

A Neighborhood Rough Set Attribute Reduction Method Based on Attribute Importance

Peiyu Su¹, Feng Qin², Fu Li^{1*}

¹Faculty of Mathematics and Statistics, Qinghai Nationalities University, Xining, China

²Faculty of Mathematics and Statistics, Jiangxi Normal University, Nanchang, China

Email: *fl0971@163.com

How to cite this paper: Su, P.Y., Qin, F. and Li, F. (2023) A Neighborhood Rough Set Attribute Reduction Method Based on Attribute Importance. *American Journal of Computational Mathematics*, 13, 578-593. <https://doi.org/10.4236/ajcm.2023.134031>

Received: November 4, 2023

Accepted: November 28, 2023

Published: December 1, 2023

Copyright © 2023 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Attribute reduction is a hot topic in rough set research. As an extension of rough sets, neighborhood rough sets can effectively solve the problem of information loss after data discretization. However, traditional greedy-based neighborhood rough set attribute reduction algorithms have a high computational complexity and long processing time. In this paper, a novel attribute reduction algorithm based on attribute importance is proposed. By using conditional information, the attribute reduction problem in neighborhood rough sets is discussed, and the importance of attributes is measured by conditional information gain. The algorithm iteratively removes the attribute with the lowest importance, thus achieving the goal of attribute reduction. Six groups of UCI datasets are selected, and the proposed algorithm SAR is compared with L_2 -ELM, LapTELM, CTSVM, and TBSVM classifiers. The results demonstrate that SAR can effectively improve the time consumption and accuracy issues in attribute reduction.

Keywords

Rough Sets, Attribute Importance, Attribute Reduction

1. Introduction

Rough set theory [1] is an analytical theory for dealing with uncertain information, but it has significant limitations as it is built on strict equivalence relations. Many scholars have conducted research and extensions, proposing probability rough sets [2], variable precision rough sets [3], neighborhood rough sets [4], and more. Neighborhood rough set effectively addresses the issue of missing information and handling numerical data caused by discretizing data through the granulation of the domain based on neighborhood relations.

Attribute reduction [5] [6], as an important problem in rough set theory, fo-

cuses on deleting redundant attributes while ensuring that the partitioning ability of the attribute set on the domain remains unchanged. Redundant attributes are crucial factors that affect the speed and accuracy of attribute reduction algorithms. Many scholars have conducted research on attribute reduction algorithms based on this foundation. For example, Yao Sheng [7] removes redundant attributes by calculating the rank correlation coefficient between conditional attributes. Liu Fang [8] introduces a tolerance relation matrix and uses the boundary region as inspiration to obtain the optimal attribute reduction set through matrix operations. Wei Bipeng [9] proposed an attribute reduction method for incomplete decision information tables based on \leq -dominance relation. Jing Yunchao [10] presents an incremental matrix reduction algorithm that increases attributes and refines attribute values. Jing [11] designs an incremental attribute reduction method for dynamic data mining, considering changes in both instance objects and attribute sets. In the process of attribute reduction, attribute importance represents the degree of attribute classification change before and after removing certain attributes and has also received widespread attention. For instance, Wei [12] proposed attribute reduction based on granular computing to calculate attribute importance and facilitate attribute reduction. Hu [13] and others propose a forward greedy attribute reduction algorithm based on attribute importance. Liu [14] and others improved the traditional attribute reduction algorithm proposed by Hu *et al.* and put forward the FHARA algorithm. Entropy, as a measure of disorder and uncertainty, is used to calculate conditional information measures in literature [15] [16].

The attribute reduction algorithm proposed in literature [12] is a greedy-based approach that consistently performs high-dimensional operations during attribute selection. It also involves repetitive calculations, resulting in a significant time consumption. Therefore, in this paper, conditional information entropy is used as a measure to calculate the importance of conditional attributes. The obtained importance values are then used to rank all the conditional attributes. Redundant attributes are subsequently removed until the final reduction result is obtained. The structure of the paper is as follows: the first section reviews some basic concepts of neighborhood rough sets and attribute reduction; the second section presents the neighborhood rough set attribute reduction method based on attribute importance; the third section compares the proposed method with L_2 -ELM, LapTELM, CTSVM, and TBSVM classifiers; and finally, the fourth section provides a summary.

2. Preliminary Knowledge

To facilitate the research work, this section introduces the basic knowledge of neighborhood rough sets, attribute reduction methods, and other related concepts.

2.1. Neighborhood Rough Sets

The classical rough set must be discretized before processing the actual sample

data, in which the integrity of the data and the authenticity of the discretization results cannot be guaranteed. Neighborhood rough sets, as an extension of rough sets, can handle continuous or mixed-type data in real-life situations and effectively address the problem of information loss after data discretization, which is a common issue in classical rough set models.

Let $S = \{U, AT = A \cup D, \delta\}$ be a neighborhood rough set decision system, where U is the set of all objects, A is the set of conditional attributes, and D is the set of decision attributes, with $A \cap D = \emptyset$.

Definition 1 [4]: Given a decision system S and a neighborhood radius δ , for $\forall x_i \in U$, its δ neighborhood is defined as:

$$\delta(x_i) = \{x \in U \mid d(x, x_i) \leq \delta\} \tag{1}$$

Here, $\delta \geq 0$, $d(x, x_i)$ represents the distance between x and x_i , and common distance calculation methods include Manhattan distance, Euclidean distance, and norm distance functions. In this paper, the Euclidean distance function is used to calculate $d(x, x_i)$. The δ -neighborhood information granule, denoted as $\delta(x_i)$, δ -neighborhood granule, is denoted as x_i neighborhood granule.

Property 1: Given a decision system S and a neighborhood radius δ , for any subset $B \subseteq A$, of conditional attributes, if $\delta_1 < \delta_2$, then $\delta_1(x) \subseteq \delta_2(x)$.

Proof: For any $\forall x_i \in \delta_1(x)$, we have $d(x, x_i) < \delta_1 < \delta_2$. If $d(x, x_i) < \delta_1$, then $\delta_1(x) \subseteq \delta_2(x)$.

Definition 2 [4]: Given a decision system S and the decision attribute set D , the decision class set $\{X_1, X_2, \dots, X_n\}$ is obtained. Then, the lower and upper approximations of the decision attribute D with respect to the attribute subset B are defined as follows:

$$\begin{aligned} \underline{R}_B(D) &= \bigcup_{i=1}^n \underline{R}_B(X_i) \\ \overline{R}_B(D) &= \bigcup_{i=1}^n \overline{R}_B(X_i) \end{aligned} \tag{2}$$

where:

$$\begin{aligned} \underline{R}_B(X) &= \{x_i \in U \mid \delta(x_i) \subseteq X_i\} \\ \overline{R}_B(X) &= \{x_i \in U \mid \delta(x_i) \cap X_i \neq \emptyset\} \end{aligned} \tag{3}$$

The boundary of the decision system is defined as:

$$BN(D) = \overline{R}_B(D) - \underline{R}_B(D) \tag{4}$$

The positive region and negative region of the neighborhood decision system are defined as:

$$Pos_B(D) = \underline{R}_B(D), Neg_B(D) = U - \overline{R}_B(D) \tag{5}$$

The dependency degree of decision attribute D on conditional attribute B is defined as:

$$\gamma_B(D) = \frac{|Pos_B(D)|}{|U|} \tag{6}$$

Property 2: Given a decision system S and a neighborhood radius δ , for any subset $B \subseteq A$ of conditional attributes, the following properties hold:

- 1) For any $\forall B_1 \subseteq B_2$, $R_{\delta, B_1}(X) \supseteq R_{\delta, B_2}(X)$;
- 2) If $\delta_1 < \delta_2$, then $Pos_{B, \delta_2} \subset Pos_{B, \delta_1}$;
- 3) For any $B_1 \subseteq B_2 \subseteq \dots \subseteq A$, have $Pos_{B_1}(D) \subseteq Pos_{B_2}(D) \subseteq \dots \subseteq Pos_{B_A}(D)$, then $\gamma_{B_1} \leq \gamma_{B_2} \leq \dots \leq \gamma_{B_A}$.

2.2. Attribute Reduction

Attribute reduction is a core problem in rough set theory research. Its core idea is to remove redundant attributes while maintaining the classification ability of the information system, thereby extracting key attributes, and simplifying the information system. Depending on the measurement method of attribute importance, commonly used measurement criteria for attribute reduction include attribute dependency-based and conditional information entropy-based attribute reduction.

Definition 3 [6]: Given a decision system $DS = \{U, AT, V, f\}$, for any conditional attribute subset $B \subseteq A$ and $B \neq \emptyset$, the indistinguishable relationship of attribute subset B on the domain U is defined as:

$$IND(B) = \{(x, y) \in U \times U \mid \forall a \in B, f(x, a) = f(y, a)\} \tag{7}$$

Here, x and y are any two samples on the domain U , and the indistinguishable relationship determined by the attribute subset can be seen as dividing the domain U into multiple partitions, denoted as $U/IND(B)$.

Definition 4 [6]: Given a knowledge base $K = (U, R)$ and an equivalence relation $P \subseteq R$ on the knowledge base, for any $G \subseteq P$, if G satisfies the following conditions:

- 1) G is independent, meaning that every element in G is indispensable.
- 2) $IND(D) = IND(P)$, meaning that it does not affect the partition of the knowledge base.

Then G is called a reduction of P , denoted as $G \in Red(P)$. Here, $Red(P)$ represents the set of all reductions of P . If G satisfies the following condition:

$$IND(P - \{G\}) \neq IND(P) \tag{8}$$

Then G is called indispensable in P . The set of all indispensable knowledge sets in P is called the core of P , denoted as $Core(P)$.

Definition 5 [6]: Given a decision system $DS = \{U, AT, V, f\}$, where $AT = A \cup D$, $B \subseteq A$, if the conditional attribute subset B satisfies the following conditions:

- 1) $\gamma_B(D) = \gamma_A(D)$, meaning $Pos_B(D) = Pos_A(D)$;
- 2) For any $\forall a \in B$, $\gamma_B(D) > \gamma_{B - \{a\}}(D)$

Then, attribute subset B is called a relative reduction of the conditional attribute set A . To represent attribute reduction based on dependency functions using information measures, Hu defined the notion of information droplets in

decision systems.

Definition 6 [15] [16]: Given a decision system S , $U = \{x_1, x_2, \dots, x_n\}$, and $B \subseteq A$, the conditional information entropy of D with respect to B is defined as:

$$H(D|B) = -\frac{1}{n} \sum_{i=1}^n \log \frac{|[x_i]_B \cap [x_i]_D|}{|[x_i]_B|} \tag{9}$$

Here, $[x_i]_B$ represents the fuzzy equivalence class, $[x_i]_D$ represents the decision class. For $\forall a \in B$, B is a dependency-based reduction if and only if $H(D|B) = H(D|A)$ and $H(D|B) > H(D|B - \{a\})$.

3. Attribute Reduction Method for Neighborhood Rough Sets Based on Attribute Importance

As an effective measure of uncertainty, information entropy can be improved and optimized on the basis of combining with classical rough sets, which can deal with mixed data effectively. During the attribute reduction process based on attribute importance, the conditional information entropy is considered as a measure of attribute importance, which can effectively enhance the discriminative power of attributes. The larger the conditional information entropy, the higher the attribute importance in the neighborhood rough set, and vice versa.

Definition 7: Given a neighborhood decision system S , for $\forall B \subseteq A$, if $H(D|B) = H(D|A)$ and $H(D|B) > H(D|B - \{a\})$, then the conditional attribute subset B is a reduction of the conditional attribute set A . B_1, B_2, \dots, B_n are all the possible reductions of A .

Definition 8: In a neighborhood rough set decision system $S = \{U, AT = A \cup D, \delta\}$, for $\forall B \in A$, the conditional information entropy with respect to the conditional attribute B is defined as:

$$H(B) = -\frac{1}{n} \sum_{i=1}^n \log \frac{1}{\delta_B(x_i)} \tag{10}$$

Definition 9: In the neighborhood rough set decision system $S = \{U, AT = A \cup D, \delta\}$, for any decision attribute set B , the process of selecting a subset B that has the same discernibility effect on objects as the attribute set AT is called attribute reduction. The conditional information entropy of D with respect to B is defined as follows:

$$H(D|B) = -\frac{1}{n} \sum_{i=1}^n \log \frac{\delta_B(x_i) \cap [x_i]_D}{\delta_B(x_i)} \tag{11}$$

Definition 10: Given a neighborhood decision system S , for $\forall a \subseteq B$, the attribute importance of a with respect to B is defined as follows:

$$Sig(a, B, D) = H(D|B) - H(D|B \cup \{a\}) \tag{12}$$

Conclusion 1: In a neighborhood rough set decision system $S = \{U, AT = A \cup D, \delta\}$, for any decision attribute sets B , if $B_1 \subseteq B_2$, then $H(D|B_1) \leq H(D|B_2)$.

Proof: Since $B_1 \subseteq B_2$, then $\delta_{B_1}(x_i) \supseteq \delta_{B_2}(x_i)$. It follows from Equation (11) that $H(D|B_1) - H(D|B_2) \leq 0$, which implies $H(D|B_1) \leq H(D|B_2)$.

Conclusion 2: In a neighborhood rough set decision system $S = \{U, AT = A \cup D, \delta\}$, for $\forall a \subseteq B$, the attribute importance of a with respect to B , denoted as $Sig(a, B, D)$, is higher when the conditional information entropy $H(D|B)$ is larger. Conversely, when $H(D|B)$ is smaller, the attribute a is considered less important.

Example 1: Given a neighborhood decision information system $S = \{U, AT = A \cup D, \delta\}$, you can see this table (Table 1), where $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, $A = \{a_1, a_2, a_3\}$, $D = \{d_1, d_2, d_3\}$, $\delta_{a_1} = 0.134$, $\delta_{a_2} = 0.172$, $\delta_{a_3} = 0.154$. The neighborhood decision information table is as follows:

According to the decision attribute D , the domain U can be divided into $D_1 = \{x_1, x_5\}$, $D_2 = \{x_2, x_4\}$, and $D_3 = \{x_3, x_6\}$. The conditional information entropy is calculated as follows:

$$H(D|B) = -\frac{1}{n} \sum_{i=1}^n \log \frac{\delta_B(x_i) \cap [x_i]_D}{\delta_B(x_i)} = 3.2749 \tag{13}$$

Then $H(D|B \cup \{a_1\}) = 2.5874$, $H(D|B \cup \{a_2\}) = 2.3854$, $H(D|B \cup \{a_3\}) = 3.1028$. The formula $Sig(a, B, D) = H(D|B) - H(D|B \cup \{a\})$ has been used to determine the attribute importance $Sig(a_1, B, D) = 0.6875$, $Sig(a_2, B, D) = 0.8895$, $Sig(a_3, B, D) = 0.1721$ of conditional attributes a_1, a_2, a_3 . Therefore, based on the attribute importance values, the order of importance is $a_2 > a_1 > a_3$, and attribute a_3 can be considered as a redundant attribute in the reduction process.

The traditional neighborhood rough set attribute reduction algorithm uses greedy search strategy to select attributes, and adds new candidate condition attributes with cycles to produce new attribute subsets. However, this method requires many cycles and has high time complexity. To address neighborhood rough set attribute sets, the attribute reduction algorithm based on attribute importance is proposed, utilizing conditional information entropy as a measure of importance. The algorithm uses forward search and heuristic information, and the detailed description is as follows:

Table 1. Neighborhood decision information system.

U	a_1	a_2	a_3	d
x_1	0.20	0.25	0.25	1
x_2	0.40	0.40	0.50	2
x_3	0.30	0.25	0.45	3
x_4	0.45	0.30	0.50	2
x_5	0.45	0.50	0.55	1
x_6	0.60	0.45	0.55	3

Algorithm 1: Attribute Reduction Method for Neighborhood Rough Sets Based on Attribute Importance (SAR).

Input: Decision information system $S = \{U, AT = A \cup D, \delta\}$, radius δ .

Output: Reduction set B

Step 1: Set $B = \emptyset$.

Step 2: Calculate the conditional information entropy $H(D|B)$ for the entire neighborhood decision system.

Step 3: **DO**

1) Calculate the conditional information entropy $H(D|B \cup \{a_i\})$ for each conditional attribute $a_i (i = 1, 2, \dots, n)$;

2) Calculate the attribute importance of a_i ,

$$Sig(a_i, B, D) = H(D|B) - H(D|B \cup \{a_i\});$$

3) Select the attribute with the highest attribute importance

$$Sig(a_i, B, D) = \max(a_i);$$

4) If $Sig(a_i, B, D) > Sig(B, D)$, add a_i to the reduction set B and go to Step 3. Otherwise, go to Step 4.

Step 4: Output the reduction result.

This paper utilizes conditional information entropy to calculate the attribute importance. In the reduction process, attribute importance is calculated only once. The algorithm improves efficiency and reduces computational complexity when dealing with datasets with a large number of attributes, while also enhancing the accuracy of reduction.

4. Experimental Analysis

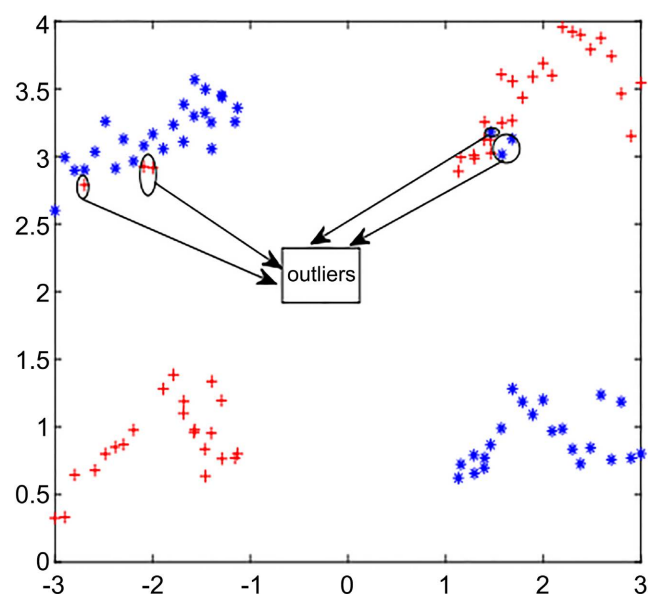
To validate the effectiveness of the proposed SAR algorithm, 8 UCI datasets were selected for experimentation. You can see this table (Table 2). The datasets were downloaded from the UCI Machine Learning Repository. The experimental environment consisted of a computer running Windows 10, with an Intel Core i5-6300HQ CPU @ 2.30 GHz processor, and the MATLAB R2021a programming language.

A 10-fold cross-validation method was employed in the experiments. Ten different radii were selected: 0.02, 0.04, ..., 0.2. The data was divided into 10 subsets, with each subset used as the test set while the remaining 9 subsets were used as the training set in a rotating manner. The average classification accuracy of the 10 experiments was calculated as the performance metric for the attributes. The SAR algorithm proposed in this paper was compared with L_2 -ELM, Lap-TELM, CTSVM, and TBSVM classifiers in terms of classifying the test set samples.

Initially, 100 artificial datasets were divided into positive and negative classes, represented by + and *, respectively. To examine the impact of outliers on the classification performance, 6 outliers were introduced into the artificial datasets, with an equal distribution of 3 positive and 3 negative outliers; you can see this figure (Figure 1):

Table 2. Dataset description.

No.	Dataset Name	Samples	Conditional Attributes
1	WDBC	569	30
2	Vote	432	16
3	QSAR	1055	41
4	Cancer	699	9
5	spectf	267	44
6	Pima	768	8

**Figure 1.** Distribution of artificial dataset with outliers.

4.1. Time Comparison for Attribute Reduction

During the experiment, the time consumption for attribute reduction was recorded for SAR, L_2 -ELM, LapTELM, CTSVM, and TBSVM classifiers. The comparison results are shown in the following table, you can see this table (**Table 3**):

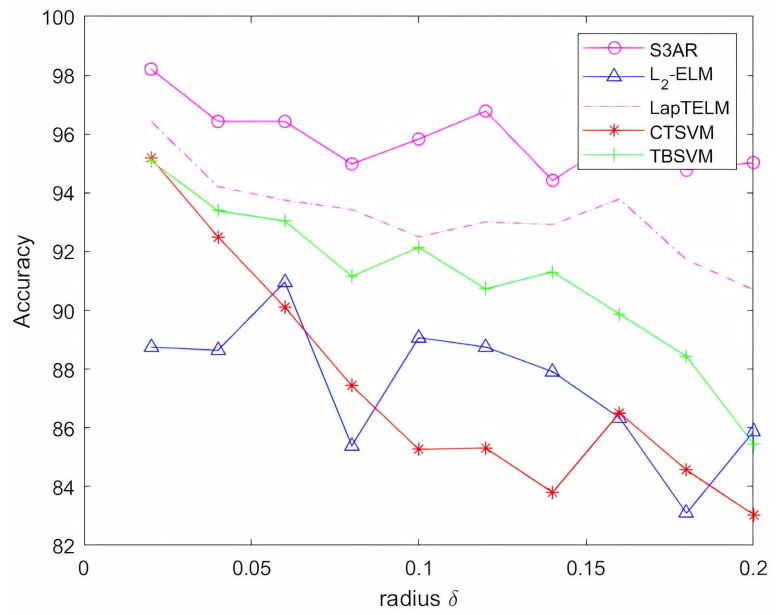
4.2. Accuracy Comparison for Attribute Reduction

To demonstrate the effectiveness of SAR in terms of classification accuracy, SAR was compared with L_2 -ELM, LapTELM, CTSVM, and TBSVM classifiers on the test set samples. The experimental results are as follows, you can see this figure (**Figure 2**):

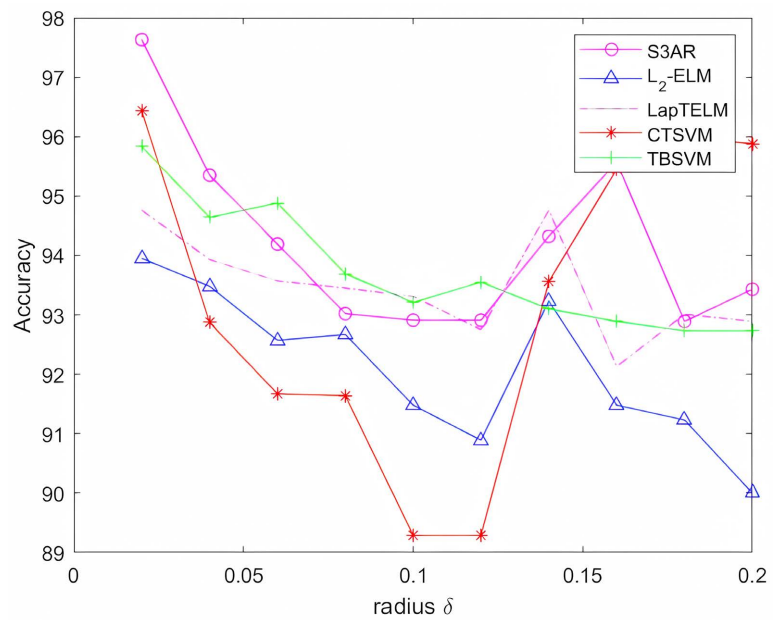
The experimental results show that SAR achieves higher classification accuracy than L_2 -ELM, LapTELM, CTSVM, and TBSVM classifiers for most radii. Therefore, SAR can effectively improve the classification accuracy of attribute reduction. To further validate the advantages of SAR, experiments were conducted on artificial datasets using the five algorithms. You can see this figure (**Figure 3**), the results are as follows:

Table 3. Time comparison for five algorithms' attribute reduction.

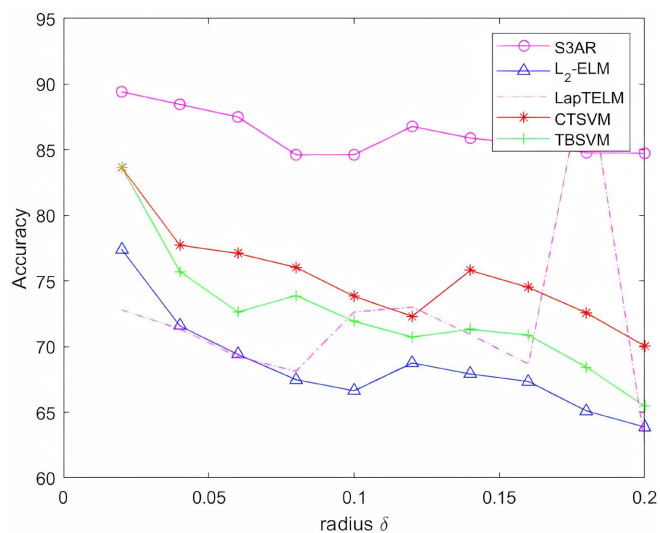
	L_2 -ELM	LapTELM	CTSVM	TBSVM	S3AR
WDBC	1.545	2.555	3.114	1.864	0.451
Vote	1.478	5.607	2.501	6.867	0.455
QSAR	3.101	4.523	7.060	8.302	3.457
Cancer	4.274	5.881	3.151	7.359	1.274
spectf	0.991	4.388	1.582	5.539	1.522
Pima	1.967	4.857	3.810	2.689	0.699



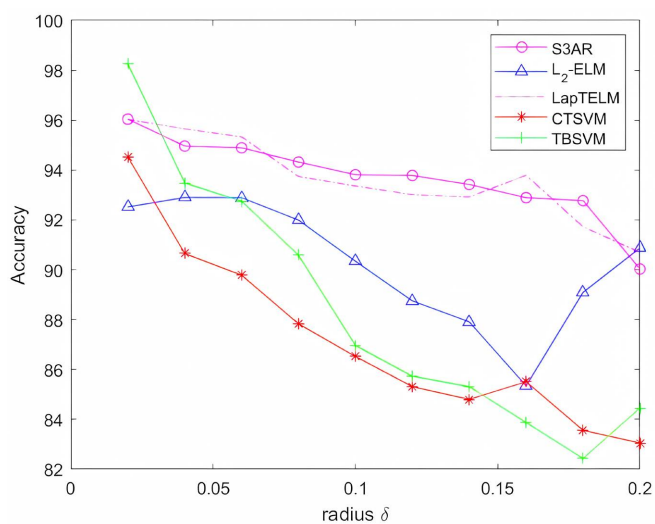
(a)



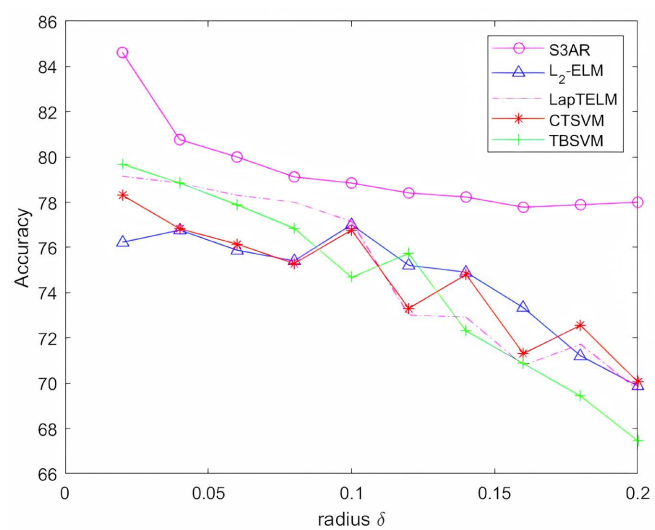
(b)



(c)



(d)



(e)

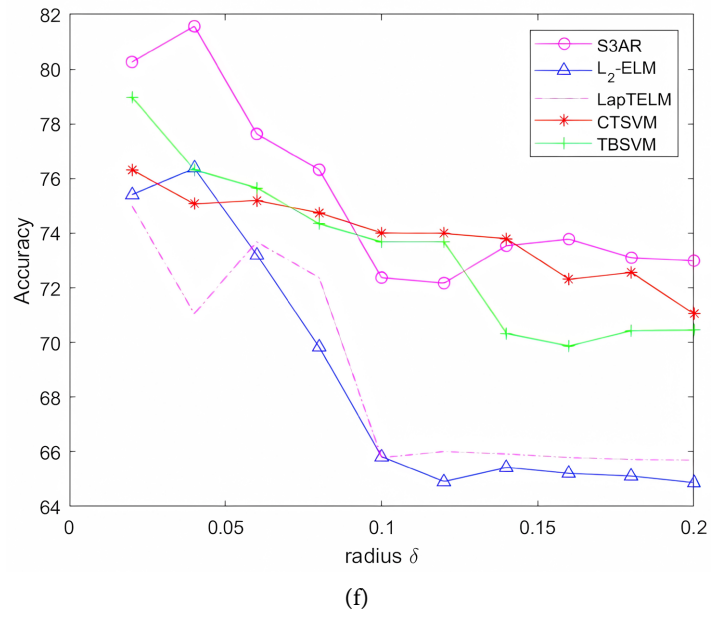
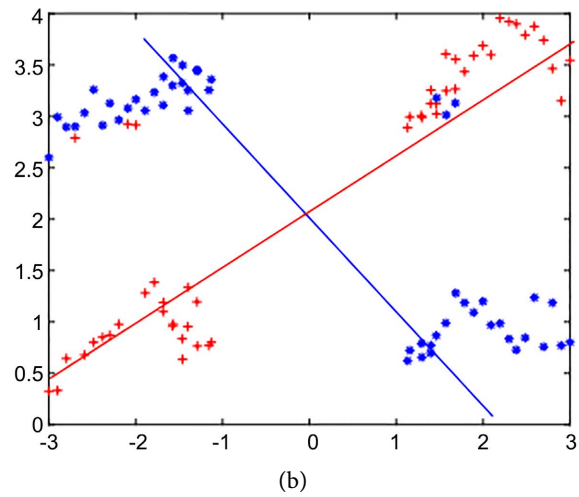
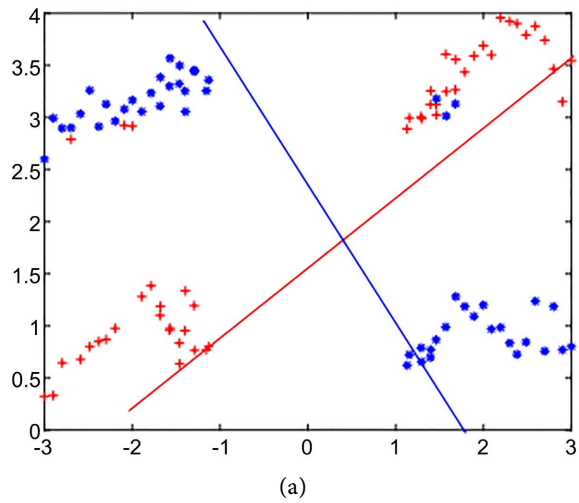
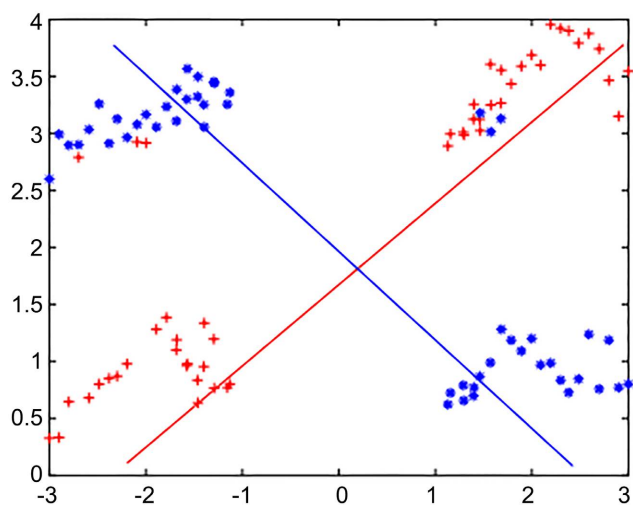
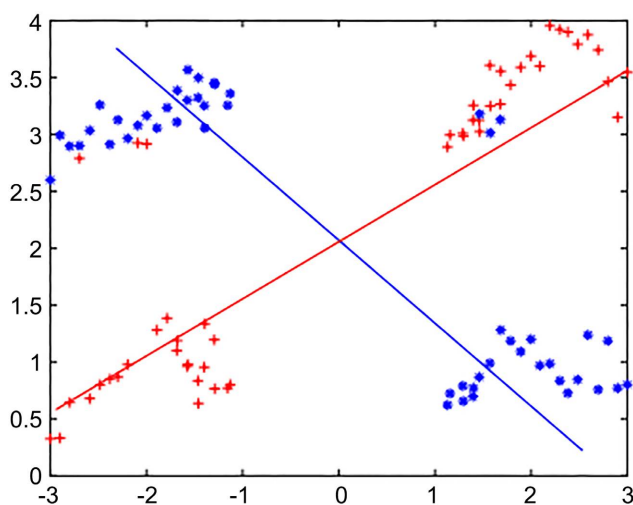


Figure 2. Accuracy comparison of five algorithms at different radii. (a) WDBC; (b) Vote; (c) QSAR; (d) Cancer; (e) Spectf; (f) Pima.

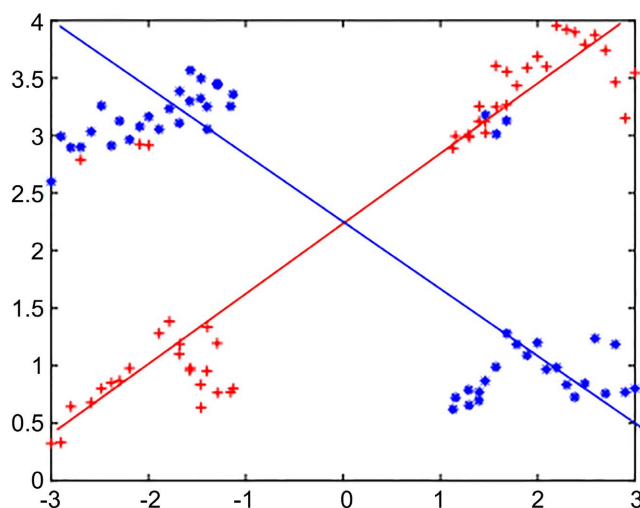




(c)



(d)

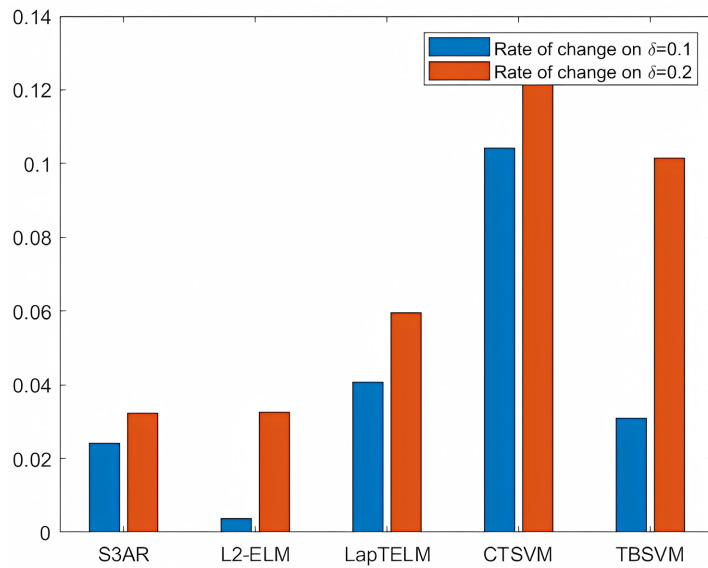


(e)

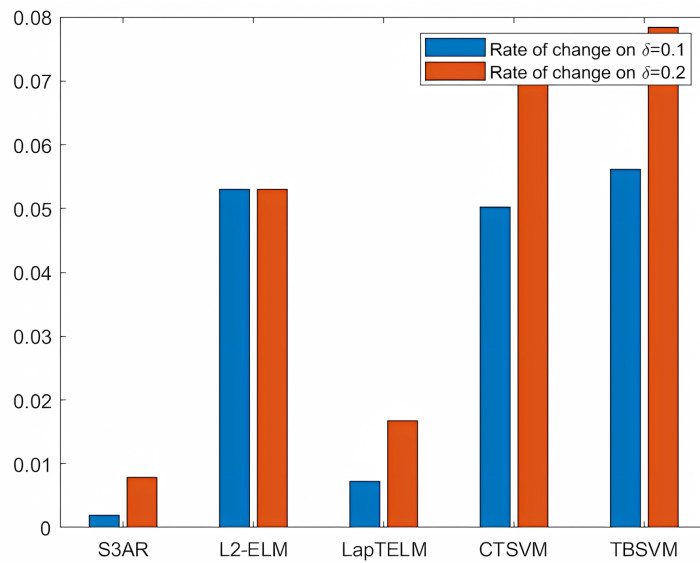
Figure 3. Accuracy comparison of five algorithms at different radii. (a) L_2 -ELM; (b) LapTELM; (c) CTSVM; (d) TBSVM; (e) S3AR.

From the above figure, it can be observed that SAR still exhibits good classification performance on artificial datasets with outliers. To better illustrate the stability of SAR in the classification process, experiments were conducted to analyze the variation rate of SAR, L_2 -ELM, LapTELM, CTSVM, and TBSVM classifiers at radii of $\delta=0.1$ and $\delta=0.2$. You can see this figure (Figure 4), the results are as follows:

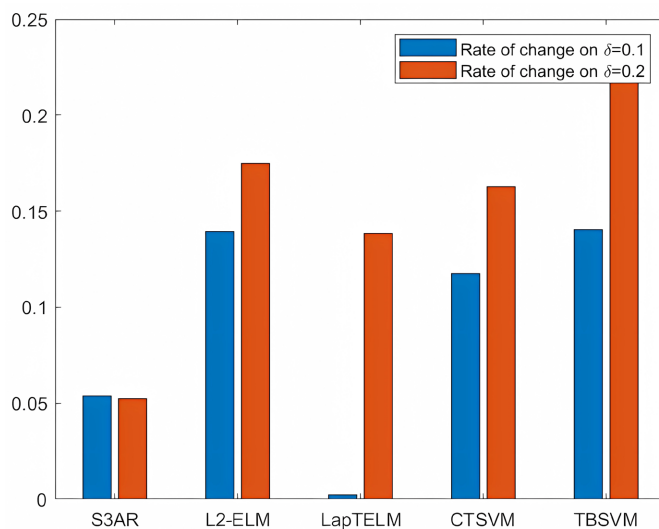
The above experiments provide a comparison of the classification accuracy and precision of the five algorithms on six datasets. Additionally, the variation rates at radii of $\delta=0.1$ and $\delta=0.2$ were compared across different datasets. The experimental results indicate that, under the same conditions of algorithm, radius, and metric, SAR performs as well as L_2 -ELM, LapTELM, CTSVM, and TBSVM classifiers in terms of classification performance.



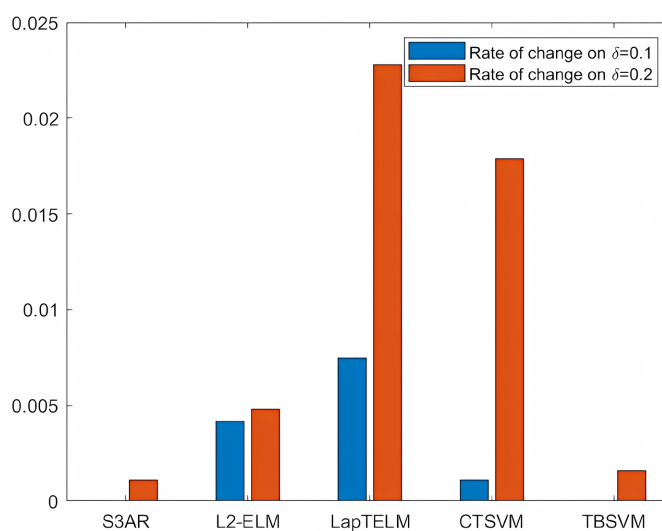
(a)



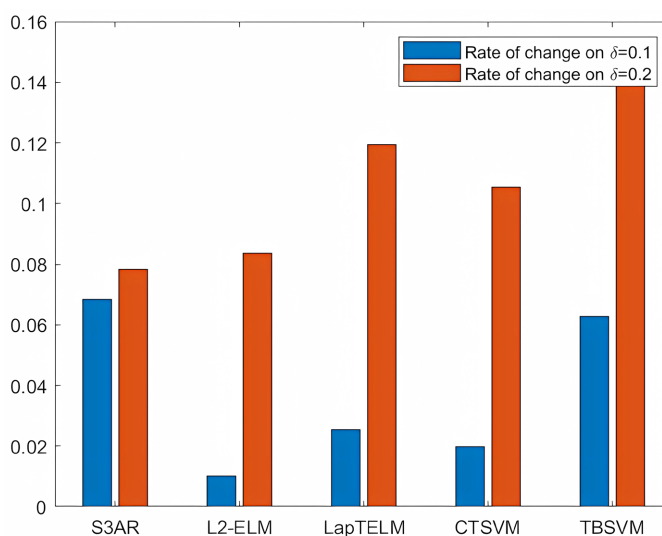
(b)



(c)



(d)



(e)

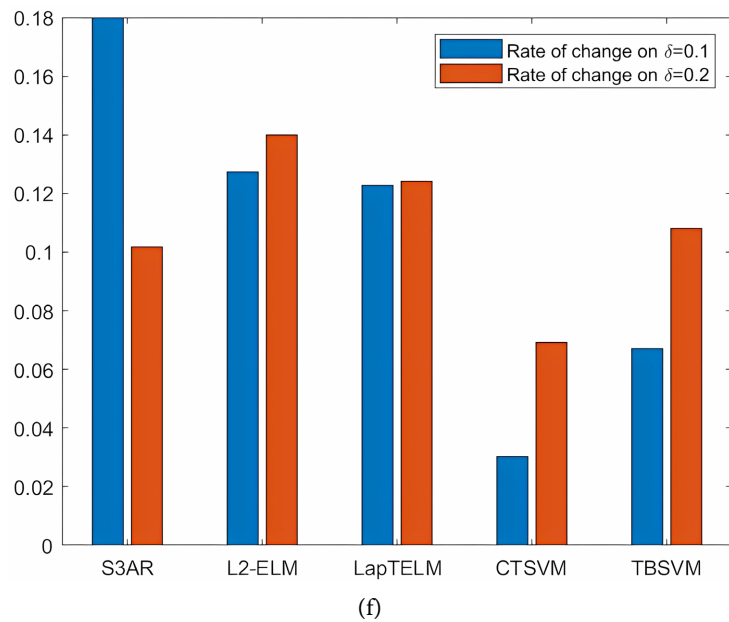


Figure 4. Comparison of the change rates of $\delta = 0.1$ and $\delta = 0.2$ under different UCI datasets. (a) WDBC; (b) Vote; (c) QSAR; (d) Cancer; (e) Spectf; (f) Pima.

5. Conclusion

This paper proposes a neighborhood rough set attribute reduction method based on attribute importance. It utilizes conditional information entropy to calculate attribute importance and obtain the final attribute reduction results. By comparing the time consumption and classification accuracy of SAR, L_2 -ELM, LapTELM, CTSVM, and TBSVM classifiers, the proposed SAR method shows good classification performance with lower time consumption. Compared with other algorithms, S3AR algorithm has a faster running speed, but the time cost is still very high. In the future, we will focus on how to optimize the algorithm to reduce the time cost.

Funding

This work was supported by the National Natural Science Foundation of China under grant numbers 11971210 and 12371459.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Pawlak, Z. (1982) Rough Sets. *International Journal of Computer and Information Sciences*, **11**, 341-356. <https://doi.org/10.1007/BF01001956>
- [2] Pawlak, Z., Wong, S.K.M. and Ziarko, W. (1988) Rough Sets: Probabilistic versus Deterministic Approach. *International Journal of Man-Machine Studies*, **29**, 81-95. [https://doi.org/10.1016/S0020-7373\(88\)80032-4](https://doi.org/10.1016/S0020-7373(88)80032-4)

-
- [3] Ziarko, W. (1993) Variable Precision Rough Set Model. *Journal of Computer and System Sciences*, **46**, 39-59. [https://doi.org/10.1016/0022-0000\(93\)90048-2](https://doi.org/10.1016/0022-0000(93)90048-2)
- [4] Lin, T.Y. (1989) Neighborhood Systems and Approximation in Relational Databases and Knowledge Bases. *Proceedings of the Fourth International Symposium on Methodologies of Intelligent Systems*, Charlotte, February 1988, 75-86.
- [5] Ju, H., et al. (2017) Cost-Sensitive Rough Set: A Multi-Granulation Approach. *Knowledge-Based Systems*, **123**, 137-153. <https://doi.org/10.1016/j.knosys.2017.02.019>
- [6] Wang, G., Yao, Y. and Yu, H. (2009) A Survey on Rough Set Theory and Its Applications. *Journal of Computer Science*, **32**, 1229-1246. <https://doi.org/10.3724/SP.J.1016.2009.01229>
- [7] Yao, S., Wang, J., Xu, F., et al. (2018) Uncertainty Measurement and Attribute Reduction Group of Inconsistent Neighborhood Rough Sets. *Journal of Chinese Computer Systems*, **39**, 700-706.
- [8] Liu, F. and Li, T. (2016) Attribute Reduction Method for Incomplete Information System Based on Boundary Region. *Computer Science*, **43**, 242-245, 284.
- [9] Wei, B., Lv, Y. and Li, J. (2014) Attribute Reduction of Rough Set Model under 4-Dominance Relation. *Journal of Chinese Intelligent Systems*, **9**, 251-258.
- [10] Jing, Y., Jing, L. and Wang, B. (2020) Incremental Attribute Reduction Algorithm Based on Attribute Value and Attribute Change. *Journal of Shandong University (Natural Science)*, **55**, 62-68.
- [11] Jing, Y., Li, T., Fujita, H., et al. (2018) An Incremental Attribute Reduction Method for Dynamic Data Mining. *Information Fusion*, **465**, 202-218. <https://doi.org/10.1016/j.ins.2018.07.001>
- [12] Wei, W., Wu, X., Liang, J., et al. (2018) Discernibility Matrix Based Incremental Attribute Reduction for Dynamic Data. *Knowledge-Based Systems*, **140**, 142-157. <https://doi.org/10.1016/j.knosys.2017.10.033>
- [13] Hu, Q., Yu, D. and Xie, Z. (2008) Numerical Attribute Reduction Based on Neighborhood Granulation and Rough Approximation. *Journal of Software*, **19**, 640-649. <https://doi.org/10.3724/SP.J.1001.2008.00640>
- [14] Liu, Y., et al. (2014) Quick Attribute Reduction Algorithm for Neighborhood Rough Set Model. *Information Sciences*, **271**, 65-81. <https://doi.org/10.1016/j.ins.2013.08.022>
- [15] Hu, Q., Yu, D., Xie, Z., et al. (2006) Fuzzy Probabilistic Approximations Spaces and Their Information Measures. *IEEE Transactions on Fuzzy Systems*, **14**, 191-201. <https://doi.org/10.1109/TFUZZ.2005.864086>
- [16] Hu, Q., Yu, D. and Xie, Z. (2006) Information-Preserving Hybrid Data Reduction Based on Fuzzy Rough Techniques. *Pattern Recognition Letters*, **27**, 414-423. <https://doi.org/10.1016/j.patrec.2005.09.004>