\rvert$	*epsilon*	$\lvert\mathcal{B}\rvert$	*alpha*	$\lvert\mathcal{B}\rvert$	*delta*	$\lvert\mathcal{B}\rvert$
Leaf	0.002	10	0.15	9	0.05	9	0.75	10	0.4	9	0.3	7
Ionosphere	0.001	9	0.1	6	0.1	9	1.0	4	0.75	5	0.05	4
Vehicle	0.001	13	0.1	14	0.0	10	0.7	17	0.7	8	0.15	10
Vowel	0.002	6	0.1	9	0.05	8	0.35	8	0.5	7	0.3	7
German	0.002	6	0.05	12	0.0	9	0.55	12	0.05	6	0.7	10
LSVT	0.001	6	0.6	10	0.3	18	0.05	18	0.8	7	0.85	6
Leukemia	0.001	2	0.05	2	0.1	10	0.85	3	0.4	7	0.55	3
Mfeat	0.003	56	0.35	27	0.05	21	0.5	23	0.6	22	0.55	13
PD	0.001	25	0.65	21	0.1	6	0.7	31	0.5	41	0.9	9
Segmentation	0.001	14	0.05	9	0.05	13	0.9	10	0.8	6	0.6	6
Wall	0.006	8	0.05	14	0.2	15	1.0	4	0.9	3	0.05	3
Waveform2	0.003	26	0.05	8	0.05	31	0.45	15	0.2	24	0.35	13
Average		16.08		11.75		13.25		12.92		12.08		7.58

Table 3.3: Execution time comparison of ARIFW and other algorithms on U_{ori}

ID	Datasets	NIFS	WNRS	W-MGMN	FMIFRFS	ARPD	ARIFW
1	Leaf	0.0028	0.0534	1.4969	0.1242	0.0186	0.0152
2	Ionosphere	0.0033	0.1806	3.4146	0.1257	0.0192	0.0186
3	Vehicle	0.0074	0.2942	14.757	0.7542	0.3798	0.3759
4	Vowel	0.0051	0.2253	22.159	1.2874	0.2784	0.2519
5	German	0.0048	0.6966	21.973	0.6743	0.5128	0.5236
6	LSVT	0.0743	0.3816	2.7744	1.9096	0.7639	0.7483
7	Leukemia	13.433	6.6504	38.491	231.83	229.15	209.56
8	Mfeat	6.8312	33.086	60.789	36.674	23.382	17.947
9	PD	2.9826	61.321	307.52	103.65	108.68	28.183
10	Segmentation	1.0652	4.0692	214.63	10.687	5.7058	4.6415
11	Wall	1.0913	29.779	598.78	15.774	8.0352	7.8904
12	Waveform2	1.6991	56.158	626.06	57.093	72.449	43.865
Average		2.2667	16.075	109.40	38.382	37.448	26.168

The comparison results of classification accuracy for attribute reduction algorithms on fixed decision tables are presented in Table 3.4. In 12 cases, the reducts obtained by the proposed algorithm achieve better classification performance than the original datasets. Moreover, when using KNN, the average classification accuracy of the proposed algorithm reaches 82.18

For datasets with low initial classification accuracy and large numbers of objects, methods based on intuitionistic fuzzy rough sets consistently return reducts with superior classification performance compared to those based on fuzzy rough sets and weighted neighborhood rough sets. This highlights the advantage of intuitionistic fuzzy rough sets, where the incorporation of the non-membership function effectively reduces the influence of noisy objects in the data.

It can be observed that, on most datasets, the reducts generated by the ARIFW algorithm yield better classification results than those produced by other algorithms, even though ARIFW selects fewer attributes. Furthermore, the proposed algorithm achieves the highest classification accuracy in 8 out of 12 cases, demonstrating its effectiveness in selecting important attributes across diverse datasets.

Table 3.4: Comparison of classification accuracy between ARIFW and other algorithms on U_{ori}.

Datasets	Raw data	NIFS	WNRS	W-MGMN	FMIFRFS	ARPD	ARIFW
Leaf	71.76 ± 10.12	74.12 ± 12.39	75.88 ± 8.5	$\mathbf{79.41 \pm 9.58}$	72.35 ± 11.78	74.12 ± 9.56	73.53 ± 9.58
Ionosphere	76.57 ± 9.22	82.19 ± 8.94	80.59 ± 8.46	82.81 ± 5.21	84.02 ± 11.23	84.05 ± 9.73	$\mathbf{84.97 \pm 9.94}$
Vehicle	67.85 ± 4.23	67.40 ± 5.82	67.60 ± 4.61	71.87 ± 7.05	68.08 ± 4	67.84 ± 5.51	$\mathbf{74.46 \pm 6.91}$
Vowel	48.81 ± 12.36	61.53 ± 11.32	$\mathbf{65.44 \pm 5.71}$	59.14 ± 8.42	50.43 ± 13.99	52.43 ± 15.45	59.57 ± 9.55
German	73.20 ± 4.21	72.00 ± 6.13	72.80 ± 4.02	74.80 ± 4.12	75.20 ± 3.12	73.60 ± 3.2	$\mathbf{75.40 \pm 4.9}$
LSVT	74.30 ± 14.9	79.29 ± 11.67	$\mathbf{90.48 \pm 10.81}$	79.05 ± 12.77	86.90 ± 9.95	88.81 ± 10.49	$\mathbf{90.48 \pm 7.82}$
Leukemia	86.67 ± 13.54	91.67 ± 12.91	96.67 ± 10	90.00 ± 22.91	$\mathbf{97.50 \pm 7.5}$	96.67 ± 10	$\mathbf{97.50 \pm 7.5}$
Mfeat	92.50 ± 1.8	91.00 ± 1.9	92.00 ± 2.86	93.30 ± 1.62	94.50 ± 1.91	94.90 ± 1.64	$\mathbf{95.30 \pm 1.73}$
PD	79.11 ± 8.42	75.41 ± 5.97	$\mathbf{86.54 \pm 7.38}$	79.11 ± 8.42	83.88 ± 8.9	83.88 ± 6.98	83.36 ± 5.47
Segmentation	93.77 ± 2.45	92.82 ± 2.37	93.00 ± 3.08	93.51 ± 2.11	93.69 ± 2.37	93.43 ± 2.11	$\mathbf{93.78 \pm 2.33}$
Wall	63.78 ± 9.39	68.37 ± 8.91	64.59 ± 8.29	71.30 ± 7.82	74.63 ± 6.44	75.00 ± 4.47	$\mathbf{76.80 \pm 5.15}$
Waveform2	78.20 ± 2.02	78.28 ± 2.21	72.96 ± 2.63	79.40 ± 1.92	$\mathbf{81.72 \pm 1.3}$	79.96 ± 2.19	81.04 ± 2.8
Average	74.54 ± 7.72	77.84 ± 7.55	79.88 ± 6.36	79.48 ± 7.66	80.24 ± 6.87	80.39 ± 6.78	$\mathbf{82.18 \pm 6.14}$
Win/Draw/Loss	12/0/0	10/0/2	8/1/3	11/0/1	9/1/2	10/0/2	*best*

Next, the IARIF-AO algorithm is compared with AIFSA-FKD [82], W-MGMA [86], and IARPD-AO. These algorithms are evaluated on datasets obtained by incrementally adding object subsets U_1 to U_5 into U_{ori}. The experimental results are analyzed as follows.

- *Runtime:* Incremental algorithms exhibit superior performance compared to their corresponding algorithms operating on fixed decision tables, even though they handle larger datasets. Among the algorithms based on the IFRSs model, the proposed algorithm achieves faster execution in most incremental stages. This advantage mainly stems from the characteristics of the weighted neighborhood relation, which effectively reduces the computational space.

- *Reduct size:* Across all cases, the reducts obtained by IARPD-AO and IARIF-AO achieve the smallest sizes in 20 and 43 cases, respectively. These results further highlight the effectiveness of the proposed method in eliminating redundant attributes from decision tables during incremental steps.

- *Classification accuracy:* Algorithms based on intuitionistic fuzzy rough set models generally outperform the others on datasets with low initial classification accuracy and large numbers of objects. Intuitively, these datasets often contain objects with distributions that differ significantly from the majority of other objects, which degrades the performance of classification models. The IARIF-AO demonstrates clear superiority by achieving the highest classification accuracy in 44 out of 60 incremental cases.

3.4.2 Performance of the IARIF-DO Algorithm

This subsection conducts several experiments to evaluate the effectiveness of the incremental algorithm on decision tables with object set deletion. The experiments are performed on standard benchmark datasets described in Table 3.5.

Table 3.5: Datasets used for evaluating IARIF-DO and several algorithms

ID	Datasets	Objects	$\lvert U_{dec} \rvert$	Attributes	Classes	Source
1	Leaf	340	170	15	30	UCI
2	Ionosphere	351	176	34	2	UCI
3	Vehicle	846	423	18	4	OpenML
4	Vowel	990	495	12	11	OpenML
5	German	1000	500	20	2	UCI
6	LSVT	126	63	310	2	OpenML
7	Leukemia	72	36	7129	2	UCI
8	Mfeat	2000	1000	76	10	OpenML
9	PD	756	378	754	2	UCI
10	Segmentation	2310	1155	19	7	OpenML
11	Wall	5456	2728	24	4	OpenML
12	Waveform2	5000	2500	40	3	OpenML

First, the dissertation compares the algorithms NIFS [82], WNRS [29], W-MGMN [86], FMIFRFS [56], ARPD, and ARIFW by extracting an optimal reduct from the original datasets.

Table 3.6 presents the sizes of the obtained reducts and the corresponding parameter settings of each algorithm. For most datasets with low initial classification accuracy, such as Ionosphere, Vehicle, Vowel, German, LSVT, Wall, and Waveform2, the NIFS, WNRS, and W-MGMN algorithms still produce reducts with relatively large sizes. In contrast, algorithms based on the intuitionistic fuzzy rough set model generally extract smaller reducts. When considering all cases, ARPD and ARIFW tend to obtain the smallest reducts. However, the average reduct size produced by ARIFW is smaller than that of ARPD.

Table 3.6: Execution time comparison of ARIFW and other algorithms on U

Datasets	NIFS		WNRS		W-MGMN		FMIFRFS		ARPD		ARIFW	
	k	$\|red\|$	$delta$	$\|red\|$	$delta$	$\|red\|$	$epsilon$	$\|red\|$	$alpha$	$\|red\|$	$delta$	$\|red\|$
Leaf	0.001	12	0.05	9	0.95	11	0.7	13	0.55	13	0.85	11
Ionosphere	0.001	14	0.05	6	0.1	13	0.9	9	0.55	12	0.45	4
Vehicle	0.001	14	0.1	13	0.2	15	0.65	12	0.1	14	0.15	11
Vowel	0.008	10	0.05	7	0.55	11	0.4	8	0.9	6	0.3	7
German	0.005	10	0.1	19	0.15	14	0.05	15	0.9	8	0.75	7
LSVT	0.002	9	0.8	22	0.45	34	0.65	35	0.75	6	0.75	8
Leukemia	0.001	4	0.9	5	0.5	19	0.7	8	0.6	5	0.45	3
Mfeat	0.002	54	0.45	59	0.05	21	0.55	35	0.45	33	0.75	15
PD	0.003	46	0.45	18	0.15	30	0.3	38	0.35	41	0.5	8
Segmentation	0.004	10	0.05	9	0.85	18	0.95	9	0.85	7	0.3	6
Wall	0.005	13	0.1	20	0.2	15	0.05	7	0.9	4	0.05	3
Waveform2	0.002	27	0.05	9	0.45	15	0.5	17	0.2	25	0.35	13
Average		18.58		16.33		18.00		17.17		14.50		8.00

Table 3.7 presents the execution time of attribute reduction algorithms on fixed decision tables. It is evident that as the number of objects in the decision table increases, the proposed algorithm demonstrates increasingly superior computational efficiency compared with intuitionistic fuzzy set–based approaches. Specifically, the average execution time of the proposed algorithm is 56.450 seconds, whereas FMIFRFS and ARPD require 97.748 seconds and 92.472 seconds, respectively.

Table 3.7: Execution time comparison of ARIFW and other algorithms on U

ID	Datasets	NIFS	WNRS	W-MGMN	FMIFRFS	ARPD	ARIFW
1	Leaf	0.0079	0.2311	5.5813	0.6643	0.0756	0.0430
2	Ionosphere	0.0084	0.7941	8.1018	0.2356	0.1619	0.0507
3	Vehicle	0.0248	1.6584	35.647	3.0443	1.7041	1.5401
4	Vowel	0.0186	1.3825	56.821	4.7608	0.7494	0.9378
5	German	0.0204	2.3336	59.884	3.2374	2.8125	2.3473
6	LSVT	0.2605	2.7798	19.387	1.6951	1.1382	1.4905
7	Leukemia	28.880	37.931	92.566	380.93	472.54	372.56
8	Mfeat	16.722	95.189	185.40	136.63	112.73	59.473
9	PD	6.4523	211.73	714.05	211.70	218.54	47.528
10	Segmentation	1.9302	18.528	617.08	28.053	8.2013	7.3485
11	Wall	2.8172	63.196	1328.4	102.22	32.730	27.505
12	Waveform2	3.7391	236.16	1421.8	299.80	258.29	156.58
Average		5.0776	55.993	378.73	97.748	92.472	56.450

According to Table 3.8, on most datasets, the reducts generated by the ARIFW algorithm yield better classification performance than those returned by the other algorithms, even though ARIFW selects fewer attributes. Moreover, the proposed method achieves the highest classification accuracy in 9 out of 12 cases, demonstrating its ability to effectively select important attributes and to handle datasets with diverse characteristics.

Finally, the dissertation presents several comparisons of the performance of the IARIF-DO algorithm with the incremental algorithms AIFSD-FKD [82], W-MGMD [86], and IARPD-RO. These algorithms are evaluated on datasets obtained by sequentially removing object subsets U_1 to U_5 from U. The results are consistent with those of the IARIF-AO

Table 3.8: Comparison of classification accuracy between ARIFW and other algorithms on U.

Datasets	Raw data	NIFS	WNRS	W-MGMN	FMIFRFS	ARPD	ARIFW
Leaf	60.59 ± 6.34	59.41 ± 5.55	69.12 ± 3.01	67.94 ± 4.64	61.47 ± 5.33	61.47 ± 6.76	**69.41 ± 4.78**
Ionosphere	83.76 ± 6.39	84.89 ± 6.02	88.90 ± 6.89	86.04 ± 5.34	90.33 ± 5.27	88.32 ± 7.61	**90.62 ± 6.08**
Vehicle	70.45 ± 1.85	69.39 ± 3.04	69.86 ± 2.02	70.57 ± 2.75	70.10 ± 2.91	70.22 ± 2.83	**71.29 ± 3.19**
Vowel	51.52 ± 12.38	53.03 ± 11.65	**64.85 ± 5.8**	52.02 ± 11.3	53.54 ± 6.94	52.12 ± 6.94	**64.85 ± 5.8**
German	73.00 ± 2.79	69.40 ± 4.43	69.60 ± 3.9	69.10 ± 2.88	73.50 ± 2.94	71.50 ± 2.77	**74.70 ± 2.97**
LSVT	83.97 ± 14.31	80.06 ± 9.07	89.68 ± 6.39	86.47 ± 5.07	86.90 ± 9.95	90.58 ± 7.5	**91.15 ± 9.39**
Leukemia	84.46 ± 13.51	90.36 ± 6.35	**98.57 ± 10**	91.43 ± 9.48	**98.57 ± 4.29**	**98.57 ± 4.29**	**98.57 ± 4.29**
Mfeat	80.70 ± 1.36	81.50 ± 1.99	80.15 ± 1.52	82.15 ± 1.64	81.65 ± 1.99	**83.15 ± 1.69**	82.60 ± 1.84
PD	79.22 ± 6.07	74.18 ± 7.02	**84.25 ± 5.89**	78.44 ± 4.34	79.63 ± 4.58	79.76 ± 3.61	82.52 ± 3.98
Segmentation	95.28 ± 1.52	95.41 ± 1.5	95.71 ± 1.88	95.19 ± 1.32	95.63 ± 1.67	95.71 ± 1.53	**96.32 ± 1.44**
Wall	77.26 ± 5.91	77.27 ± 6.36	78.41 ± 5.35	75.81 ± 5.88	81.56 ± 3.78	83.14 ± 4.01	**83.32 ± 4.56**
Waveform2	80.08 ± 1.46	80.44 ± 1.51	78.38 ± 1.69	**83.24 ± 1.18**	82.86 ± 1.44	81.82 ± 1.54	82.58 ± 1.06
Average	76.69 ± 6.16	76.28 ± 5.37	80.62 ± 4.53	79.06 ± 7.87	78.20 ± 4.65	79.70 ± 4.26	**82.33 ± 4.12**
Win/Draw/Loss	12/0/0	12/0/0	9/2/1	11/0/1	10/1/1	10/1/1	*best*

algorithm, as the proposed method demonstrates the ability to obtain smaller reducts while achieving superior classification accuracy.

3.5. Conclusion of Chapter 3

This chapter presents the intuitionistic fuzzy weighted neighborhood rough set model and several of its important properties. Based on this model, a distance measure between two intuitionistic fuzzy weighted neighborhood families is constructed, serving as the foundation for designing an attribute reduction algorithm on fixed decision tables. To address practical data scenarios, Chapter 3 also develops update mechanisms for information granules and proposes two incremental algorithms for efficient processing on dynamic decision tables. Experimental results demonstrate that the proposed methods achieve superior classification accuracy on most datasets compared with approaches based on fuzzy rough sets and weighted neighborhood rough sets, while also reducing execution time relative to methods based on intuitionistic fuzzy rough sets.

It can be observed that IFWNRSs effectively overcome the limitations that remain in certain models belonging to the first extension branch of rough set theory. On this basis, the proposed model is regarded as an effective tool for efficiently solving the attribute reduction problem across various types of data, and it exhibits high practical feasibility, particularly in the context of big data.

CONCLUSION AND FUTURE RESEARCH DIRECTIONS

A. Main contributions of the dissertation

This dissertation focuses on the attribute reduction problem in decision tables with the aim of improving the performance of machine learning models while reducing the complexity of rule sets. Rough set theory is regarded as an important foundational tool for the development of attribute reduction methods. However, methods based on traditional rough set models still have many limitations when applied to numerical and continuous decision tables. This practical challenge has resulted in the development of two major extension branches of rough set theory to address these limitations. Through comprehensive investigation, analysis, and comparison, this dissertation identifies several limitations of existing extended methods and makes the following main contributions.

First, to address the limitations of fuzzy rough set–based approaches in handling noisy data as well as the computational inefficiency of intuitionistic fuzzy rough set models, the dissertation proposes the α, β-level intuitionistic fuzzy set model as a generalization of intuitionistic fuzzy sets. Several key properties of the proposed model are established to clarify its advantages. Based on this model, the dissertation develops a number of attribute reduction algorithms with the following contributions:

- An α, β-level intuitionistic fuzzy partition distance measure is constructed to redefine an effective reduct on decision tables. In addition, an attribute significance measure is developed to identify highly important attributes. Based on these measures, a filter-based attribute reduction algorithm for fixed decision tables (ARPD) is proposed.

- The distance measure for α, β-level intuitionistic fuzzy partitions is further extended to efficiently handle decision tables with object set addition and deletion, leading to the development of two incremental algorithms for dynamic decision tables in practical data scenarios.

Second, to overcome the limitations of several extensions within the neighborhood rough set branch, the dissertation proposes the intuitionistic fuzzy weighted neighborhood rough set (IFWNRSs) model, which evaluates the influence of condition attributes on object decisions while providing a more detailed characterization of object roles within information granules. Based on this model, several attribute reduction algorithms are proposed with the following notable contributions:

- A distance measure between two intuitionistic fuzzy weighted neighborhood families is constructed to redefine an effective reduct on decision tables and to establish an attribute significance measure. Based on these measures, an attribute reduction algorithm for fixed decision tables (ARIFW) with polynomial time complexity is proposed.

- The distance measure between intuitionistic fuzzy weighted neighborhood families is extended to handle decision tables with object set addition and deletion, thereby enabling the development of two incremental attribute reduction algorithms for dynamic decision tables.

B. Future research directions

The current scope of the dissertation primarily focuses on decision tables that change with object set addition and deletion. In future work, the research will be extended to attribute reduction methods for decision tables with changing attribute sets. This scenario commonly occurs in real-world applications, where data attributes are continuously updated over time

Second, in the proposed models, decision classes are still represented as crisp sets, which limits the ability to accurately measure the information of each class. To address this issue, future research will focus on developing models that transform decision classes into intuitionistic fuzzy numbers. This approach is expected to provide more flexible and effective measures for identifying important attributes.

Third, in the intuitionistic fuzzy weighted neighborhood rough set model, the use of a fixed radius threshold may lead to difficulties in datasets with high-density distributions, where information granules may contain an excessive number of objects, negatively affecting reduction results. To overcome this limitation, future work will develop new models that restrict the number of objects with significant influence within each information granule, thereby providing a stronger foundation for developing more effective attribute reduction algorithms.

LIST OF THE PUBLICATIONS RELATED TO THE DISSERTATION

A. Published

1. Pham Viet Anh, Nguyen Ngoc Thuy, Vu Duc Thi, and Nguyen Long Giang, "On distance-based attribute reduction with α, β-level intuitionistic fuzzy sets", IEEE Access, vol. 11, pp. 138095–138107, 2023. (SCIE Q1 IF 3.4).

2. Pham Viet Anh, Nguyen Ngoc Thuy, Le Hoang Son, Tran Hung Cuong, and Nguyen Long Giang, "Incremental attribute reduction with α, β-level intuitionistic fuzzy sets", International Journal of Approximate Reasoning, vol. 176, p. 109326, 2025. (SCIE Q1 IF 3.2).

3. Nguyen Long Giang, Pham Viet Anh, Janos Demetrovics, and Vu Duc Thi, "Attribute reduction based on rough set theory and its extensions: A review", Journal of Computer Science and Cybernetics, vol. 41, no. 3, pp. 105-123, 2025.

4. Phạm Việt Anh, Nguyễn Long Giang, Nguyễn Ngọc Thủy, Nguyễn Thế Thủy, Phạm Đình Khánh, "Về một thuật toán gia tăng tìm tập rút gọn trên bảng quyết định khi loại bỏ tập đối tượng", Các công trình nghiên cứu và phát triển CNTT và truyền thông, Hà Nội, số 2, tr. 58-65, 2023.

5. Phạm Việt Anh, Nguyễn Long Giang, Nguyễn Ngọc Thủy, Cao Chính Nghĩa, Vũ Đức Thi, "Rút gọn thuộc tính dựa trên mô hình tập thô mờ trực cảm sử dụng độ đo khoảng cách mở rộng và lát cắt α", Kỷ yếu Hội nghị Khoa học Công nghệ Quốc Gia lần thứ XV: Nghiên cứu cơ bản và ứng dụng công nghệ thông tin, Hà Nội, 11/2022, tr. 320-331, 2023.

6. Pham Viet Anh, Nguyen Long Giang, Nguyen Ngoc Thuy, Le Van Dung, and Phung Hong Quan, "Hybrid filter-wrapper attribute reduction method with the uncertainty classification degree", in Advances in Data Science and Optimization of Complex Systems - Proceedings of the International Conference on Applied Mathematics and Computer Science - ICAMCS, vol. 1569, pp. 298-309, 2025.

B. Waiting review

7. Pham Viet Anh, Nguyen Ngoc Thuy, Nguyen Long Giang, "Incremental attribute reduction on dynamic decision tables with intuitionistic fuzzy weighted neighborhood rough sets", Đang chờ phản biện.