

Attribute Reduction with Fuzzy Weighted Neighborhood Rough Sets

Pham Viet Anh

HaUI Institute of Technology,
Hanoi University of Industry, Hanoi, Vietnam.
E-mail: anhpv@hau.edu.vn, vietanh.hn.4078@gmail.com
ORCID: <https://orcid.org/0000-0001-6628-3531>

Nguyen Ngoc Thuy

University of Sciences,
Hue University, Hue, Vietnam.
E-mail: nnthuy.cs@hueuni.edu.vn
ORCID: <https://orcid.org/0000-0003-0541-0108>

Dinh Thi Thanh Uyen

GeoInformatic Research Center,
Thai Nguyen University of Agriculture and Forestry, Thainguyn, Vietnam.
Corresponding author: dinhthithanhuyen@tuaf.edu.vn, ncs.dinhthithanhuyen@ictu.edu.vn
ORCID: <https://orcid.org/0009-0008-6271-7492>

Nguyen Chien Thang

School of Information and Communications Technology,
Hanoi University of Industry, Hanoi, Vietnam.
E-mail: thangnc@hau.edu.vn
ORCID: <https://orcid.org/0000-0003-1243-3906>

Nguyen Long Giang

Institute of Information Technology,
Vietnam Academy of Science and Technology, Hanoi, Vietnam.
E-mail: nlgiang@ioit.ac.vn
ORCID: <https://orcid.org/0000-0001-6184-1469>

(Received on March 10, 2026; Revised on July 2, 2026; Accepted on August 18, 2026)

Abstract

Neighborhood rough sets (NRSs) have been widely recognized as an efficient tool for solving attribute reduction problems in numerical data. In constructing neighborhood information granules, many existing extensions introduce weighting mechanisms for conditional attributes to quantify their influence on the decision attribute, while other studies focus on assigning weights to individual objects within granules. Nevertheless, the internal distribution of objects inside each granule, particularly the density of objects in the same decision class, has not been sufficiently characterized. Therefore, the generated information granules may not accurately assess the actual contribution of individual objects. To address this issue, a generalized weighted neighborhood rough set model has recently been proposed by integrating both object weights and attribute weights. In this study, we investigate a special case of that framework by proposing an alternative approach to assigning weights to objects that better captures local class-consistency and structural characteristics within neighborhood granules. Based on the newly constructed information granules, we develop the Fuzzy Weighted Neighborhood Rough Set (FWNRS) model as a basis for designing an efficient attribute reduction method. Experimental evaluations on multiple benchmark datasets indicate that the proposed method is effective and outperforms several state-of-the-art approaches in terms of reduct size as well as classification performance.

Keywords- Fuzzy weighted neighborhood rough sets, Decision tables, Object weights, Attribute reduction, Attribute weights.

Abbreviations	Description
RS	Rough Set
FRS	Fuzzy Rough Set
NRS	Neighborhood Rough Set
KNRS	k -nearest Neighborhood Rough Set
WNRS	Weighted Neighborhood Rough Set
WKNRS	Weighted k -nearest Neighborhood Rough Set
WNPRS	Weighted Neighborhood Probability Rough Set
DWNRS	Divergence-based Weighted Neighborhood Rough Set
GWNRS	Generalized Weighted Neighborhood Rough Set
FWNRS	Fuzzy Weighted Neighborhood Rough Set
NIG	Neighborhood Information Granule
WNIG	Weighted Neighborhood Information Granule
FWNIG	Fuzzy Weighted Neighborhood Information Granule
$[u]_A^{\varepsilon, w}$	Weighted neighborhood information granule of object u with respect to attribute subset A
$G(A)$	Family of weighted neighborhood information granules with respect to attribute subset A
$\varpi_u(v)$	Weighted of v in the weighted neighborhood information granule $[u]_A^{\varepsilon, w}$
$[u]_A^{\varepsilon, w}$	Fuzzy weighted neighborhood information granule of u with respect to attribute subset A
$\widehat{G}(A)$	Family of fuzzy weighted neighborhood information granules with respect to attribute subset A

1. Introduction

The rough set (RS) framework is widely regarded as a robust theoretical foundation that has been extensively utilized in various fields, comprising data analysis, knowledge discovery as well as attribute reduction. This model is designed to approximate sets of objects based on information granules formed by indiscernibility relations between pairs of samples within a given feature set (Pawlak, 1982). However, because traditional RSs rely on a strict equivalence relation, they are typically suitable for categorical data type and less efficient when processing continuous or numerical data. Therefore, several alternative and more advanced approaches have been presented by expanding Pawlak's RS model, including the neighborhood rough set (NRS) model (Jiang et al., 2020; Ju et al., 2023), the fuzzy rough set (FRS) model (Tsang et al., 2016; Wang et al., 2017; Yang et al., 2017), and the intuitionistic fuzzy rough set (IFRS) model (Jain et al., 2020; Tan et al., 2022; Anh et al., 2023; Anh et al., 2025). These models have subsequently been widely applied to various practical problems, including clinical decision-making (Sun et al., 2026), university admission (Ejegwa et al., 2024), and higher education (Ejegwa et al., 2025).

Neighborhood rough sets (NRSs) have been extensively studied as an efficient framework for solving the attribute reduction problem in numerical decision information systems. Unlike the classical RS model, which is based on indiscernibility relations, the NRS model adopts neighborhood relations. Accordingly, each object (sample) in the universe is characterized through a neighborhood information granule (NIG) comprising objects in the ε -neighborhood of the given object (Hu et al., 2008a). From this model, numerous effective attribute selection methods have been introduced, such as the dependency degree (Hu et al., 2008a, 2008b), the neighborhood-based decision error rate (Hu et al., 2010) and discrimination index (Wang et al., 2018). From the entropy and Lebesgue measures, Sun et al. (2019) introduced a heuristic method for feature reduction to efficiently handle incomplete and mixed datasets. According to other approaches, Yang et al. (2019) and Wang et al. (2019) also developed novel feature reduction methods based on pseudo-label NRSs and k -nearest NRSs. Next, Yang et al. (2021) addressed this problem using distance metric learning for optimizing the structure of NIGs and developing two feature reduction methods. Moreover, Zhang et al. (2024) proposed conditional neighborhood-based entropy as a crucial measure for attribute extraction from heterogeneous data. Furthermore, methods based on the NRS model have been developed in various ways (Chen et al., 2019; Hu et al., 2022; Pham et al., 2025) and have proven to be successful in performing

attribute reduction on numerical data. However, NRSs model does not consider the effect of each condition attribute on the decisions of the objects. In particular, this model assumes that the weights of each conditional attribute are the same. This may lead to an incorrect representation of the relationship between the conditional features and the decision feature. Hence, some vital attributes are often overlooked in the reduction process.

In response, Hu et al. (2021) introduced the weighted NRS (WNRS) model, which calculates weights for all condition attributes to reflect their varying levels of importance and forms weighted neighborhood information granules (WNIGs). Following a similar approach, Xie et al. (2022) calculated attribute weights through data clustering and entropy theory, subsequently proposing the weighted neighborhood probability rough set (WNPRS) model and introducing some crucial properties. Additionally, Thuy & Wongthanavas (2024) established attribute weights based on fuzzy divergence and proposed the divergence-based WNRS (DWNRS) model. Although experimental results have indicated that the WNRS model yields high efficiency in improving the attribute reduction process, these models primarily focus on the number of objects in each WNIG. This indicates that every object in the WNIG is considered to have equal importance. However, data in practice always has a diverse distribution, i.e., each object that lies inside the information granule has a different role. Therefore, accurately assessing the importance degree of objects in the WNIG is essential.

Recently, Wang & Zhao (2024) proposed the weighted k -nearest NRS (WKNRS) model. This model evaluates the quality and assigns weights to the nearest neighbor objects through the standard deviation of their distance values within the same class. Thus, this method can reduce the effect of noisy objects in the attribute reduction problem. It is evident that developing weighting mechanisms for conditional attributes and for objects has led to considerable success in attribute reduction methods. However, to date, the simultaneous integration of both mechanisms has not yet been investigated. To address the issues discussed above, we have recently introduced a generalized WNRS (GWNRS) model and demonstrated that all existing NRS models may be regarded as special cases of this model (Thuy et al., 2025). Building upon this foundation, this study focuses on analyzing a special case of the GWNRS model, with particular emphasis on assigning weights to objects within information granules in order to establish a new type of granule. Specifically, the weight of each object in a granule is represented by the distribution density of other objects that belong to the same decision class and have a neighborhood relation with it. On the basis of these granules, we present a novel concept known as the fuzzy weighted neighborhood rough set (FWNRS) model to construct the fuzzy weighted dependency function as a measure of classification certainty. Accordingly, we define an efficient reduct set and propose a condition attribute significance measure. Finally, we construct a filter-wrapper-based attribute reduction method for identifying an optimal condition attribute subset.

The subsequent sections of the paper are structured as follows. In Section 2, we survey several representative NRS models used for attribute reduction. Section 3 will recall the crucial concepts associated with decision information systems and NRS theory. Section 4 introduces the FWNRS model and the fuzzy weighted dependency degree. Section 5 introduces the attribute reduction method through the FWNRS model. Section 6 illustrates the experimental evaluation, and Section 7 highlights the limitations of the proposed approach and suggests promising research directions.

2. Related Works

This section will present in more detail some remaining limitations of the models that have been applied to the attribute reduction problem, especially recent models that consider weighting mechanisms such as WNRS, WPNRS, WKNRS and GWNRS. From these analyses, this study will clarify the existing research

gaps, thereby forming the basis for developing a new model.

To address the limitations of traditional NRSs, WNRS, WNPRS, and DWNRS are among the early weighted neighborhood rough set models developed to construct weight assignment mechanisms for condition attributes in decision tables. However, these models mainly focus on evaluating the importance of attributes, while they have not fully considered a weight assignment mechanism for objects. This makes the constructed neighborhood information granules unable to clearly reflect the different roles and contributions of each object within a granule. In addition, the attribute weight assignment mechanism of WNPRS requires a data discretization step before computation, which may increase the execution time and lead to the risk of information loss. Meanwhile, the DWNRS model establishes attribute weights based on fuzzy divergence, which requires the determination of the fuzzy equivalence classes of each attribute. Therefore, for datasets containing a large number of objects, determining attribute weights in DWNRS may become inefficient in terms of the runtime.

Table 1. Summary of the characteristics and drawbacks of each model.

Model	Attribute weight	Object weight	Limitations
NRS, KNRS	No	No	- The lack of mechanisms for assigning weights to condition attributes leads to the generated information granules not accurately reflecting the degree of relevance between each condition attribute and the decision attribute. - The lack of mechanisms for assigning object weights leads to the information granule structure not clearly characterizing the role of each object in the granule. - Focus on consistent objects, while objects in the boundary region are often ignored.
WNRS	Yes	No	- The lack of a mechanism for assigning object weights leads to the structure of WNIGs not clearly characterizing the role of each object in the granule. - Focus on consistent objects of the decision table, while objects in the boundary region are often ignored.
WNPRS	Yes	No	- The lack of a mechanism for assigning object weights leads to the structure of NIGs not clearly characterizing the role of each object in the granule. - The mechanism for assigning attribute weights requires a data discretization step before computation, which may increase the execution time and lead to the risk of information loss.
DWNRS	Yes	No	- The lack of a mechanism for assigning object weights leads to the structure of NIGs not clearly characterizing the role of each object in the granule. - The mechanism for assigning attribute weights requires the determination of the fuzzy equivalence classes of each attribute.
WKNRS	No	Yes	The lack of a mechanism for assigning weights to condition attributes may lead to important attributes being overlooked during the reduction process.
GWNRS	No	Yes	The mechanism for assigning weights to condition attributes in the model does not reflect the influence of the decision attribute.
FRS	No	Yes	- The lack of mechanisms for assigning weights to condition attributes leads to the generated information granules not accurately reflecting the degree of relevance between each condition attribute and the decision attribute. - Information granules are constructed based on fuzzy equivalence relations, which only reflect the degree of similarity between the object under consideration and other objects in the universe.

Recently, the WKNRS model was proposed based on evaluating the quality and assigning weights to nearest neighbor objects through the standard deviation of distance values. However, this model has not considered an attribute weight assignment mechanism for condition attributes, thereby reducing the ability to identify important attributes during the reduction process. To overcome the above limitations, the GWNRS model was proposed by simultaneously integrating both attribute and object weight assignment mechanisms, thereby achieving high effectiveness in the attribute reduction problem. Building upon these advantages, our study focuses on analyzing a specific case of GWNRS and proposes the FWNRS model with several new extensions compared with GWNRS, as follows.

(1) In terms of the attribute weight assignment mechanism, the proposed model integrates a weight assignment mechanism that considers the relationship between condition attributes and the decision attribute. Although GWNRS is regarded as a generalized model, the determination of attribute weights in this model is mainly based on the standard deviation of condition attribute values, without fully considering the influence of the decision attribute. Therefore, the efficiency of the model in evaluating the role of condition attributes may be reduced.

(2) For object weight assignment, our model establishes a new mechanism with an extension over GWNRS. Specifically, this mechanism simultaneously considers the relevance and reliability of objects in the weighted neighborhood information granule, aiming to reflect their roles and degree of influence in greater detail.

(3) Based on the proposed model, this study combines a hybrid approach to select attribute subsets that are highly effective for the classification process. This is a crucial approach in current deep learning and machine learning problems.

From the analysis of the limitations of NRS models and WNRS models in the attribute reduction problem, we present **Table 1** to summarize the main characteristics and remaining drawbacks of each model. The content summarized in the table serves as an important basis for constructing an effective model in this study.

3. Basic Notion

This section provides a summary of key concepts, including information systems, decision tables, and a review of several existing NRS models (Hu et al., 2008a; Hu et al., 2008b; Hu et al., 2021). This theoretical framework is a widely used tool for addressing attribute reduction problems.

An *information system* is typically expressed as an ordered pair $I = (U, B)$, in which U is the set comprising samples (objects), and B represents a finite and non-empty set of features (attributes). Accordingly, each attribute $b \in B$ determines a map $b: U \rightarrow V_b$, where V_b denotes the domain of b , also called the value set of b . For any $u \in U$ and $b \in B$, $b(u)$ is considered the *attribute value of object u* with respect to (w.r.t.) attribute b .

Next, an *information system* $I = (U, B)$, in which the attribute set B is partitioned into two separate subsets, comprising conditional attributes C and decision attributes D , is referred to as a *decision table*, or *decision information system*, denoted by $I = (U, C, D)$. Note that in this study, we suppose that for any $I = (U, C, D)$, the set of values $V_b \in [0,1]$ for every $b \in C$. This can be easily obtained by normalizing the input decision tables, such as using min-max normalization.

Let $I = (U, C, D)$ be a decision table, $u, v \in U$, and $A \subseteq C$ be a conditional attribute subset. Assume that $\varepsilon \in [0,1]$ is a neighborhood radius. Accordingly, the binary relation $R^\varepsilon(A)$, called the *neighborhood relation* on U determined by A , is determined as follows:

$$R^\varepsilon(A) = \{(u, v) \in U \times U : \Delta_A(u, v) \leq \varepsilon\} \quad (1)$$

where, $\Delta_A(u, v)$ is the distance between u and v on A .

Here, if $(u, v) \in R^\varepsilon(A)$, then v and u are ε -neighborly induced by A . For each object $u \in U$, the set $[u]_A^\varepsilon = \{v \in U : (u, v) \in R^\varepsilon(A)\}$ is called the *neighborhood information granule* (NIG) of u w.r.t A . Clearly, $[u]_A^\varepsilon$

is an ordinary set comprising all objects whose distance to u is no greater than ε .

According to the concept of NIGs, consider a subset of objects $K \subseteq U$, *upper* and *lower approximations* of K w.r.t the relation $R^\varepsilon(A)$ are respectively determined as follows:

$$\overline{N}_A(K) = \{u \in U: [u]_A^\varepsilon \cap K \neq \emptyset\} \quad (2)$$

and

$$\underline{N}_A(K) = \{u \in U: [u]_A^\varepsilon \subseteq K\} \quad (3)$$

To assess the effect of each conditional attribute on the decision of each object, Hu et al. (2021) initially presented a *weighted distance* measure. Specifically, consider objects u, v in U and a conditional attribute subset $A \subseteq C$. Then, the *weighted distance* between u and v w.r.t A , denoted by $\Delta_A^w(u, v)$, is computed as follows:

$$\Delta_A^w(u, v) = \sqrt{\sum_{b \in A} w(b)^2 \cdot |b(u) - b(v)|} \quad (4)$$

where, $w(b)$ represents the weight assigned to attribute b , determined as follows. Given a decision table $I = (U, C, D)$ with condition attribute set $C = \{b_1, b_2, \dots, b_m\}$ and object set $U = \{u_1, u_2, \dots, u_n\}$, the partition coefficients of all condition attributes $\zeta = (\zeta(b_1), \zeta(b_2), \dots, \zeta(b_m))^T$ are determined by solving the optimization problem $\zeta^* = \operatorname{argmin} \|B\zeta - Y\|^2 + \|\zeta\|^2$, where B is a coefficient matrix defined by

$$B = \begin{bmatrix} b_1(u_1) & b_2(u_1) & \cdots & b_m(u_1) \\ b_1(u_2) & b_2(u_2) & \cdots & b_m(u_2) \\ \vdots & \vdots & \ddots & \vdots \\ b_1(u_n) & b_2(u_n) & \cdots & b_m(u_n) \end{bmatrix},$$

and $Y = (D(u_1), D(u_2), \dots, D(u_n))^T$ is a decision vector.

From the partition coefficients of attributes as presented below, the weight of a condition attribute $b \in C$ is defined as

$$w(b) = \frac{|C| \times |\zeta(b)|}{\sum_{b \in C} |\zeta(b)|} \quad (5)$$

The weights of condition attributes are determined using their partition coefficients with decisions. In particular, a stronger correlation with the decision leads to a higher assigned weight. Then, internal relevance between condition attributes and decisions are high.

Next, for any $u \in U$, a *weighted neighborhood information granule* (WNIG) of object u w.r.t A is determined as follows:

$$[u]_A^{\varepsilon, w} = \{v \in U: \Delta_A^w(u, v) \leq \varepsilon\} \quad (6)$$

It can be readily observed that $\sum_{b \in C} w(b) = |C|$ and the average value of the attribute weights is equal to 1. Hence, $w(b) > 1$ indicates that attribute b has an influence on the decision that is higher than the average level. When incorporated into the distance measure, these weights amplify the differences between objects on the corresponding attributes and have a stronger impact on the process of determining neighborhood relations. Therefore, the construction of WNIGs becomes stricter w.r.t attributes that play important roles.

It is clear that the family containing all these WNIGs constitutes a cover $G(A) = \{[u]_A^{\varepsilon, w} : u \in U\}$ on U . From the WNIGs, Hu et al. (2021) introduced the *weighted neighborhood rough set* (WNRS) model through the following upper and lower approximations. Specifically, given an object subset $K \subseteq U$, upper and lower approximations of K w.r.t $G(A)$ are respectively defined as

$$\overline{WNR}_A(K) = \{u \in U : [u]_A^{\varepsilon, w} \cap K \neq \emptyset\} \quad (7)$$

and

$$\underline{WNR}_A(K) = \{u \in U : [u]_A^{\varepsilon, w} \subseteq K\} \quad (8)$$

4. Fuzzy Weighted Neighborhood Rough Sets (FWNRSs)

In this part, we propose a novel model referred to as fuzzy weighted neighborhood rough sets (FWNRSs). In FWNRSs, each information granule is constructed by combining attribute weights and object weights, enabling the model not only to evaluate the role of each object in the granule but also to reflect the influence of conditional attributes on the decision attribute. Consequently, FWNRSs are expected to deliver superior performance in the attribute reduction problem.

4.1 Fuzzy Weighted Neighborhood Granules

To commence, we can see that the objects in $[u]_A^{\varepsilon, w}$ have the equal role in the decision-making process for the object u . Nonetheless, these objects can serve different roles in assessing u . Hence, characterizing the importance degree of each object, also known as assigning weights, is essential. On this basis, we will construct the information granule through establishing weight to each object within it. First, we introduce a novel concept referred to as the relevance between two objects.

Definition 1. Given $I = (U, C, D)$, $u, v \in U$, and $A \subseteq C$, the *relevance* between u and v on A , denoted by $Sim_A(u, v)$, is calculated as follows:

$$Sim_A(u, v) = |1 - \Delta_A^w(u, v)| \quad (9)$$

It can be seen that the relevance between two objects u and v is strongly influenced by the distance between them. The closer v is to u , the greater the relevance between these two objects, and vice versa. If v is an object belonging to the WNIG $[u]_A^{\varepsilon, w}$, then $\Delta_A^w(u, v) \in [0, 1]$. This means that $Sim_A(u, v) = 1 - \Delta_A^w(u, v)$ and $Sim_A(u, v) \in [0, 1]$. Hence, in the WNIG $[u]_A^{\varepsilon, w}$, the relevance is used to describe the degree of similarity between the objects within the granule and object u . Objects whose characteristics are more similar to those of u are more likely to belong to the same decision class as u . In other words, these objects are considered to play a more important role in determining the decision class of object u . As illustrated in **Figure 1**, object u_5 has a higher relevance to u_1 than objects u_3 , u_8 , and u_9 , because its distance to u_1 is smaller. Therefore, u_5 is considered to be highly similar to u_1 and has a strong influence on evaluating the classification of u_1 . Conversely, although objects u_3 , u_8 , and u_9 also belong to the weighted neighborhood information granule of u_1 , they have a very low similarity to u_1 because their distances to u_1 are relatively large. Therefore, the relevance of these objects to u_1 is very low, and they have little influence on evaluating the classification of u_1 . When $\varepsilon = 1$, objects u_8 and u_9 , which lie on the boundary of the WNIG $[u_1]_A^{\varepsilon, w}$, are considered to have no relevance to u_1 , i.e., $Sim_A(u_1, u_8) = Sim_A(u_1, u_9) = 0$.

Clearly, the relevance is a vital component reflecting the influence of objects within a WNIG. However, to evaluate the role of these objects in greater detail, we further introduce a new concept to examine their contribution to the classification process within a WNIG.

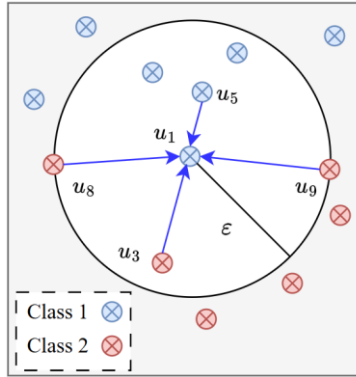


Figure 1. The relevance of objects in the WNIG $[u_1]_A^{\varepsilon,w}$.

Definition 2. Let $[u]_A^{\varepsilon,w}$ be the WNIG of u induced by A and v be an object in $[u]_A^{\varepsilon,w}$. The *reliability* of v w.r.t the classification of u in $[u]_A^{\varepsilon,w}$, denoted by $Rel_A(u, v)$, is determined as follows:

$$Rel_A(u, v) = \begin{cases} 1 - \mu_A(v) & \text{if } N_v > 0 \\ 0 & \text{if } N_v = 0 \end{cases} \quad (10)$$

where, $\mu_A(v)$ denotes the standard deviation of weighted distance values from object v to the remaining objects in $[u]_A^{\varepsilon,w}$ that have the same decision class as v and N_v is denotes the number of objects in $[u]_A^{\varepsilon,w}$ that belong to the same decision class as v .

In the case where $N_v > 0$, since $v \in [u]_A^{\varepsilon,w}$, we have $\Delta_A^w(u, v) \in [0,1]$. It follows that $\mu_A(v) \in [0,0.5]$. Hence, $Rel_A(u, v) \in [0.5,1]$. The reliability of object v with respect to the classification of object u represents the distribution of all objects belonging to the same class as v within the information granule $[u]_A^{\varepsilon,w}$. As illustrated in **Figure 2**, although objects u_7 and u_6 have the same relevance to object u_1 because their distances to u_1 are equal, their degrees of influence on the classification of u_1 are different. This difference arises because the objects belonging to the same class are distributed differently around u_7 and u_6 . Specifically, the objects belonging to the same class as u_7 are concentrated and densely distributed, whereas those around u_6 are sparsely distributed. Therefore, the decision information of u_7 is supported by more objects of the same class, making u_7 more reliable than u_6 in contributing the classification of u_1 . Accordingly, u_7 has higher reliability than u_6 in classifying object u_1 .

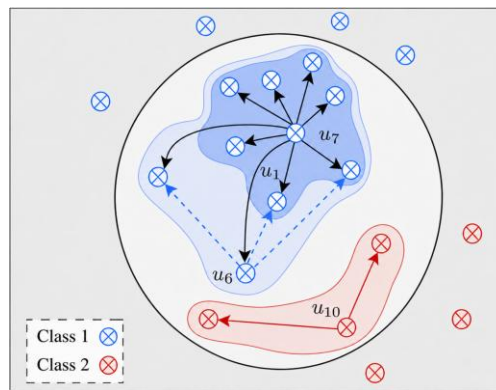


Figure 2. The reliability of objects in the WNIG $[u_1]_A^{\varepsilon,w}$.

Definition 3. Let $[u]_A^{\varepsilon,w}$ be a WNIG of u induced by A . Then, the weight of each object v in $[u]_A^{\varepsilon,w}$, denoted by $\varpi_u^A(v)$, is calculated as follows:

$$\varpi_u^A(v) = \text{Sim}_A(u, v) \times \text{Rel}_A(u, v) \quad (11)$$

It can be easily seen that the weight assigned to each object v in the WNIG $[u]_A^{\varepsilon,w}$ can be regarded as combined information measure, simultaneously reflecting the relevance and reliability of that object in $[u]_A^{\varepsilon,w}$. For objects v with high relevance and reliability, their role in $[u]_A^{\varepsilon,w}$ becomes more significant and has a greater influence on the decision of object u . Conversely, when either of these two components is small, the weight of object v decreases, causing v to play a negligible role or have no contribution to the information granule $[u]_A^{\varepsilon,w}$. For objects v outside the weighted neighborhood information granule $[u]_A^{\varepsilon,w}$, we assign $\varpi_u^A(v) = 0$.

By aggregating all evaluations, we construct a new information granule in which each object is assigned a weight representing its membership value. This granule may be viewed as a fuzzy set over U and is termed the fuzzy weighted neighborhood information granule (FWNIG), denoted as:

$$[u]_A^{\varepsilon,w} = \left\{ \frac{\varpi_u^A(u_1)}{u_1}, \frac{\varpi_u^A(u_2)}{u_2}, \dots, \frac{\varpi_u^A(u_{|U|})}{u_{|U|}} \right\} \quad (12)$$

Consequently, $\widehat{G(A)} = \{[u]_A^{\varepsilon,w} : u \in U\}$ is defined as the family containing all FWNIGs.

Figure 3 illustrates the difference between the proposed information granules and those in the WNRS model. It can be observed that, within the FWNIG of u_1 , object u_2 is considered more crucial than object u_3 in assessing u_1 . In contrast, in the WNRS model, these two objects are regarded as having equal importance. Therefore, the proposed approach enables greater focus on denser objects while effectively reducing the influence of noisy objects.

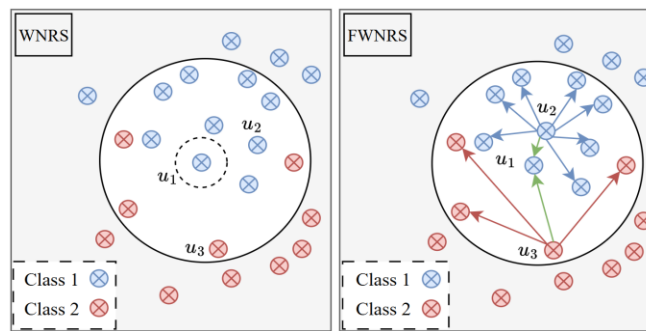


Figure 3. Differences between information granules in WNRS and FWNRS.

Table 2. The decision table according to example 1.

U	b_1	b_2	b_3	b_4	D
u_1	0.72	0.93	0.65	0.52	0
u_2	0.43	0.12	0.52	0.44	0
u_3	0.21	0.84	0.97	0.62	0
u_4	0.95	0.89	0.21	0.81	1
u_5	0.47	0.72	0.53	0.87	1
u_6	0.14	0.58	0.59	0.21	1

Example 1. Given a decision table $I = (U, C, D)$ with $C = \{b_1, \dots, b_4\}$ and $U = \{u_1, \dots, u_6\}$ as **Table 2** below, we determine $\widetilde{G}(\mathcal{C})$ with the neighborhood radius $\varepsilon = 0.70$ as follows:

Based on Equation (5), we obtain the weights of condition attributes: $w(b_1) = 0.803$, $w(b_2) = 1.518$, $w(b_3) = 0.211$, $w(b_4) = 1.467$. Since $w(b_1)$ and $w(b_3)$ are less than 1, the impacts of b_1 and b_3 are reduced in construction WNIGs. On the contrary, $w(b_2)$ and $w(b_4)$ are more than 1, the impacts of b_2 and b_4 are raised in construction WNIGs.

Consider the object u_1 , we are easy to compute the weighted distance between u_1 and other objects with respect to C : $\Delta_C^w(u_1, u_1) = 0.000$, $\Delta_C^w(u_1, u_2) = 1.257$, $\Delta_C^w(u_1, u_3) = 0.461$, $\Delta_C^w(u_1, u_4) = 0.477$, $\Delta_C^w(u_1, u_5) = 0.637$, $\Delta_C^w(u_1, u_6) = 0.841$. Then, $[u_1]_C^{\varepsilon, w} = \{u_1, u_3, u_4, u_5\}$ and it follows that $\mu_C(u_1) = 0.231$. Similarly, we also have: $\mu_C(u_3) = 0.231$, $\mu_C(u_4) = 0.239$ and $\mu_C(u_5) = 0.239$.

Next, the weights of objects can be computed by using Equation (9), as below:

$$\varpi_{u_1}^C(u_1) = (1 - 0.000) \times (1 - 0.231) = 0.769, \varpi_{u_1}^C(u_3) = (1 - 0.461) \times (1 - 0.231) = 0.415, \\ \varpi_{u_1}^C(u_4) = (1 - 0.477) \times (1 - 0.239) = 0.398, \varpi_{u_1}^C(u_5) = (1 - 0.637) \times (1 - 0.239) = 0.276.$$

Therefore, we have:

$$[\widetilde{u_1}]_C^{\varepsilon, w} = \left\{ \frac{0.769}{u_1}, \frac{0.000}{u_2}, \frac{0.415}{u_3}, \frac{0.398}{u_4}, \frac{0.276}{u_5}, \frac{0.000}{u_6} \right\}.$$

By conducting similar computations, we obtain the remaining FWNIGs with respect to C :

$$[\widetilde{u_2}]_C^{\varepsilon, w} = \left\{ \frac{0.000}{u_1}, \frac{1.000}{u_2}, \frac{0.000}{u_3}, \frac{0.000}{u_4}, \frac{0.000}{u_5}, \frac{0.000}{u_6} \right\}, \\ [\widetilde{u_3}]_C^{\varepsilon, w} = \left\{ \frac{0.415}{u_1}, \frac{0.000}{u_2}, \frac{0.769}{u_3}, \frac{0.243}{u_4}, \frac{0.404}{u_5}, \frac{0.000}{u_6} \right\}, \\ [\widetilde{u_4}]_C^{\varepsilon, w} = \left\{ \frac{0.415}{u_1}, \frac{0.000}{u_2}, \frac{0.769}{u_3}, \frac{0.243}{u_4}, \frac{0.404}{u_5}, \frac{0.000}{u_6} \right\}, \\ [\widetilde{u_5}]_C^{\varepsilon, w} = \left\{ \frac{0.279}{u_1}, \frac{0.000}{u_2}, \frac{0.409}{u_3}, \frac{0.398}{u_4}, \frac{0.761}{u_5}, \frac{0.000}{u_6} \right\}, \\ [\widetilde{u_6}]_C^{\varepsilon, w} = \left\{ \frac{0.000}{u_1}, \frac{0.000}{u_2}, \frac{0.000}{u_3}, \frac{0.000}{u_4}, \frac{0.000}{u_5}, \frac{1.000}{u_6} \right\}.$$

4.2 Proposed FWNRSs Model

Based on the FWNIGs framework, we continue to propose a novel model called fuzzy weighted neighborhood rough sets (FWNRSs).

Let $I = (U, C, D)$ be a decision table, $K \subseteq U$ be a subset of objects, and $A \subseteq C$ be a subset of condition attribute. The lower and upper approximations of K based on the FWNIGs induced by A are respectively determined as follows:

$$\underline{FWN}_A(K) = \left\{ u \in U : \frac{|[\widetilde{u}]_A^{\varepsilon, w} \cap K|}{|[\widetilde{u}]_A^{\varepsilon, w}|} \geq \alpha \right\} \quad (13)$$

and

$$\overline{FWN}_A(K) = \left\{ u \in U : \frac{|\widehat{[u]_A^{\varepsilon,w}} \cap K|}{|\widehat{[u]_A^{\varepsilon,w}}|} \geq \beta \right\} \quad (14)$$

where, $0 \leq \beta < \alpha \leq 1$.

Compared with traditional models, the proposed model extends the approximation sets by incorporating two parameters α and β , which allow for expanding the number of objects to be considered and eliminating weakly classified objects.

Then, the fuzzy weighted boundary region of K w.r.t $\widehat{[u]_A^{\varepsilon,w}}$ is defined as

$$FBN_A(K) = \overline{FWN}_A(K) - \underline{FWN}_A(K) \quad (15)$$

Next, we shall present the fuzzy weighted positive region as well as the fuzzy weighted boundary region of D on the attribute subset A as follows:

$$FPOS_A(D) = \bigcup_{K \in U/D} \underline{FWN}_A(K) \quad (16)$$

and

$$FBN_A(D) = \overline{FWN}_A(D) - \underline{FWN}_A(D) \quad (17)$$

where, $\overline{FWN}_A(D) = \bigcup_{K \in U/D} \overline{FWN}_A(K)$ and $\underline{FWN}_A(D) = \bigcup_{K \in U/D} \underline{FWN}_A(K)$.

Given $I = (U, C, D)$ and $A \subseteq C$, then the fuzzy weighted dependency function of D on A is defined as

$$Dep(A|D) = \frac{|FPOS_A(D)|}{|U|} \quad (18)$$

Example 2. Continuing example 1 with $\alpha = 0.55$, the lower approximations of D based on the FWNIGs with respect to C are $\underline{FWN}_C(D_1) = \{u_1, u_2, u_3\}$ and $\underline{FWN}_C(D_2) = \{u_3, u_4, u_5\}$. Hence, we obtain $FPOS_A(D) = \{u_1, u_2, u_3, u_4, u_5, u_6\}$ and $Dep(C|D) = 1$.

Proposition 1. Let $I = (U, C, D)$ be a decision table with $A \subseteq C$. Then, the following properties of FWNRSs hold:

- 1) If $\beta = 0$, then $U = \overline{FWN}_A(D)$,
- 2) If $\beta = 0$, then $U - FPOS_A(D) = FBN_A(D)$,
- 3) $FPOS_A(D) \cap FBN_A(D) = \emptyset$,
- 4) $FPOS_A(D) \cup FBN_A(D) = \overline{FWN}_A(D)$.

Proof. 1) According to Equation (14) with $\beta = 0$, for any $v \in K$, we have $v \in \overline{FWN}_A(K)$. This leads to $K \subseteq \overline{FWN}_A(D)$ for any $K \in U/D$. Hence, $\bigcup_{K \in U/D} K \subseteq \bigcup_{K \in U/D} \overline{FWN}_A(K) = \overline{FWN}_A(D)$. Meanwhile, $\bigcup_{K \in U/D} K = U$, we obtain $U \subseteq \overline{FWN}_A(D)$. Additionally, since $\overline{FWN}_A(D) \subseteq U$, it follows that $U = \overline{FWN}_A(D)$.

2) According to Equation (15) and statement 1.1, we have $FBN_A(D) = U - \underline{FWN}_A(D)$. Moreover, $\underline{FWN}_A(D) = \bigcup_{K \in U/D} \underline{FWN}_A(K) = FPOS_A(D)$. Hence, we can obtain $U - FPOS_A(D) = FBN_A(D)$.

3) According to Equation (15), we have $FBN_A(D) = \overline{FWN}_A(D) - \underline{FWN}_A(D)$, it follows that $FBN_A(D) = \overline{FWN}_A(D) - FPOS_A(D)$. Therefore, $FBN_A(D) \cap FPOS_A(D) = \emptyset$.

4) According to Equation (15), we have $FBN_A(D) = \overline{FWN}_A(D) - \underline{FWN}_A(D) = \overline{FWN}_A(D) - FPOS_A(D)$. Therefore, we obtain $FBN_A(D) \cup FPOS_A(D) = \overline{FWN}_A(D)$.

Proposition 2. Given a decision table $I = (U, C, D)$ and two condition attribute subsets $A, B \subseteq C$ satisfying $A \subseteq B$, the following holds.

- 1) $[u]_B^{\varepsilon, w} \subseteq [u]_A^{\varepsilon, w}$, for all $u \in U$,
- 2) If $\beta = 0$, then $\forall K \subseteq U, \overline{FWN}_B(K) \subseteq \overline{FWN}_A(K)$,
- 3) If $\alpha = 1$, then $\forall K \subseteq U, \underline{FWN}_A(K) \subseteq \underline{FWN}_B(K)$,
- 4) $FPOS_A(D) \subseteq FPOS_B(D)$ and $Dep(A|D) \leq Dep(B|D)$.

Proof. 1) Since $A \subseteq B$, based on Equation (4), we obtain $\sqrt{\sum_{b \in A} w^2(b) \cdot (b(u) - b(v))^2} \leq \sqrt{\sum_{b \in B} w^2(b) \cdot (b(u) - b(v))^2}$, for all $u, v \in U$. Hence, $[u]_B^{\varepsilon, w} \subseteq [u]_A^{\varepsilon, w}$ for all $u \in U$, it follows that $[u]_B^{\varepsilon, w} \subseteq [u]_A^{\varepsilon, w}$.

2) From $\beta = 0$, for any $u \in \overline{FWN}_B(K)$, we have $\frac{|[u]_B^{\varepsilon, w} \cap K|}{|[u]_B^{\varepsilon, w}|} \geq 0$, it follows that $|[u]_B^{\varepsilon, w} \cap K| \geq 0$. Meanwhile, according to statement 2.1, since $A \subseteq B$, we obtain $[u]_B^{\varepsilon, w} \subseteq [u]_A^{\varepsilon, w}$. Consequently, $|[u]_A^{\varepsilon, w} \cap K| \geq 0$, which implies that $u \in \overline{FWN}_A(K)$. Therefore, $\overline{FWN}_B(K) \subseteq \overline{FWN}_A(K)$.

3) From $\alpha = 1$, for any $u \in \underline{FWN}_A(K)$, we have $\frac{|[u]_A^{\varepsilon, w} \cap K|}{|[u]_A^{\varepsilon, w}|} \geq 1$, which implies that $[u]_A^{\varepsilon, w} \subseteq K$. Meanwhile, according to statement 2.1, since $A \subseteq B$, we obtain $[u]_B^{\varepsilon, w} \subseteq [u]_A^{\varepsilon, w}$. Thus, $[u]_B^{\varepsilon, w} \subseteq K$ and $\frac{|[u]_B^{\varepsilon, w} \cap K|}{|[u]_B^{\varepsilon, w}|} = 1$. Therefore, $u \in \underline{FWN}_B(K)$ and $\underline{FWN}_A(K) \subseteq \underline{FWN}_B(K)$.

4) We have $FPOS_A(D) = \cup_{K \in U/D} \underline{FWN}_A(K) \subseteq \cup_{K \in U/D} \underline{FWN}_B(K) = FPOS_B(D)$ from statement 2.3. Therefore, we obtain $Dep(A|D) \leq Dep(B|D)$.

Statement 2.1 indicates the antimonotonic property between the size of the condition attribute subset and the cardinality of the FWNIG. Accordingly, the smaller the attribute subset is, the larger the cardinality of the generated FWNIG, and vice versa. Therefore, $|[u]_C^{\varepsilon, w}|$ reaches the minimum value on the decision table $I = (U, C, D)$.

Proposition 3. Let $I = (U, C, D)$ be a decision table, $K \subseteq U$ and $A \subseteq C$. If $\alpha = 1$ and $\beta = 0$, then $\underline{FWN}_A(K) = \underline{WNR}_A(K)$ and $\overline{FWN}_A(K) = \overline{WNR}_A(K)$.

Proof. From Equation (13) and $\alpha = 1$, we have $\underline{FWN}_A(K) = \left\{ u \in U : \frac{|\widehat{[u]_A^{\varepsilon,w}} \cap K|}{|\widehat{[u]_A^{\varepsilon,w}}|} \geq 1 \right\}$. This leads to $\underline{FWN}_A(K) = \left\{ u \in U : |\widehat{[u]_A^{\varepsilon,w}} \cap K| \geq |\widehat{[u]_A^{\varepsilon,w}}| \right\}$. Hence, we obtain $\underline{FWN}_A(K) = \left\{ u \in U : \widehat{[u]_A^{\varepsilon,w}} \subseteq K \right\} = \underline{WNR}_A(K)$. Similarly, for $\beta = 0$, $\overline{FWN}_A(K) = \left\{ u \in U : \frac{|\widehat{[u]_A^{\varepsilon,w}} \cap K|}{|\widehat{[u]_A^{\varepsilon,w}}|} \geq 0 \right\} = \left\{ u \in U : \widehat{[u]_A^{\varepsilon,w}} \cap K \neq \emptyset \right\} = \overline{WNR}_A(K)$.

According to Proposition 3, WNRS can be considered a special case of FWNRS. In addition, when $\alpha = 1$, $\beta = 0$, and all weights in the FWNRS model are set to 1, the traditional NRS model can also be viewed as a special case of FWNRS. Intuitively, FWNRS enables attribute reduction methods to achieve better generalization across a wide range of data scenarios.

5. Attribute Reduction based on FWNRSs

This section presents a main approach based on our model for attribute reduction. The aim is to extract an optimal subset of condition attributes, known as a reduct set, that preserves the crucial information equivalent to that generated by the complete set of condition attributes of the table. In particular, we shall define a reduct for the decision table through the fuzzy weighted dependency function.

Definition 4. Let $I = (U, C, D)$ be a decision table. A condition attribute subset A is called a *Dep-reduct* of attribute set C relative to D if A satisfies the following conditions:

- 1) $Dep(A|D) \geq Dep(C|D)$,
- 2) $\forall b \in A, Dep(A \setminus \{b\}|D) < Dep(A|D)$.

Based on this definition, it can be observed that a decision table may have multiple reducts. Similar to classical rough set models, finding all reducts is known to have exponential time complexity w.r.t the number of condition attributes. Nevertheless, in practical applications, it is unnecessary to find all reducts. Instead, the emphasis is on extracting the optimal one according to a specific criterion. To design an efficient attribute reduction method, we start by defining the significance of an attribute w.r.t a given condition attribute subset.

Definition 5. Given $I = (U, C, D)$ with $A \subseteq C$ and $b \in C \setminus A$, the *significance* degree of b w.r.t A , denoted $Sig(b, A)$, is defined as follows.

$$Sig(b, A) = Dep(A \cup \{b\}|D) - Dep(A|D) \quad (19)$$

The significance of a condition attribute relative to a given attribute subset is measured by the alteration in the certainty degree resulting from adding that attribute to the subset. Intuitively, the larger the value of $Sig(b, A)$, the more vital attribute b becomes. Note that the fuzzy weighted dependency function does not exhibit monotonicity with respect to the cardinality of the conditional attribute subset. Condition attributes that have positive significance values is regarded as important and contribute substantially to the classification performance. On the contrary, the attribute is regarded as unimportant and noisy in the decision table. Next, we design a filter-wrapper method to extract a subset of conditional attributes in decision tables. The method is illustrated by the pseudocode below.

Algorithm 1. Filter Wrapper Attribute Reduction based on FWNRSs (ARFNR)

Input: a decision table $I = (U, C, D)$, a stopping threshold δ and a neighborhood radius ε

Output: an optimal attribute subset Red

```

1. determine the weight of each condition attribute according to Equation (5);
2. for  $b \in C$  do
3.   calculate  $\widetilde{G(\{b\})}$ ;
4.   calculate  $Dep(\{b\}|D)$  according to Equation (18);
5. end for
6.  $Red = \{b_0\}$  which satisfies  $Dep(\{b_0\}|D) = \max_{b \in C} \{Dep(\{b\}|D)\}$ ;
7.  $W_1 \leftarrow Red$ ;  $q = 1$ ;
// Filter stage
8. while  $C \setminus Red \neq \emptyset$  do
9.   for each  $b \in C \setminus Red$  do
10.    calculate  $Dep(Red \cup \{b\}|D)$  according to Equation (18);
11.    calculate  $Sig(b, Red)$  according to Equation (19);
12.   end for
13.   select attribute  $b_0$  such that  $Sig(b_0, Red) = \max_{b \in C \setminus Red} \{Sig(b, Red)\}$ ;
14.    $Red \leftarrow Red \cup \{b_0\}$ ;
15.    $W_{++q} \leftarrow Red$ ;
16.   if  $Sig(b_0, Red) \leq \delta$  then
17.     break;
18.   else
19.     go to step 8;
20.   end if
21. end while
// Wrapper stage
22. for  $(i = 1; i \leq q; i++)$  do
23.   calculate the classification accuracy for  $W_i$  by applying a classifier (ten-fold);
24. end for
25.  $Red = W_{best}$  with  $W_{best} \in \{W_1, W_2, \dots, W_q\}$  is the condition subset that has the best classification performance;
26. return  $Red$ 

```

We now examine the runtime of the ARFNR algorithm. The analysis focuses on the filter and wrapper phases. In step 1, the algorithm determines the attribute weights through computations on a coefficient matrix consisting of $|U|$ rows and $|C|$ columns. Thus, the runtime of this step is $O(|U| \times |C|)$. In step 2-5, for each condition attribute $b \in C$, the algorithm constructs $\widetilde{G(\{b\})}$ and calculates $Dep(\{b\}|D)$ based on the WNIGs. Specifically, the construction of $\widetilde{G(\{b\})}$ requires assigning weights to $|U|$ WNIGs, and the calculation of $Dep(\{b\}|D)$ is also performed based on these information granules. Therefore, the runtime for one condition attribute is $O(|U|^2)$. Since this process is repeated for all condition attributes, the runtime of steps 2-5 is $O(|U|^2 \times |C|)$. Next, the runtime of steps 6-7 for identifying the attribute with the highest significance is $O(|C|)$. In the filter stage (steps 8-22), the algorithm adopts a greedy strategy, selecting the most significant attribute at each iteration to add to the reduct that satisfies a given threshold δ . In each iteration of the while loop, the processing time is equivalent to that of the for loop from steps 2 to 5. Therefore, the running time of the entire while loop is $O(|U|^2 \times |C|^2)$.

Suppose that the runtime of the classifier applied in the wrapper phase is $O(T)$. Consequently, the wrapper phase (steps 22-26) is conducted in $O(|C| \times T)$ time. Overall, the runtime of the ARFNR method is $O(|U|^2 \times |C|^2) + O(|C| \times T)$.

Example 3. Applying the ARFNR algorithm with the threshold of $\delta = 0.01$ to find a reduct in **Table 2**.

First, we initialize W and Red as empty sets. Then, we can easily compute the certainty classification degree from each condition attribute in the decision table, the same as example 2, and obtain: $Dep(\{b_1\}|D) = 0.33$, $Dep(\{b_2\}|D) = 0.50$, $Dep(\{b_3\}|D) = 0.00$, $Dep(\{b_4\}|D) = 0.83$. It is easy to see that $Dep(\{b_4\}|D) = 0.83$ is the largest. Thus, we get $Red = \{b_4\}$ and $W = \{W_1 = \{b_4\}\}$. In the while loop, we have:

$$Dep(Red \cup \{b_1\}|D) = 0.67 \Rightarrow Sig\{b_1, Red\} = 0.67 - 0.83 < 0.00,$$

$$Dep(Red \cup \{b_2\}|D) = 1.00 \Rightarrow Sig\{b_2, Red\} = 1.00 - 0.83 = 0.17,$$

$$Dep(Red \cup \{b_3\}|D) = 0.83 \Rightarrow Sig\{b_3, Red\} = 0.83 - 0.83 = 0.00.$$

Here, the attribute b_2 has the maximal significance, we obtain $Red = \{b_4, b_2\}$ and $W = \{W_1 = \{b_4\}, W_2 = \{b_4, b_2\}\}$. Since $Sig\{b_2, Red\} > 0.01$, the algorithm continues to the next iteration.

$$Dep(Red \cup \{b_1\}|D) = 1.00 \Rightarrow Sig\{b_1, Red\} = 1.00 - 1.00 = 0.00,$$

$$Dep(Red \cup \{b_3\}|D) = 1.00 \Rightarrow Sig\{b_3, Red\} = 1.00 - 1.00 = 0.00.$$

Since $Sig\{b_1, Red\} = Sig\{b_3, Red\} = 0.00$, the algorithm stops. Therefore, the algorithm returns the family of reducts is $W = \{W_1, W_2\}$.

It can easily see that weighs of attribute b_2 and b_4 are greater than 1. This indicates that these attributes are important in the decision table. This conclusion is further supported by the reduct obtained by the algorithm, which contains both attributes.

6. Experiments

In the previous part, we introduced a method for attribute reduction through the FWNRS model. In this part, we shall perform a series of experiments to evaluate the robustness and effectiveness of the ARFW method by comparing it with several state-of-the-art methods in terms of the cardinality of reducts, the runtime as well as the classification accuracy.

6.1 Experiment Setup

The algorithms used for comparison include the neighborhood discrimination index-based algorithm (HANDI) (Wang et al., 2018), the KNRS-based algorithm (NNRS) (Wang et al., 2019), the WNRS-based algorithm (WNRS) (Hu et al., 2021), and the WKNRS-based algorithm (WKNRS) (Wang & Zhao, 2024). These methods are coded in Python and executed on a personal computer equipped with an Intel Xeon E3-1545M v5 CPU (2.90 GHz, 8 cores) and 16 GB of RAM. The twelve benchmark datasets used in the experiments are adopted from the UCI Machine Learning Repository and are listed in **Table 3**. Before reduction, all numerical attributes in the datasets are first normalized into the interval $[0, 1]$.

Our method involves two main parameters, ε and δ , with the first parameter ε being used to construct fuzzy weighted neighborhood granules. Different values of ε may result in different reducts. Consequently, this parameter has a substantial effect on the performance of the method. To ensure fair comparisons, the parameter ε of HANDI, WNRS, and the proposed method is varied from 0 to 1 with a step size of 0.05 to find an optimal reduct for each dataset. For the WKNRS and NNRS methods, the parameter k is varied from 1 to 18 to identify the optimal reduct. The second parameter δ is used as the stopping threshold in the Filter phase. This parameter affects the cardinality of reduct returned by the algorithm. As suggested by

extensive experimental evaluations, we fix $\delta = 0.01$ for all datasets. In addition, to calculate the lower approximation w.r.t conditional attribute subset A , we set $\alpha = \max_{D_i \in U/D} \left\{ \frac{|D_i \cap [u]_A^{\varepsilon, w}|}{|[u]_A^{\varepsilon, w}|} \right\}$ for each object u .

Table 3. Summary of datasets used in the experiments.

ID	Datasets	Abbreviation	Objects	Attributes	Classes
1	Glass	Glass	214	9	6
2	Robot-failures	Robot	164	90	5
3	Triazines	Triazines	186	60	2
4	Heart Disease	Heart	270	13	2
5	Leaf	Leaf	340	15	30
6	SPECTF Heart	SPECTF	349	44	2
7	Libras Movement	Movement	360	90	15
8	Vehicle Silhouettes	Vehicle	846	18	4
9	SRBCT	SRBCT	83	2308	4
10	WarpAR10P	WarpAR10P	130	2400	10
11	Tumors C	Tumors	60	7129	2
12	Prostate	Prostate	102	12600	2

The four methods are applied to the training set to find a reduct. Afterward, we use the k -nearest neighbor (k-NN) rule with $k = 5$ and the support vector machine (SVM) to assess the classification accuracy of the reducts on the testing set. The final results are obtained by calculating the mean as well as standard deviation of the accuracies over the ten folds.

6.2 Experiment Results

Tables 4 to 5 present the cardinality of the reducts, the runtime (seconds) and the corresponding parameter of each method, with the results yielding the smallest reduct sizes highlighted in bold. Note that the parameter values of ε (epsilon) and k reported in **Tables 4 and 5** are those at which the reduct obtained by each method achieves the best classification accuracy. Clearly, compared to the raw data, all methods efficiently reduce the number of conditional attributes. Nonetheless, methods that employ the assignment of object weights, such as WKNRS and ARFNR, generally obtain smaller condition attribute sets than the other methods, leading to competition between these two methods in terms of reduct size.

Table 4. Size of reducts, runtime and corresponding parameters with KNN classifier.

Datasets	HANDI			NNRS			WNRS			WKNRS			ARFNR		
	ε	size	time	k	size	time	ε	size	time	k	size	time	ε	size	time
Glass	0.1	6.7	0.013	2	8.1	0.243	0.4	5.0	0.886	14	2.3	0.951	0.5	2.9	2.074
Robot	0.05	1.0	0.012	2	1.0	0.201	0.5	8.1	16.30	13	3.9	10.57	0.1	4.5	29.91
Triazines	0.9	17.1	0.168	3	22.1	6.727	0.9	15.2	35.32	7	5.1	7.486	0.75	3.8	23.28
Heart	0.1	8.6	0.032	7	10.6	0.634	0.25	9.8	3.078	15	2.1	1.186	0.75	5.1	9.354
Leaf	0.1	1.0	0.006	9	4.0	0.564	0.35	10.3	11.16	13	7.3	6.320	0.9	5.1	20.96
SPECTF	0.1	13.4	0.244	3	5.7	2.709	0.95	5.1	19.01	3	1.0	1.936	0.15	3.8	68.58
Movement	0.0	3.0	0.120	2	13.1	36.88	0.25	7.9	75.21	16	9.4	57.86	0.8	7.9	265.5
Vehicle	0.05	10.4	0.526	3	3.1	3.083	0.25	12.7	109.6	17	4.1	12.06	0.25	8.0	429.6
SRBCT	0.45	10.0	2.687	3	4.2	11.57	1.0	6.1	35.07	17	5.8	144.9	0.95	2.9	46.36
WarpAR10P	0.15	1.0	0.383	3	3.4	14.71	0.2	4.3	98.91	15	8.7	442.3	0.7	3.5	194.9
Tumors	0.0	2.0	3.110	3	3.4	19.81	0.7	4.2	57.43	12	3.0	185.2	0.0	1.2	19.78
Prostate	0.35	13.6	68.81	3	2.0	29.84	0.3	2.1	120.3	16	3.4	667.6	0.5	1.4	277.8
Average	0.20	7.32	6.343	3.58	6.73	10.58	0.50	7.57	48.52	13.16	4.68	128.2	0.53	4.18	115.7

First, we analyze the runtime of the methods. It can be seen that HANDI and NNRS are the two methods with the fastest execution time among the compared methods. The reason is that these algorithms do not

establish weighting mechanisms and only perform computations on information granules with relatively simple structures. For the three remaining algorithms, the methods using object weights usually require longer execution time. This is clearly shown when WKNRS and FWNRS process datasets containing a large number of objects. In each iteration of attribute selection, the proposed algorithm usually has a high runtime because it has to simultaneously perform the integration of object weights and attribute weights. However, on some high-dimensional datasets, including SRBCT, WarpAR10P, Tumors, and Prostate, the proposed algorithm has a lower running time than WKNRS. This is because the number of attributes in the reduct obtained by our algorithm is smaller, thereby reducing the computational cost in the attribute selection process.

Table 5. Size of reducts, runtime and corresponding parameters with SVM classifier.

Datasets	HANDI			NNRS			WNRS			WKNRS			ARFNR		
	ε	size	time	k	size	time	ε	size	time	k	size	time	ε	size	time
Glass	0.1	6.7	0.013	4	5.6	0.178	0.0	2.0	0.366	17	2.6	1.085	0.05	3.6	1.909
Robot	0.05	1.0	0.012	4	1.0	0.209	0.85	4.9	7.934	2	2.8	3.572	0.05	2.5	18.54
Triazines	0.3	15.0	0.146	5	25.1	7.759	0.85	16.2	38.16	2	15.8	5.970	0.8	6.8	25.42
Heart	0.85	7.3	0.027	7	10.6	0.634	0.5	11.3	4.906	17	2.4	1.480	0.5	4.6	9.328
Leaf	0.05	1.0	0.005	8	3.4	0.479	0.05	6.1	6.292	11	6.9	5.831	0.25	3.0	19.77
SPECTF	0.1	13.4	0.244	8	6.0	2.901	0.2	10.5	22.03	2	3.9	3.885	0	2.3	31.44
Movement	0.0	3.0	0.120	2	13.1	36.88	0.25	7.9	75.21	17	8.8	41.15	1	8.9	267.7
Vehicle	0.05	10.4	0.526	2	2.7	2.462	0.15	13.6	117.5	17	4.1	12.06	0.25	9.7	411.5
SRBCT	0.25	7.2	1.955	2	4.6	12.79	0.9	6.4	67.40	14	5.2	130.6	0.9	3.4	47.63
WarpAR10P	0.45	1.0	0.382	2	4.0	18.20	0.7	7.3	237.7	13	7.6	436.2	0.8	5.3	217.2
Tumors	0.35	11.2	17.41	3	3.4	19.81	0.25	2.6	31.58	18	2.6	160.1	0.85	2.0	53.85
Prostate	0.6	48.4	243.1	4	2.0	34.62	0.3	2.1	118.9	16	3.4	667.6	0.3	1.4	221.3
Average	0.26	10.47	22.00	4.25	6.79	11.41	0.42	7.58	60.66	12.17	5.51	122.5	0.48	4.46	110.5

Among the 24 comparison cases, ARFNR achieves the smallest reduct size in 9 cases, whereas WKNRS attains this result in only 6 cases. On both the KNN and SVM models, the ARFNR algorithm achieves average reduct sizes of 4.18 and 4.46, respectively, which are significantly smaller than those obtained by the other methods. **Figures 4** and **5** illustrate comparisons of the number of conditional attributes obtained by the three algorithms HANDI, WNRS, and ARFNR for parameter values ranging from 0 to 1 with a step size of 0.05. We can observe that the parameters have a significant impact on the size of the reducts returned by the methods. Notably, the red curve representing the number of attributes returned by FWNRS generally lies below the other curves. This implies that, for any given radius threshold, the proposed algorithm is able to eliminate redundant attributes more effectively, particularly on high-dimensional datasets, including SRBCT, Tumor, and Prostate.

Tables 6 and **7** show the classification performance of the methods, where the Raw data denotes the classification accuracy obtained by applying all condition attributes for evaluation. For each method, the classification accuracies reported in **Tables 6** and **7** correspond to the highest accuracies obtained by tuning the number of neighbors in the NNRS, WKNRS methods and the neighborhood radius ε for the remaining methods on each dataset.

We start by evaluating the classification performance of condition attribute subsets returned by the methods using the KNN classifier. From **Table 6**, the reducts generated by the four algorithms have the potential to significantly improve classification accuracy compared to the raw datasets. It can be observed quickly that ARFNR demonstrates superior classification accuracy compared with the other methods on most datasets. Specifically, there is only one case out of twelve in which our algorithm achieves classification accuracy equal to that of the HANDI algorithm. Besides, the Win/Draw/Loss ratio of ARFNR compared to HANDI,

NNRS, WNRS, and WKNRS is 11/1/0, 10/0/3, 12/0/0, and 12/0/0, respectively. In terms of average accuracy comparison, our method attained 74.73%, outperforming the other methods, HANDI, NNRS, WNRS, and WKNRS, with corresponding values of 67.78%, 69.75%, 69.92%, and 68.68%.

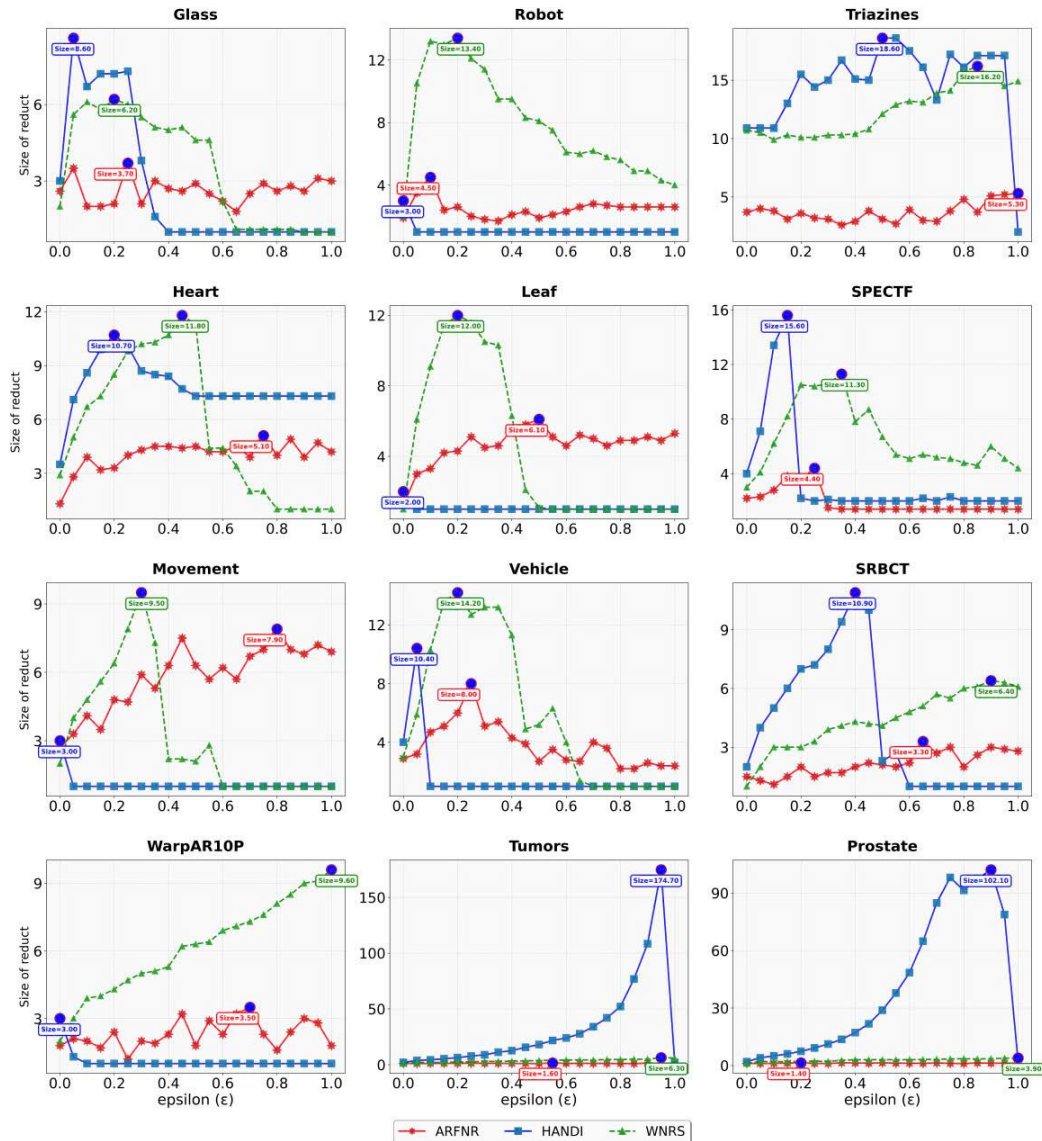


Figure 4. Reduct size of algorithms with varying parameters based on KNN.

Notably, on datasets characterized by initially low classification accuracy, including Glass, Robot-failures, Leaf, Movement, and SRBCT, the ARFNR algorithm demonstrates outstanding performance, despite the very small size of the reducts returned. This indicates that the method is effective in eliminating noisy attributes while retaining those with high significance. This advantage stems from the combination of attribute weights and object weights in the FWNRS model, which helps generate meaningful information granules and ensures that the attribute evaluation measure does not overlook important attributes.

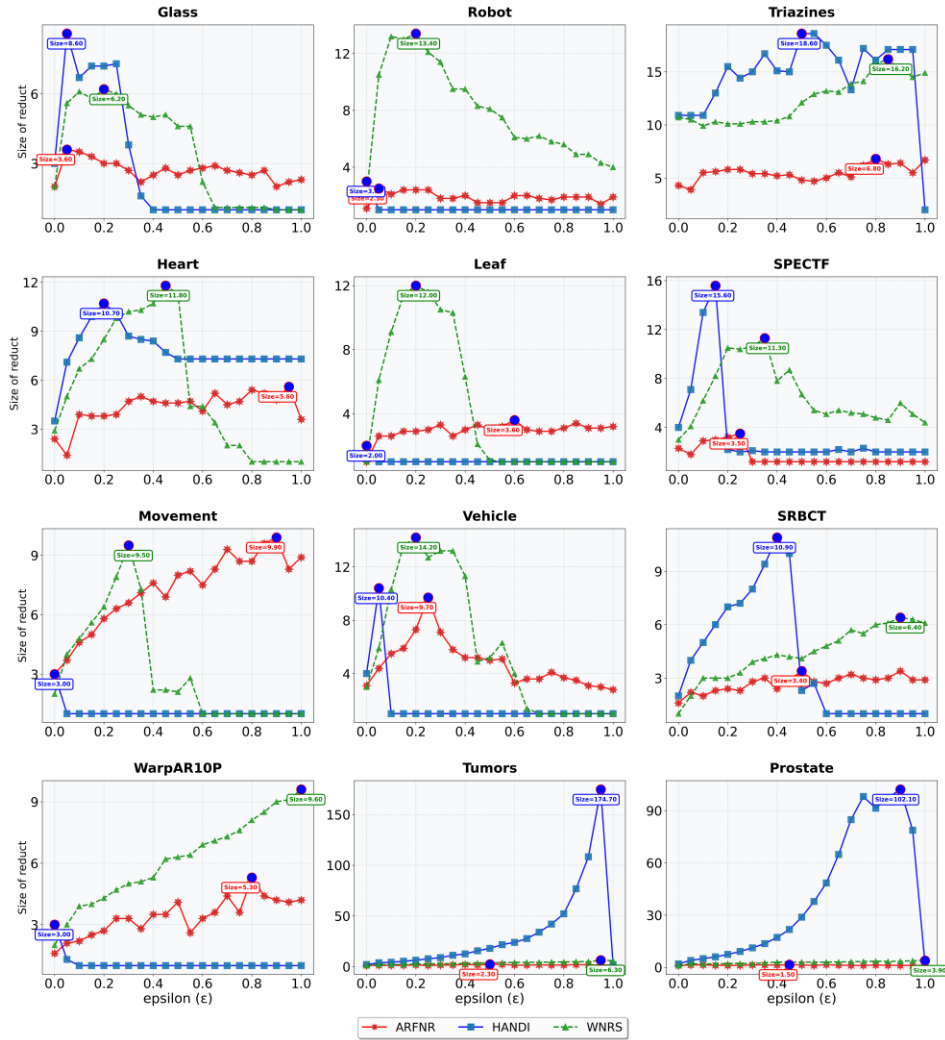


Figure 5. Reduct size of algorithms with varying parameters based on SVM.

Table 6. Classification accuracy of methods based on KNN.

ID	Datasets	Raw data	HANDI	NNRS	WNRS	WKNRS	ARFNR
1	Glass	63.20±10.54	65.06±8.88	62.75±10.84	65.04±10.1	69.16±7.48	72.49±6.54
2	Robot-failures	55.00±16.17	59.67±12.71	68.93±9.21	62.39±9.4	60.29±10.37	66.47±10.39
3	Triazines	74.15±5.8	75.82±7.68	77.95±8.31	75.82±8.4	75.26±5.17	81.75±4.27
4	Heart	84.07±4.29	85.56±6.64	84.82±5.07	85.93±5.74	79.63±8.94	88.52±4.43
5	Leaf	27.06±5.68	37.65±7.31	47.06±7.2	46.18±13.94	50.59±8.75	55.00±7.34
6	SPECTF	80.81±9.32	84.24±6.62	89.68±4.72	85.1±4.21	78.23±7.49	87.41±4.48
7	Movement	58.89±5.37	51.11±7.07	53.61±7.3	55.83±7.91	60.56±5.37	63.61±5.92
8	Vehicle	71.27±5.21	71.86±3.19	72.91±5.12	73.51±5.04	72.68±5.55	76.94±4.82
9	SRBCT	57.64±15.03	72.50±12.96	68.61±10.48	73.45±11.57	69.58±14.94	76.94±7.33
10	WarpAR10P	40.00±14.86	51.54±8.15	47.69±8.73	56.15±12.59	55.39±13.47	61.54±6.28
11	Tumors	68.33±14.59	75.00±11.78	76.67±11.65	73.33±8.6	71.67±8.05	75.00±11.78
12	Prostate	73.46±8.35	83.36±13.98	86.27±8.36	86.36±10.36	81.46±11.71	91.09±8.75
Average		62.82±9.6	67.78±8.91	69.75±8.08	69.92±8.99	68.68±8.94	74.73±6.86
Win/Draw/Loss		12/0/0	11/1/0	10/0/3	12/0/0	12/0/0	

The outcomes show consistency when using the SVM classifier to assess the effectiveness of methods. As shown in **Table 7**, our method achieved the highest accuracy in 8 out of 12 datasets when the best results are considered. Moreover, the average accuracy of reducts returned from ARFNR reached 80.73%, while remaining methods achieved respective averages of 74.54%, 72.67%, 77.04%, and 72.52%. More particularly, the Win/Draw/Loss comparative ratios between our algorithm with HANDI, NNRS, WNRS, and WKNRS algorithms are 7/2/3, 10/1/1, 10/0/2, and 12/0/0, respectively. These results show that ARFNR outperforms the other methods in both the cardinality of the obtained reducts and in classification accuracy.

Table 7. Classification accuracy of methods based on SVM.

ID	Datasets	Raw data	HANDI	NNRS	WNRS	WKNRS	ARFNR
1	Glass	63.09±7.09	66.36±7.21	64.52±10.79	67.29±6.44	63.55±8.47	73.38±8.56
2	Robot-failures	49.49±8.31	54.78±12.4	57.98±14.30	49.52±6.38	53.01±9.84	61.14±9.24
3	Triazines	85.53±6.19	88.16±6.72	88.54±3.57	88.80±6.68	83.98±6.88	90.41±5.94
4	Heart	92.22±5.08	91.85±4.55	91.85±6.25	92.22±5.08	81.85±9.31	91.85±5.47
5	Leaf	16.47±4.43	34.12±7.62	35.59±11.04	36.47±14.82	39.12±8.67	52.35±9.28
6	SPECTF	86.51±5.8	83.96±4.05	87.09±6.10	85.67±4.04	77.35±6.61	83.66±2.8
7	Movement	78.89±6.83	56.94±7.77	60.56±12.06	75.28±8.93	77.78±6.93	79.44±7.32
8	Vehicle	75.52±5.05	75.76±3.88	64.53±6.08	75.87±4.29	62.64±8.92	77.29±5.55
9	SRBCT	97.64±4.99	98.98±3.51	88.06±11.17	94.17±6.17	89.03±10.93	96.39±5.83
10	WarpAR10P	73.85±15.04	54.62±9.9	61.54±10.26	73.85±16.7	71.54±13.59	77.69±12.27
11	Tumors	80.00±10.54	90.00±11.65	86.66±10.54	88.33±11.25	83.33±13.61	90.00±11.65
12	Prostate	72.64±11.59	99.00±3.16	86.18±10.65	97.00±4.83	87.09±9.52	97.09±4.69
Average		72.65±7.58	74.54±6.87	72.76±9.40	77.04±7.97	72.52±9.44	80.73±7.38
Win/Draw/Loss		11/0/1	7/2/3	10/1/1	10/0/2	12/0/0	

6.3 Sensitivity Analysis of Parameter

In this section, we analyze the impact of the ε parameter on the performance of the proposed method. According to the illustrations in **Figures 4 to 7**, several important observations can be drawn as follows.

(1) **The reduct size:** On most datasets, the reduct sizes returned by our algorithm tend to increase gradually, then decrease or remain nearly unchanged as the neighborhood radius ε increases. When ε is small, the weighted neighborhood information granules contain fewer objects, thereby decreasing the number of objects considered in the construction of the FWNIGs. As a result, these information granules become finer, making the stopping condition of the algorithm easier to satisfy. For medium values of ε , the reduct sizes of ARFNR exhibit only slight fluctuations, with a significantly smaller fluctuation amplitude than those of HANDI and WNRS. This indicates that our method is less sensitive to variations in the neighborhood radius ε . Notably, on high-dimensional datasets containing a large number of attributes, including SRBCT, WarpAR10P, Tumors, and Prostate, the reduct sizes show almost no significant variation as ε changes. This result demonstrates that ARFNR maintains robust stability when handling high-dimensional datasets.

(2) **The classification performance:** Obviously, the classification performance of the returned reducts is significantly affected by variations in the parameter ε . On most low-dimensional datasets, including Glass, Robot-failures, Triazines, and Leaf the classification accuracy exhibits slight fluctuations across different values of ε and achieves optimal performance at a certain point. Meanwhile, the classification accuracy on the SPECTF and Vehicle datasets tends to decrease markedly when ε increases to a medium range of values. This is because when the neighborhood radius becomes larger, the weighted neighborhood information granules may include more noisy objects. Therefore, the algorithm may encounter many difficulties in identifying important attributes. However, for some datasets where improving classification accuracy is difficult, such as Heart and Movement, selecting a larger neighborhood radius allows the weighted neighborhood information granules to include more objects that are useful for classification. Consequently,

the algorithm can identify an efficient reduct. These observations indicate that the selection of an appropriate value of ε depends strongly on the distribution characteristics of the data. For high-dimensional datasets, including SRBCT, WarpAR10P, Tumors, and Prostate, the classification accuracy of the reducts fluctuates only slightly when ε changes. This can be explained by the fact that the reducts obtained on these datasets show almost no significant changes in their attribute composition when the neighborhood radius is adjusted. Therefore, classification performance is maintained relatively stable over different ranges of values of ε .

(3) The runtime: The effect of parameter ε on the runtime of the method shows a trend similar to that on the reduct size. When ε lies in a small range of values, the stopping condition of the algorithm is more easily satisfied. Therefore, the number of iterations in the attribute selection process is usually smaller, leading to a lower execution time. Conversely, when ε increases to a larger range of values, the number of selected attributes tends to increase. In this case, the algorithm needs to consider and evaluate more important attributes, which increases the computational cost and prolongs the runtime.

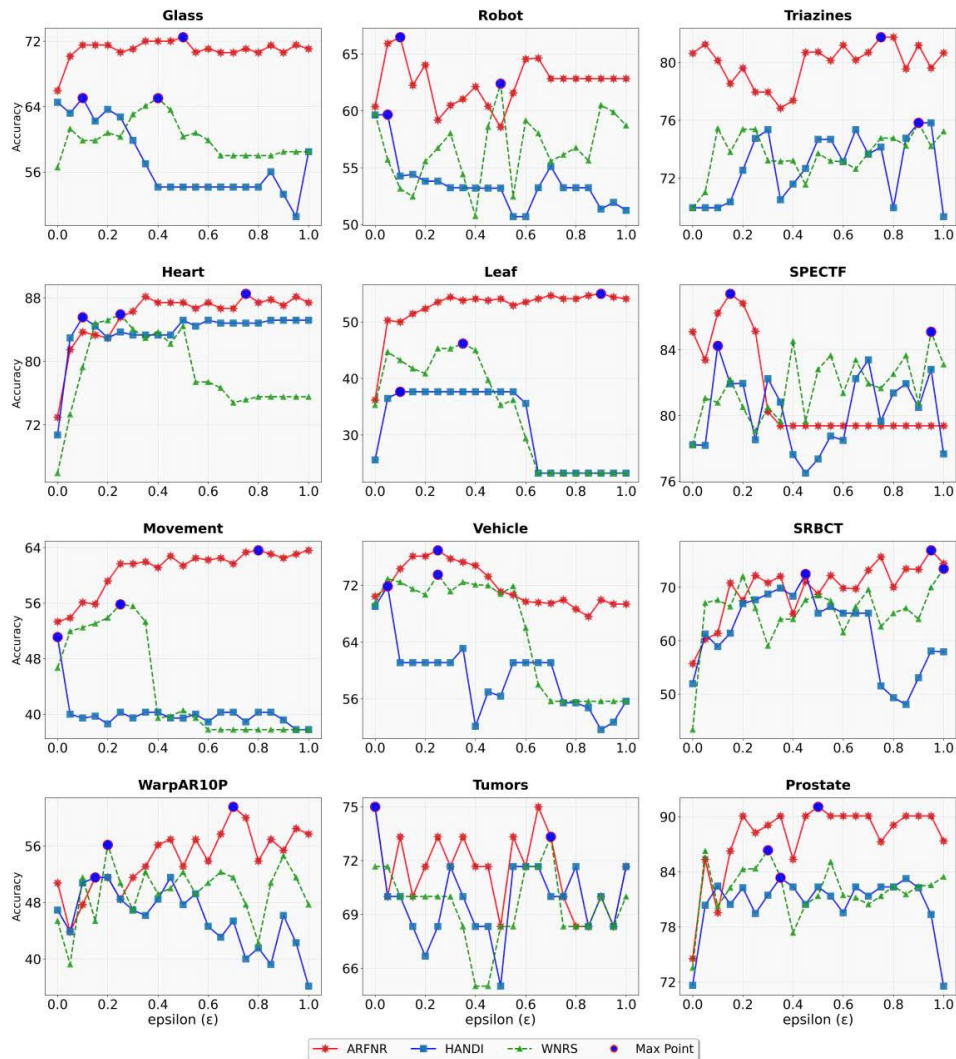


Figure 6. Classification performance of algorithms with varying parameters based on KNN.

An examination of **Figures 6 and 7** shows that the red line, representing the ARFNR algorithm, tends to lie above the other lines over a wide range of neighborhood radius values. This highlights the effectiveness of ARFNR compared with the remaining methods. However, for some datasets exhibiting low original classification performance, including leaf, movement, and vehicle, the performance of the algorithms tends to decline as the radius threshold increases. This is because neighborhood information granules may include a large number of noisy objects, which negatively affect the effectiveness of attribute evaluation measures. In such cases, the models tend to prioritize smaller radius thresholds in order to mitigate the impact of noise and improve performance for each classifier model.

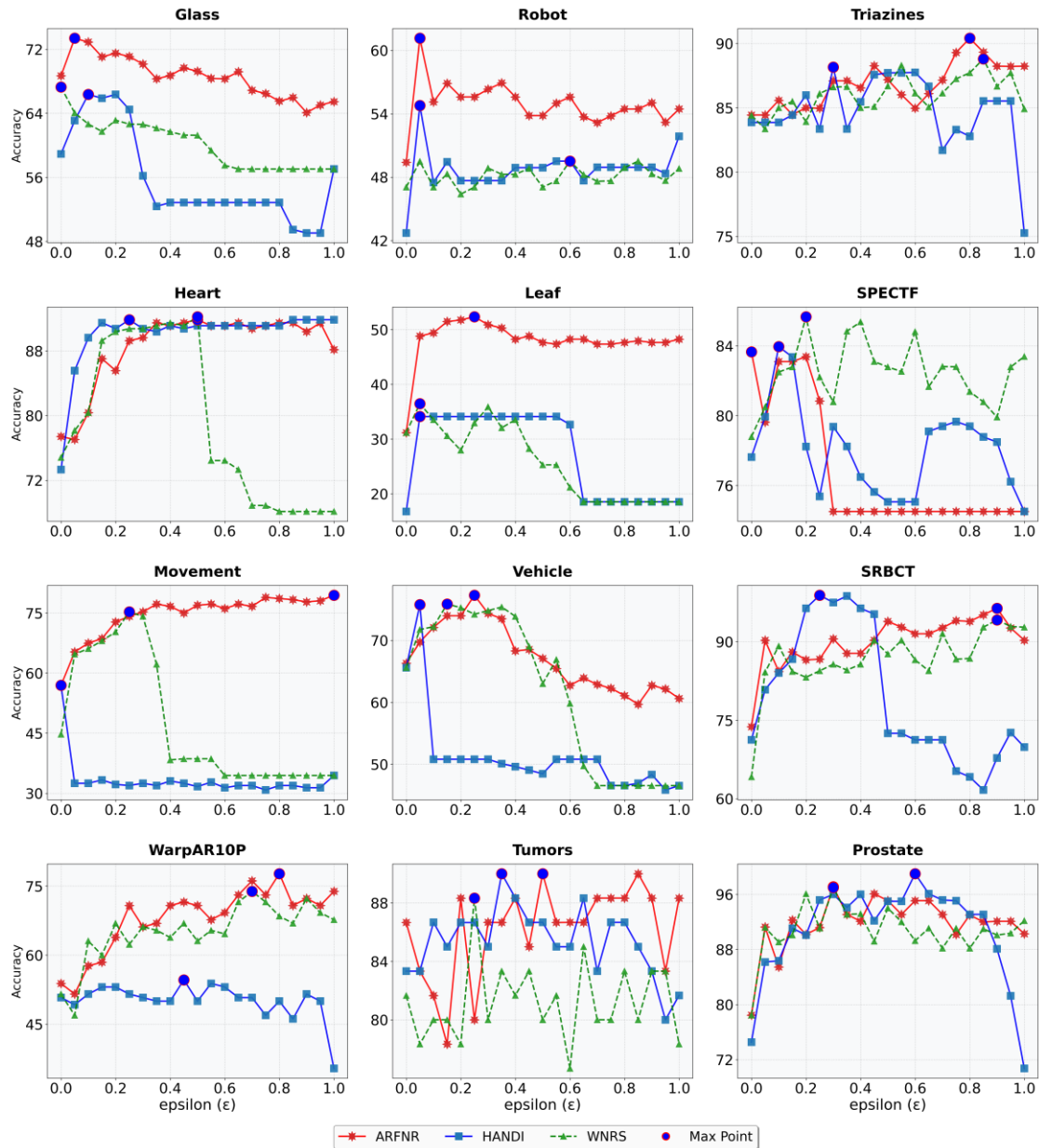


Figure 7. Classification performance of algorithms with varying parameters based on SVM.

6.4 Statistical Analysis

To analyze the statistical performance of the four methods, we first conduct the Friedman test (Friedman, 1940) to examine whether there are significant differences in their efficiency. We denote M as the number of datasets, L as the number of methods, and k_j as the average rank of the j -th method. The Friedman test statistic is defined as follows:

$$\chi_F^2 = \frac{12M}{L(L+1)} \left(\sum_{j=1}^L k_j^2 - \frac{L(L+1)^2}{4} \right) \quad (20)$$

and

$$\varphi_F = \frac{(M-1)\chi_F^2}{L(M-1)-\chi_F^2} \quad (21)$$

where, χ_F^2 follows a χ^2 distribution with degrees of freedom $(L-1)$ and φ_F obeys the F -distribution with $(L-1)$ and $(L-1)(M-1)$ degrees of freedom.

In this experiment, since $L = 5$ and $M = 12$, the distribution of φ_F has 4 and 44 degrees of freedom. Based on the Friedman test, the average rankings of all methods across the 12 datasets are presented in **Table 8**, where the test values for the KNN and SVM classifiers are $\varphi_F = 7.051$ and $\varphi_F = 5.027$, respectively. Clearly, these values are much higher than the critical value $\varphi_F(4, 44) = 2.58$ at the significance level $\alpha = 0.05$. Thus, the null hypothesis that all methods have the equivalent classification performance is refuted for KNN and SVM classifiers. This result shows that there are statistically significant differences in classification performance among the compared attribute reduction methods for each classifier.

Table 8. Statistical test-based comparison of five algorithms using two classifiers.

Classifier	Mean ranking					χ_F^2	φ_F
	HANDI	NNRS	WNRS	WKNRS	ARFNR		
KNN	3.917	3.083	2.875	3.833	1.292	21.483	7.051
SVM	3.125	3.583	2.583	4.083	1.625	17.250	5.027

Next, the Nemenyi post hoc test (Demšar, 2006) is employed to identify the specific differences among the methods. In which, the critical difference (CD) is computed as follows:

$$CD_\alpha = q_\alpha \sqrt{\frac{L(L+1)}{6M}} \quad (22)$$

where, q_α is the critical tabulated value.

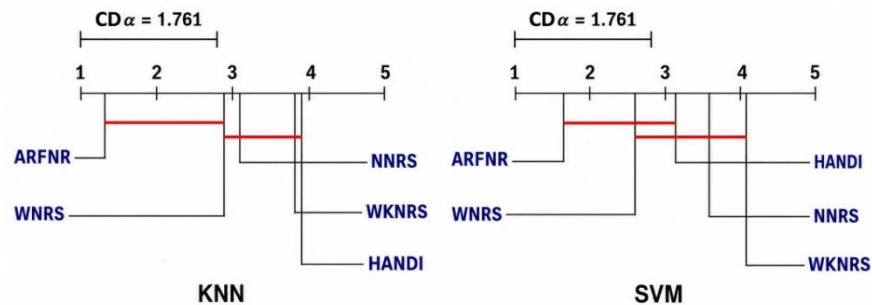


Figure 8. Nemenyi test results for attribute reduction methods using KNN and SVM.

Based on study of Demšar (2006), we obtain $q_\alpha = 2.728$ and the corresponding critical difference $CD_\alpha = 1.761$. As shown in **Figure 8**, ARFNR differs significantly from HANDI, NNRS and WKNRS with the KNN classifier. Under the SVM classifier, ARFNR shows significant differences from NNRS and WKNRS, whereas its differences from WNRS and HANDI are not statistically significant. However, **Table 8** shows that ARFNR achieves a better average ranking than the other methods on both KNN and SVM classifiers.

We also conduct a Paired Samples Test to assess the differences between ARFNR and the comparison algorithms at a 95% confidence level. As shown in **Tables 9** and **10**, the corresponding p-values (two-tailed) were 3.08E-04, 0.007, 1.70E-5, and 2.00E-06 for KNN, 0.041, 0.002, 0.026 and 4.00E-06 for SVM. The results indicate that ARFNR outperforms the other algorithms.

Table 9. Paired samples t-test results with KNN.

CID: Confidence interval of the difference		Paired differences					t	df	Significance (2-tailed)
		Mean	Standard Deviation	Standard Error mean	95% CID				
					Lower	Upper			
Pair 1	ARFNR-HANDI	6.949	4.656	1.344	3.991	9.918	5.170	11	3.08E-04
Pair 2	ARFNR-NNRS	4.984	5.245	1.514	1.652	8.317	3.292	11	0.007
Pair 3	ARFNR-WNRS	4.806	2.310	0.667	3.338	6.273	7.208	11	1.70E-05
Pair 4	ARFNR-WKNRS	6.047	2.363	0.682	4.545	7.548	8.865	11	2.00E-06

Table 10. Paired samples t-test results with SVM.

CID: Confidence interval of the difference		Paired Differences					t	df	Significance (2-tailed)
		Mean	Standard Deviation	Standard Error mean	95% CID				
					Lower	Upper			
Pair 1	ARFNR-HANDI	6.347	9.511	2.746	0.304	12.390	2.312	11	0.041
Pair 2	ARFNR-NNRS	8.133	7.203	2.079	3.556	12.709	3.911	11	0.002
Pair 3	ARFNR-WNRS	3.852	5.178	1.495	0.562	7.141	2.577	11	0.026
Pair 4	ARFNR-WKNRS	8.368	3.470	1.002	6.164	10.573	8.355	11	4.00E-06

7. Conclusion

This study introduces a new type of NIG constructed by integrating conditional attribute weights and object weights in each granule. This approach allows a detailed evaluation of the role of each object in the NIG and the influence of condition attributes on decision attributes. From the proposed information granule, we develop a FWNRS model through the concepts of lower and upper approximations. Accordingly, we construct a fuzzy weighted neighborhood dependency function and define an optimal reduct as the basis for designing an attribute reduction method based on a hybrid filter-wrapper approach. The experimental evaluation on benchmark datasets indicates that our algorithm achieves smaller reduct sizes and higher classification performance than existing NRS-based methods.

However, the simultaneous integration of the weighting assignment mechanisms in the FWNRS model increases the computational complexity, thereby creating difficulties for the proposed algorithm when handling high-dimensional datasets. In addition, due to the diversity of data characteristics, the model needs to apply the grid search technique to determine an appropriate parameter value. This may increase the execution time of the algorithm, since the grid search process requires evaluating multiple different parameter values. To address this difficulty, in future studies, we will investigate parallel computing methods that allow the algorithm to process multiple parameter values simultaneously in order to search for an optimal reduct more efficiently. In addition, we also aim to develop a formula capable of automatically determining an appropriate parameter value for each type of data, thereby reducing the dependence on the grid search process and improving the computational efficiency of the model.

Conflicts of Interest

The authors affirm that no conflicts of interest are associated with this work.

Acknowledgments

The authors of this paper would like to express their sincere appreciation to the anonymous reviewers for their insightful comments, which have contributed to the improvement in the quality of this study.

AI Disclosure

The author(s) declare that no generative AI was used in the writing of this article.

References

- Anh, P.V., Thuy, N.N., Giang, N.L., Khanh, P.D., & Thuy, N.T. (2023). A novel incremental attribute reduction algorithm based on intuitionistic fuzzy partition distance. *Computer Systems Science and Engineering*, 47(3), 2791-2988. <https://doi.org/10.32604/csse.2023.042068>
- Anh, P.V., Thuy, N.N., Son, L.H., Cuong, T.H., & Giang, N.L. (2025). Incremental attribute reduction with α, β -level intuitionistic fuzzy sets. *International Journal of Approximate Reasoning*, 176, 109326. <https://doi.org/10.1016/j.ijar.2024.109326>
- Chen, H., Li, M., Fan, X., & Lou, X. (2019). Feature selection for imbalanced data based on neighborhood rough sets. *Information Sciences*, 483, 1-20. <https://doi.org/10.1016/j.ins.2019.01.041>
- Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7(1), 1-30.
- Ejegwa, P.A., Anum, M.T., & Isife, K.I. (2024). A new method of distance measure between intuitionistic fuzzy sets and its application in admission procedure. *Journal of Uncertain Systems*, 17(2), 2440005. <https://doi.org/10.1142/S1752890924400051s>
- Ejegwa, P.A., Onyeke, I.C., & Adah, V. (2025). Recognition principle for course allocations in higher institutions based on intuitionistic fuzzy correlation coefficient. *Journal of Fuzzy Extension and Applications*, 6(1), 94-108. <https://doi.org/10.22105/jfea.2024.448400.1411>
- Friedman, M. (1940). A comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, 11(1), 86-92. <https://doi.org/10.1214/aoms/1177731944>
- Hu, M., Tsang, E.C.C., Gou, Y., Chen, D., & Xu, W. (2022). Attribute reduction based on overlap degree and k -nearest-neighbor rough sets in decision information systems. *Information Sciences*, 584, 301-324. <https://doi.org/10.1016/j.ins.2021.10.063>
- Hu, M., Tsang, E.C.C., Guo, Y., Chen, D., & Xu, W. (2021). A novel approach to attribute reduction based on weighted neighborhood rough sets. *Knowledge-Based Systems*, 220, 106908. <https://doi.org/10.1016/j.knosys.2021.106908>
- Hu, Q., Liu, J., & Yu, D. (2008a). Mixed feature selection based on granulation and approximation. *Knowledge-Based Systems*, 21(4), 294-304. <https://doi.org/10.1016/j.knosys.2007.07.001>
- Hu, Q., Pedrycz, W., Yu, D., & Liang J. (2010). Selecting discrete and continuous features based on neighborhood decision error minimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40(1), 137-150. <https://doi.org/10.1109/TSMCB.2009.2024166>
- Hu, Q., Yu, D., Liu, J., & Wu, C. (2008b). Neighborhood rough set based heterogeneous feature subset selection. *Information Sciences*, 178(18), 3577-3594. <https://doi.org/10.1016/j.ins.2008.05.024>
- Jain, P., Tiwari, A.K., & Som, T. (2020). A fitting model based intuitionistic fuzzy rough feature selection. *Engineering Applications of Artificial Intelligence*, 89, 103421. <https://doi.org/10.1016/j.engappai.2019.103421>

- Jiang, Z., Liu, K., Yang, X., Yu, H., Fujita, H., & Qian, Y. (2020). Accelerator for supervised neighborhood based attribute reduction. *International Journal of Approximate Reasoning*, 119, 122-150. <https://doi.org/10.1016/j.ijar.2019.12.013>
- Ju, H., Ding, W., Yang, X., & Gu, P. (2023). Bi-directional adaptive neighborhood rough sets based attribute subset selection. *International Journal of Approximate Reasoning*, 160, 108966. <https://doi.org/10.1016/j.ijar.2023.108966>
- Pawlak, Z. (1982). Rough sets. *International Journal of Computer & Information Sciences*, 11(5), 341-356. <https://doi.org/10.1007/BF01001956>
- Pham, V.A., Nguyen, L.G., Ngoc, N.T., Van, D.L., & Quan, P.H. (2025). Hybrid filter-wrapper attribute reduction method with the uncertainty classification degree. In: Thi, H.A.L., Le, H.M., Nguyen, Q.T. (eds) *Advances in Data Science and Optimization of Complex Systems*. Springer Nature Switzerland, Cham, pp. 298-310. https://doi.org/10.1007/978-3-032-00267-9_27
- Sun, H., Chen, X., Sun, B., Zhao, X., Chu, X., & Cai, J. (2026). Community partitioning and node belonging prediction method based on neighborhood attribute reduction and fuzzy rough set with application in clinical decision making. *Applied Intelligence*, 54, 246. <https://doi.org/10.1007/s10489-026-07258-2>
- Sun, L., Wang, L., Qian, Y., Xu, J., & Zhang, S. (2019). Feature selection using Lebesgue and entropy measures for incomplete neighborhood decision systems. *Knowledge-Based Systems*, 186, 104942. <https://doi.org/10.1016/j.knosys.2019.104942>
- Tan, A., Shi, S., Wu, W.-Z., Li, J., & Pedrycz, W. (2022). Granularity and entropy of intuitionistic fuzzy information and their applications. *IEEE Transactions on Cybernetics*, 52(1), 192-204. <https://doi.org/10.1109/TCYB.2020.2973379>
- Thuy, N.N., & Wongthanavas, S. (2024). Attribute reduction with fuzzy divergence-based weighted neighborhood rough sets. *International Journal of Approximate Reasoning*, 173, 109256. <https://doi.org/10.1016/j.ijar.2024.109256>
- Thuy, N.N., Anh, T.D., & Thanh, L.M. (2025). Generalized weighted neighborhood rough sets. *Information Sciences*, 707, 122020. <https://doi.org/10.1016/j.ins.2025.122020>
- Tsang, E.C.C., Hu, Q., & Chen, D. (2016). Feature and instance reduction for PNN classifiers based on fuzzy rough sets. *International Journal of Machine Learning and Cybernetics*, 7(1), 1-11. <https://doi.org/10.1007/s13042-014-0232-6>
- Wang, C., Hu, Q., Wang, X., Chen, D., Qian, Y., & Dong, Z. (2018). Feature selection based on neighborhood discrimination index. *IEEE Transactions on Neural Networks and Learning Systems*, 29(7), 2986-2999. <https://doi.org/10.1109/TNNLS.2017.2710422>
- Wang, C., Qi, Y., Shao, M., Hu, Q., Chen, D., Qian, Y., & Lin, Y. (2017). A fitting model for feature selection with fuzzy rough sets. *IEEE Transactions on Fuzzy Systems*, 25(4), 741-753. <https://doi.org/10.1109/TFUZZ.2016.2574918>
- Wang, C., Shi, Y., Fan, X., & Shao, M. (2019). Attribute reduction based on k -nearest neighborhood rough sets. *International Journal of Approximate Reasoning*, 106, 18-31. <https://doi.org/10.1016/j.ijar.2018.12.013>
- Wang, N., & Zhao, E. (2024). A new method for feature selection based on weighted k -nearest neighborhood rough set. *Expert Systems with Applications*, 238(Part F), 122324. <https://doi.org/10.1016/j.eswa.2023.122324>
- Xie, J., Hu, B.Q., & Jiang, H. (2022). A novel method to attribute reduction based on weighted neighborhood probabilistic rough sets. *International Journal of Approximate Reasoning*, 144, 1-17. <https://doi.org/10.1016/j.ijar.2022.01.010>
- Yang, X., Chen, H., Li, T., Wan, J., & Sang, B. (2021). Neighborhood rough sets with distance metric learning for feature selection. *Knowledge-Based Systems*, 224, 107076. <https://doi.org/10.1016/j.knosys.2021.107076>

- Yang, X., Liang, S., Yu, H., Gao, S., & Qian Y. (2019). Pseudo-label neighborhood rough set: measures and attribute reductions. *International Journal of Approximate Reasoning*, 105, 112-129. <https://doi.org/10.1016/j.ijar.2018.11.010>
- Yang, Y., Chen, D., Wang, H., Tsang, E.C.C., & Zhang, D. (2017). Fuzzy rough set based incremental attribute reduction from dynamic data with sample arriving. *Fuzzy Sets and Systems*, 312, 66-86. <https://doi.org/10.1016/j.fss.2016.08.001>
- Zhang, P., Li, T., Yuan, Z., Luo, C., Liu, K., & Yang, X. (2024). Heterogeneous feature selection based on neighborhood combination entropy. *IEEE Transactions on Neural Networks and Learning Systems*, 35(3), 3514-3527. <https://doi.org/10.1109/TNNLS.2022.3193929>

Original content of this work is copyright © Ram Arti Publishers. Uses under the Creative Commons Attribution 4.0 International (CC BY 4.0) license at <https://creativecommons.org/licenses/by/4.0/>

Publisher's Note- Ram Arti Publishers remains neutral regarding jurisdictional claims in published maps and institutional affiliations.