THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 25, Number 4/2024, pp. 359–370
DOI: 10.59277/PRA-SER.A.25.4.12

INTUITIONISTIC FUZZY SVM BASED ON KERNEL GRAY RELATIONAL ANALYSIS

Zhe JU, Qingbao ZHANG

Shenyang Aerospace University, College of Sciences, Shenyang, P.R. China, 110136
Corresponding author: Zhe JU, E-mail: juzhe@sau.edu.cn

Abstract. Fuzzy Support Vector Machine (FSVM) is a machine learning algorithm that combines fuzzy logic with Support Vector Machine (SVM) to deal with the uncertainty and fuzziness in classification and regression problems. This algorithm improves the performance of traditional SVM by introducing fuzzy membership degrees, making it more robust when handling datasets with noise or uncertainty. Although the existing FSVM algorithms can overcome the influence of noise to a certain extent, they cannot effectively distinguish outliers or abnormal values from boundary support vectors. To solve this problem, this study proposes an Intuitionistic Fuzzy Support Vector Machine algorithm (KGRA-IFSVM) based on Kernel Grey Relational Analysis (KGRA). This approach utilizes gray relational analysis in the kernel space to calculate the gray relational degree between each sample and its K isomorphic neighboring points, and takes the average value as the membership degree of the sample. Then, the same approach is used to compute the gray relational degree between each sample and its K heterogeneous neighboring points, and the average value is taken as the non-membership degree of the sample. Finally, each sample is assigned with an appropriate fuzzy value based on intuitionistic fuzzy sets using a specific scoring function. Test results on UCI datasets show that KGRA-IFSVM has better classification performance and stronger noise resistance.

***Keywords**:* fuzzy support vector machine, intuitionistic fuzzy number, kernel grey relational analysis, fuzzy member function.

1. INTRODUCTION

With the continuous development of information technology, the scale, complexity, and diversity of data continue to increase. How to extract useful information from these massive amounts of data has become one of the key issues [1]. As a powerful tool, machine learning has been widely used in the field of data mining [2]. Data mining mainly includes tasks such as classification, clustering, and anomaly detection [3]. Among them, classification problems are widely existed in practical applications and are important research contents in the field of machine learning and data mining. In recent years, many classification learning algorithms have been proposed successively, including decision tree (DT), artificial neural network (ANN), support vector machine (SVM), naive Bayes classifier (NB), and K-nearest neighbor (KNN) [4]. Among them, SVM are superior to other machine learning algorithms, such as ANN, in terms of generalization ability [5]. The latter has problems of overfitting and local minimums [5]. SVM can overcome the problems of local minimums and the curse of dimensionality in traditional machine learning algorithms [6]. In a variety of applications such as face detection, feature extraction, gene prediction, and other classification problems, the performance of support vector machines is better than most other learning techniques [5]. However, SVM treats all samples equally, ignoring the influence of outliers and noise on the construction of the optimal hyperplane, which makes it perform poorly in imbalanced classification problems or problems with irregular data distribution [6]. In response to this, scholars have proposed the fuzzy support vector machine algorithm (FSVM).

FSVM is a machine learning algorithm that combines the advantages of fuzzy logic and SVM. It improves the standard SVM by introducing the concept of fuzzy logic [7]. In FSVM, fuzzy logic is used to deal with the uncertainty and fuzziness in the data [7]. By assigning a fuzzy membership degree to each sample, FSVM can more flexibly deal with the weights of the samples and the determination of the classification boundary [8]. This makes FSVM have better robustness in processing complex datasets with noise, uncertainty and outliers. The key to FSVM is how to automatically or adaptively determine the fuzzy membership degree of training data [9]. Lin and Wang [10] calculated the membership degrees of training samples based on the distances between samples and class centers in the input space. However, when the training data is not spherically distributed, the classification contribution of each sample may not be correctly represented. For nonlinear classification problems, Jiang *et al.* [11] performed kernel extension on the fuzzy membership degree calculation formula in [10], and the fuzzy membership degree was calculated in the feature space. Based on [11], Wang *et al.* [12] calculated the outlier factor of instances in the kernel space by utilizing the similarity of sampling points and their neighborhood densities. The membership degree of majority instances was set as the reciprocal of the outlier factor, while the membership degree of minority instances was set to 1 to highlight their importance. However, the distribution characteristics of minority instances were not taken into account. Tang *et al.* [13] calculated the fuzzy membership degree of instances based on the grey relational degree between instances and class centers. Although this approach can effectively distinguish boundary support vectors from outliers, the model is sensitive to noise because the class centers are prone to deviation caused by noise or outliers. The above fuzzy membership degree calculation approaches are all determined based on the relationship between instances and a certain class, rather than the relationship between classes, which leads to inaccurate sample distribution information. To address this issue, Ha and Wang [14] combined the membership degree and non-membership degree determined by the instance distribution information with the intuitionistic fuzzy set to calculate the classification contribution value of the instance. Among them, the membership degree was calculated based on the distance between the instance and the class center, while the non-membership degree was calculated based on the ratio of the number of heterogeneous points of the instance to the number of all instances within its neighborhood. However, if the class center is calculated inaccurately or affected by noisy data, the membership degree of the sample may deviate, which may further affect the classification performance.

Obviously, most of the existing FSVM algorithms measure the classification contribution value (membership degree) of each sample based on the Euclidean distance between samples, and the classification performance of the model is largely limited by the distribution of samples. In this regard, based on the literature [13] [14], this study proposes a new method to calculate the classification contribution value of samples. In the high-dimensional feature space, the kernel grey relational analysis approach is utilized to calculate the grey relational degrees between the instances and their K_a isomorphic neighboring points, and the mean values are taken as the membership degree of the instances. Similarly, the grey relational degrees between the instances and their K_b heterogeneous neighboring points are calculated, and the mean values are taken as the non-membership degree of the instances. Finally, the classification contribution of different instances is obtained through the scoring function. Compared with the literature [13], [14], the calculation of instance classification contribution in this paper no longer relies on the calculation of class centers, which can effectively distinguish outliers from noise and better cope with the impact of noise or outliers.

The main contributions of this paper are as follows:

- A new method for calculating the degree of membership is proposed, which utilizes grey relational analysis in the kernel space to calculate the degree of membership and non-membership of samples.
- Based on the concept of intuitionistic fuzzy set, the training samples are converted into intuitionistic fuzzy numbers, and a new score function for intuitionistic fuzzy numbers is introduced to measure the contribution of intuitionistic fuzzy numbers. Finally, a new FSVM is constructed according to the score value of each training sample.

The rest of this paper is structured as follows: Section 2 briefly introduces the basic principles of FSVM and the limitations of a membership degree calculation method. Section 3 first introduces the basic principles of kernel grey relational analysis and intuitionistic fuzzy set, and then proposes a new membership degree calculation method based on them. The rationality of this method is analyzed on an irregular artificial dataset. In Section 4, tests are performed on eight UCI datasets to verify the effectiveness of the KGRA-IFSVM algorithm. Section 5 concludes the paper.

2. FUZZY SUPPORT VECTOR MACHINE

When performing binary classification on data containing noise or outliers, SVM assigns the same misclassification cost to each sample, which results in significant bias in the decision boundary [15]. Unlike SVM, FSVM assigns different fuzzy membership degrees to instances based on a specified fuzzy membership function, distinguishing the classification contributions of different samples through the magnitude of fuzzy membership degrees [16]. Assuming an initial binary classification dataset is: $D = \{(x_i, y_i)|\, i = 1,2,\cdots,N\}$, where $x_i \in R^m$ represents an m-dimensional input vector and $y_i \in \{-1,1\}$ represents the class label. By introducing the fuzzy membership degree, the dataset is expanded to $D_s = \{(x_i, y_i, s_i)|\, i = 1,2,\cdots,N\}$, where s_i represents the membership degree of the ith instance. The objective function of FSVM can be expressed as:

$$\min_{\omega,b} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N} s_i\varepsilon_i \tag{1}$$

$$s.t.\ \ y_i(\omega^T\varphi(x_i) + b) + \varepsilon_i \ge 1, \varepsilon_i \ge 0, \forall i = 1,2,\cdots,N.$$

The membership degree s_i of the instance x_i is incorporated into the objective function in the FSVM optimization problem (1), which is the only difference from the original SVM optimization problem. The fuzzy value s_i represents the weight of instance x_i in the objective function, reflecting the importance of the corresponding instance to the classification hyperplane [17]. Then, by assigning smaller fuzzy values to noise and outliers, the classification hyperplane becomes more reasonable. Obviously, with the help of a well-defined membership function, FSVM can effectively handle outliers and noise. Therefore, defining an appropriate fuzzy membership function becomes a critical issue for improving the generalization performance of FSVM.

In reference [10], the membership function is defined as:

$$s_i = \begin{cases} 1 - \dfrac{|x_i - o^+|}{r^+ + \delta}, y_i = +1 \\ 1 - \dfrac{|x_i - o^-|}{r^- + \delta}, y_i = -1 \end{cases}, \quad r^+ = \max_{\{x_i: y_i = 1\}} |x_i - o^+|, \quad r^- = \max_{\{x_i: y_i = -1\}} |x_i - o^-| \tag{2}$$

o^+ and o^- are the class centres of the minority instances and the majority instances, respectively.

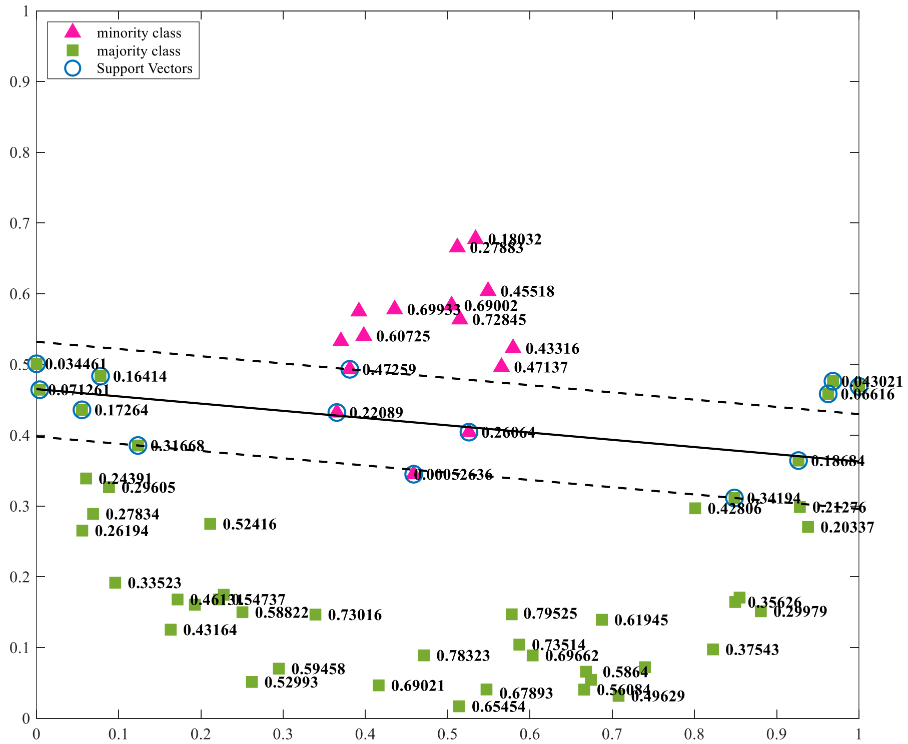

Fig. 1 – Fuzzy membership values of instances calculated by class center.

Obviously, Eq. (2) is a membership degree calculation method based on the distance between instances and class centres. As shown in Fig. 1, we have analysed the limitations of this method on an artificial dataset. When confronted with such annular-like data, the performance of the algorithm is greatly affected. This is because the support vectors are mostly distributed near the classification hyperplane and far away from the class centre. The membership degree calculation approach based on the class centre assigns very small membership degrees to the support vectors near the classification hyperplane, making it easy for the support vectors to be misjudged as noise or outliers. Furthermore, the importance of samples at the same distance from the class centre may not necessarily be the same.

3. INTUITIONISTIC FUZZY SVM BASED ON KERNEL GREY RELATIONAL ANALYSIS

Based on the above analysis, we find that when facing datasets with irregular distribution or complicated distribution, the membership degree calculation method based on Euclidean distance may be invalid. To address this issue, based on a new score function of intuitionistic fuzzy numbers and kernel grey relational analysis, this study proposes a new method to calculate the degree of membership.

3.1. Kernel grey relational analysis

The basic idea of grey relational analysis is to judge whether the connection is close based on the similarity of the geometric shape of the sequence curve. The value is only related to the geometric shape of the sequence, and has nothing to do with its spatial relative position [18]. As for the actual data, they do not necessarily have typical distribution laws, and there are often uncertain quantitative relationships between attributes. Traditional distance approaches cannot accurately measure the closeness between samples [18]. Based on the above characteristics, this paper adopts grey relational degree to replace Euclidean distance to calculate the similarity between samples, so as to reduce the limitations on the spherical distribution of samples and the dependence on the spatial relative positional relationship.

To establish the relationship between each sample and its K nearest neighbors, it is necessary to analyze the grey relational degree between the comparison sequence and the reference sequence, thereby determining the differences and connections between them. First, the data need to be dimensionless processed. $\gamma\big(\varphi(x_i),\varphi(x_k)\big)$ is used to represent the grey relational degree between the target sample sequence x_i and the kth nearest neighbor sequence x_k in the kernel space, while $\gamma[\varphi(x_i)(l),\varphi(x_k)(l)]$ represents the relationship between x_i and x_k at l [19, 20]. The equation is as follows:

$$\gamma\big(\varphi(x_i),\varphi(x_k)\big)=\frac{1}{m}\sum_{l=1}^{m}\gamma[\varphi(x_i)(l),\varphi(x_k)(l)] \tag{3}$$

Here, $\varphi(x_i)$ represents the Gaussian kernel mapping of sample x_i, $\varphi(x_i)(l)$ represents the l^{th} element of x_i in high-dimensional space, and m represents the number of sample features, $0<\gamma\big(\varphi(x_i),\varphi(x_k)\big)\le 1$.

According to the basic principles of grey relational analysis, the grey relational degree depends on the closeness of the two sequence curves. The smaller $|\varphi(x_i)(l)-\varphi(x_k)(l)|$ is, the larger $\gamma[\varphi(x_i)(l),\varphi(x_k)(l)]$ becomes [19, 20]. The equation is as follows:

$$\gamma[\varphi(x_i)(l),\varphi(x_k)(l)]=\frac{\min\limits_k\min\limits_l|\varphi(x_i)-\varphi(x_k)|+\rho\max\limits_k\max\limits_l|\varphi(x_i)-\varphi(x_k)|}{|\varphi(x_i)(l)-\varphi(x_k)(l)|+\rho\max\limits_k\max\limits_l|\varphi(x_i)-\varphi(x_k)|},\quad l=1,2,\cdots m \tag{4}$$

where ρ is the resolution factor, $\rho\in[0,1]$.

To measure the importance of the target sample x_i, the average value of the grey relational degree between the target sample and its K nearest neighbors is taken as the degree of attribution between the target sample and the category. The calculation formula is as follows:

$$\gamma_{ik}=\frac{1}{K}\sum_{k=1}^{K}\gamma\big(\varphi(x_i),\varphi(x_k)\big) \tag{5}$$

3.2. Categorical contribution of instances based on intuitionistic fuzzy numbers

Fuzzy set is the basis of constructing FSVM. However, the membership degree of elements in fuzzy set is only a real number, which can only represent one of the degrees of support (affirmation), opposition (negation) and hesitation (uncertainty) in practical applications such as decision-making, but cannot represent the degree of all the three at the same time. In this regard, Atanassov [21] proposed an intuitionistic fuzzy set based on three aspects of information: membership degree, non-membership degree, and hesitation degree, so that the intuitionistic fuzzy set can describe and characterize the essence of fuzziness in the objective world more delicately than the traditional fuzzy set.

For a binary classification problem, instances can be transformed into intuitionistic fuzzy numbers using intuitionistic fuzzy sets [22, 23]:

$$D_s = \{(x_i, y_i, \mu_i, v_i) |\ i = 1,2,\cdots,N\} \tag{6}$$

Here, μ_i represents the membership degree of x_i, while v_i represents the non-membership degree of x_i. For each given intuitionistic fuzzy number (u_i, v_i), a scoring function can be used to measure the classification contribution of each training sample.

For each given intuitionistic fuzzy number (μ_i, v_i), the score function can be used to measure the classification contribution of each training sample. we define the scoring function as [22, 23]:

$$H(x_i) = \begin{cases} 0, & u(x_i) < v(x_i) \\ \dfrac{1 - v(x_i)}{2 - u(x_i) - v(x_i)}, & \text{others} \end{cases} \tag{7}$$

where $v(x_i) = [1 - \mu(x_i)]\rho(x)$, $\rho(x)$ represents the average value of the grey relational grade between an instance and its K_b heterogeneous neighboring points. $\mu(x)$ represents the average value of the grey relational grade between an instance and its K_a isomorphic neighboring points.

Finally, the classification contribution values of the instances are defined as:

$$s_i(x_i^-) = \frac{H(x_i^-)}{IR} = \begin{cases} 0, & u(x_i^-) < v(x_i^-) \\ \dfrac{1 - v(x_i^-)}{2 - u\left(x_i^-\right) - v\left(x_i^-\right)}, & \text{others} \end{cases} \tag{8}$$

$$u(x_i^-) = \gamma_{ik}^- = \frac{1}{K_2}\sum_{k=1}^{K_2} \gamma\left(\varphi(x_i^-), \varphi(x_k^-)\right), \quad K_2 = \sqrt{N_2} \tag{9}$$

$$v(x_i^-) = (1 - u(x_i^-))\frac{1}{K_1}\sum_{k=1}^{K_1} \gamma\left(\varphi(x_i^-), \varphi(x_k^+)\right), \quad K_1 = \sqrt{N_1} \tag{10}$$

where IR is the ratio of the number of samples in the majority category to the number of samples in the minority category; x_i^- – the majority instances, x_k^- – the k^{th} isomorphic adjacent point of x_i^-, and x_k^+ – the k^{th} heterogeneous adjacent point of x_i^-; N_1 – the number of minority instances, N_2 is the number of majority instances

$$s_i(x_i^+) = u(x_i^+) = \gamma_{ik}^+ = \frac{1}{K_1}\sum_{k=1}^{K_1} \gamma\left(\varphi(x_i^+), \varphi(x_k^+)\right), \quad K_1 = \sqrt{N_1} \tag{11}$$

where x_i^+ – the minority instances, and x_k^+ – the k^{th} isomorphic adjacent point of x_i^+; N_1 – the number of minority instances.

To explore the rationality of this approach, the proposed approach is used to calculate the fuzzy membership values of samples in the artificial dataset, and the fuzzy membership values of the instances corresponding to Fig. 1 are marked out. The results are shown in Fig. 2. As shown in Fig. 2, the fuzzy membership values of the samples on the periphery of the dataset and the support vectors near the classification hyperplane are not too small as shown in Fig. 1, so they will not be misjudged as noise or outliers. Instead, it follows the distribution characteristics of instances. Therefore, the fuzzy membership function designed based on Kernel grey relational analysis and intuitionistic fuzzy numbers score function is more reasonable.

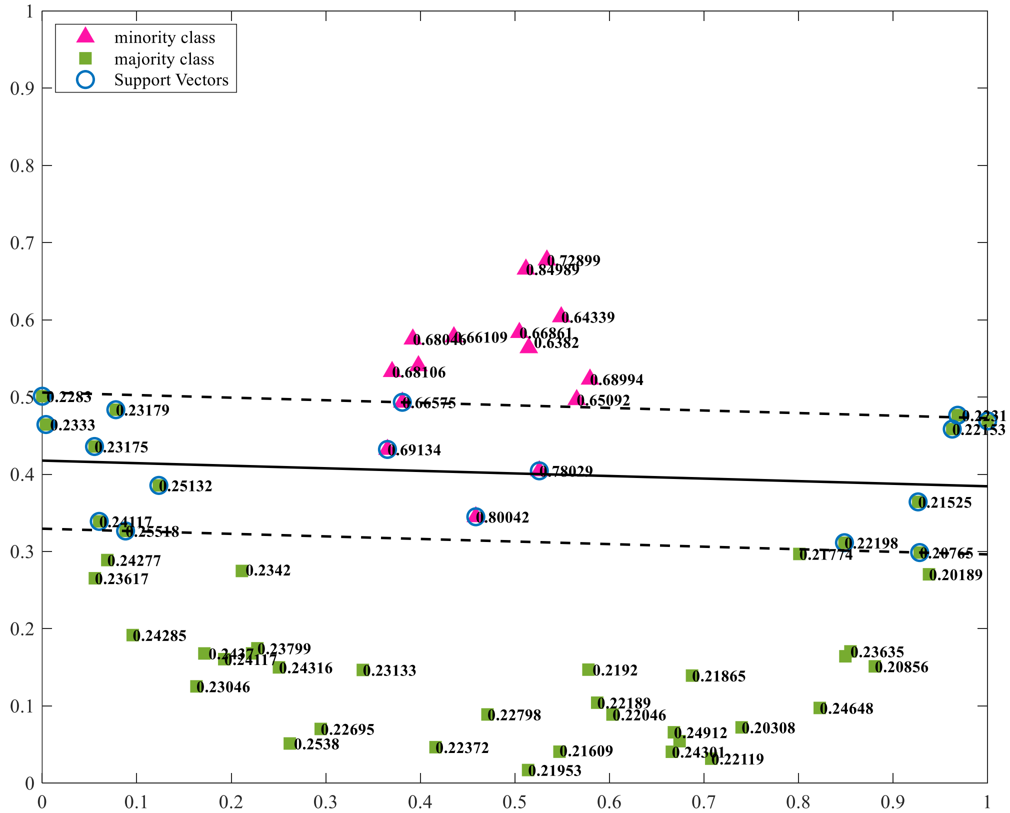

Fig. 2 – Fuzzy membership values of instances calculated by KGRA.

Integrating the fuzzy membership values $s_i(x_i^-)$ and $s_i(x_i^+)$, calculated based on intuitionistic fuzzy numbers and kernel grey relational analysis, into the objective function (1) of FSVM, a new objective function is obtained as:

$$\min_{\omega,b}\frac{1}{2}\|\omega\|^2+C\sum s_i(x_i^+)\varepsilon_i+C\sum s_i(x_i^-)\varepsilon_i \tag{12}$$

$$s.t.\ y_i(\omega^T\varphi(x_i)+b)+\varepsilon_i\geq 1,\varepsilon_i\geq 0,\forall i=1,2,\cdots,N$$

Until now, we propose an intuitionistic fuzzy support vector machine algorithm based on kernel grey relational analysis. The training process of this model is summarized as Algorithm 1.

Algorithm 1 KGRA-IFSVM algorithm

Input: Training dataset $D=\{(x_i,y_i)|\ i=1,2,\cdots,N\}$, Initializing the penalty factor C and the Gaussian kernel parameter g ($C=2^0,g=2^{-15}$).

Output: Predicted labels

Step 1: The training set is divided into D_1 and D_2, where D_1 is the set of minority instances and D_2 is the set of majority instances. The number of samples in D_1 and D_2 are represented by N_1 and N_2 respectively;

Step 2: Calculate the number of isomorphic neighbor points K_1 and the number of heterogeneous neighbor points K_2;

Step 3: For the instance $x_i^+\in D_1$, calculate $s_i(x_i^+)$ of the minority instance according to (11);

Step 4: For the instance $x_i^-\in D_2$, calculate $u(x_i^-)$ and $v(x_i^-)$ according to (9)–(10);

Step 5: Calculate $s_i(x_i^-)$ of the majority instance according to (8);

Step 6: Train the model on the training set.

Step 7: Verify the performance of KGRA-IFSVM using the test set.

---

4. VALIDATION

In this section, we utilize benchmark datasets to explore the performance and superiority of our proposed algorithm. Ten-fold cross-validation is adopted for ten times to select all parameters of the algorithm. The Gaussian kernel function $K(x_1, x_2) = \exp(-g\|x_1 - x_2\|^2)$ is used for all datasets. All tests were conducted on a PC with an AMD Ryzen 5 Microsoft Surface (R) Edition processor running at 2.10 GHz, and the algorithms were implemented on MATLAB 2023b.

4.1. Evaluation metrics

The classification performance of traditional classification models is evaluated based on accuracy. When dealing with unbalanced classification problems, it can only guarantee the classification accuracy of the majority instances while ignoring the influence of the minority class, resulting in very poor classification accuracy of the minority instances [24]. To effectively evaluate the classification performance of the model, AUC, G-mean (GM) and F1 Score (F1) are taken as the evaluation indicators of the model [25, 26]. Most of them are based on the confusion matrix, as shown in Table 1.

$$\mathrm{GM} = \sqrt{Sensitivity \times Specificity} \tag{13}$$

$$\mathrm{F1} = \frac{2 \times Sensitivity \times Precision}{Sensitivity + Precision} \tag{14}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{15}$$

$$Specificity = \frac{TN}{TN + FP} \tag{16}$$

Table 1

Confusion matrix

	Predicted class	
Actual class	Positive	Negative
Positive	TP (true positive)	FN (false negative)
Negative	FP (false positive)	TN (true negative)

The ROC curve for calculating AUC takes the false positive rate as the horizontal axis and the recall rate as the vertical axis, which is a curve used to measure the performance of the binary classification model, reflecting the trade-off between the classifier's coverage of positive examples and its coverage of negative examples. The area enclosed by the ROC curve and the X-axis can be used as a comprehensive measurement indicator, namely AUC. The closer the value of AUC is to 1, the better the overall performance of the model is [25].

4.2. Test results on 8 benchmark datasets

Considering that the contribution of KGRA-IFSVM to classification tasks may be limited to specific datasets, the scope of this study has been broadened. In this section, we selected 8 datasets from UCI repositories (https://archive.ics.uci.edu/) and obtained a series of binary classification datasets with different imbalance ratios by using different combinations of categories from the datasets to evaluate the performance of the proposed KGRA-IFSVM model in handling imbalanced classification problems. Table 2 gives a complete description of each dataset, where the minority class and majority class columns contain the complete list of class combinations for these datasets. All attributes are normalized to the interval [0,1]. 70% of the minority instances are randomly selected as the training set, while the remaining 30% are used as the test set. The same process is applied to the majority instances. Next, the classification performance of KGRA-IFSVM on different datasets in Table 2 will be verified based on the flowchart of Algorithm 1. During the test process, the model was also compared with algorithms such as FSVM [10], KLOF-FSVM [12], GRD-FSVM [13], and K-IFSVM [14], and the results are shown in Table 3. (We select the optimal combination of penalty factor C and kernel parameter g in $[2^0, 2^{15}]$ and $[2^{-15}, 2^0]$ based on the ten-fold cross-validation for ten times to ensure that the AUC value of the model in the test set is the largest, and record the Gm value and F1 value at this time. In addition to the inconsistency of the membership calculation method, the test environment of the comparative algorithms is consistent.)

Table 2

Details of the imbalanced datasets

Name	Attributes	Instances	minority instances	majority instances	IR	Minority Class	Majority Class
E. coli	7	272	52	220	4.23	class3	Class1,2
Vertebral-column	6	310	100	210	2.10	Class2	Class1
Vehicle	18	416	199	217	1.09	Class1	Class2
Yeast3	8	1484	49	1435	29.29	Class6,10	others
mammographic	5	961	445	516	1.16	Class2	Class1
Glass	9	214	29	185	6.38	Class6	others
Cardiotocography-1	21	584	252	332	1.32	Class7	Class6
Cardiotocography-2	21	854	275	579	2.11	Class3,4,5,9	Class2

Table 3

Classification results obtained by six algorithms on eight datasets, where the best values are in bold

Measure	AUC	Gm	F1	AUC	Gm	F1
	Dataset: E. coli			Dataset: Vertebral-column		
FSVM	0.9666±0.0037	0.9394±0.0052	0.9079±0.0092	0.9296±0.0026	0.8263±0.0086	0.7699±0.0116
KLOF-FSVM	0.9700±0.0041	0.9314±0.0064	0.8969±0.0110	0.9354±0.0023	0.8339±0.0052	0.7716±0.0059
GRD-FSVM	0.9698±0.0039	0.9324±0.0075	0.8980±0.0128	0.9349±0.0019	0.8302±0.0072	0.7684±0.0087
K-IFSVM	0.9678±0.0027	0.9404±0.0041	0.9123±0.0063	0.9326±0.0020	0.8303±0.0027	0.7775±0.0038
KGRA-IFSVM	**0.9764±0.0042**	**0.9457±0.0055**	**0.9126±0.0082**	**0.9374±0.0042**	**0.8561±0.0051**	**0.7896±0.0065**
	Dataset: Vehicle			Dataset: Yeast3		
FSVM	0.9994±0.0002	0.9785±0.0035	0.9773±0.0036	0.9866±0.0024	0.8169±0.0168	0.6678±0.0164
KLOF-FSVM	**0.9995±0.0001**	0.9877±0.0015	0.9870±0.0016	0.9872±0.0005	0.6161±0.0111	0.4972±0.0133
GRD-FSVM	0.9995±0.0008	0.9887±0.0015	0.9880±0.0016	0.9867±0.0004	0.5772±0.0171	0.4618±0.0240
K-IFSVM	0.9974±0.0003	0.9641±0.0020	0.9630±0.0019	0.9859±0.0012	0.7543±0.0164	0.6063±0.0174
KGRA-IFSVM	**0.9995±0.0001**	**0.9887±0.0010**	**0.9880±0.0010**	**0.9881±0.0020**	**0.9213±0.0105**	**0.7344±0.0153**
	Dataset: mammographic			Dataset: Glass		
FSVM	0.8844±0.0019	0.8126±0.0020	0.8050±0.0023	0.9771±0.0077	0.9088±0.0245	0.8904±0.0269
KLOF-FSVM	0.8904±0.0012	0.8169±0.0030	0.8083±0.0032	**0.9850±0.0078**	**0.9260±0.0117**	**0.8905±0.0163**
GRD-FSVM	0.8892±0.0015	0.8170±0.0029	0.8059±0.0030	**0.9850±0.0078**	**0.9260±0.0117**	**0.8905±0.0163**
K-IFSVM	0.8822±0.0010	0.8125±0.0021	0.8046±0.0020	0.9743±0.0077	0.9162±0.0089	0.8673±0.0190
KGRA-IFSVM	**0.8905±0.0008**	**0.8249±0.0015**	**0.8143±0.0016**	**0.9850±0.0078**	**0.9260±0.0117**	**0.8905±0.0163**
	Dataset: Cardiotocography-1			Dataset: Cardiotocography-2		
FSVM	0.9930±0.0007	0.9655±0.0024	0.9612±0.0026	0.9892±0.0012	0.9490±0.0053	0.9306±0.0065
KLOF-FSVM	0.9936±0.0007	0.9589±0.0027	0.9540±0.0029	0.9895±0.0012	0.9501±0.0042	0.9317±0.0044
GRD-FSVM	0.9934±0.0006	0.9607±0.0038	0.9561±0.0043	0.9894±0.0013	0.9498±0.0042	0.9325±0.0054
K-IFSVM	0.9927±0.0004	0.9650±0.0025	0.9607±0.0029	0.9882±0.0010	0.9486±0.0051	0.9295±0.0060
KGRA-IFSVM	**0.9937±0.0007**	**0.9676±0.0019**	**0.9633±0.0020**	**0.9907±0.0011**	**0.9560±0.0040**	**0.9362±0.0047**

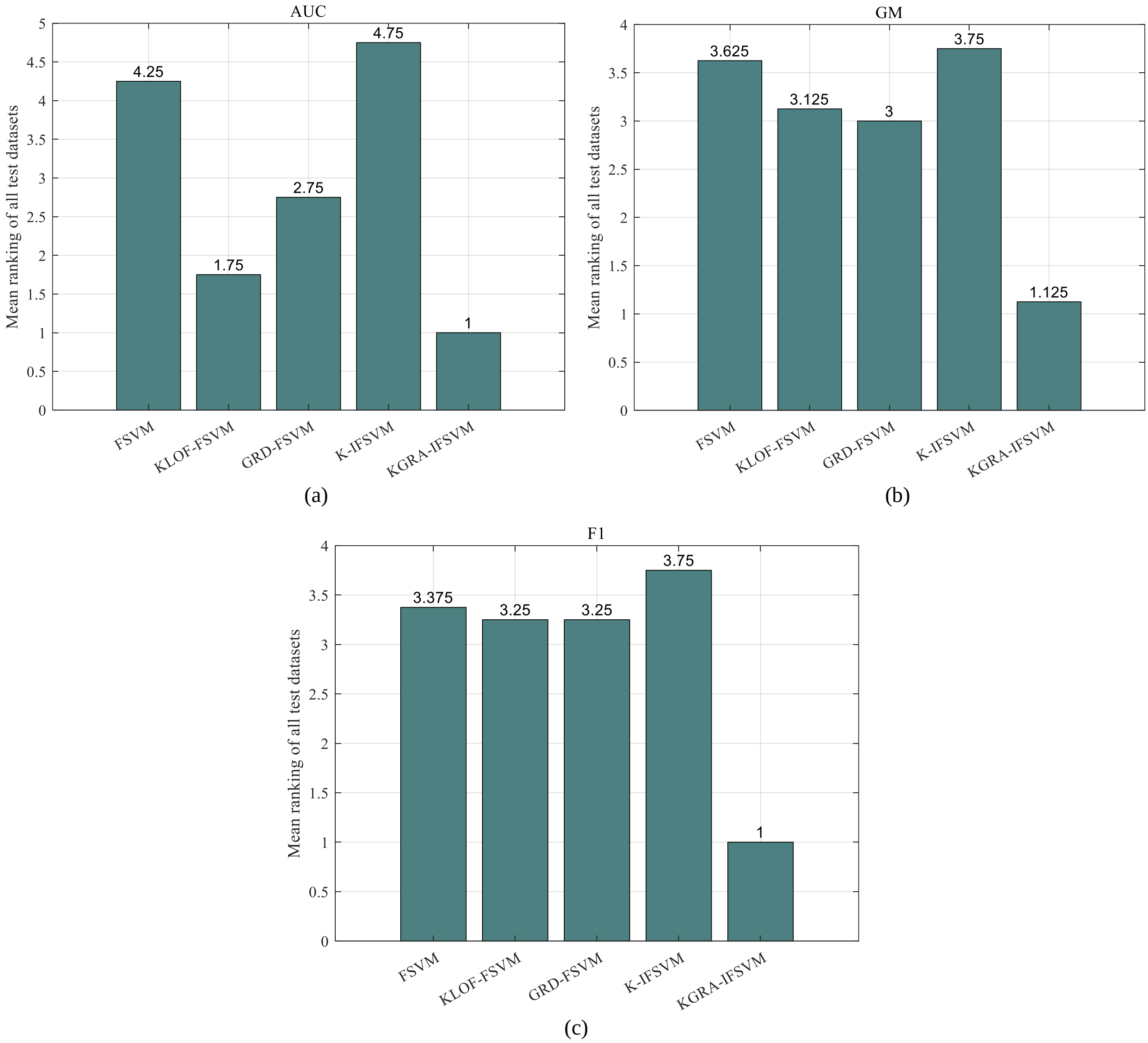

Fig. 3 –Average ranking of the model's classification performance.

As can be seen from the final test results in Table 3, the algorithm proposed in this paper has achieved the best classification results under the evaluation metrics of AUC, GM, and F1. Since the new algorithm KGRA-IFSVM has shown good classification performance on datasets with different distribution characteristics, it is more likely to capture the general rules behind the data compared with other algorithms, rather than just overfitting a specific dataset, which increases the credibility of the model's generalization ability. In addition, through multiple cross-validations, the model has been fully trained and evaluated on different partitions of training and test sets. The robustness of this method further reduces the risks of random errors and overfitting. Fig. 3 shows the average rank values of different classification algorithms on all datasets. It can be seen from Fig. 3a that the average rankings of AUC values of FSVM and K-IFSVM on all datasets are the lowest, so their comprehensive classification performance is the weakest, which is because the fuzzy membership degrees of these three approaches are calculated based on the distance between instances and class centers. However, the presence of outliers or noise can cause the class centers to shift, thereby affecting the assessment of the fuzzy membership degrees of samples. Compared to FSVM and K-IFSVM, KLOF-FSVM exhibit superior comprehensive classification performance because they consider the compactness around each instance when calculating the fuzzy membership degree, thus enabling them to obtain more accurate instance distribution information. The fuzzy membership degree of GRD-FSVM is also calculated based on the class center, but instead of computing the Euclidean distance between an instance and the class center, it is measured according to the similarity of geometric shapes between the target sample sequence and the class center sequence. Although it is still sensitive to noise and outliers, it is more adaptable to various datasets with irregular distribution characteristics. As can be seen from Fig. 3a, the comprehensive classification performance of GRD-FSVM is second only to the new algorithm and is rarely affected by possible outliers. This may be because there are very few noises or outliers in these datasets themselves. Later, we will verify the noise immunity of these algorithms by adding different levels of Gaussian white noise to the datasets. The higher the F1 value is, the higher the classification accuracy of the model for minority classes will be. As can be seen from Figure 3(c), the classification performance of KLOF-FSVM and GRD-FSVM for minority instances is second only to the new algorithm, because KLOF-FSVM sets the membership degree of minority instances to 1, while GRD-FSVM is not affected by the distribution characteristics of the dataset. The average ranking of F1 values of K-IFSVM on all datasets is the lowest, which indicates that the classification accuracy of intuitionistic fuzzy support vector machine on minority instances is lower than that of FSVM. The KGRA-IFSVM algorithm combines the advantages of GRD-FSVM and K-IFSVM. It can describe and characterize the fuzzy nature of classification problems more precisely through intuitionistic fuzzy sets, and can also enable the model to be applied to datasets with different distribution characteristics through the grey relational analysis method. Theoretically speaking, the complexity of the algorithm in this paper is equivalent to that of FSVM, KLOF-FSVM, GRD-FSVM, and K-IFSVM, which is $O(l^2)$, where l is the number of samples. However, KGRA-IFSVM requires some extra computation time when calculating the classification contribution value s_i of the samples. FSVM, GRD-FSVM, and K-IFSVM only need to analyze the relevant information between each sample and the class center when calculating s_i, while KGRA-IFSVM needs to additionally analyze the correlation between each sample and its K_a isomorphic nearest neighbors as well as K_b heterogeneous nearest neighbors. Obviously, KGRA-IFSVM takes longer time to train on a larger training set.

In addition, following the approach in reference [27], this study selected to add Gaussian white noise ranging from – 6 dB to 8 dB in the “E. coli” dataset, further analyzing the noise resistance of KGRA-IFSVM. The test results are presented in Table 4, and a visual analysis is conducted as shown in Fig. 4. The noise resistance of GRD-FSVM is significantly better than FSVM, which also indicates that replacing the Euclidean distance between samples with grey relational degree can overcome the influence of noise to a certain extent. When $-2 \leq \text{SNR} \leq 4$, the noise resistance of K-IFSVM is superior to FSVM and GRD-FSVM, which confirms the effectiveness of introducing intuitionistic fuzzy sets. KGRA-IFSVM exhibits the optimal noise resistance, which benefits from its combination of the advantages of GRD-FSVM and K-IFSVM.

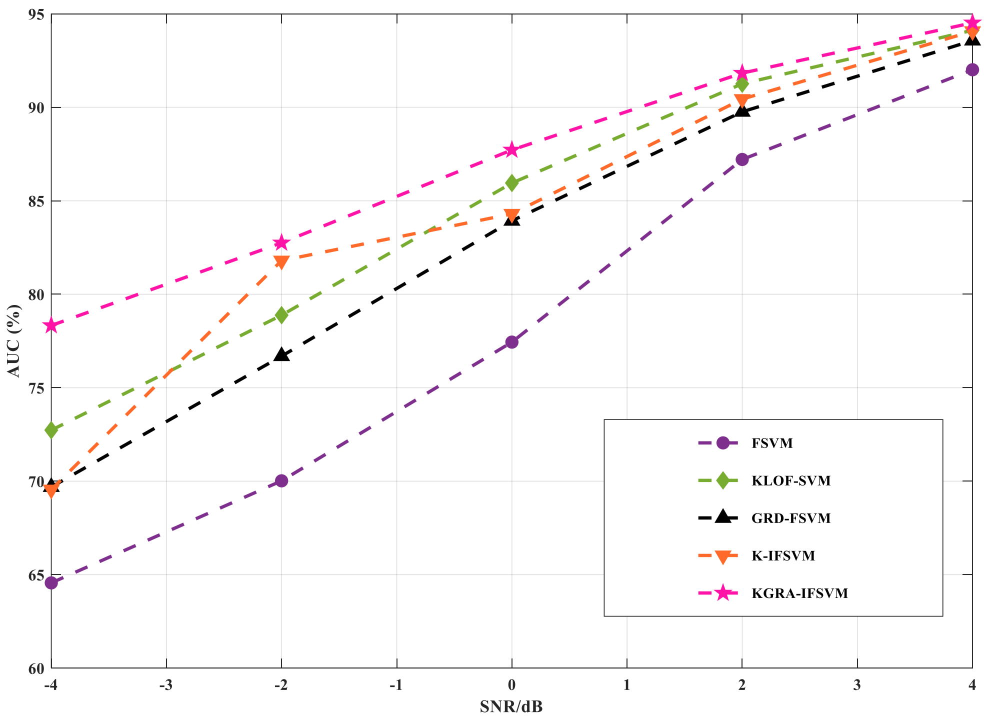

Fig. 4 – Comparison of performance under different noise conditions.

Table 4

Comparison of performance under different noise conditions

	-4dB	-2dB	0dB	2dB	4dB
FSVM	0.6456	0.7002	0.7744	0.8721	0.9201
KLOF-FSVM	0.7273	0.7888	0.8595	0.9126	0.9412
GRD-FSVM	0.6969	0.7669	0.8393	0.8976	0.9357
K-IFSVM	0.6955	0.8180	0.8430	0.9044	0.9407
KGRA-IFSVM	0.7833	0.8275	0.8772	0.9183	0.9452

5. CONCLUSION

This study proposes an intuitionistic fuzzy support vector machine algorithm based on kernel grey relational analysis. Firstly, the limitations of the fuzzy membership degree calculation approach based on the distance between instances and class centers are analyzed on artificial datasets. It is found that the class center may shift due to the presence of outliers, which leads to inaccurate estimation of sample weights. To address this issue, the study calculates the membership degree of instances based on the grey relational degree between the instances and their K nearest neighbors. To obtain more accurate sample distribution information and the relationship between classes, the intuitionistic fuzzy set is further introduced into the calculation of sample weights. By calculating the non-membership degree, we can obtain heterogeneous information of instances. Test results on 8 UCI datasets show that the proposed algorithm has superior classification performance.

REFERENCES

[1] Kilic U, Essiz ES, Keles MK. Binary anarchic society optimization for feature selection. Romanian Journal of Information Science and Technology. 2023; 26: 351–364.

[2] Zhu LT, Ouyang B, Zhang XB, Luo ZH. Progress on application of machine learning to multiphase reactors. Chem. Eng. Prog. 2021; 40: 1699–1714.

[3] Yang F, Wang P, Wang Y, Wang W. LVS: An efficient semantic window query algorithm on time series. Journal of Computer Research and Development. 2015; 52: 19–27.

[4] Su YL, Li WJ, Liu XY, Ding JP, Liu SX, Li HN, Li GF. A review of text classification methods based on graph neural networks. Computer Engineering and Applications. to appear.

[5] Verma RN, Deo R, Srivastava R, Subbarao N, Singh GP. A new fuzzy support vector machine with pinball loss. Discover Artificial Intelligence. 2023; 3: 14.

[6] Huang X, Zhang S, Lin C, Xia J. Quantum fuzzy support vector machine for binary classification. Comput. Syst. Sci. Eng. 2023; 45: 2783–2794.

[7] Wang Z, Xu T, Wang L. Teaching evaluation method based on fuzzy support vector machine algorithm. Mobile Information Systems. 2022: 7354972.

[8] Wang QL, Zheng J, Zhang WJ. Dense fuzzy support vector machine to binary classification for imbalanced data. Journal of Intelligent & Fuzzy Systems. 2023; 45: 9643–9653.

[9] Dhanasekaran Y, Murugesan P. Improved bias value and new membership function to enhance the performance of fuzzy Support Vector Machine. Expert Systems with Applications. 2022; 208: 118003.

[10] Lin CF, Wang SD, Fuzzy support vector machines. IEEE Transactions on Neural Networks. 2002; 13: 464–471.

[11] Jiang XF, Zhang Y, Jian CL. Fuzzy SVM with a new fuzzy membership function. Neural Computing & Applications. 2006; 15: 268–276.

[12] Wang K, An J, Yu Z, Yin X, Ma C. Kernel local outlier factor-based fuzzy support vector machine for imbalanced classification. Concurrency and Computation: Practice and Experience. 2021; 33: e6235.

[13] Tang WM, Zhang Y, Wang P. Fuzzy SVM with a new fuzzy membership function based on gray relational grade. In: Proceedings of the International Conference on Information Engineering and Applications. 2012; pp. 561–569.

[14] Ha M, Wang C, Chen J. The support vector machine based on intuitionistic fuzzy number and kernel function. Soft Computing. 2013; 17: 635–641.

[15] Li WW, Chen X, Zhang HW, Huang ZQ, Jia XY. Heterogeneous defect prediction based on simultaneous semantic alignment. Journal of Software. 2023; 34: 2669–2689.

[16] Wang KF, An J, Wei Z, Cui C, Ma XH, Ma C, Bao HQ. Deep learning-based imbalanced classification with fuzzy support vector machine. Frontiers in Bioengineering and Biotechnology. 2022; 9: 802712.

[17] Kchouri M, Harum N, Obeid A, Hazimeh H. Fuzzy support vector machine based fall detection method for traumatic brain injuries. Int. J. Adv. Comp. Sci. Appl. 2022; 13: 302–314.

[18] Yi B, Zhu J, Zhang S. Fuzzy support vector machine improvement method based on grey correlation degree of double reference point. Computer Engineering and Science. 2016; 38: 1244–1251.
[19] Ba Y, Liu X, Yin Z, Li X, Yin L, Zheng W. Effect of urbanization on aerosol optical depth over Beijing: Land use and surface temperature analysis. Urban Climate. 2023; 51: 101655.
[20] Zhu X, Dong X, Wang S, Zhang. Identification of airport similar days based on K-prototype cluster and grey correlation analysis. Journal of Physics: Conference Series. 2023; 2491: 012003.
[21] Atanassov KT. Intuitionistic fuzzy sets. Fuzzy Sets Syst. 1986; 20: 87–96.
[22] Rezvani S, Wang X, Pourpanah F. Intuitionistic fuzzy twin support vector machines. IEEE Transactions on Fuzzy Systems. 2019; 27: 2140–2151.
[23] Laxmi S, Gupta SK, Kumar S. Intuitionistic fuzzy least square twin support vector machines for pattern classification. Annals of Operations Research. 2024; 339(3): 1329–1378.
[24] Tao X, Li Q, Ren C, Guo W, He Q, Liu R, Zou J. Affinity and class probability-based fuzzy support vector machine for imbalanced data sets. Neural Networks. 2020; 122: 289–307.
[25] Bammou Y, Benzougagh B, Abdessalam O, Brahim I, Kader S, Spalevic V, Sestras P, Ercişli S. Machine learning models for gully erosion susceptibility assessment in the Tensift catchment, Haouz plain, Morocco for sustainable development. Journal of African Earth Sciences. 2024; 213: 105229.
[26] Günen MA. Comparison of histogram-curve fitting-based and global threshold methods for cloud detection. International Journal of Environmental Science and Technology. 2024; 21: 5823–5848.
[27] Zou F, Zhang H, Sang S, Li X, He W, Liu X, Chen Y. An anti-noise one-dimension convolutional neural network learning model applying on bearing fault diagnosis. Measurement. 2021; 186: 110236.

Received July 2, 2024