Electronic Research Archive

https://www.aimspress.com/journal/era

ERA, 34(10): 7392–7419.
DOI: 10.3934/era.2026319
Received: 17 June 2026
Revised: 04 August 2026
Accepted: 12 August 2026
Published: 27 August 2026

Research article

Attribute reduction based on optimized granular ball neighborhood rough sets

Yufei Wang* **and Nan Zhang**

School of Computer and Control Engineering, Yantai University, Yantai 264005, China

* **Correspondence:** Email: wangyufei_0218@163.com.

Abstract: Neighborhood rough sets (NRS) are commonly applied to attribute reduction because they can process numerical data directly without discretization. Granular ball neighborhood rough sets extend NRS by adaptively determining neighborhood radii according to the data distribution. However, existing granular ball generation and reduction algorithms can be further enhanced in terms of granule representation, radius computation, and attribute evaluation. To address these issues, an optimized granular ball neighborhood rough set model (OGNRS) and a corresponding attribute reduction algorithm are proposed. In this model, a scoring function integrating density and centroid-based distance is employed to select centers, and a median absolute deviation-based strategy is introduced to determine the radii of granular balls. In addition, the minimal-redundancy-maximal-relevance criterion is incorporated into the reduction process to provide search guidance and reduce redundancy among selected attributes. Experiments on 14 UCI datasets demonstrate that OGNRS achieves competitive overall performance in terms of classification accuracy, reduct size, and computational efficiency.

Keywords: rough set; granular computing; granular ball; attribute reduction

1. Introduction

Data with high dimensionality often includes numerous redundant or irrelevant attributes that can affect data processing and analysis. Such redundancy not only increases computational cost but also degrades the generalization performance of models. As a fundamental task in rough set theory [1], attribute reduction is designed to select a minimum set of attributes that retains the classification capability of the complete dataset [2, 3]. To obtain more effective and efficient reducts, numerous rough set-based attribute reduction algorithms have been developed over the past decades [4, 5]. Wu et al. [6] introduced a neighborhood equivalence relation, addressing the classification uncertainty caused by overlapping neighborhood granules. Based on this relation, they further designed a corresponding reduction algorithm. Dai et al. [7] proposed the concepts of the simplified decision

table and neighbor inconsistent pairs to accelerate the attribute selection process. However, their effectiveness is closely tied to the data types supported, as classical rough set models rely on equivalence classes defined on discrete data [8–10].

Because continuous numerical data are prevalent in real-world applications, the neighborhood rough set (NRS) model introduced by Hu et al. [11] has attracted considerable research interest. In NRS, neighborhood relation replaces the equivalence class as the fundamental granules for approximation computation. Compared with data discretization methods [12, 13], NRS processes continuous data without information loss. Building on this foundation, NRS has been continuously extended and applied in recent years, among which attribute reduction has attracted particular attention. Hu et al. [14] developed a fast search forward attribute reduction algorithm (FARNeMF), which employs attribute significance measured by dependency degree. Furthermore, NRS has been extended to handle more complex data. Xu et al. [15] proposed an algorithm employing mutual information and conditional mutual information to evaluate attribute significance for hybrid data with imbalanced distributions. Dai et al. [16] developed an algorithm that uses complementary entropy to construct an uncertainty measure for evaluating attribute significance in heterogeneous data.

The neighborhood radius in NRS is a critical parameter that significantly affects reduction results. To alleviate the dependence on a manually specified radius, Liu et al. [17] introduced the sparrow search (SSA) to optimize adaptive neighborhood radii and proposed the SSA-based adaptive neighborhood rough set (SSAANRS) for attribute reduction. Granular ball computing (GBC) [18] offers another data-driven solution by adaptively determining local neighborhood scales from the data distribution rather than setting a unified radius manually. GBC represents data using granular balls, each defined by a center and a radius adapted to the local data distribution. This structure naturally provides data-dependent neighborhood scales through the generated granular balls. Motivated by this, Xia et al. [19] incorporated GBC into the NRS framework and constructed the granular ball neighborhood rough set (GBNRS). In GBNRS, neighborhood radii are adaptively determined. Building on this framework, Zeng and Liu [20] introduced information entropy-based attribute weighting and developed the attribute reduction algorithm (IEWGBNRS). Xia et al. [21] developed the granular ball rough set (GBRS), capable of handling continuous data processing while simultaneously supporting equivalence class-based knowledge representation. Zhang et al. [22] further incorporated attribute weights into granular-ball construction and dependency evaluation and developed the WGBRS attribute reduction algorithm.

The effectiveness of GBC fundamentally depends on the quality of the generated granular balls. Jia et al. [23] developed a framework built upon the principle of justifiable granularity, which maximizes the quality while ensuring alignment with the data distribution. Zhang et al. [24] proposed a separability degree index to measure the quality of granular balls by evaluating the separability between different decision classes, which serves as the basis for guiding the generation process. In addition, researchers have focused on the design of generation strategies to further enhance their quality. Pan et al. [25] developed a framework that incorporates adaptive granularity tuning, where weight factors are used to compute the center and radius. Xia et al. [26] proposed an adaptive method that eliminates the need for manual parameter setting while accounting for overlap reduction among granular balls. Xie et al. [27] proposed GBG++, which employs an attention mechanism to improve generation stability and integrates an outlier detection strategy to identify local outliers.

Despite the progress made by existing methods, two limitations remain. First, existing center

selection strategies directly use the centroid, which lacks interpretability, and existing radius computation strategies are sensitive to extreme distances. Second, granular-ball-based attribute reduction algorithms rely on the positive region as heuristic information, overlooking redundancy among selected attributes [28]. In response to these challenges, the primary contributions of this paper can be outlined as follows:

1) By selecting from the k objects nearest to the centroid the object with the highest density, an optimized center selection strategy is proposed to improve granular ball interpretability.
2) Based on median absolute deviation (MAD), a radius computation method is introduced to ensure sufficient coverage while alleviating overlap between heterogeneous granular balls.
3) An optimized granular ball neighborhood rough set model (OGNRS) is constructed, in which the minimal-redundancy-maximal-relevance (mRMR) criterion is incorporated as heuristic information to guide attribute reduction toward more discriminative and less redundant reducts.

The rest of this paper is structured as follows. The related work on GBNRS is discussed in Section 2. The proposed OGNRS model is presented in Section 3, followed by the corresponding attribute reduction algorithm in Section 4. Section 5 reports the experimental analysis, and the conclusions are drawn in Section 6.

2. Related work

This section reviews the relevant knowledge. The basic concepts of GBC are first introduced in Section 2.1, and the formal definitions of GBNRS that integrate GBC into the neighborhood rough set framework are then presented in Section 2.2.

2.1. Granular ball computing

GBC is a multigranularity representation method that covers the object space using granular balls. To understand its mechanism, we first introduce the basic granular ball structure.

Definition 2.1. *(See [19]) Let O be a finite set of objects, and let $O_i = \{o_1, o_2, \ldots, o_n\} \subseteq O$ be a nonempty subset containing n objects. A granular ball $\mathcal{GB}_i$ is generated on O_i with center ctr_i and radius rad_i as follows:*

$$ctr_i = \frac{1}{n}\sum_{j=1}^{n} o_j, \quad rad_i = \frac{1}{n}\sum_{j=1}^{n} \Delta(o_j, ctr_i), \tag{2.1}$$

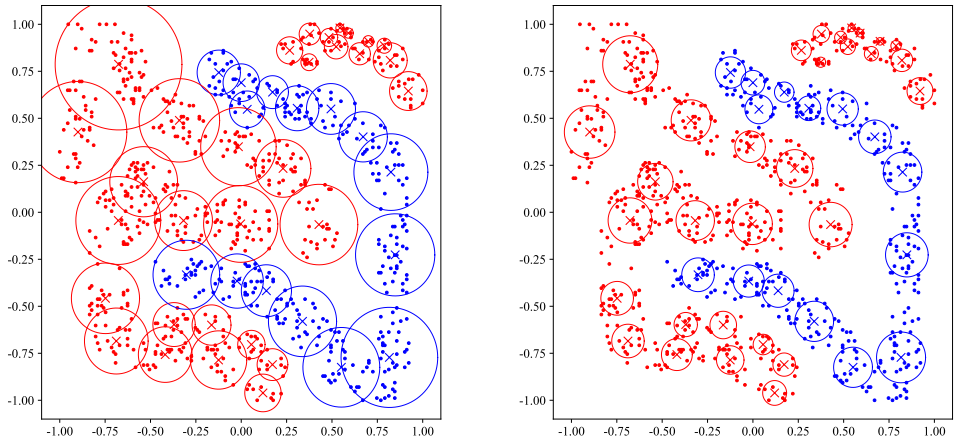
where ctr_i and rad_i represent the centroid of O_i and the average distance from objects in O_i to ctr_i, and $\Delta(o_j, ctr_i)$ measures the distance between point o_j and ctr_i.

Aside from the average-distance radius defined above, the radius in certain granular ball models is defined as the maximum distance from the center to all objects within the ball, ensuring complete object coverage:

$$rad_{max} = \max_{o_j \in O_i} \{\Delta(o_j, ctr_i)\}. \tag{2.2}$$

The average-distance radius rad_{avg} and the maximum-distance radius rad_{max} are two common radius strategies in granular ball generation. The corresponding generation results are illustrated in Figure 1,

using the fourclass dataset as an example. This is a typical binary classification dataset, where objects belonging to the positive and negative classes are represented by blue and red points.



(a) $ctr + rad_{\max}$ **(b)** $ctr + rad_{\text{avg}}$

Figure 1. Granular balls generated with existing radius strategies: (a) maximum-distance radius $rad_{\max}$ with centroid ctr; (b) average-distance radius rad_{avg} with centroid ctr.

A granular ball is structurally determined once its center and radius are defined. To further characterize a granular ball, it is assigned the label of its majority class. The purity of the granular ball is then defined as the proportion of majority-class objects within it, serving as a measure of its quality.

Definition 2.2. *(See [19]) Let $\mathcal{GB}_i$ be a granular ball generated from O_i, and let $\mathcal{M}_i \subseteq O_i$ denote the set of objects in $\mathcal{GB}_i$ belonging to the majority class. The purity of $\mathcal{GB}_i$ is defined as*

$$\mathcal{P}(\mathcal{GB}_i) = \frac{|\mathcal{M}_i|}{|\mathcal{GB}_i|}, \tag{2.3}$$

where $|\mathcal{GB}_i|$ denotes the total number of objects in $\mathcal{GB}_i$, and $|\mathcal{M}_i|$ denotes the number of majority-class objects. Therefore, $0 < \mathcal{P}(\mathcal{GB}_i) \leq 1$.

2.2. Granular ball neighborhood rough sets

In GBNRS, the neighborhood structure is formalized through a granular ball-based indiscernibility relation, which is introduced as follows.

Definition 2.3. *(See [21]) Let $\mathcal{DS} = (O, \mathcal{A}, \mathcal{D})$ be a decision system, where O denotes a finite set of objects, and $\mathcal{A}$ and $\mathcal{D}$ denote the conditional attribute set and the decision attribute, respectively. For any $o_x, o_y \in O$ and $\mathcal{B} \subseteq \mathcal{A}$, the granular ball-based indiscernible relation $\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}$ of the conditional attribute subset $\mathcal{B}$ is defined as*

$$\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}} = \{(o_x, o_y) \mid \exists \mathcal{GB}_j(ctr_j, rad_j), \Delta_{\mathcal{B}}(o_x, ctr_j) \leq rad_j,\ \Delta_{\mathcal{B}}(o_y, ctr_j) \leq rad_j,\ o_x, o_y \in O\}, \tag{2.4}$$

where ctr_j and rad_j denote the center and radius of $\mathcal{GB}_j$, and $\Delta_{\mathcal{B}}(o_x, o_y)$ measures the distance between o_x and o_y under conditional attribute subset $\mathcal{B}$.

For any $\mathcal{B} \subseteq \mathcal{A}$, $\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}$ is an equivalence relation on $\mathcal{O}$, and $[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}} = \{o_y \mid (o_x, o_y) \in \mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}},\ o_x, o_y \in \mathcal{O}\}$ in $\mathcal{O}/\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}$ is an equivalence class generated by GBC. Based on the partition $\mathcal{O}/\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}$ consisting of all such equivalence classes, the approximations are defined below.

Definition 2.4. *(See [21]) Let $\mathcal{DS} = (\mathcal{O}, \mathcal{A}, \mathcal{D})$. For any $\mathcal{B} \subseteq \mathcal{A}$ and $\mathcal{T} \subseteq \mathcal{O}$, the lower and upper approximations of $\mathcal{T}$ induced by $\mathcal{B}$ are defined as*

$$\underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}}\mathcal{T} = \bigcup\{[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}} \mid [o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}} \in \mathcal{O}/\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}},\ [o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}} \subseteq \mathcal{T}\}, \tag{2.5}$$

$$\overline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}}\mathcal{T} = \bigcup\{[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}} \mid [o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}} \in \mathcal{O}/\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}},\ [o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}} \cap \mathcal{T} \neq \emptyset\}. \tag{2.6}$$

When the target set $\mathcal{T}$ is taken as a decision class $\mathcal{D}_i$, the approximations of the decision attribute $\mathcal{D}$ are obtained by combining all decision classes.

Definition 2.5. *(See [21]) Let $\mathcal{DS} = (\mathcal{O}, \mathcal{A}, \mathcal{D})$. For any $\mathcal{B} \subseteq \mathcal{A}$, let $\mathcal{O}/\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_c\}$ denote the partition of $\mathcal{O}$ into c decision classes. The lower and upper approximations of $\mathcal{D}$ induced by $\mathcal{B}$ are defined as*

$$\underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}}\mathcal{D} = \bigcup_{i=1}^{c} \underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}}\mathcal{D}_i, \tag{2.7}$$

$$\overline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}}\mathcal{D} = \bigcup_{i=1}^{c} \overline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}}\mathcal{D}_i. \tag{2.8}$$

The positive and boundary regions can be derived once the approximations of $\mathcal{D}$ are established. Specifically, the positive region consists of objects correctly classified under $\mathcal{B}$, and the boundary region contains those whose classification remains uncertain.

Definition 2.6. *(See [21]) Let $\mathcal{DS} = (\mathcal{O}, \mathcal{A}, \mathcal{D})$. For any $\mathcal{B} \subseteq \mathcal{A}$, the positive and boundary regions of $\mathcal{D}$ are defined as*

$$\mathrm{POS}_{\mathcal{B}}(\mathcal{D}) = \underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}}\mathcal{D}, \tag{2.9}$$

$$\mathrm{BN}_{\mathcal{B}}(\mathcal{D}) = \overline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}}\mathcal{D} - \underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}}\mathcal{D}. \tag{2.10}$$

3. Optimized granular ball neighborhood rough sets

Building on the concepts of GBC and GBNRS introduced above, this section presents an optimized extension. Because the performance of the GBNRS model on attribute reduction is determined by the quality of the granular ball generation, a novel granular ball generation algorithm with optimized center selection and radius computation strategies is first proposed in Section 3.1. The generated granular balls then serve as the foundation for the OGNRS constructed in Section 3.2. Specifically, the optimized center and MAD-based radius determine the construction of optimized granular balls, which are subsequently used to establish the corresponding granular ball-based relation and rough approximations.

3.1. Optimized granular ball generation

The center and radius are two key parameters in granular-ball construction. The center determines the location of a granular ball, whereas the radius defines its spatial extent. Accordingly, the center selection strategy is introduced first, followed by the radius computation method. To improve center representativeness and interpretability, a density-distance scoring function is introduced to select the ball center from the k nearest neighbors of the centroid. These neighbors are first collected as the candidate center set, which is formally defined as follows.

Definition 3.1. *Let $O_i = \{o_1, o_2, \ldots, o_n\} \subseteq O$ be a nonempty subset containing n objects. Given a parameter k ($1 \le k \le n$), the k objects in O_i nearest to the centroid are selected as the candidate center set, denoted by $S_i = \{o_1^s, o_2^s, \ldots, o_k^s\}$.*

Based on the set S_i, the density of each candidate center is computed to measure how densely it is surrounded by the objects within the granular ball. The Gaussian kernel density estimation [29, 30] is adopted with an adaptive bandwidth σ_i^s, and the density is defined as follows.

Definition 3.2. *For any nonempty subset $O_i = \{o_1, o_2, \ldots, o_n\} \subseteq O$ with candidate center set $S_i = \{o_1^s, o_2^s, \ldots, o_k^s\}$, the density of any candidate center $o_i^s \in S_i$ is defined as*

$$\rho(o_i^s) = \sum_{j=1}^{n} \exp\left(-\frac{\Delta(o_i^s, o_j)^2}{2(\sigma_i^s)^2}\right), \tag{3.1}$$

where σ_i^s is the average distance from o_i^s to its k nearest neighbors inside the granular ball.

Although the centroid geometrically represents the objects, it is a calculated value. Therefore, it is susceptible to outliers and may deviate from the high-density region of the data distribution. This limitation motivates the introduction of a density-distance scoring function to select a more representative center from the candidate set S_i.

Definition 3.3. *For any nonempty subset $O_i = \{o_1, o_2, \ldots, o_n\} \subseteq O$ with candidate center set $S_i = \{o_1^s, o_2^s, \ldots, o_k^s\}$, the score of any candidate center $o_i^s \in S_i$ is defined as*

$$\text{score}(o_i^s) = \hat{\rho}(o_i^s) + \hat{d}(o_i^s), \tag{3.2}$$

where $\hat{\rho}(o_i^s)$ and $\hat{d}(o_i^s)$ denote the density and distance scores, both of which are normalized to the same range to reduce the effect of different scales. Specifically, $\hat{\rho}(o_i^s)$ is obtained by min-max normalization of the density as follows:

$$\hat{\rho}(o_i^s) = \frac{\rho(o_i^s) - \min_{o_q^s \in S_i} \rho(o_q^s)}{\max_{o_p^s \in S_i} \rho(o_p^s) - \min_{o_q^s \in S_i} \rho(o_q^s)}. \tag{3.3}$$

Because the candidate set S_i is already restricted to the k objects nearest to the centroid, a smaller distance to the centroid further ensures that the selected center remains close to the geometric center of the granular ball. Therefore, $\hat{d}(o_i^s)$ is defined as the inverted max-normalized distance:

$$\hat{d}(o_i^s) = 1 - \frac{\Delta(o_i^s, ctr_i)}{\max_{o_j^s \in S_i} \Delta(o_j^s, ctr_i)}. \tag{3.4}$$

Among all candidate centers, the one with the highest score is selected as the optimized center of the granular ball.

Definition 3.4. *For any nonempty subset $O_i = \{o_1, o_2, \ldots, o_n\} \subseteq O$ with candidate center set $\mathcal{S}_i = \{o_1^s, o_2^s, \ldots, o_k^s\}$, the optimized center ctr_i' is defined as*

$$ctr_i' = \arg\max_{o_i^s \in \mathcal{S}_i} \text{score}(o_i^s). \tag{3.5}$$

To intuitively illustrate the advantage of the optimized center over the centroid, a schematic diagram is provided in Figure 2. When outliers are present within a granular ball, the centroid tends to deviate from the high-density region. In contrast, the optimized center selected by the proposed scoring function remains within it. This demonstrates its higher representativeness.

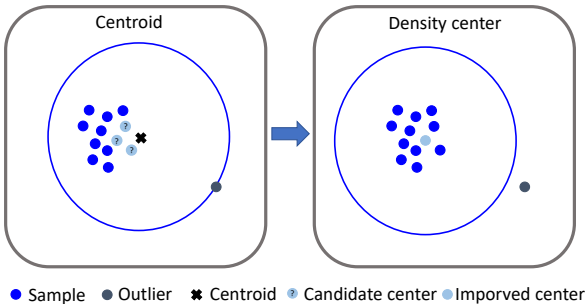


Figure 2. Schematic diagram of granular ball center selection.

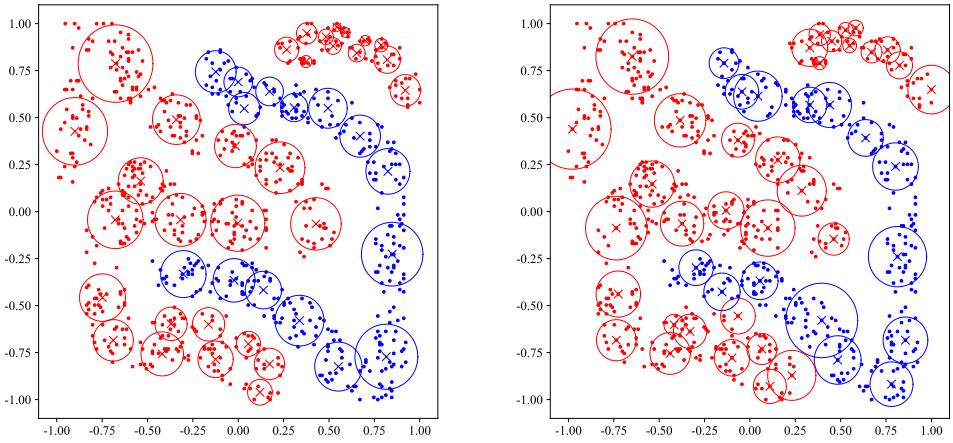
Apart from center selection, the radius computation also plays a critical role in granular ball generation. The maximum-distance radius ensures complete object coverage but often produces excessively large balls, leading to overlap between heterogeneous granular balls. Conversely, the average-distance radius reduces overlap but may leave numerous objects outside the ball boundaries, resulting in incomplete coverage. Therefore, we develop a radius computation strategy based on the MAD. The median of distances ensures robust basic coverage by excluding the influence of extreme values, and the MAD adaptively adjusts the radius according to the distance distribution.

Definition 3.5. *For any nonempty subset $O_i = \{o_1, o_2, \ldots, o_n\} \subseteq O$, an optimized granular ball OGB_i is generated on O_i with optimized center ctr_i'. Its MAD-based radius rad_i' is defined as*

$$rad_i' = \text{med}(\boldsymbol{d}_i) + \text{MAD}(\boldsymbol{d}_i), \tag{3.6}$$

where $\boldsymbol{d}_i = \left(\Delta(o_1, ctr_i'), \Delta(o_2, ctr_i'), \ldots, \Delta(o_n, ctr_i')\right)$ denotes the vector of distances from each object in O_i to ctr_i', $\text{med}(\boldsymbol{d}_i)$ is the median of the distances, and $\text{MAD}(\boldsymbol{d}_i) = \text{med}\left(|\boldsymbol{d}_i - \text{med}(\boldsymbol{d}_i)|\right)$ is the MAD.

The granular balls generated by combining the MAD-based radius with the original centroid and the optimized center are shown in Figure 3. Compared with the average-distance and maximum-distance radii, the MAD-based radius ensures sufficient object coverage while reducing overlap between heterogeneous balls. Consequently, it achieves a balance between object coverage and boundary control.



(a) $ctr + rad'$ **(b)** $ctr' + rad'$

Figure 3. Granular balls generated with the proposed strategies: (a) MAD-based radius rad' with original centroid ctr; (b) MAD-based radius rad' with optimized center ctr'.

To further accelerate granular ball splitting, an attention mechanism [27] is incorporated. Specifically, it first identifies the majority class among the unassigned objects, allowing sub-balls to be generated preferentially in the corresponding regions. This strategy facilitates efficient local assignment of the data. The complete optimized granular ball generation algorithm (OGBG) is detailed in Algorithm 1.

Algorithm 1 OGBG

Input: A decision system $\mathcal{DS} = (\mathcal{O}, \mathcal{A}, \mathcal{D})$, purity threshold τ, and nearest-neighbor parameter k.
Output: A granular ball set $\mathcal{OGBSET}^{\tau}$.

Treat the universe $\mathcal{O}$ as the initial granular ball $\mathcal{OGB}_j^i$, where i represents the number of iterations initialized to 1, and j represents ith GB generated in each iteration initialized to 1.
for each $\mathcal{OGB}_j^i \in \mathcal{OGBSET}^{\tau}$ **do**
  if $\mathcal{P}(\mathcal{OGB}_j^i) < \tau$ **then**
    $\mathcal{O}_{\text{rem}} \leftarrow \mathcal{O}_j^i, \quad \mathcal{OGB}_j^{i+1} \leftarrow \emptyset$
    while $\mathcal{O}_{\text{rem}} \neq \emptyset$ **do**
      Identify majority-class subset $\mathcal{O}_{jm}^i \subseteq \mathcal{O}_{\text{rem}}$
      Compute the centroid of $\mathcal{O}_{jm}^i$ and form $\mathcal{S}_{jm}^i$ from its k nearest objects
      $ctr'^i_{jm} \leftarrow \arg\max_{o \in \mathcal{S}_{jm}^i} \text{score}(o)$ via Definition 3.4
      Compute rad'^i_{jm} from ctr'^i_{jm} via Definition 3.5
      $gb_{jm}^{i+1} \leftarrow \{o \in \mathcal{O}_{jm}^i \mid \Delta(o, ctr'^i_{jm}) \leq rad'^i_{jm}\}$
      $\mathcal{O}_{\text{rem}} \leftarrow \mathcal{O}_{\text{rem}} - gb_{jm}^{i+1}, \quad \mathcal{OGB}_j^{i+1} \leftarrow \mathcal{OGB}_j^{i+1} \cup \{gb_{jm}^{i+1}\}$
    end while
    $\mathcal{OGBSET}^{\tau} \leftarrow (\mathcal{OGBSET}^{\tau} - \{\mathcal{OGB}_j^i\}) \cup \mathcal{OGB}_j^{i+1}$
  end if
end for
return $\mathcal{OGBSET}^{\tau}$

To provide a more intuitive understanding of the generation process, the overall procedure of Algorithm 1 is illustrated in Figure 4. For each granular ball requiring further splitting, the majority-class objects are processed first, and the optimized center together with the MAD-based radius is then used to generate a new sub-ball. The procedure is repeated until no granular ball requires further splitting.

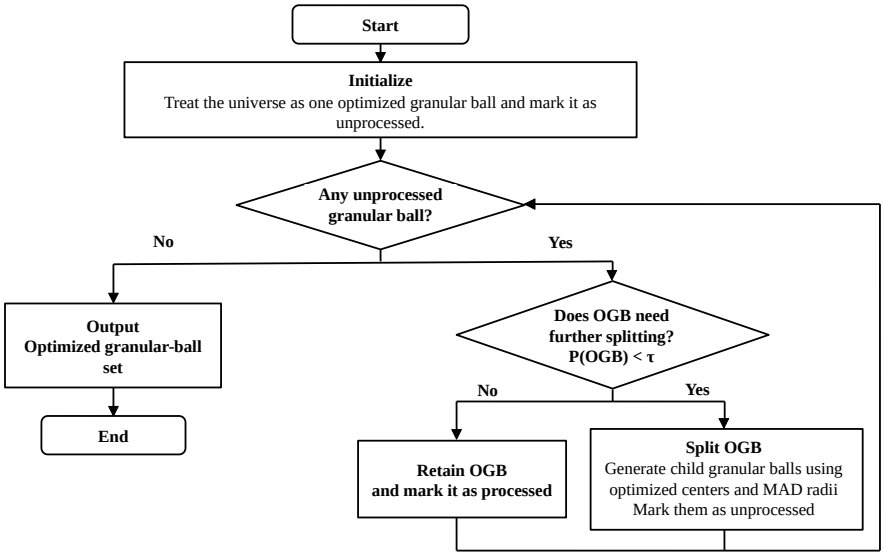


Figure 4. Flowchart of the OGBG algorithm.

Based on the above generation procedure, the computational complexity of OGBG is analyzed as follows. The total number of objects and iterations in this procedure are denoted by n and t. At each iteration, the main operations include majority-class identification, candidate-center scoring, MAD-based radius computation, and object assignment. Given that k is a small constant, the computational cost of center selection remains linear. The complexity of these operations is $O(n+r_i)$, where r_i denotes the number of unassigned objects involved in repeated distance computations. Therefore, the overall time complexity is $O\left(t\sum_{i=1}^{t}(n+r_i)\right)$. Because t and r_i are much smaller than n at each iteration, the time complexity can be approximated as $O(n)$.

3.2. Optimized granular ball neighborhood rough set

Based on the optimized granular balls generated above, the corresponding neighborhood rough set model is constructed by following the same logical sequence as GBNRS. Specifically, the granular ball-based relation is defined first, followed by the lower and upper approximations, and finally the positive and boundary regions used for attribute reduction.

Definition 3.6. *Let $\mathcal{DS} = \langle \mathcal{O}, \mathcal{A}, \mathcal{D} \rangle$. For any $o_x, o_y \in \mathcal{O}$ and $\mathcal{B} \subseteq \mathcal{A}$, the optimized granular ball-based indiscernible relation $\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}$ of the conditional attribute subset $\mathcal{B}$ is defined as*

$$\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}} = \{(o_x, o_y) \mid \exists \mathcal{OGB}_j(ctr'_j, rad'_j), \Delta_{\mathcal{B}}(o_x, ctr'_j) \leq rad'_j,\ \Delta_{\mathcal{B}}(o_y, ctr'_j) \leq rad'_j,\ o_x, o_y \in \mathcal{O}\}, \quad (3.7)$$

where ctr'_j and rad'_j are the center and radius of $\mathcal{OGB}_j$, and $\Delta_{\mathcal{B}}(o_x, o_y)$ denotes the distance between o_x and o_y with respect to conditional attribute subset $\mathcal{B}$.

For any $\mathcal{B} \subseteq \mathcal{A}$, $\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}$ induces a partition $\mathcal{O}/\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}$ on $\mathcal{O}$, where $[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}} = \{o_y \in \mathcal{O} \mid (o_x, o_y) \in \mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}\}$ is the equivalence class of o_x. Based on this partition, the approximations are defined below.

Definition 3.7. *Let $\mathcal{DS} = \langle \mathcal{O}, \mathcal{A}, \mathcal{D} \rangle$. For any $\mathcal{B} \subseteq \mathcal{A}$ and $\mathcal{T} \subseteq \mathcal{O}$, the lower and upper approximations of $\mathcal{T}$ induced by $\mathcal{B}$ are defined as:*

$$\underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}}\mathcal{T} = \bigcup \{[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}} \mid [o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}} \in \mathcal{O}/\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}, [o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}} \subseteq \mathcal{T}\}, \tag{3.8}$$

$$\overline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}}\mathcal{T} = \bigcup \{[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}} \mid [o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}} \in \mathcal{O}/\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}, [o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}} \cap \mathcal{T} \neq \emptyset\}. \tag{3.9}$$

When the target set $\mathcal{T}$ is taken as a decision class $\mathcal{D}_i$, the approximations extend to the decision attribute $\mathcal{D}$ as follows.

Definition 3.8. *Let $\mathcal{DS} = \langle \mathcal{O}, \mathcal{A}, \mathcal{D} \rangle$. For any $\mathcal{B} \subseteq \mathcal{A}$, let $\mathcal{O}/\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_c\}$ denote the partition of $\mathcal{O}$ into c decision classes. The lower and upper approximations of $\mathcal{D}$ induced by $\mathcal{B}$ are defined as:*

$$\underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}}\mathcal{D} = \bigcup_{i=1}^{c} \underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}}\mathcal{D}_i, \tag{3.10}$$

$$\overline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}}\mathcal{D} = \bigcup_{i=1}^{c} \overline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}}\mathcal{D}_i. \tag{3.11}$$

Based on the above approximations, the corresponding regions used for knowledge characterization can be further defined as follows.

Definition 3.9. *Let $\mathcal{DS} = \langle \mathcal{O}, \mathcal{A}, \mathcal{D} \rangle$. For any $\mathcal{B} \subseteq \mathcal{A}$, the positive and boundary regions of $\mathcal{D}$ induced by $\mathcal{B}$ are defined as:*

$$\mathrm{POS}_{\mathcal{B}}(\mathcal{D}) = \underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}}\mathcal{D}, \tag{3.12}$$

$$\mathrm{BN}_{\mathcal{B}}(\mathcal{D}) = \overline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}}\mathcal{D} - \underline{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}}\mathcal{D}. \tag{3.13}$$

Although NRS, GBNRS, and OGNRS all construct rough approximations based on neighborhood granules, they differ in the form of these granules and in how their neighborhood scales are determined. In NRS, the neighborhood of each object is determined by a fixed radius. GBNRS extends NRS by introducing data-dependent granular ball neighborhoods. Each granular ball is characterized by a center and a radius, which together determine the corresponding local neighborhood. This replaces the globally specified neighborhood radius with adaptive local neighborhood scales determined through granular ball construction. Building on this adaptive neighborhood mechanism, OGNRS further refines the local granulation by introducing optimized center selection and MAD-based radius computation. The relationships among NRS, GBNRS, and OGNRS in terms of neighborhood construction are summarized in Table 1.

Table 1. Relationships among NRS, GBNRS, and OGNRS.

Comparison item	Model	Description
Neighborhood granule	NRS	$\delta_{\mathcal{B}}(o_x)$, centered at o_x with a globally specified radius δ
	GBNRS	$\mathcal{GB}_i(ctr_i, rad_i)$, characterized by a data-dependent center and radius
	OGNRS	$\mathcal{OGB}_i(ctr'_i, rad'_i)$, characterized by an optimized center and MAD-based radius
Neighborhood relation	NRS	$o_y \in \delta_{\mathcal{B}}(o_x) \iff \Delta_{\mathcal{B}}(o_x, o_y) \le \delta$
	GBNRS	$o_x \, \mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}} \, o_y$ if o_x and o_y belong to the same $\mathcal{GB}_i(ctr_i, rad_i)$
	OGNRS	$o_x \, \mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}} \, o_y$ if o_x and o_y belong to the same $\mathcal{OGB}_i(ctr'_i, rad'_i)$
Rough approximation	NRS	Lower: $\delta_{\mathcal{B}}(o_x) \subseteq \mathcal{T}$; Upper: $\delta_{\mathcal{B}}(o_x) \cap \mathcal{T} \neq \emptyset$
	GBNRS	Lower: $[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}} \subseteq \mathcal{T}$; Upper: $[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{GB}}} \cap \mathcal{T} \neq \emptyset$
	OGNRS	Lower: $[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}} \subseteq \mathcal{T}$; Upper: $[o_x]_{\mathcal{IR}_{\mathcal{B}}^{\mathcal{OGB}}} \cap \mathcal{T} \neq \emptyset$

4. Attribute reduction based on optimized granular ball neighborhood rough sets

Based on the optimized granular ball neighborhood rough set model established above, this section further develops an attribute reduction algorithm for selecting informative attributes. The optimized granular balls provide a more reliable local granulation for constructing neighborhood relations and computing the positive region. In particular, the optimized center better reflects the local data distribution, and the MAD-based radius reduces sensitivity to extreme distances and alleviates overlap between heterogeneous balls, thereby providing a more reliable basis for subsequent attribute evaluation. In granular ball-based reduction algorithms, the positive region is commonly adopted as heuristic information. However, the individually best attributes do not necessarily form the optimal combination because highly relevant attributes may contain redundant information with respect to those already selected. To address this limitation, the mRMR criterion is introduced to guide the heuristic search by jointly considering the relevance of candidate attributes to the decision attribute and their redundancy with the selected attributes.

Given two random variables u and v with probability density functions $f(u)$, $f(v)$ and joint density $f(u, v)$, their mutual information [28] is

$$MI(u, v) = \iint f(u, v) \log \frac{f(u, v)}{f(u)f(v)} \, du \, dv. \tag{4.1}$$

Suppose $\mathcal{A}' = \{a_1, a_2, \ldots, a_r\}$ is the selected conditional attribute subset, and $\mathcal{D}$ is the decision attribute. The max-relevance criterion [28] seeks to maximize the mutual information between individual attributes and $\mathcal{D}$:

$$\max Rel(\mathcal{A}', \mathcal{D}), \qquad Rel(\mathcal{A}', \mathcal{D}) = \frac{1}{|\mathcal{A}'|} \sum_{a_i \in \mathcal{A}'} MI(a_i, \mathcal{D}). \tag{4.2}$$

However, attributes selected under the max-relevance criterion are typically highly redundant, and removing any of them does not diminish discriminative power. The min-redundancy criterion [28] is therefore added to select mutually exclusive attributes:

$$\min Red(\mathcal{A}'), \qquad Red(\mathcal{A}') = \frac{1}{|\mathcal{A}'|^2} \sum_{a_i, a_j \in \mathcal{A}'} MI(a_i, a_j). \tag{4.3}$$

By incorporating the above two constraints, the mRMR criterion [28] is established to simultaneously optimize relevance and redundancy:

$$\max \phi(Rel, Red), \qquad \phi = Rel - Red. \tag{4.4}$$

In light of the above criterion, an attribute reduction algorithm based on optimized granular ball neighborhood rough sets (OGNRS) is developed, as detailed in Algorithm 2.

Algorithm 2 Attribute reduction based on OGNRS

Input: A decision system $\mathcal{DS} = (\mathcal{O}, \mathcal{A}, \mathcal{D})$, purity threshold τ, and parameter k.
Output: A reduct $\mathcal{A}'$.

$\mathcal{A}' \leftarrow \{\arg\max_{a_i \in \mathcal{A}} Rel(a_i, \mathcal{D})\}$
Generate $\mathcal{OGBSET}^{\tau}(\mathcal{A}')$ via Algorithm 1 and set $\mathrm{POS}_{\mathrm{num}} \leftarrow \mathrm{POS}_{\mathrm{num}}(\mathcal{A}')$
while $|\mathcal{A}'| < |\mathcal{A}|$ **do**
    $c_M \leftarrow \arg\max_{a_i \in \mathcal{A} - \mathcal{A}'} \left(Rel(a_i, \mathcal{D}) - \frac{1}{|\mathcal{A}'|} \sum_{a_j \in \mathcal{A}'} MI(a_i, a_j) \right)$
    $\mathcal{B} \leftarrow \mathcal{A}' \cup \{c_M\}$
    Generate $\mathcal{OGBSET}^{\tau}(\mathcal{B})$ via Algorithm 1 and compute $\mathrm{POS}_{\mathrm{num}}(\mathcal{B})$
    if $\mathrm{POS}_{\mathrm{num}}(\mathcal{B}) \geq \mathrm{POS}_{\mathrm{num}}$ **then**
        $\mathcal{A}' \leftarrow \mathcal{B}, \quad \mathrm{POS}_{\mathrm{num}} \leftarrow \mathrm{POS}_{\mathrm{num}}(\mathcal{B})$
    else
        break
    end if
end while
return $\mathcal{A}'$

To further clarify the reduction mechanism of OGNRS, the overall procedure is illustrated in Figure 5. This reduction process mainly consists of two stages: mRMR-based candidate attribute selection and positive region-based attribute acceptance. In the first stage, candidate attributes are evaluated by considering both their relevance to the decision attribute and their redundancy with the selected attributes. In the second stage, the selected candidate is accepted only when the corresponding positive region size does not decrease. These two stages are performed iteratively until the reduction procedure terminates.

Based on the above reduction procedure, the computational complexity of OGNRS is analyzed as follows. Let a dataset consist of n objects and m attributes, and assume that the final reduct contains r attributes. The time complexity of OGNRS is primarily determined by three stages. First, computing the mutual information between each candidate attribute and the decision attribute requires a complexity of $O(mn)$. Second, the attribute selection proceeds for r iterations. In the ith iteration, calculating the redundancy between the $m - i$ remaining candidates and the $i - 1$ already selected attributes requires a complexity of $O((m - i)(i - 1))$. The sum of these computational costs over r iterations results in a complexity of $O(mr^2)$. Third, granular ball generation and positive region computation over all r iterations require a time complexity of $O(rn)$. Consequently, the overall time complexity of OGNRS is $O(mn + mr^2 + rn)$. Because r is much smaller than m, the time complexity simplifies to $O(mn)$.

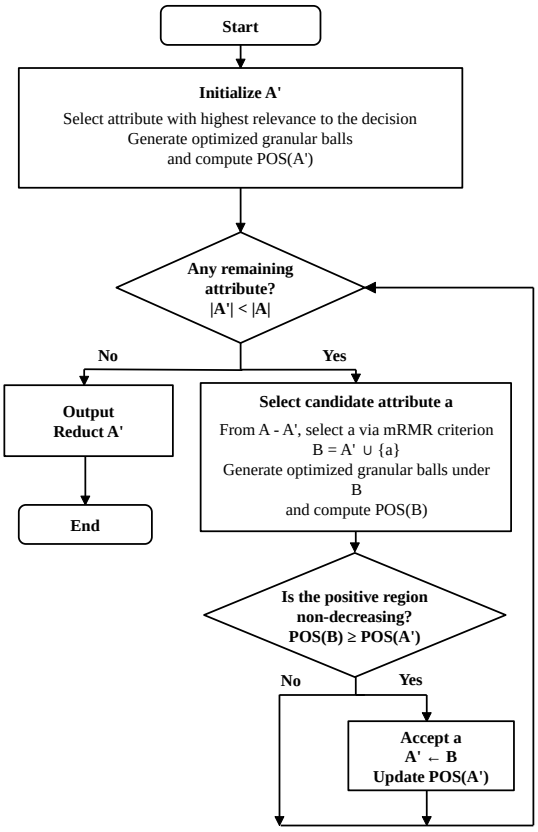


Figure 5. Flowchart of the OGNRS attribute reduction algorithm.

The time complexities of OGNRS and several related attribute reduction algorithms are summarized in Table 2, where q denotes the number of SSA iterations, and $c = |\mathcal{O}/\mathcal{D}|$ denotes the number of decision classes. According to the complexity analysis, the time complexity of OGNRS is $O(mn + mr^2 + rn)$ and can be approximated as $O(mn)$ when the number of selected attributes is much smaller than the total number of attributes. Based on the complexity expressions reported for the compared algorithms, OGNRS has a lower dependence on the numbers of attributes and objects than NRS, SSAANRS, and GBRS. Although a lower complexity of $O(n)$ is reported for GBNRS, OGNRS maintains an approximately linear complexity with respect to both the number of objects and attributes.

Table 2. Time complexities of the compared algorithms.

Algorithm	Time complexity
NRS [11]	$O(m^2 n \log n)$
SSAANRS [17]	$O(qmn^2)$
GBNRS [19]	$O(n)$
IEWGBNRS [20]	$O(mc) + O(\max(m, n))$
GBRS [21]	$O(mn^2)$
WGBRS [22]	$O(m^2)$
OGNRS	$O(mn)$

5. Experimental analysis

In this section, a series of comparative experiments and ablation studies are performed to validate the performance of OGNRS. Specifically, the experiments include: 1) parameter sensitivity analysis; 2) classification accuracy comparison; 3) reduct size comparison; 4) statistical significance analysis; 5) ablation study; 6) computational efficiency analysis; 7) robustness analysis. The experimental platform is a 64-bit Windows 10 system with an Intel Core i5-8265U CPU (1.60/1.80 GHz) and 8 GB RAM. All algorithms are coded in Python 3.8 using PyCharm 2020.1. Fourteen UCI datasets (https://archive.ics.uci.edu) are used, and their basic information is presented in Table 3. All datasets are normalized before experiments to ensure comparability across different attribute scales.

Table 3. Experimental datasets.

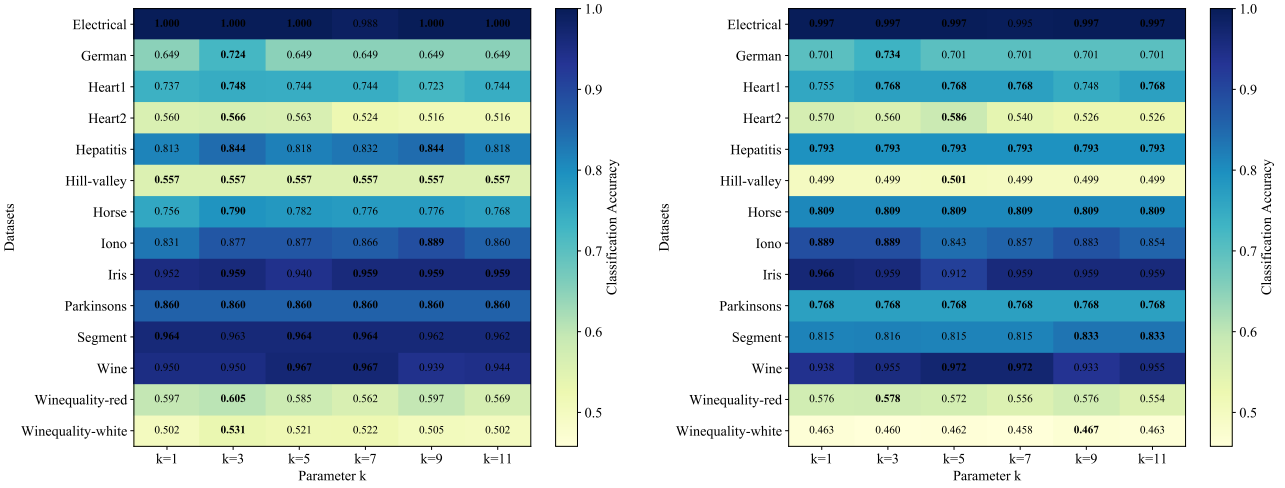
No.	Datasets	Attributes	Size	Classes
1	Electrical	12	10,000	2
2	German	19	1000	2
3	Heart1	13	270	2
4	Heart2	13	303	5
5	Hepatitis	19	155	2
6	Hill-valley	101	606	2
7	Horse	23	368	2
8	Iono	34	351	2
9	Iris	4	151	3
10	Parkinsons	24	195	2
11	Segment	20	2310	7
12	Wine	13	178	3
13	Winequality-red	12	1599	5
14	Winequality-white	11	4898	7

In the experiments, we compare the OGNRS with six algorithms: NRS [11], SSAANRS [17], GBNRS [19], IEWGBNRS [20], GBRS [21], and WGBRS [22]. Parameters of the compared algorithms are set as suggested in the original studies. For granular ball-based algorithms, including GBNRS, IEWGBNRS, GBRS, WGBRS, and OGNRS, the purity threshold is set to 1. This ensures that each granular ball contains an identical label, which is consistent with the concept of positive regions in NRS.

5.1. Parameter sensitivity analysis

To evaluate the sensitivity of OGNRS to the parameter k and determine an appropriate setting for the subsequent experiments, k is varied over $\{1, 3, 5, 7, 9, 11\}$. The parameter k determines the number of objects nearest to the centroid that are considered as candidate centers during optimized granular ball generation. A smaller k restricts center selection to a relatively local region around the centroid, thereby reducing the candidate search space, but it may also limit the possibility of identifying a more representative high-density object. In contrast, a larger k provides a broader candidate set for center

selection, but it may include objects farther from the centroid and increase the computational cost of candidate evaluation. Therefore, the influence of k is evaluated in terms of both classification accuracy and reduct size on the 14 benchmark datasets. As shown in Figure 6, different values of k lead to different classification accuracies under the K-nearest neighbors (KNN), support vector machines (SVM), and classification and regression tree (CART) classifiers. The horizontal axis represents the candidate values of k, and the vertical axis represents the 14 benchmark datasets. The color of each cell in the heatmap indicates the corresponding classification accuracy, with darker shades denoting higher accuracy.



(a) KNN **(b)** SVM

(c) CART

Figure 6. Classification accuracy with different values of k on three classifiers.

The aggregated results in Table 4 further illustrate the influence of k on OGNRS. When $k = 1$, OGNRS achieves an overall average classification accuracy of 76.33% and the smallest average reduct size of 3.14 attributes. Increasing k to 3 improves the overall average classification accuracy to 77.41%, which is the highest among the tested settings, and the average reduct size increases

moderately to 3.50 attributes. When k is further increased to 5, 7, 9, and 11, the overall average classification accuracies decrease to 76.36%, 76.08%, 75.85%, and 75.76%, respectively. The corresponding average reduct sizes are 3.36, 3.50, 3.79, and 3.36 attributes. These results show that enlarging the candidate-center set does not continuously improve the performance of OGNRS. Although a larger k provides more candidate objects for optimized center selection, an excessively broad candidate range may include objects that are less representative of the local data distribution, thereby affecting granular ball construction and the subsequent reduction process.

Based on the above results, $k = 3$ is adopted as a fixed parameter in all subsequent experiments. This setting achieves the highest overall average classification accuracy while maintaining a relatively small reduct size. Although $k = 1$ yields the smallest average reduct size, its overall average classification accuracy is 1.08 percentage points lower than that obtained with $k = 3$. Therefore, the choice of $k = 3$ reflects a trade-off between classification performance and reduct size rather than pursuing the minimum reduct size alone.

Table 4. Classification accuracy and reduct size of OGNRS under different values of k.

k	KNN accuracy (%)	SVM accuracy (%)	CART accuracy (%)	Average accuracy (%)	Average reduct size
1	76.63±4.27	75.28±3.34	77.07±4.63	76.33	**3.14**
3	**78.39±4.36**	**75.61±3.46**	**78.24±4.45**	**77.41**	3.50
5	77.34±3.98	74.99±3.24	76.75±4.65	76.36	3.36
7	76.93±4.07	74.93±3.21	76.37±4.66	76.08	3.50
9	76.97±4.56	74.94±3.41	75.64±4.61	75.85	3.79
11	76.49±4.79	74.85±3.46	75.94±4.36	75.76	3.36

5.2. Classification accuracy comparison

The classification performance of the attribute subsets selected by different reduction algorithms is evaluated using KNN, SVM, and CART. For KNN, the number of nearest neighbors is set to five. For SVM, the radial basis function kernel is adopted and γ is set to `auto`. CART uses the default parameter settings. The corresponding average accuracy and standard deviation obtained from ten-fold cross-validation are reported in Tables 5–7, with the best results highlighted in bold. The Original column reports the classification performance obtained using all attributes without reduction.

As shown in Tables 5–7, OGNRS achieves the highest average classification accuracy under KNN and CART, reaching 78.54% and 78.15%, respectively. These results demonstrate that OGNRS performs competitively compared with other attribute reduction algorithms. Under SVM, OGNRS obtains an average accuracy of 75.75%. Although this value is slightly lower than the best average result of 76.15%, OGNRS achieves the highest classification accuracy on several datasets, such as German, Heart1, and Hepatitis. For the Electrical dataset, OGNRS obtains the same classification accuracy as NRS and GBRS because they select the same attribute subset. The competitive classification performance of OGNRS can be attributed to the combination of mRMR-guided attribute selection and the OGNRS. Specifically, the mRMR-guided selection strategy identifies a discriminative attribute subset by reducing redundant attributes, whereas the OGNRS provides a more appropriate granular representation for positive region computation. The integration of these two components enables OGNRS to maintain competitive classification performance across

different classifiers.

Table 5. Classification accuracies of all compared algorithms with the KNN classifier (%).

Datasets	Original	NRS	SSAANRS	GBNRS	IEWGBNRS	GBRS	WGBRS	OGNRS
Electrical	91.43±0.56	**99.97±0.06**	62.21±0.94	91.43±0.56	91.70±0.71	**99.97±0.06**	99.65±0.23	**99.97±0.06**
German	70.47±2.83	70.66±4.38	65.87±2.80	72.07±2.97	**72.77±3.43**	65.86±2.97	67.27±4.31	72.37±3.20
Heart1	75.82±5.99	75.82±5.99	60.09±8.53	**78.20±6.17**	75.79±3.62	71.75±7.23	71.02±5.13	74.76±3.00
Heart2	54.30±4.14	55.28±5.72	50.63±8.01	53.28±4.33	54.24±6.74	51.30±5.93	50.32±4.20	**56.62±7.19**
Hepatitis	79.88±8.07	81.75±7.20	79.83±5.53	**85.08±7.09**	83.75±9.55	85.00±7.94	78.50±9.87	84.38±6.68
Hill-valley	53.23±5.41	52.06±6.03	54.73±3.26	54.04±6.85	54.71±5.48	53.23±5.41	53.23±5.41	**55.70±7.12**
Horse	79.57±6.55	76.53±7.84	56.17±7.70	80.08±4.99	80.36±5.55	59.35±7.99	52.53±10.27	**81.19±4.84**
Iono	85.43±3.71	86.57±3.84	87.71±4.44	88.00±4.92	88.57±4.43	88.57±5.42	**89.43±6.52**	87.71±6.00
Iris	**95.90±4.63**	95.24±6.14	93.29±5.96	95.24±6.14	**95.90±5.50**	**95.90±4.63**	**95.90±4.63**	**95.90±5.50**
Parkinsons	93.76±5.12	**94.87±2.24**	79.47±7.01	90.26±7.42	91.74±4.76	89.16±7.81	91.79±6.21	86.03±5.88
Segment	95.71±1.53	95.67±1.42	85.06±1.92	95.54±1.47	96.28±1.21	95.76±0.72	95.45±1.32	**96.32±0.99**
Wine	95.52±3.35	94.38±4.97	93.86±5.24	95.46±3.37	**97.22±3.73**	94.93±4.70	93.17±3.50	94.97±4.62
Winequality-red	59.01±4.62	59.01±4.62	52.26±3.50	59.01±4.62	58.01±3.27	57.38±3.20	57.07±3.65	**60.52±4.54**
Winequality-white	**57.08±1.88**	**57.08±1.88**	51.93±1.80	**57.08±1.88**	**57.08±1.88**	44.31±3.18	44.31±3.18	53.05±1.39
Average	77.65±4.17	78.21±4.45	69.51±4.76	78.20±4.48	78.44±4.28	75.18±4.80	74.26±4.89	**78.54±4.36**

Table 6. Classification accuracies of all compared algorithms with the SVM classifier (%).

Datasets	Original	NRS	SSAANRS	GBNRS	IEWGBNRS	GBRS	WGBRS	OGNRS
Electrical	98.23±0.36	**99.69±0.14**	68.55±1.00	98.23±0.36	98.51±0.38	**99.69±0.14**	99.66±0.12	**99.69±0.14**
German	71.27±0.57	71.37±1.63	70.07±0.21	70.07±5.33	70.77±1.30	70.07±5.33	70.07±0.21	**73.37±2.23**
Heart1	73.79±6.13	74.14±5.94	61.78±0.97	73.79±6.13	73.76±5.72	74.08±4.50	74.08±4.50	**76.76±5.40**
Heart2	58.27±2.35	**58.60±4.05**	53.97±1.08	57.96±2.35	57.95±5.71	53.97±5.50	53.97±1.08	57.96±2.35
Hepatitis	**79.25±2.18**	**79.25±2.18**	**79.25±2.18**	**79.25±2.18**	**79.25±2.18**	**79.25±2.18**	**79.25±2.18**	**79.25±2.18**
Hill-valley	50.08±1.76	50.08±1.76	49.91±2.25	50.08±1.76	**51.24±3.68**	50.08±1.76	50.08±1.76	50.08±1.76
Horse	**80.92±5.37**	**80.92±5.37**	58.04±1.14	**80.92±5.37**	**80.92±5.37**	58.04±7.70	58.04±1.14	**80.92±5.37**
Iono	87.43±3.71	83.71±4.96	83.14±5.34	86.57±3.84	81.71±3.88	86.86±4.81	87.43±5.74	**88.86±5.64**
Iris	95.90±5.50	95.90±5.50	87.14±7.94	**96.57±5.56**	95.90±5.50	95.90±5.50	95.90±5.50	95.90±5.50
Parkinsons	75.26±3.82	**83.47±3.97**	83.03±5.11	**83.47±3.97**	81.45±5.76	79.87±6.89	75.26±1.93	76.79±3.54
Segment	88.13±2.14	88.52±1.43	70.98±1.66	88.39±4.88	**89.35±1.91**	82.42±6.25	84.02±1.92	81.55±1.42
Wine	**98.89±2.69**	92.19±7.93	93.27±4.83	92.16±7.10	96.63±4.46	93.27±5.09	95.49±4.18	95.49±4.86
Winequality-red	56.63±2.84	56.63±2.38	53.69±3.74	56.63±2.84	56.51±2.68	55.25±4.41	56.70±3.04	**57.82±2.93**
Winequality-white	**51.56±1.51**	**51.56±1.51**	46.25±0.60	**51.56±1.51**	**51.56±1.51**	44.86±0.05	44.86±0.05	46.03±0.83
Average	76.12±2.92	**76.15±3.48**	68.51±2.72	76.12±3.80	76.11±3.57	73.12±4.29	73.20±2.38	75.75±3.15

A comparison of the reduct sizes obtained by different algorithms is presented in Table 8. OGNRS selects 3.50 attributes on average, demonstrating effective attribute reduction while maintaining competitive classification performance. The relatively small reduct size of OGNRS may benefit from the mRMR-guided attribute selection strategy, which helps limit the selection of redundant attributes. Although SSAANRS obtains the smallest average reduct size of 2.50 attributes, its overall classification performance is lower than that of OGNRS, as shown in Tables 5–7. Therefore, OGNRS achieves a reasonable balance between reduct size and classification performance. In comparison, NRS and GBNRS produce relatively larger reducts on average, whereas GBRS and WGBRS retain

more attributes overall. This may be partly related to its forward greedy selection strategy, which retains an attribute when the positive region size does not decrease and may consequently introduce redundant attributes. This phenomenon is particularly evident on the high-dimensional datasets, with Hill-valley as a typical example.

Table 7. Classification accuracies of all compared algorithms with the CART classifier (%).

Datasets	Original	NRS	SSAANRS	GBNRS	IEWGBNRS	GBRS	WGBRS	OGNRS
Electrical	**99.97±0.06**	**99.97±0.06**	57.37±1.15	**99.97±0.06**	**99.97±0.06**	**99.97±0.06**	**99.97±0.06**	**99.97±0.06**
German	68.57±2.77	67.77±4.51	59.96±5.27	65.26±4.21	69.27±3.76	68.97±2.77	69.27±2.97	**70.87±2.77**
Heart1	66.90±5.29	65.53±7.56	60.07±9.88	71.26±0.80	68.24±4.96	71.34±5.29	68.63±2.98	**72.01±6.17**
Heart2	46.69±5.32	46.01±6.02	33.77±8.26	50.96±3.10	44.70±7.41	46.02±8.45	37.41±7.40	**56.61±7.19**
Hepatitis	70.17±0.57	82.42±6.07	72.71±7.56	81.08±8.78	74.79±10.91	85.00±7.95	77.33±12.64	**86.29±5.54**
Hill-valley	57.33±6.52	52.05±3.43	52.56±7.77	55.20±6.05	**57.85±4.74**	54.53±5.94	55.70±6.11	49.61±7.56
Horse	76.81±5.69	76.01±4.85	62.45±6.09	73.86±5.21	80.09±6.65	57.47±9.12	53.36±5.88	**80.92±5.37**
Iono	**90.00±4.02**	87.14±4.47	84.29±4.82	87.71±4.25	84.00±6.54	88.00±5.08	89.71±5.88	87.71±4.25
Iris	93.90±4.82	93.24±5.29	91.90±7.84	93.90±5.50	**94.62±4.02**	94.57±3.18	93.90±5.67	93.90±5.67
Parkinsons	84.61±4.09	85.45±7.82	75.84±9.63	85.05±4.41	87.66±7.73	**88.71±8.94**	88.05±5.39	79.84±4.73
Segment	96.11±1.27	96.75±1.33	87.87±1.48	**97.06±1.34**	96.88±0.98	96.80±1.16	95.67±1.49	96.19±0.66
Wine	91.01±4.41	90.42±3.52	88.17±5.82	88.10±5.52	91.60±5.09	92.12±3.63	92.71±4.32	**93.30±5.09**
Winequality-red	64.08±4.70	63.27±2.89	58.82±2.55	64.52±5.42	62.33±2.97	61.83±3.86	61.39±1.88	**65.46±5.42**
Winequality-white	62.69±1.52	62.75±2.03	59.77±1.99	**63.16±1.20**	63.12±1.38	46.64±2.90	46.64±2.90	61.42±2.21
Average	76.35±3.65	76.34±4.28	67.54±5.72	76.94±3.99	76.79±4.80	75.14±4.88	73.55±4.68	**78.15±4.48**

Table 8. Sizes of the attribute subsets selected by the compared algorithms.

Dataset	NRS	SSAANRS	GBNRS	IEWGBNRS	GBRS	WGBRS	OGNRS
Electrical	1	1	13	9	1	2	1
German	12	2	13	11	5	3	9
Heart1	12	3	7	6	4	6	7
Heart2	12	3	9	9	3	4	2
Hepatitis	7	3	6	7	4	6	3
Hill-valley	9	2	7	8	100	100	1
Horse	9	2	10	8	3	1	3
Iono	8	3	9	7	5	4	2
Iris	3	3	3	2	4	4	3
Parkinsons	5	1	6	6	5	5	1
Segment	12	3	16	8	8	6	4
Wine	6	2	4	6	4	6	3
Winequality-red	11	3	11	9	5	7	5
Winequality-white	11	4	11	11	1	1	5
Average	8.43	2.50	8.93	7.64	10.86	11.07	3.50

5.3. Statistical significance analysis

To provide a more detailed view of the relative performance of the compared algorithms, Tables 9–11 report the datasetwise rankings of the seven attribute reduction algorithms under KNN, SVM, and CART, respectively. The rankings are determined according to the average classification

accuracy on each dataset, with a lower rank indicating better performance. When two or more algorithms obtain the same average accuracy, their tied positions are replaced by the corresponding average rank.

Table 9. Classification accuracy ranking of all compared algorithms on KNN.

Dataset	NRS	SSAANRS	GBNRS	IEWGBNRS	GBRS	WGBRS	OGNRS
Electrical	2	7	6	5	2	4	2
German	4	6	3	1	7	5	2
Heart1	2	7	1	3	5	6	4
Heart2	2	6	4	3	5	7	1
Hepatitis	5	6	1	4	2	7	3
Hill-valley	7	2	4	3	5.5	5.5	1
Horse	4	6	3	2	5	7	1
Iono	7	5.5	4	2.5	2.5	1	5.5
Iris	5.5	7	5.5	2.5	2.5	2.5	2.5
Parkinsons	1	7	4	3	5	2	6
Segment	4	7	5	2	3	6	1
Wine	5	6	2	1	4	7	3
Winequality-red	2.5	7	2.5	4	5	6	1
Winequality-white	2	5	2	2	6.5	6.5	4
Average rank	3.79	6.04	3.36	2.71	4.29	5.18	2.64

Table 10. Classification accuracy ranking of all compared algorithms on SVM.

Dataset	NRS	SSAANRS	GBNRS	IEWGBNRS	GBRS	WGBRS	OGNRS
Electrical	2	7	6	5	2	4	2
German	2	5.5	5.5	3	5.5	5.5	1
Heart1	2	7	5	6	3.5	3.5	1
Heart2	1	6	2.5	4	6	6	2.5
Hepatitis	4	4	4	4	4	4	4
Hill-valley	4	7	4	1	4	4	4
Horse	2.5	6	2.5	2.5	6	6	2.5
Iono	5	6	4	7	3	2	1
Iris	4	7	1	4	4	4	4
Parkinsons	1.5	3	1.5	4	5	7	6
Segment	2	7	3	1	5	4	6
Wine	6	4.5	7	1	4.5	2.5	2.5
Winequality-red	3.5	7	3.5	5	6	2	1
Winequality-white	2	4	2	2	6.5	6.5	5
Average rank	2.96	5.79	3.68	3.54	4.64	4.36	3.04

The detailed ranking results show that OGNRS obtains the best average rank under KNN at 2.64 and under CART at 2.86, but it ranks second under SVM at 3.04. To further determine whether significant performance differences exist among the compared attribute reduction algorithms, the Friedman test is

conducted based on their rankings over the 14 datasets.

Table 11. Classification accuracy ranking of all compared algorithms on CART.

Dataset	NRS	SSAANRS	GBNRS	IEWGBNRS	GBRS	WGBRS	OGNRS
Electrical	3.5	7	3.5	3.5	3.5	3.5	3.5
German	5	7	6	2.5	4	2.5	1
Heart1	6	7	3	5	2	4	1
Heart2	4	7	2	5	3	6	1
Hepatitis	3	7	4	6	2	5	1
Hill-valley	6	5	3	1	4	2	7
Horse	3	5	4	2	6	7	1
Iono	5	6	3.5	7	2	1	3.5
Iris	6	7	4	1	2	4	4
Parkinsons	4	7	5	3	1	2	6
Segment	4	7	1	2	3	6	5
Wine	5	6	7	4	3	2	1
Winequality-red	3	7	2	4	5	6	1
Winequality-white	3	5	1	2	6.5	6.5	4
Average rank	4.32	6.43	3.50	3.43	3.36	4.11	2.86

Let N denote the number of datasets, M denote the number of compared algorithms, and R_j denote the average rank of the jth algorithm. The Friedman statistic is first calculated based on the average ranks and then transformed into the corresponding F statistic as follows:

$$\chi_F^2 = \frac{12N}{M(M+1)}\left[\sum_{j=1}^{M} R_j^2 - \frac{M(M+1)^2}{4}\right], \tag{5.1}$$

$$F_F = \frac{(N-1)\chi_F^2}{N(M-1)-\chi_F^2}. \tag{5.2}$$

As shown in Table 12, the p-values are calculated based on the corresponding F_F statistics with $M-1$ and $(M-1)(N-1)$ degrees of freedom. In this study, $N = 14$, $M = 7$, and all p-values are below the significance level of 0.05. Therefore, the null hypothesis is rejected under all three classifiers, indicating statistically significant overall differences among the seven attribute reduction algorithms.

Table 12. Statistical significance test results.

Classifier	χ_F^2	F_F	p-value
KNN	29.613	7.078	4.85×10^{-6}
SVM	22.065	4.631	4.49×10^{-4}
CART	26.337	5.938	3.82×10^{-5}

To further examine the pairwise differences among the compared algorithms, the Nemenyi post-hoc test is conducted. The critical difference is calculated as

$$CD = q_\alpha \sqrt{\frac{M(M+1)}{6N}}, \tag{5.3}$$

where q_α is the critical value corresponding to the significance level α. For $M = 7$, $N = 14$, and $\alpha = 0.05$, $q_{0.05} = 2.949$, resulting in $CD = 2.408$. When the absolute difference between the average ranks of two algorithms is not less than the critical difference, their performance difference is considered statistically significant.

According to the Nemenyi test, OGNRS shows statistically significant differences from SSAANRS under KNN, SVM, and CART and from WGBRS under KNN. In contrast, the differences between OGNRS and the remaining baseline algorithms do not exceed the critical difference.

5.4. Ablation study

The effectiveness of each component in the OGNRS is evaluated through two groups of ablation studies on granular ball generation and attribute selection. For the granular ball generation part, we compare four variants: *ctr* + *rad*, *ctr*′ + *rad*, *ctr* + *rad*′, and *ctr*′ + *rad*′. These variants are evaluated under the same positive region-based forward greedy reduction framework without mRMR. Among them, *ctr* + *rad* serves as the baseline employing the centroid and the average distance-based radius. *ctr*′ + *rad* replaces the centroid with the optimized center. *ctr* + *rad*′ replaces the average-distance radius with the MAD-based radius. *ctr*′ + *rad*′ uses both the optimized center and the MAD-based radius.

Table 13. Classification accuracy comparison of different granular ball center strategies (%).

Dataset	**KNN**		**SVM**		**CART**	
	ctr + *rad*	*ctr*′ + *rad*	*ctr* + *rad*	*ctr*′ + *rad*	*ctr* + *rad*	*ctr*′ + *rad*
Electrical	91.43±0.56	**92.34±0.27**	98.23±0.36	**98.28±0.31**	**99.97±0.06**	**99.97±0.06**
German	69.97±2.77	**70.47±2.83**	70.17±0.57	**71.27±0.80**	69.07±4.32	**69.18±3.95**
Heart1	69.68±9.34	**74.08±4.50**	**73.79±6.13**	**73.79±6.13**	60.40±9.89	**64.77±7.65**
Heart2	55.60±4.72	**57.60±2.65**	**57.96±2.35**	57.29±3.96	**49.95±7.04**	47.66±7.56
Hepatitis	74.67±11.14	**82.54±5.57**	**79.25±2.18**	**79.25±2.18**	**75.38±7.23**	69.54±11.03
Hill-valley	**53.56±6.33**	53.40±5.97	**50.08±1.76**	**50.08±1.76**	**58.17±4.52**	57.36±6.33
Horse	77.12±5.97	**79.57±6.55**	**80.92±5.37**	**80.92±5.37**	72.73±4.87	**78.47±7.04**
Iono	**86.29±5.54**	85.14±3.33	86.86±4.81	**87.14±5.15**	88.00±5.08	**91.14±6.45**
Iris	**95.90±5.50**	95.24±6.14	95.90±5.50	**96.57±5.56**	93.24±5.29	**93.90±5.67**
Parkinsons	91.79±5.74	**92.26±5.37**	78.84±3.71	**80.39±3.97**	**87.66±7.95**	87.16±9.71
Segment	95.45±1.34	**96.02±1.47**	87.87±1.51	**88.48±1.67**	96.19±0.66	**96.93±1.22**
Wine	93.89±7.22	**97.19±2.81**	92.16±7.10	**98.30±2.60**	91.60±5.09	**92.12±3.63**
Winequality-red	**58.63±4.46**	58.01±3.27	56.32±2.84	**56.51±2.68**	62.64±3.86	**62.70±4.35**
Winequality-white	56.95±1.98	**57.28±1.81**	51.50±1.38	**51.89±1.52**	**63.45±2.31**	62.26±1.90
Average	76.50±5.19	**77.94±3.75**	75.70±3.26	**76.44±3.12**	76.32±4.87	**76.65±5.47**

The baseline *ctr* + *rad* is compared with the optimized center variant *ctr*′ + *rad*, with the corresponding results reported in Table 13. The average accuracy increases from 76.50% to 77.94% on KNN, from 75.70% to 76.44% on SVM, and from 76.32% to 76.65% on CART. Similar improvements across all three classifiers are observed for the MAD-based radius variant *ctr* + *rad*′ in Table 14. These results indicate that the two components improve the granular ball structure from different aspects, thereby improving classification accuracy. The optimized center enhances the representativeness of the granular ball, whereas the MAD-based radius provides a more robust

computation of ball coverage. When both components are combined as *ctr′ + rad′*, the corresponding results across all three classifiers are reported in Table 15. Specifically, the average accuracy reaches 77.67% on KNN, 76.20% on SVM, and 77.02% on CART. We therefore conclude that the two components improve the granular ball structure from different aspects and that their combination improves the average classification accuracy over the baseline under all three classifiers.

Table 14. Classification accuracy comparison of different granular ball radius strategies (%).

Dataset	KNN		SVM		CART	
	ctr + rad	*ctr + rad′*	*ctr + rad*	*ctr + rad′*	*ctr + rad*	*ctr + rad′*
Electrical	91.43±0.56	**91.82±0.65**	98.23±0.36	**98.30±0.35**	**99.97±0.06**	**99.97±0.06**
German	69.97±2.77	**71.17±3.49**	70.17±0.57	**71.07±1.05**	69.07±4.32	**69.17±5.56**
Heart1	69.68±9.34	**76.83±6.52**	**73.79±6.13**	**73.79±5.93**	60.40±9.89	**69.26±7.27**
Heart2	**55.60±4.72**	53.63±3.97	**57.96±2.35**	57.62±3.10	**49.95±7.04**	48.33±7.32
Hepatitis	74.67±11.14	**77.92±7.37**	**79.25±2.18**	**79.25±2.18**	75.38±7.23	**75.96±9.44**
Hill-valley	**53.56±6.33**	53.22±5.52	**50.08±1.76**	**50.08±1.76**	58.17±4.52	**59.33±5.94**
Horse	77.12±5.97	**80.11±6.56**	**80.92±5.37**	**80.92±5.37**	72.73±4.87	**78.18±5.92**
Iono	**86.29±5.54**	84.57±3.18	86.86±4.81	**88.29±4.32**	88.00±5.08	**89.43±5.12**
Iris	**95.90±5.50**	**95.90±5.50**	**95.90±5.50**	**95.90±5.50**	93.24±5.29	**93.90±4.82**
Parkinsons	91.79±5.74	**92.26±6.66**	78.84±3.71	**79.39±3.94**	**87.66±7.95**	86.61±7.45
Segment	95.45±1.34	**95.67±1.45**	87.87±1.51	**88.48±1.37**	96.19±0.66	**96.49±1.43**
Wine	93.97±7.22	**94.38±4.97**	92.16±7.10	**96.11±4.24**	**91.60±5.09**	88.76±8.92
Winequality-red	58.63±4.46	**59.26±2.95**	56.32±2.84	**57.38±2.42**	62.64±3.86	**64.70±3.35**
Winequality-white	56.95±1.98	**57.08±1.88**	51.50±1.38	**51.56±1.51**	**63.45±2.31**	63.39±1.59
Average	76.50±5.19	**77.42±4.33**	75.70±3.26	**76.30±3.07**	76.32±4.87	**77.39±5.30**

Table 15. Classification accuracy comparison for the joint effect of the center and radius strategies (%).

Dataset	KNN		SVM		CART	
	ctr + rad	*ctr′ + rad′*	*ctr + rad*	*ctr′ + rad′*	*ctr + rad*	*ctr′ + rad′*
Electrical	91.43±0.56	**93.06±0.44**	98.23±0.36	**98.53±0.42**	**99.97±0.06**	**99.97±0.06**
German	69.97±2.77	**70.07±5.33**	70.17±0.57	**71.17±1.14**	**69.07±4.32**	68.48±4.21
Heart1	69.68±9.34	**73.41±5.29**	73.79±6.13	**74.14±5.94**	60.40±9.89	**63.13±8.79**
Heart2	**55.60±4.72**	55.28±4.41	55.96±6.67	**58.59±3.23**	**49.95±7.04**	48.70±9.04
Hepatitis	74.67±11.14	**80.54±8.78**	**79.25±2.18**	**79.25±2.18**	75.38±7.23	**75.96±10.98**
Hill-valley	**53.56±6.33**	53.40±6.01	**50.08±1.76**	**50.08±1.76**	**58.17±4.52**	56.02±5.81
Horse	77.12±5.97	**79.85±5.69**	**80.92±5.37**	**80.92±5.37**	72.73±4.87	**77.90±5.21**
Iono	**86.29±5.54**	84.86±4.44	86.86±4.81	**87.71±4.25**	88.00±5.08	**89.43±4.96**
Iris	**95.90±5.50**	**95.90±5.50**	**95.90±5.50**	**95.90±5.50**	93.24±5.29	**94.62±4.02**
Parkinsons	91.79±5.74	**93.26±4.09**	**78.84±3.71**	76.29±3.20	87.66±7.95	**88.71±8.94**
Segment	95.45±1.34	**95.63±1.45**	87.87±1.51	**88.48±1.47**	96.19±0.66	**96.75±1.06**
Wine	93.97±7.22	**97.19±3.75**	92.16±7.10	**97.19±3.75**	91.60±5.09	**92.16±4.41**
Winequality-red	58.63±4.46	**59.38±3.98**	56.32±2.84	**56.76±2.45**	62.64±3.86	**63.70±3.74**
Winequality-white	**56.95±1.98**	55.48±1.74	51.50±1.38	**51.79±1.15**	**63.45±2.31**	62.69±1.29
Average	76.50±5.19	**77.67±4.35**	75.56±3.56	**76.20±2.99**	76.32±4.87	**77.02±5.18**

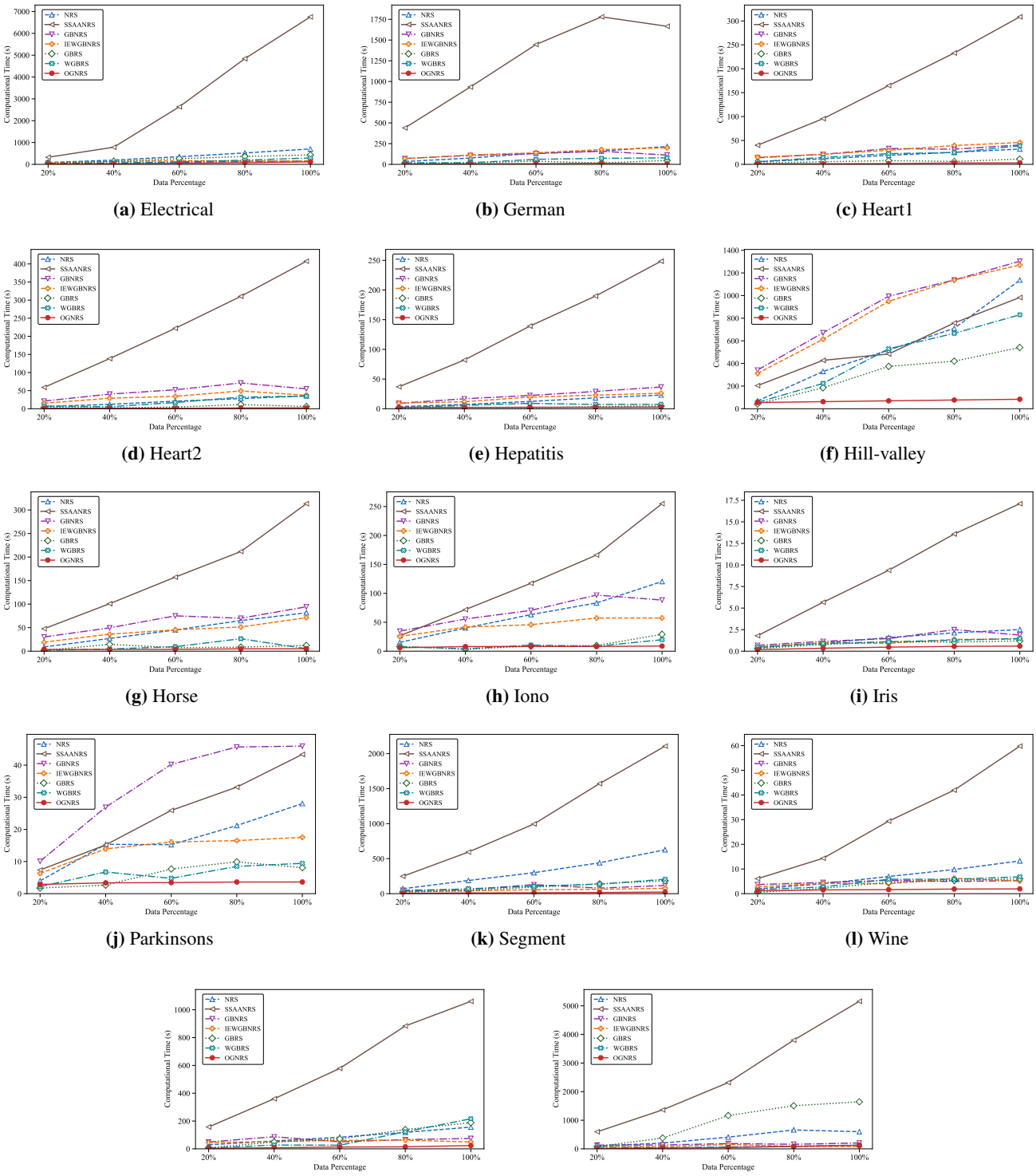
The effect of incorporating the mRMR criterion is evaluated in Table 16. The average accuracy increases from 77.67% to 78.54% on KNN, and from 77.02% to 78.15% on CART. These improvements indicate that jointly considering attribute relevance and redundancy contributes to the selection of a more effective reduct. In contrast, the average accuracy under SVM changes slightly, from 76.20% to 75.75%. This is because SVM determines the decision boundary by maximizing the margin and therefore inherently suppresses redundant attributes. Consequently, the redundancy control provided by mRMR provides only limited improvement on the SVM classifier.

Table 16. Classification accuracy comparison of different attribute selection criteria (%).

Dataset	KNN		SVM		CART	
	Forward greedy	mRMR-guided	Forward greedy	mRMR-guided	Forward greedy	mRMR-guided
Electrical	93.06±0.44	**99.97±0.06**	98.53±0.42	**99.69±0.14**	**99.97±0.06**	**99.97±0.06**
German	70.07±5.33	**72.37±3.20**	71.17±1.14	**73.37±2.23**	68.48±4.21	**70.87±2.77**
Heart1	73.41±5.29	**74.76±3.00**	74.14±5.94	**76.76±5.40**	63.13±8.79	**72.01±6.17**
Heart2	55.28±4.41	**56.62±7.19**	**58.59±3.23**	57.96±2.35	48.70±9.04	**56.61±7.19**
Hepatitis	80.54±8.78	**84.38±6.68**	**79.25±2.18**	**79.25±2.18**	75.96±10.98	**86.29±5.54**
Hill-valley	53.40±6.01	**55.70±7.12**	**50.08±1.76**	**50.08±1.76**	**56.02±5.81**	49.61±7.56
Horse	79.85±5.69	**81.19±4.84**	**80.92±5.37**	**80.92±5.37**	77.90±5.21	**80.92±5.37**
Iono	84.86±4.44	**87.71±6.00**	87.71±4.25	**88.86±5.64**	**89.43±4.96**	87.71±4.25
Iris	**95.90±5.50**	**95.90±5.50**	**95.90±5.50**	**95.90±5.50**	**94.62±4.02**	93.90±5.67
Parkinsons	**93.26±4.09**	86.03±5.88	76.29±3.20	**76.79±3.54**	**88.71±8.94**	79.84±4.73
Segment	95.63±1.45	**96.32±0.99**	**88.48±1.47**	81.55±1.42	**96.75±1.06**	96.19±0.66
Wine	**97.19±3.75**	94.97±4.62	**97.19±3.75**	95.49±4.86	92.16±4.41	**93.30±5.09**
Winequality-red	59.38±3.98	**60.52±4.54**	56.76±2.45	**57.82±2.93**	63.70±3.74	**65.46±5.42**
Winequality-white	**55.48±1.74**	53.05±1.39	**51.79±1.15**	46.03±0.83	**62.69±1.29**	61.42±2.21
Average	77.67±4.35	**78.54±4.36**	**76.20±2.99**	75.75±3.15	77.02±5.18	**78.15±4.48**

5.5. Computational efficiency analysis

The computational efficiency of OGNRS is evaluated by comparing the reduction times of NRS, SSAANRS, GBNRS, IEWGBNRS, GBRS, and WGBRS. These comparisons are shown in Figure 7, which are conducted under different data percentages across all datasets. The horizontal axis represents the dataset percentage, with the vertical axis indicating the reduction time. Overall, OGNRS exhibits lower computational cost than the other algorithms on most datasets, demonstrating its efficiency in attribute reduction. This efficiency improvement can be attributed to two factors. First, GBNRS and IEWGBNRS rely on iterative k-means clustering, which requires repeated distance computations over all objects until convergence. In contrast, OGNRS employs an attention-guided strategy that prioritizes majority-class objects during granular ball generation. Distance computations are restricted to the attended objects rather than the entire dataset. As a result, the computational cost is progressively reduced during the iterative construction process. Second, the MAD-based radius computation strategy effectively reduces overlap among granular balls of different classes while maintaining object coverage. As a result, OGNRS avoids the overlap detection and de-overlap operations required by algorithms such as GBRS and WGBRS.



(a) Electrical **(b)** German **(c)** Heart1

(d) Heart2 **(e)** Hepatitis **(f)** Hill-valley

(g) Horse **(h)** Iono **(i)** Iris

(j) Parkinsons **(k)** Segment **(l)** Wine

(m) Winequality-red **(n)** Winequality-white

Figure 7. Running time comparison under different data percentages.

5.6. Robustness analysis

To evaluate robustness to attribute noise, all seven algorithms are tested at noise ratios of 0%, 5%, and 10%. For each noise level, all algorithms are evaluated on the same noisy datasets, and the classifier and cross-validation settings remain unchanged. The classification accuracies obtained with KNN, SVM, and CART are presented in Table 17 as the average accuracy ± standard deviation across the 14 datasets, with the 0% condition serving as the clean-data baseline.

Table 17. Classification accuracy under different attribute noise levels (%).

Classifier	Noise ratio	NRS	SSAANRS	GBNRS	IEWGBNRS	GBRS	WGBRS	OGNRS
KNN	0%	78.21±4.45	69.51±4.76	78.20±4.48	78.44±4.28	75.18±4.80	74.26±4.89	**78.54±4.36**
	5%	77.63±3.53	68.14±3.85	77.01±3.04	77.66±3.08	74.46±3.38	74.97±3.20	**77.67±3.48**
	10%	**78.33±3.36**	68.11±3.97	77.95±2.81	77.73±3.13	73.73±4.23	77.30±3.88	77.88±3.18
SVM	0%	76.15±3.48	68.51±2.72	75.97±3.93	**76.25±3.45**	73.12±4.29	73.20±2.38	75.75±3.15
	5%	**76.41±2.66**	67.89±1.59	75.79±2.54	76.24±2.52	72.40±2.59	73.56±2.25	74.91±2.02
	10%	75.65±2.34	67.16±1.63	**76.33±2.68**	75.78±2.33	73.22±2.57	74.09±2.16	75.23±2.47
CART	0%	76.34±4.28	67.54±5.72	76.94±3.99	76.79±4.80	75.14±4.88	73.55±4.68	**78.15±4.48**
	5%	76.08±3.14	66.77±4.34	75.72±3.74	77.11±3.82	73.36±3.64	75.05±3.22	**77.82±3.35**
	10%	75.47±3.43	66.40±4.69	75.86±3.56	76.39±3.15	73.82±3.86	75.23±3.59	**77.87±3.44**

Because absolute classification accuracy does not fully characterize robustness to attribute noise, the average absolute accuracy change relative to the original data is further calculated over the 14 datasets. Let $Acc_{d,0}$ and $Acc_{d,\eta}$ denote the accuracies on dataset d under 0% and η attribute noise, respectively. A smaller value indicates lower sensitivity to attribute noise and greater robustness. The metric is defined as

$$\mathrm{Avg}\left|\Delta Acc_{\eta}\right| = \frac{1}{14}\sum_{d=1}^{14}\left|Acc_{d,0} - Acc_{d,\eta}\right|. \tag{5.4}$$

The average absolute accuracy change results presented in Table 18 indicate that OGNRS achieves the smallest value in three of the six classifier–noise settings, namely KNN at 10% attribute noise and CART at both 5% and 10% noise levels. Combined with the classification accuracy results in Table 17, OGNRS achieves the highest average accuracy under KNN at 5% noise and under CART at both noise levels. In particular, under CART, OGNRS consistently achieves both the highest average accuracy and the smallest average absolute accuracy change at the two noise levels. These results indicate that OGNRS exhibits competitive robustness to moderate attribute noise, with particularly stable performance under CART, although its robustness varies across classifiers and noise levels.

Table 18. Average absolute accuracy change under attribute noise.

Classifier	Noise ratio	NRS	SSAANRS	GBNRS	IEWGBNRS	GBRS	WGBRS	OGNRS
KNN	5%	**1.01**	1.77	1.96	1.53	3.66	3.98	1.25
	10%	1.35	3.21	1.41	1.64	4.66	3.74	**1.33**
SVM	5%	0.59	1.87	0.99	**0.56**	4.50	3.98	0.85
	10%	**0.67**	2.66	0.84	1.09	3.22	3.33	0.87
CART	5%	1.87	2.30	2.12	1.80	4.24	3.83	**1.70**
	10%	2.35	2.93	2.51	2.75	4.28	3.30	**2.25**

6. Conclusions

This paper develops the OGNRS model for attribute reduction by integrating optimized center selection, MAD-based radius computation, and mRMR-guided attribute evaluation. These components jointly enhance the quality of the generated granular balls and the performance of the reduction process. Experiments demonstrate that OGNRS achieves competitive classification accuracy while selecting fewer attributes with relatively low computational cost. Despite these advantages, several aspects of OGNRS remain to be further explored. Although the approximate time complexity of OGNRS is $O(mn)$ when the reduct size is much smaller than the number of attributes, the absolute computational cost of granular ball construction, density estimation, and attribute evaluation may still increase as the data scale or dimensionality grows. In addition, the present framework does not yet support incremental updating for dynamically evolving data. In such scenarios, an adaptive strategy for selecting the nearest neighbor parameter k may also be desirable to accommodate changes in the data distribution. Future work will therefore focus on scalable implementations for large-scale and high-dimensional data as well as incremental learning with adaptive k selection for dynamic data.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

The authors declare that there are no conflicts of interest.

References

1. Z. Pawlak, Rough sets, *Int. J. Comput. Inf. Sci.*, **11** (1982), 341–356. https://doi.org/10.1007/BF01001956
2. M. Modrzejewski, Feature selection using rough sets theory, in *Machine Learning: ECML-93*, **667** (1993), 213–226. https://doi.org/10.1007/3-540-56602-3_138
3. R. W. Swiniarski, A. Skowron, Rough set methods in feature selection and recognition, *Pattern Recognit. Lett.*, **24** (2003), 833–849. https://doi.org/10.1016/S0167-8655(02)00196-4

4. N. Parthaláin, Q. Shen, R. Jensen, A distance measure approach to exploring the rough set boundary region for attribute reduction, *IEEE Trans. Knowl. Data Eng.*, **22** (2010), 305–317. https://doi.org/10.1109/TKDE.2009.119
5. V. K. H. Turaga, S. Chebrolu, Rapid and optimized parallel attribute reduction based on neighborhood rough sets and MapReduce, *Expert Syst. Appl.*, **260** (2025), 125323. https://doi.org/10.1016/j.eswa.2024.125323
6. S. Wu, L. Wang, S. Ge, Z. Hao, Y. Liu, Neighborhood rough set with neighborhood equivalence relation for feature selection, *Knowl. Inf. Syst.*, **66** (2024), 1833–1859. https://doi.org/10.1007/s10115-023-01999-z
7. J. Dai, Q. Hu, H. Hu, D. Huang, Neighbor inconsistent pair selection for attribute reduction by rough set approach, *IEEE Trans. Fuzzy Syst.*, **26** (2018), 937–950. https://doi.org/10.1109/TFUZZ.2017.2698420
8. W. Ziarko, Variable precision rough set model, *J. Comput. Syst. Sci.*, **46** (1993), 39–59. https://doi.org/10.1016/0022-0000(93)90048-2
9. W. Ziarko, Probabilistic approach to rough sets, *Int. J. Approximate Reasoning*, **49** (2008), 272–284. https://doi.org/10.1016/j.ijar.2007.06.014
10. D. Slezak, W. Ziarko, The investigation of the Bayesian rough set model, *Int. J. Approximate Reasoning*, **40** (2005), 81–91. https://doi.org/10.1016/j.ijar.2004.11.004
11. Q. Hu, D. Yu, J. Liu, C. Wu, Neighborhood rough set based heterogeneous feature subset selection, *Inf. Sci.*, **178** (2008), 3577–3594. https://doi.org/10.1016/j.ins.2008.05.024
12. L. Shen, F. E. H. Tay, A discretization method for rough sets theory, *Intell. Data Anal.*, **5** (2001), 431–438. https://doi.org/10.3233/IDA-2001-5506
13. R. Ali, M. H. Siddiqi, S. Lee, Rough set-based approaches for discretization: A compact review, *Artif. Intell. Rev.*, **44** (2015), 235–263. https://doi.org/10.1007/s10462-014-9426-2
14. Q. Hu, H. Zhao, D. Yu, Efficient symbolic and numerical attribute reduction with neighborhood rough sets, *Pattern Recognit. Artif. Intell.*, **21** (2008), 732–738.
15. W. Xu, Z. Yuan, Z. Liu, Feature selection for unbalanced distribution hybrid data based on *k*-nearest neighborhood rough set, *IEEE Trans. Artif. Intell.*, **5** (2024), 229–243. https://doi.org/10.1109/TAI.2023.3237203
16. J. Dai, W. Chen, L. Xia, Feature selection based on neighborhood complementary entropy for heterogeneous data, *Inf. Sci.*, **682** (2024), 121261. https://doi.org/10.1016/j.ins.2024.121261
17. C. Liu, B. Lin, D. Miao, A novel adaptive neighborhood rough sets based on sparrow search algorithm and feature selection, *Inf. Sci.*, **679** (2024), 121099. https://doi.org/10.1016/j.ins.2024.121099
18. S. Xia, Y. Liu, X. Ding, G. Wang, H. Yu, Y. Luo, Granular ball computing classifiers for efficient, scalable and robust learning, *Inf. Sci.*, **483** (2019), 136–152. https://doi.org/10.1016/j.ins.2019.01.010
19. S. Xia, H. Zhang, W. Li, G. Wang, E. Giem, Z. Chen, GBNRS: A novel rough set algorithm for fast adaptive attribute reduction in classification, *IEEE Trans. Knowl. Data Eng.*, **34** (2022), 1231–1242. https://doi.org/10.1109/TKDE.2020.2997039

20. S. Zeng, D. Liu, IEWGBNRS: A novel attribute reduction approach based on information entropy-weighted granular-ball neighborhood rough set, *J. Northwest Univ. (Nat. Sci. Ed.)*, **56** (2026), 128–140. https://doi.org/10.16152/j.cnki.xdxbzr.2026-01-012
21. S. Xia, C. Wang, G. Wang, X. Gao, W. Ding, J. Yu, et al., GBRS: A unified granular ball learning model of Pawlak rough set and neighborhood rough set, *IEEE Trans. Neural Networks Learn. Syst.*, **36** (2025), 1719–1733. https://doi.org/10.1109/TNNLS.2023.3325199
22. Z. Zhang, J. Song, H. Li, E. C. C. Tsang, A weighted granular-ball rough set: model and attribute reduction, in *2024 International Conference on Machine Learning and Cybernetics (ICMLC)*, (2024), 433–438. https://doi.org/10.1109/ICMLC63072.2024.10935042
23. Z. Jia, Z. Zhang, W. Pedrycz, Generation of granular balls for clustering based on the principle of justifiable granularity, *IEEE Trans. Cybern.*, **55** (2025), 1687–1700. https://doi.org/10.1109/TCYB.2025.3534195
24. J. Zhang, K. Sun, B. Huang, T. Wang, X. Wang, Attribute reduction based on a rapid variable granular ball generation model, *Expert Syst. Appl.*, **265** (2025), 126030. https://doi.org/10.1016/j.eswa.2024.126030
25. J. Pan, G. Lang, Q. Xiao, T. Yang, A framework of granular ball generation for classification via granularity tuning, *Appl. Intell.*, **55** (2025), 63. https://doi.org/10.1007/s10489-024-05904-1
26. S. Xia, X. Dai, G. Wang, X. Gao, E. Giem, An efficient and adaptive granular ball generation method in classification problem, *IEEE Trans. Neural Networks Learn. Syst.*, **35** (2024), 5319–5331. https://doi.org/10.1109/TNNLS.2022.3203381
27. Q. Xie, Q. Zhang, S. Xia, F. Zhao, C. Wu, G. Wang, et al., GBG++: A fast and stable granular ball generation method for classification, *IEEE Trans. Emerging Top. Comput. Intell.*, **8** (2024), 2022–2036. https://doi.org/10.1109/TETCI.2024.3359091
28. H. Peng, F. Long, C. Ding, Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, *IEEE Trans. Pattern Anal. Mach. Intell.*, **27** (2005), 1226–1238. https://doi.org/10.1109/TPAMI.2005.159
29. A. Rodriguez, A. Laio, Clustering by fast search and find of density peaks, *Science*, **344** (2014), 1492–1496. https://doi.org/10.1126/science.1242072
30. F. Liu, Q. Zhang, S. Xia, Q. Xie, W. Liao, S. Zhang, A granular ball generation method based on local density for classification, *Inf. Sci.*, **717** (2025), 122295. https://doi.org/10.1016/j.ins.2025.122295

AIMS Press