Heliyon 9 (2023) e20085

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Identification effect of least square fitting method in archives management

Caichang Ding [a,b], Hui Liang [c], Na Lin [c], Zenggang Xiong [a,b], Zhimin Li [a,b], Peilong Xu [d,e,*]

[a] School of Computer and Information Science, Hubei Engineering University, Xiaogan, 432000, China
[b] Institute for AI Industrial Technology Research, Hubei Engineering University, Xiaogan, 432000, China
[c] Binzhou Polytechnic, Binzhou, 256600, China
[d] State Key Laboratory of Bio-Fibers and Eco Textiles, Qingdao University, Qingdao, 266071, China
[e] Department of Computer Engineering, Pukyong National University, 45, Yongso-ro, Nam-Gu, Busan 48513, Korea

A R T I C L E I N F O

Keywords:
Least squares fitting method
Archives management
Wavelet kernel function
Cholesky algorithm
Classification

A B S T R A C T

Archives management plays an important role in the current information age. Solving the problem of identifying and classifying archives is essential for promoting the development of archives management. The Least Squares Support Vector Machine (LS-SVM) is obtained by introducing the least squares fitting method into SVM, which is good at solving nonlinear classification. A new wavelet function is used to improve the classifier. At the same time, the cross-validation method is used to optimize the kernel parameters. Finally, the fuzzy theory and LS-SVM are combined to obtain Fuzzy Least Squares Support Vector Machines (FLS-SVM). This FLS-SVM classifier can use the distance between the data points and the classification hyperplane to classify the data in the non-separable region. The performance of FLS-SVM is verified by simulation experiments. The experimental results show that the classification accuracy of FLS-SVM classifier in archive data sets is 98.7%, and the loss rate is only 0.26%. When the wavelet function is used as the kernel function, the average accuracy of the classifier reaches 98.38%. Experiments show that the proposed method has good classification performance. It verifies the feasibility and effectiveness of the least squares fitting method in file management identification and classification.

1. Introduction

In the current information age, massive amounts of archive information are constantly being generated, and the difficulty of managing them is increasing. So its classification is an urgent problem to be solved. According to the differences in the categories of archive information, different classification methods can be used to accurately classify them. However, files with multidimensional attributes are difficult to achieve correct classification in the current classifier [1]. Kalra et al. used the method of weight inclusion to enable the generated domain vocabulary set to perform the document classification task. The vocabulary set consisted only of important words with high weight. The value domain was constructed by the highest term of words in the document and the inverse document frequency. Finally, the sum of the associated weights was used to classify the documents [2]. This research correctly

* Corresponding author.
E-mail addresses: 202356105@pukyong.ac.kr, xpl@qdu.edu.cn (P. Xu).

https://doi.org/10.1016/j.heliyon.2023.e20085
Received 3 April 2023; Received in revised form 11 September 2023; Accepted 11 September 2023
Available online 12 September 2023

classified the files of multidimensional attributes. It applied the least squares fitting method to support vector machines (SVM) to obtain least squares support vector machines (LS-SVM) [3]. This kind of classifier has strong adaptability and fitting ability, and has certain advantages in solving nonlinear problems. But it also has obvious improvement space in classification accuracy and running time [4]. Mishra et al. used different wavelet changes to extract features and select coefficients of brain tumors. They classified these features and coefficients by SVM classifier. The results showed that the features extracted by different wavelet transforms are helpful to improve the classification accuracy of SVM classifier [5]. In addition, this study also optimized the kernel function and constructed a new wavelet kernel function. Fuzzy theory and LS-SVM are combined to obtain a fuzzy least squares support vector machine (FLS-SVM). Through FLS-SVM, the classification problem of unclassifiable regions is successfully solved. This method further improves the classification accuracy of the classifier and shortens the classification time of files. Unclassifiable regions refer to sample sets where different samples have the same label or where one sample has multiple labels. The purpose of this research is to further improve the classification accuracy and classification time of the classifier. It also aims to solve the classification problem of the non-separable regions of the classifier. There are three main objectives in this study:

(1) The first objective is to apply and optimize the least squares fitting method in SVM. The study first enhances the resolution ability of the original SVM through the least squares method to solve nonlinear classification problems. The least squares method, as a mathematical optimization technique, energetically evaluates and minimizes the deviation between model predictions and actual observation results. Using this method, the study aims to improve the accuracy of archive classification.
(2) The second goal is to develop a new type of wavelet kernel function, coupled with a wavelet kernel function, that will further optimize LS-SVM. As a mathematical tool, wavelet functions can perform information analysis in both frequency and time domains simultaneously, and new wavelet kernel functions can improve the performance of classifiers.
(3) The third goal is to optimize the computational performance of SVM, while improving model accuracy and reducing the computational time required. In summary, the research goal is to propose an algorithm that can more accurately and quickly classify complex multidimensional archives. This goal plays an extremely important role in managing and processing large amounts of archival data in the modern society of information explosion.

2. Related work

In the environment of increasing archival information, many scholars have done more research on the issue of archival classification. Gayathri and other researchers proposed a classification method based on the concept of ontology. This method used the terms of understanding concepts to search the relevant contents of documents. It also improved the classification accuracy through the description of domain ontology and semantic documents. Relevant experiments proved that the classification accuracy of this method is superior to other existing supervising machine algorithms [6]. Chen's team proposed a tagged instance framework to solve the problem of target classification. This framwork integrated two clustering algorithms and selection methods, and then constructed five classifiers. Through comparative experiments, the classification accuracy of the classifier was verified. The training data set for classification was constructed under the cost constraint [7]. Seifollahi et al. constructed a two-stage topic modeling algorithm. The algorithm used word embedding and co-occurrence to determine the distribution of topic words. In the second stage, the distribution of file topics was determined by extracting topic words. Through the experiment of file classification, the results showed that this method has high classification accuracy and solves the problem that the classification model ignores the association between words [8]. Azarbonyad et al. combined different information sources to classify archive categories. This method reflected the similarity between categories and archives by combining different information sources. It selected classification labels of documents by sorting learning. Experiments have proved that this method has a classification accuracy of 96% in classification. Under the condition of poor classifier performance, co-occurrence information is of great help for classification [9]. Wijaya constructed a data set of words and phrases and trained it with naive Bayes methods. It took a combination of words and phrases to train Malay and Indonesian, respectively. Through classification experiments, the results showed that this method improved the classification accuracy of the two languages and accurately distinguished the Malay and Indonesian documents [10].

SVM has also received relatively mature research on nonlinear classification issues and has been applied to various nonlinear fields. Solikin and other scholars utilized SVM for the classification of seabed sediments. They completed the sediment classification by comparing the acoustic characteristics with physical characteristics of the sediment. Through the classification experiment, the overall accuracy of SVM is 80.25% and the Kber coefficient is 0.2031. It has a qualified detection rate in the classification and detection of submarine sediments [11]. Zheng and other researchers constructed a detection model for schizophrenia patients based on SVM classification algorithm combined with multimodal MRI detection method. The method was combined with the existing test cases to design relevant comparative experiments. The study revealed that the detection of schizophrenia in patients had specific effects that distinguished their brain from that of healthy individuals, addressing the limitation of MRI in overcoming the brain environment [12]. Cervantes's team introduced the SVM algorithm in practical application, and summarized the future challenges and development trend of SVM. SVM should be combined with other methods in areas with poor performance to improve the classification ability of SVM and optimize the parameters of the classifier [13]. Praveena et al. proposed a detection and classification model of EEG signals based on mixed features and SVM. The Renyi entropy and Teager energy operator were utilized to extract composite features from EEG signals in the model, followed by principal component analysis to reduce the dimension. Lastly, SVM was used to classify EEG signals for epilepsy. Experimental results showed that this method improved the classification accuracy of epileptic seizures by 7%–30% [14]. Zhang et al. used texture features to train the optimized SVM on PET images to classify malignant lung nodules. In this study, a thresholding technique was employed to segment the volume of isolated lung nodules present in PET images. Texture analysis software then extracted the texture features of these nodules. Through comparative experiments, the classification effect of this method was improved by 20% compared with the MTV method. The accuracy of lung nodule classification was improved [15].

Zhang et al. proposed a sparse least squares classifier method based on alternating minimization to solve the above problem. Based on the reconstructed row and column kernel matrices, sparse induced zero norm approximation functions are introduced into the LS-SVM model. By solving two unconstrained quadratic optimization problems or two linear system of equations alternatively, AMSLC is able to predict the category labels of a given instance. Additionally, it can extract the minimum number of crucial instances and features to achieve an interpretable classification. The experimental results indicate that AMSLC methods typically achieve the best prediction accuracy and interpretable classification with the least number of important instances and features [16]. Zhang et al. proposed an optimal hybrid prediction framework to ensure accurate prediction of long-term carbon price sequences. It utilizes Empirical Mode Decomposition (EMD) to decompose the raw data and develops a new optimal combination to predict various decomposition parts. This combination involves the integration of ARMA type models and the old cross validation and PSO algorithm of LS-SVM. The effectiveness of the optimal EMD was demonstrated using the RMA test. Then the SSVMs framework was validated using empirical examples from the European Union Emissions Trading System (EU ETS) [17]. Chen et al. proposed a solubility prediction model based on adaptive PSO algorithm and LS-SVM. Unlike traditional PSO algorithms, APSO algorithm improves the problem of easily falling into local optima. LS-SVM is optimized through the regularization parameters and kernel function adjustment parameters of APSO. The CO_2 solubility of 8 polymers in this model was predicted at a wide range of temperatures and pressures. The mean absolute relative deviation (MARD), Root-mean-square error (RMSE), and Coefficient of determination were 0.2130, 0.0120, and 0.9853, respectively. In addition, the model has good scalability and can be applied to other fields such as chemistry and medicine [18].

Based on the above literature analysis, scholars have conducted continuous research on archive classification to improve its classification accuracy. In practical classification, there is still room for improvement in the classification accuracy of SVM classifiers. It is pointed out that combining SVM classifiers with other methods can improve their classification accuracy. In addition, the LS-SVM proposed above transforms the solution problem into a system of linear equations. It avoids solving Lagrange multipliers and computing kernel functions, and improves the training speed and efficiency of the model. However, this method has lower processing efficiency in data with higher ambiguity or uncertainty. Therefore, the study combines the least squares fitting method in numerical methods with SVM, and improves the kernel function to obtain a FLS-SVM. The paper aims to solve the classification problem in archival management and provide a new way for the development of archival management.

3. Construction of least square linear system in traditional SVM

3.1. Construction of least square support vector machine

LS-SVM can be obtained by introducing the least square linear system into the standard SVM [19–22]. SVM utilizes the inner product function to apply a nonlinear transformation on the input space, thus mapping it to a feature space with high-dimensionality. This allows for the establishment of a linear relationship between the input and output variables. The specific steps are shown in Fig. 1.

Determining the parameters of the mapping function is a key part in SVM. The kernel function maps a low dimensional space to a

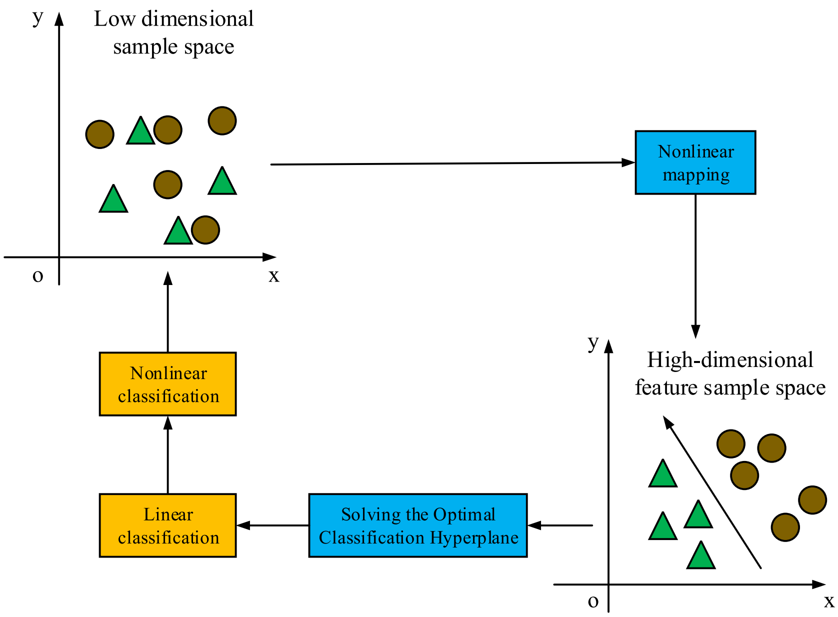


Fig. 1. Core idea of SVM.

high dimensional space, and then solves a nonlinear problem. i.e. a nonlinear problem is transformed into a linear problem in a high dimensional space. The definition of kernel function is expressed by Equation (1).

$$\begin{cases} \varphi : \begin{cases} R^n \rightarrow H \\ x \rightarrow X = \varphi(x) \end{cases} \\ K(x, x') = \varphi(x) \cdot \varphi(x^1) \end{cases} \tag{1}$$

in Equation (1), $\varphi(x) \cdot \varphi(x^1)$ represents the inner product of the high-dimensional feature space. H is the transformed high-dimensional feature space. R^n represents the original low dimensional space. $K(x, x^1)$ is the kernel. Kernel function has several common forms, such as polynomial kernel function and multilayer perceptron kernel function. The most commonly used kernel function is the radial basis kernel function, whose expression is equation (2).

$$K(x, x') = \exp\left\{ - |x - x'|^2 \Big/ \sigma^2 \right\} \tag{2}$$

in Equation (2), σ represents the width of the Gaussian kernel. Under the condition of Mercer's theorem, a new kind of wavelet kernel function is constructed by using the wavelet analysis technique. The radial basis function is extended to construct the multidimensional tensor product wavelet kernel. The constructed inner product form wavelet kernel function and tensor product wavelet kernel function are listed in equation (3).

$$\begin{cases} K(x, z) = \prod_{i=1}^{n} h((x_i - m_i)/\sigma) h((z_i - d_i)/\sigma) \\ K(x, z) = K(x - z) = \prod_{i=1}^{n} h((x_i - z_i)/\sigma) \end{cases} \tag{3}$$

$h(x)$ in Equation (3) refers to the mother wavelet function. m and d represent different translation factors. Both of the constructed kernel functions allow multidimensional support. Therefore, the Mexican hat mother wavelet is selected to construct a translation invariant wavelet kernel function. Its expression is shown in equation (4).

$$K(x, z) = \prod_{i=1}^{n} \left(\left(n - (x_i - z_i)^2 \Big/ \sigma^2\right) \exp\left(- \left((x_i - z_i)^2 \Big/ \sigma^2\right)\right)\right) \tag{4}$$

n in Equation (4) is the dimension of space points. σ of the wavelet kernel function can be selected by Cross Validation (CV), and the value of σ has a direct impact on the performance of the classifier [23–26]. The CV method divides a certain training sample into k disjoint subsets. It can be represented by k-fold. The size of each subset is basically the same, and it is trained and tested to get the training set. The algorithm determines σ according to the training set. The value of k can be 5 or 10 or the size of the whole training set. Among the two classification problems, SVM can be solved in the form of structural risk minimization. Its solution equation is Equation (5).

$$\begin{cases} \min \frac{1}{2} \|w\|^2 \\ s.t. \quad y_i\left(w^T x_i + b\right) \geq 1 \end{cases} \tag{5}$$

in Equation (5), w represents the intercept. b indicates the classification threshold. x_i indicates the input data. y_i is the ID of the output category. By introducing Lagrangian function, the dual problem of Equation (5) is derived and the original problem is transformed. The dual problem of the original problem is obtained by introducing the Lagrange function. The expression is shown in Equation (6).

$$\begin{cases} \min 0.5 \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j \left(x_i \cdot x_j\right) - \sum_{j=1}^{N} \alpha_j \\ s.t. \quad \sum_{i=1}^{N} y_i \alpha_i = 0 \end{cases} \tag{6}$$

in Equation (6), α is the Lagrange multiplier transpose matrix. α_i is the Lagrange multiplier. The solution of Equation (6) is the global optimal solution of the original problem. α^* was solved by the optimization algorithm to get $\alpha^* = (\alpha_1^*, \alpha_2^*, \cdots, \alpha_n^*)$. b^* was calculated according to the Karush-Kuhn-Tucker(KKT) condition. The optimal hyperplane of SVM is Equation (7).

$$\begin{cases} b^* = y_i - \sum_{i=1}^{N} \alpha_i^* y_i (x_i \cdot y_i) \\ y(x) = \operatorname{sgn}\left[\sum_{i=1}^{N} \alpha_i^* y_i (x_i \cdot y_i) + b^* \right] \end{cases} \tag{7}$$

in Equation (7), sgn represents the sign function. The above mentioned linearity is separable, but in actual situations, the indivisible linearity often occupies an important proportion. If the above methods are adopted, there will be errors in the classification of sample data. Therefore, relaxation variable ξ_1, whose value range is not less than 0, is introduced in this study. The constraint of relaxation variable is introduced. The expression is listed in Equation (8).

$$1-\xi_i \le y_i\left(w^T x_i + b\right) \tag{8}$$

Since the relaxation variable introduced is no less than 0, its value range can reflect the partitioning of the training set. When the value range of ξ_i is in the range (0,1), it means that the sample classification is correct. When the value range of ξ_i is not within this range, the sample classification is wrong. The relaxation variable introduced aims to maximize the sample interval and minimize the probability of misclassification. To merge two samples into one, a penalty factor should be introduced to represent the weight of the two samples. The function that introduces the penalty factor is equation (9).

$$\min J(w,\xi) = C\sum_{i=1}^{N}\xi_i + 0.5\|w\|^2 \tag{9}$$

in Equation (9), the penalty factor is represented by C. By introducing kernel $K(x,x^{'}) = \varphi(x)\cdot\varphi(x^1)$ and nonlinear transformation into the function, the C-SVM classification algorithm can be obtained, as shown in Equation (10).

$$\begin{cases} \min\limits_{\alpha} 0.5\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j K\left(x_i, x_j\right) - \sum_{j=1}^{N}\alpha_j \\ \qquad s.t.\begin{cases} \sum_{i=1}^{N} y_i\alpha_i \\ C \ge \alpha_i \ge 0 \end{cases} \end{cases} \tag{10}$$

The optimal solution α^* is $(\alpha_1^*, \alpha_2^*, \cdots, \alpha_n^*)^T$. One of the positive components is selected, and the value of the positive component is less than the penalty factor. Finally, b^* is calculated. Equation (11) is the expression of the decision function.

$$\begin{cases} b^* = y_i - \sum_{i=1}^{N} y_i\alpha_i^* K\left(x_i, x_j\right) \\ y(x) = \operatorname{sgn}\left[\sum_{i=1}^{N}\alpha_i^* y_i (x_i\cdot y_i) + b^*\right] \end{cases} \tag{11}$$

By introducing the sum of squares of the errors into the objective function of the SVM, the inequality constraint becomes an equality constraint. The solution becomes a solution of a set of linear equations. Then the classification of the LS-SVM algorithm in the original space can be represented by equation (12).

$$\begin{cases} \min J(w,\xi) = 0.5C\sum_{i=1}^{N}\xi_i^2 + 0.5\|w\|^2 \\ s.t.\left\{y_i\left[w^T\varphi(x_i) + b\right] = 1-\xi_i\right. \end{cases} \tag{12}$$

in Equation (12), $\varphi(x_i)$ represents the mapping function. N refers to the number of data samples. When constructing Lagrange functions, if w and ξ are ignored, a system of linear equations can be obtained, whose expression is Equation (13).

$$\begin{pmatrix} 0 & Y^T \\ Y & \Omega + C^{-1}I \end{pmatrix}\begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ \vec{1} \end{pmatrix} \tag{13}$$

I in Equation (13) is the identity matrix. $\vec{1}$ represents the transformation result of the identity matrix. Y is the transpose matrix of the output category object. Ω is the kernel matrix. The classification problem of the original low-dimensional space can be solved by solving Equation (13). In the process of solving, the decision function can be obtained as shown in Equation (14).

$$y(x) = \operatorname{sgn}\left[\sum_{i=1}^{N}\alpha_i y_i K(x_i\cdot y_i) + b\right] \tag{14}$$

3.2. *Fuzzy least square support vector classification mechanism construction*

SVM is proposed as a two-category classification problem. However, in the actual situation, the classification is not only binary classification, but also multi-classification. Therefore, it is an important issue to extend SVM to the maximum classification. Martínez et al. proposed the multi-classification algorithm of DAG-SVM to effectively classify slate tiles. In line with their construction ideas, research has proposed one-to-many and many-to-many classification ideas for multi-classification [27].

The training set is set as $T = \{x_k, y_k^{(i)}\}$. m classifiers are needed in the one-to-many multi-classification algorithm. When the classifier is trained, the samples of this class are all positive and the rest are negative. This method solves the problem of minimizing the loss function as shown in equation (15).

$$\begin{cases} \min C\sum_{j=1}^{m}\xi_j^i + 0.5\left(w^i\right)^T w^i \\ s.t.\begin{cases} 1-\xi_j^i \le \left(w^i\right)^T\varphi\left(x_j\right)+b^i & y_j = i \\ \xi_j^i - 1 \ge \left(w^i\right)^T\varphi\left(x_j\right)+b^i & y_j \ne i \\ 0 \le \xi_j^i \end{cases} \end{cases} \tag{15}$$

After optimizing the equation, the decision function of the number of corresponding classifiers can be obtained. Its expression is equation (16).

$$\begin{cases} \left(w^1\right)^T\varphi(x)+b^1 \\ \left(w^2\right)^T\varphi(x)+b^2 \\ \cdots \\ \left(w^m\right)^T\varphi(x)+b^m \end{cases} \tag{16}$$

The optimal hyperplane expression of classification can be obtained through the decision tree, as shown in Equation (17).

$$D_i(x) = \left(w^i\right)^T\varphi(x)+b^i = 0 \tag{17}$$

in Equation (17), $(w^i)^T$ represents the i-dimensional vector. In the sample classification process, if the expression of the optimal hyperplane is greater than 0, the tested sample belongs to this category. In the current scenario, there may be multiple instances of the same condition or no such condition. This leads to a misclassification error in the sample data, thus resulting in indistinguishability in classification. The one-to-many classification algorithm is shown in Fig. 2.

In Fig. 2, the algorithm can only distinguish the cases where $D_i(x)$ is greater than 0 and $D_i(x)$ is less than 0. The samples in the non-separable region cannot be classified correctly. In the one-to-one multi-classification algorithm, the samples are classified in pairs. For the two samples, the decision tree is expressed by equation (18).

$$D_{ij}(x) = \left(w^{ij}\right)^T\varphi(x_t)+b^{ij} \tag{18}$$

The method is tested by a certain number of classifiers and is classified by voting. If the classifier determines the sample data to be i, then the sample will be increased by one vote in i. If the classifier determines the sample data as class j, the sample will be added to class j by one vote. After the final judgment of all classifiers, which category has the most votes, the sample data belongs to that category. The one-to-one multi-classification diagram is shown in Fig. 3.

From Fig. 3, the three regions indicate that samples of the same category will be encountered in the classification algorithm. So they still cannot be correctly classified in the non-partition domain. Therefore, to solve the problem of incorrect classification, FLS-SVM was obtained by combining fuzzy theory with LS-SVM [28,29]. In FLS-SVM, a fuzzy membership function $m_{ij}(x)$ is defined for the classification hyperplane. The dimension of this function is only one dimension and its direction is perpendicular to the separation

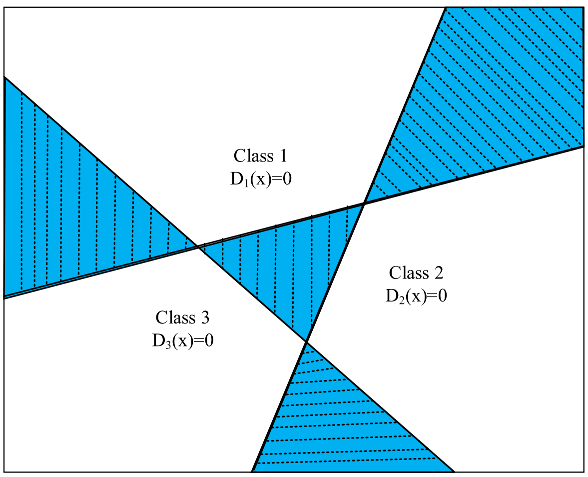


Fig. 2. Schematic diagram of one to many classification algorithm.

hyperplane. Equation (19) is the expression.

$$m_{ij}(x)=\begin{cases} 1 & D_{ij}(x)\geq 1 \\ D_{ij}(x) & D_{ij}(x<1) \end{cases} \tag{19}$$

in equation (19), if the super-flat value is not less than 1, the value of the fuzzy membership function is 1. If the hyper-flat value is less than 1, the value of the fuzzy membership function is equal to the hyper-flat value. The minimum operator is used in the fuzzy membership function to obtain the membership function of the sample. The expression is shown in equation (20).

$$m_i(x)=\min\left(1,\min_{j\neq i}D_{ij}(x)\right) \tag{20}$$

FLS-SVM correctly classifies the samples in the indivisible regions in Figs. 2 and 3 when using the minimum operator. The FLS-SVM classification plot is shown in Fig. 4. Among which, Fig. 4(a) shows FLS_ SVM one to many multi classification, and Fig. 4(b) shows FLS_ SVM one to one multi classification.

FLS-SVM can also obtain the membership function of the sample using the average operator, and its expression is shown in Equation (21).

$$m_i(x)=\frac{\sum_{j=1,j\neq i}^{m} m_{ij}(x)}{m-1} \tag{21}$$

when applying the minimum operator and the average operator, the test samples are judged to be the class with the maximum membership function when the test samples are time-divided. In solving FLS-SVM, it is important to solve the matrix equation of the linear KTT system in the dual space. In this study, the Cholesky algorithm was used. Since Ω of the matrix equation is positive definite, its expression is listed in Equation (22).

$$\alpha=\Omega^{-1}(1_N-Yb) \tag{22}$$

Y in Equation (22) represents a non-zero vector. By substituting Equation (22) into the equation matrix, the classification threshold can be solved, and its expression is Equation (23).

$$b=\left(Y^T\Omega^{-1}1_N\right)/\left(Y^T\Omega^{-1}Y\right) \tag{23}$$

After obtaining α and b, the Cholesky algorithm was used to decompose the matrix Ω to obtain the triangular matrix of the upper and lower parts, which were represented by L and L^T respectively. The triangular matrix is the decomposition factor of Cholesky's algorithm, in which the element expression is shown in Equation (24).

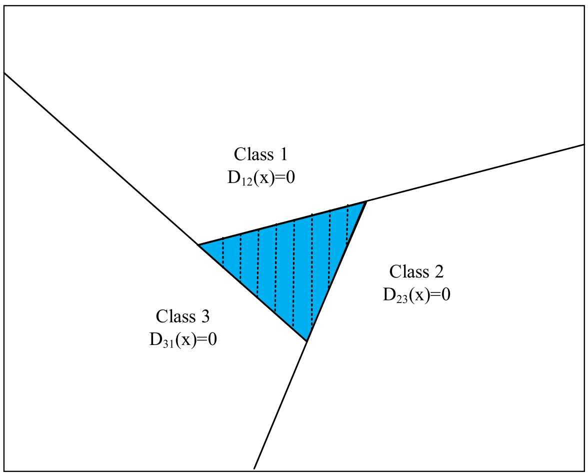
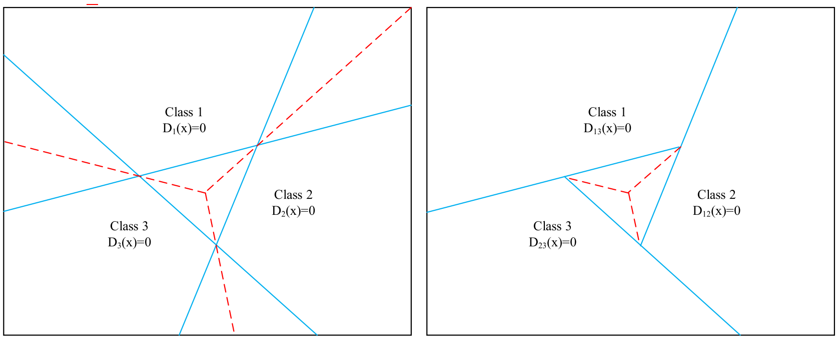


Fig. 3. One to one multi classification diagram.



(a) FLS_ SVM one to many multi classification (b) FLS_ SVM one to one multi classification

Fig. 4. FLS-SVM classification diagram.

```
# Fuzzy Least Squares Support Vector Machine (FLS-SVM) for Document Management Classification

# 1. Data Preparation and Preprocessing
- Load document management dataset
- Split the dataset into training and testing sets
- Standardize the dataset

# 2. FLS-SVM Model Construction
class FuzzyLS_SVM:
def __init__(self, fuzzification, regularization, kernel):
self.fuzzification = fuzzification
self.regularization = regularization
self.kernel = kernel

def train(self, X_train, y_train):
# 2.1 Data Fuzzification
fuzzy_X_train = self.fuzzification.fuzzify(X_train)

# 2.2 Build Least Squares model
LS_model = self.regularization.least_squares(fuzzy_X_train, y_train)

# 2.3 Generate classifier based on model results
self.classifier = self.kernel.choose_kernel(LS_model)

def predict(self, X_test):
# 2.4 Data Fuzzification
fuzzy_X_test = self.fuzzification.fuzzify(X_test)

# 2.5 Perform classification prediction
predictions = self.classifier.predict(fuzzy_X_test)
return predictions

# 3. FLS-SVM Model Parameter Settings
fuzzification = Fuzzification()  # Fuzzification method
regularization = Regularization()  # Regularization method
kernel = Kernel()  # Kernel method

# 4. Model Training and Prediction
fls_svm = FuzzyLS_SVM(fuzzification, regularization, kernel)
fls_svm.train(X_train, y_train)
predictions = fls_svm.predict(X_test)
```

Fig. 5. FLS-SVM Pseudocode in document management classification.

$$\begin{cases} l_{op} = \dfrac{q_{op} - \sum\limits_{n=1}^{p-1} l_{pn} l_{on}}{l_{pp}} \\ l_{pp} = \sqrt{q_{pp} - \sum\limits_{n=1}^{p-1} l_{pn}^2} \end{cases} \tag{24}$$

l_{op} in Equation (24) refers to the element in the triangular matrix. If the matrix Ω was close to semi-definite, the Cholesky algorithm showed instability in the decomposition process. To avoid this phenomenon, the value range of $q_{pp} - \sum_{\hat{A}\hat{o}}^{\mathring{a}-1} l_{pn}^2$ was set to no more than η. Then Ω could be defined as Equation (25).

$$\Omega a = LL^T a = Y \tag{25}$$

in Equation (25), a represents a vector of a certain dimension. Equation (25) reduces the difficulty of solving the equation, so the solution expression of b is shown in Equation (26).

$$b = \left(a^T 1_N\right) / \left(Y^T a\right) \tag{26}$$

in the FLS-SVM training process, the Cholesky decomposition algorithm has a stable and effective performance for solving linear equations. The Pseudocode of this method in document management is shown in Fig. 5.

The paper is based on least squares to improve the model. In addition to using this method, some scholars have used the same method for research. Gupta and Gupta proposed a class dispersed least squares Bi-bounded support vector machine (LS-BSVM) [30], a bipolar fuzzy least squares Bi-bounded support vector machine (BFLS-BSVM) [31], and a least squares large interval distribution machine regression (LSLI-DMR) model [32]. Borah and Gupta proposed a pair of fuzzy least squares SVM (FLS-SVM) methods based on class affinity and nonlinear transformation [33]. These methods are all improved versions based on least squares variables, and the study compares and analyzes these models. LS-BSVM is a class dispersion based method designed to handle imbalanced datasets. It uses a least squares structure to distinguish categories by determining the degree of dispersion between them. The BFLS-BSVM uses the methods of bipolar fuzzy model and LS-BSVM, aiming to solve the problem of imbalanced datasets. The LSLI-DMR model is a regression model that uses the least squares large interval method and is applied to process data distribution. The FLS-SVM method is obtained by combining class affinity and nonlinear transformation class probability. It improves model performance by addressing class imbalance issues and reducing sensitivity to outliers and noise. These models differ in algorithms and methods, and can provide solutions for different problems and data types. Each model has its unique characteristics and advantages, suitable for specific types of data analysis and processing.

4. Analysis of identification effect of least square fitting method in archives management

4.1. Performance analysis of FLS-SVM classifier

The paper focuses on the performance of the FLS-SVM classifier. In the experiment, the computer hardware configuration was used as follows: CPU is 2.4 GHz, memory is 8 GB, graphics card is GTX1060. The simulation experiments were conducted using MATLAB engineering software for verification. In the experiment, the performance of different kernel functions and hyperparameters in the algorithm was verified. The parameters of the classifier were optimized using the CV method. This chapter has compared FLS-SVM classifiers with C-SVM, V-SVM, and LS-SVM classifiers. These classifiers have good results in dichotomy problems, but more research is needed on multi-classification problems. The kernel functions used in the experiment are Radial Basis Function (RBF kernel), Poly kernel, Multi-layer Perceptron, MLP kernel, and Wavelet kernel. The dataset used in the experiment is from the multi-

Table 1
Kernel parameter values of different classifiers.

Classifier	Nuclear parameter	Wavelet kernel	RBF kernel	Poly kernel	MLP kernel
C-SVM	C	10	10	10	10
	σ^2/d	0.5	0.01	2	1
	t	/	/	3	0.1
V-SVM	C	10	10	10	10
	σ^2/d	0.5	0.01	2	1
	t	/	/	3	0.1
LS-SVM	γ	30	30	30	30
	σ^2/d	0.5	0.01	2	1
	t	/	/	3	0.1
FLS-SVM	γ	30	30	30	30
	σ^2/d	0.5	0.01	2	1
	t	/	/	3	0.1

classification dataset in the UCI database.

The selection of kernel function parameters for the classifier is determined by multiple classification experiments. The final results of determining the optimal kernel parameters for each model are shown in Table 1. The values of kernel parameters in Table 1 are calculated by the equation. The penalty factor of C-SVM and V-SVM is 10. The regularization parameter of LS-SVM and FLS-SVM is 30. The parameters of poly kernel are polynomial of degree 3, so d is 3 and t is 2. The value of σ^2 of MLP kernel is 0.1, and the value of t is 1. The value of σ^2 of RBF kernel is 0.01. The value of σ^2 of the Wavelet kernel is 0.5.

From Table 1, LS-SVM and FLS-SVM perform well when the wavelet kernel is used as the kernel function. Therefore, the experiment further compares these two forms. The experiment compared the one-to-one multiclassification with the one-to-many multi-classification. First, the kernel parameters were optimized using the CV method. The optimal parameter results are shown in Table 2.

In Table 2, the regularization parameters of LS-SVM and FLS-SVM have different values in different kernel functions. The variance parameter also has different values. The regularization parameter of RBF kernel in FLS-SVM is 10, and the variance value is 0.1. The regularization parameter of the Poly kernel in FLS-SVM is 20, and the dimension is 3. The optimal parameter values for each model are listed above. In subsequent experiments, the corresponding optimal values will be applied to each model's parameters.

The Irsi dataset in the UCI database was used for the classification experiment. The classification results are shown in Fig. 6. In Fig. 6(a), when using the minimum operator, FLS-SVM classifies the dataset through four kernel functions (Wavelet, RBF, Poly, and MLP). Among them, Wavelet kernel performs best among these kernel functions, achieving a precison of 99%. The precision of RBF and Poly kernels is about 98%, while the precision of MLP kernels is slightly lower than 90%. This indicates that when the minimum operator is used, the classification effect of wavelet on the iris dataset is the best. This may be because the Wavelet kernel utilizes the properties of wavelet transform to better extract time-domain and frequency-domain features of data, thereby enhancing the accuracy of classification. In Fig. 6(b), FLS-SVM uses the average operator to classify the dataset. The results show that the Wavelet kernel still performs well with an accuracy of 97.35%. This means that the Wavelet kernel is still the best choice when using the average operator, with good classification performance and the ability to correctly classify the dataset. In Fig. 6(c), LS-SVM uses the One-to-One to classify the dataset. The Wavelet kernel performs the best with a precison of 97.35%, while the RBF performs better than Poly and MLP. Wavelet kernel still perfoms the most excecllent. In Fig. 6(d), LS-SVM uses a one-to-many multi classification method, and the classification accuracy of the Wavelet kernel is 90.02%, while the accuracy of other kernel functions is below 90%. This indicates that the Wavelet kernel still exhibits good performance in this multi classification situation. It is more suitable for dataset classification compared to other kernel functions. Overall, in the given classification experiments, both FLS-SVM and LS-SVM shows certain classification performance, while the Wavelet kernel shows good classification performance in different situations. This may be due to the fact that the Wavelet kernel can better extract features from the Iris dataset, enabling accurate classification.

Fig. 7 shows the effect of FLS-SVM on the classification of the Iris dataset. In Fig. 7, the classifier divides the petal size into three types: large, medium, and small, with medium sized petals as the classification boundary. This means that the FLS-SVM classifier can accurately distinguish petals of different sizes and perform precise classification. This also indicates that the classifier performs stably when dealing with petals of different sizes and has good generalization ability in classification problems. This classification effect is very valuable for studying the Iris dataset. Through the FLS-SVM classifier, research can more accurately classify petal sizes into different categories, thereby better understanding and analyzing flowers. In addition, this classifier can also be applied to other similar problems to achieve a wider range of classification tasks. The runtime of each classifier in each kernel function is shown in Table 3.

In Table 3, regardless of the kernel function type, the V-SVM classifier has the longest running time. This could be attributed to the significant algorithmic complexity of V-SVM classifiers, which necessitate additional computing resources and time in processing datasets. In contrast, FLS-SVM classifiers exhibit the shortest runtime, suggesting that they possess high efficiency when handling datasets. This is very valuable for large-scale datasets and real-time applications, as it can quickly complete classification tasks in a short period of time. Within the same classifier, using the MLP kernel has the longest runtime. The MLP kernel contains more complex computational processes and may require more time to complete. On the other hand, the Wavelet kernel used in the study has the shortest runtime. The Wavelet kernel is a kernel function that can handle multi-scale features with low computational complexity, which may require less time to complete classification tasks. It should be noted that when FLS-SVM uses the Wavelet kernel as the kernel function, its data classification time is 5.912 s. This indicates that FLS-SVM using the Wavelet kernel has a certain time cost when classifying data, but overall performs well.

Table 2
Kernel parameter value optimized by CV method.

Classifier	Nuclear parameter	Wavelet kernel	RBF kernel	Poly kernel	MLP kernel
LS-SVM (one VS one)	γ	12.60	10.07	0.01	15.47
	σ^2	9.38	2.88	4.00	0.45
LS-SVM (one VS all)	γ	24.00	23.53	0.68	37.05
	σ^2	9.57	10.19	3.00	1.77
FLS-SVM (Min)	γ	24.00	10.00	20.00	8.64
	σ^2	0.60	0.01	3.00	0.80
FLS-SVM (Avg)	γ	22.31	10.00	20.00	10.28
	σ^2	0.57	0.01	3.00	0.53

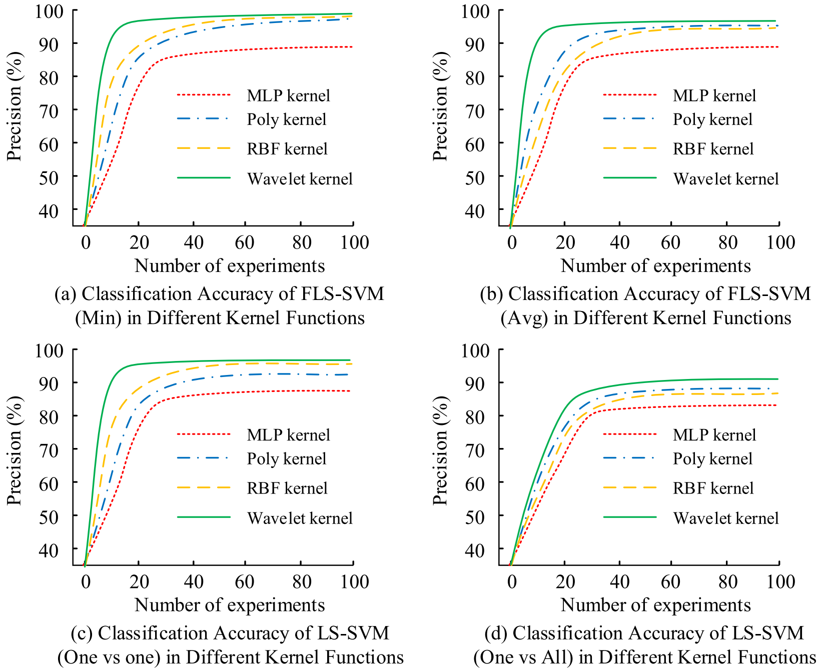


Fig. 6. Classification effect of FLS-SVM on Irsi datasets.

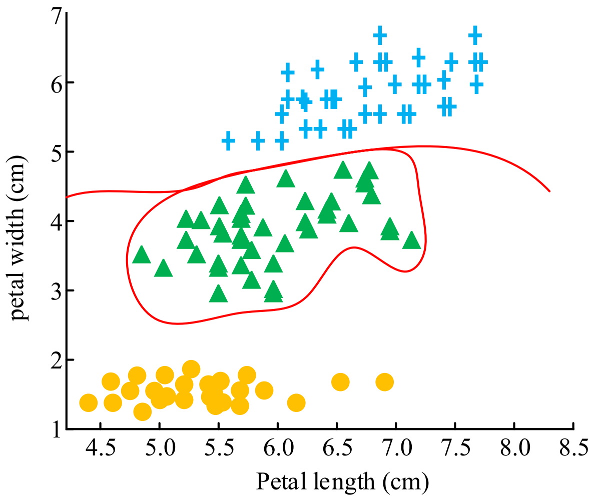


Fig. 7. Classification rendering of Iris dataset.

Table 3
Running time of classifier in dataset.

Classifier	C-SVM	V-SVM	LS-SVM	FLS-SVM
Wavelet kernel (s)	6.157	6.471	6.203	5.912
RBF kernel (s)	7.219	7.719	6.624	6.295
Poly kernel (s)	6.874	9.574	6.349	6.079
MLP kernel (s)	8.042	8.991	7.713	7.162

4.2. Identification and classification effect in archives management of FLS-SVM

The dataset in the UCI database is also used for the experiment, and the selected dataset type is archival dataset. These include Teaching Assistant Evaluation Data Set, Census Income (KDD) Data Set, Restaurant & Consumer Data Set, and Turkiye Student Evaluation Data Set. These four data sets were selected for classification. The performance of FLS-SVM was reflected by the classification accuracy and loss rate of the classifiers.

Fig. 8 reflects the classification effect of four LS-SVM multi classification algorithms in archive management. In Fig. 8(a), the classification accuracy of FLS-SVM (Min) in training reaches about 99.5%, while in actual testing, the accuracy reaches about 98.7%. This indicates that the FLS-SVM (Min) algorithm has very high classification accuracy in both the training and testing stages. This is extremely useful for the multi-classification challenge in archive management, as precise classification can facilitate efficient and effective management of various types of archives. In Fig. 8(b), the classification accuracy of FLS-SVM (Avg) in training is about 99.2%, while the accuracy in actual testing is about 97.9%. This suggests that the FLS-SVM (Avg) algorithm maintains high classification accuracy throughout both the training and testing stages, albeit with a slightly reduced performance compared to FLS-SVM (Min). Nevertheless, FLS-SVM (Avg) remains an effective classification algorithm capable of achieving high accuracy in archive management. In Fig. 8(c), the classification accuracy of LS-SVM (One VS One) in training is about 99.4%, while in actual testing, the accuracy is about 98.2%. This indicates that the LS-SVM (One VS One) algorithm can achieve high accuracy in classification during both training and testing stages. This one-to-one multi classification strategy has shown good performance in archive management, ensuring reliable and accurate classification results. In Fig. 8(d), the classification accuracy of the LS-SVM (One VS all) algorithm in training is about 99.3%, while the accuracy in actual testing is about 98%. This indicates that the LS-SVM (One VS all) algorithm can also achieve high accuracy in classification. This one-to-many multi classification strategy also provides a feasible and accurate solution to the classification problem in archive management. In summary, both FLS-SVM and LS-SVM multi classification algorithms have achieved quite high classification accuracy in archive management. These algorithms can effectively handle multi classification problems and provide accurate file classification and management.

Fig. 9 reflects the loss rate results of four LS-SVM multi classification algorithms in archive management. In Fig. 9(a), the classification loss rate of FLS-SVM (Min) in training is about 0.23%, while in actual testing, the loss rate is about 0.26%. This demonstrates the effectiveness of the FLS-SVM (Min) algorithm in achieving significantly low rates of loss during classification tasks, resulting in high classification accuracy. The minimal loss rate implies the algorithm's capability to accurately categorize data into its respective group. In Fig. 9(b), the classification loss rate of FLS-SVM (Avg) in training is about 0.48%, while in actual testing, the loss rate is about 0.51%. Although there is a slight increase compared to FLS-SVM (Min), FLS-SVM (Avg) can still achieve a lower loss rate. This indicates that the FLS-SVM (Avg) algorithm can also provide accurate classification results in classification tasks, although the loss rate is slightly higher than FLS-SVM (Min). In Fig. 9(c), the classification loss rate of LS-SVM (One VS One) in training is about 0.37%. In actual testing, the loss rate is about 0.41%. Compared with the first two algorithms, LS-SVM (One VS One) exhibits a lower loss rate in classification tasks. This further proves that the algorithm has high performance in multi classification problems in archive

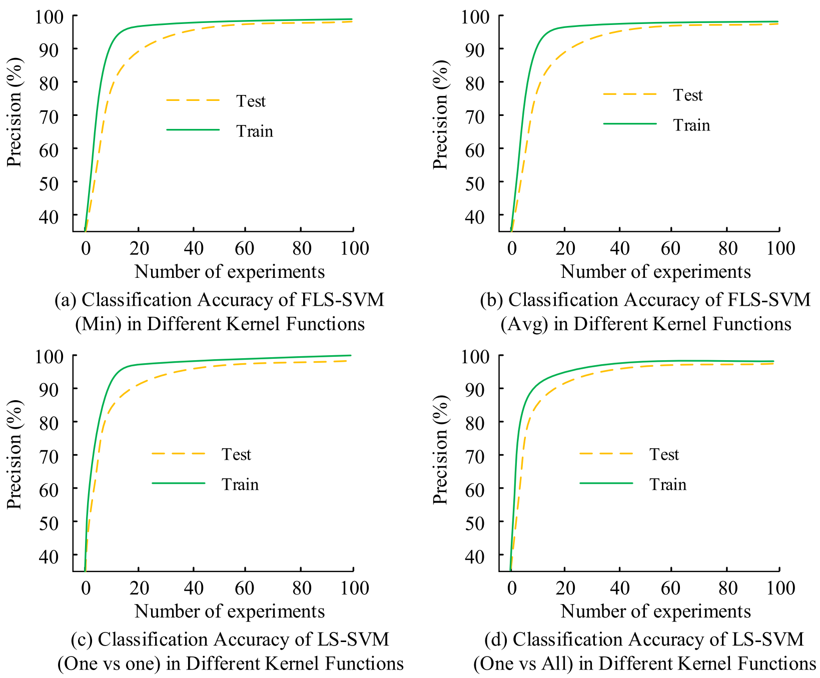
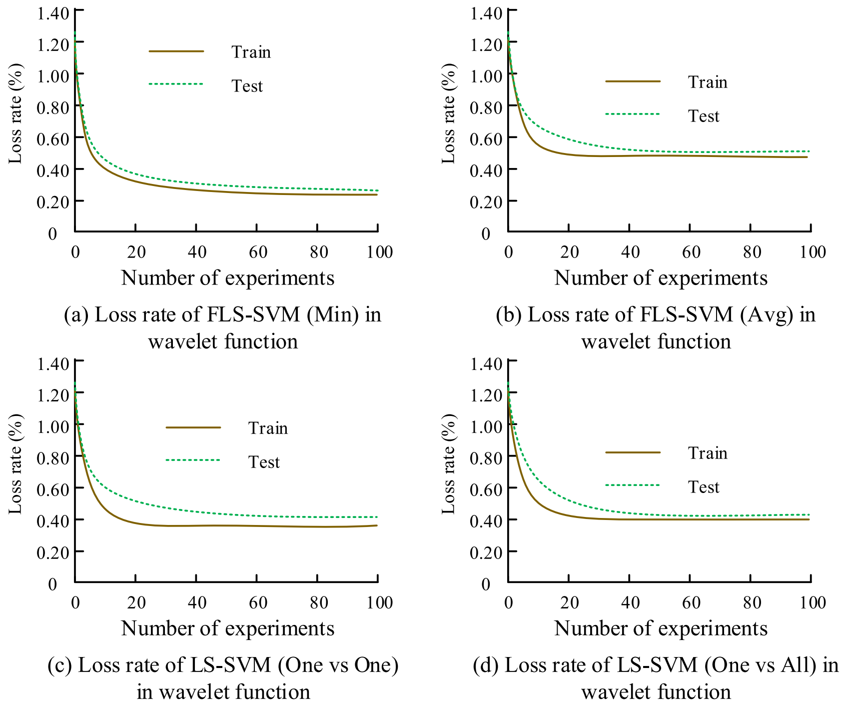


Fig. 8. Classification accuracy of FLS-SVM in archive dataset.



(a) Loss rate of FLS-SVM (Min) in wavelet function
(b) Loss rate of FLS-SVM (Avg) in wavelet function
(c) Loss rate of LS-SVM (One vs One) in wavelet function
(d) Loss rate of LS-SVM (One vs All) in wavelet function

Fig. 9. Results of FLS-SVM's loss rate in archive dataset.

management, and can effectively classify different archives. In Fig. 9(d), the classification loss rate of the LS-SVM (One VS all) algorithm in training is about 0.4%. The loss rate in actual testing is about 0.42%. Although the loss rate is relatively low, LS-SVM (One VS all) still exhibits good classification performance. By adopting a one-to-many multi classification strategy, this algorithm can achieve accurate classification results in archive management. In summary, both FLS-SVM and LS-SVM multi classification algorithms can achieve lower loss rates in archive management. These results indicate that these algorithms can accurately classify data into the correct categories to achieve efficient archive management.

Fig. 10 shows the average classification accuracy results of FLS-SVM (Min) under different kernel functions for archival data sets. In Fig. 10, when the kernel function of FLS-SVM (Min) is Wavelet kernel, the average classification accuracy is 98.38%. When the kernel function of FLS-SVM (Min) is RBF kernel, the average classification accuracy is 97.21%. When the kernel function of FLS-SVM (Min) is Poly Kernel, the average classification accuracy is 94.71%. When the kernel function of FLS-SVM (Min) is MLP kernel, the average classification accuracy is 91.04%. In summary, the four LS-SVM classification methods have high classification accuracy, among which FLS-SVM (Min) has the best classification. Among different wavelet kernel functions, the classification performance of wavelet kernel is the most outstanding. It plays an excellent role in the recognition and classification of various archive management. The study used Eg Friedman statistical software to analyze the model results. The accuracy and loss rate of the model were recorded in the experiment. The Friedman test was used to analyze the model ranking. The results are shown in Table 4.

From the test results in Tables 4 and in the Teaching Assistant Evaluation Data Set, the FLS-SVM algorithm has significant differences in accuracy and loss rate compared to other comparative algorithms ($P < 0.05$). In the Census-Income (KDD) Data Set, the FLS-SVM algorithm also shows significant differences compared to other comparative algorithms ($P < 0.05$). In the Restaurant & Consumer Data Set, the FLS-SVM algorithm shows significant differences compared to other comparative algorithms ($P < 0.05$). On the Turkiye Student Evaluation dataset, the FLS-SVM algorithm shows significant differences compared to other comparative algorithms ($P < 0.05$). The core idea of Eg Friedman is to compare the overall performance of algorithms by calculating average rankings and considering the performance of algorithms on different data sets. Therefore, the results show that the algorithm proposed in the study has a good performance.

5. Conclusion

Since least squares fitting is globally optimal, has strong adaptability and a powerful ability for fitting optimization, this study performed least squares fitting on SVM resulting in LS-SVM. By exploring the influence of different kernel parameters on the classifier, a new wavelet function was constructed and applied to FLS-SVM. The CV method was used to optimize the kernel parameters. Simulation experiments were conducted to verify the performance of the improved FLS-SVM classifier. The experimental results showed that the FLS-SVM wavelet kernel had 99% classification accuracy in the iris data set when it used the wavelet kernel for classification. In the process of archive management classification, the classification accuracy of FLS-SVM (Min) reached 98.7%. The classification loss rate of FLS-SVM (Min) was only 0.26%. When wavelet kernel was used as the classifier kernel function, the average

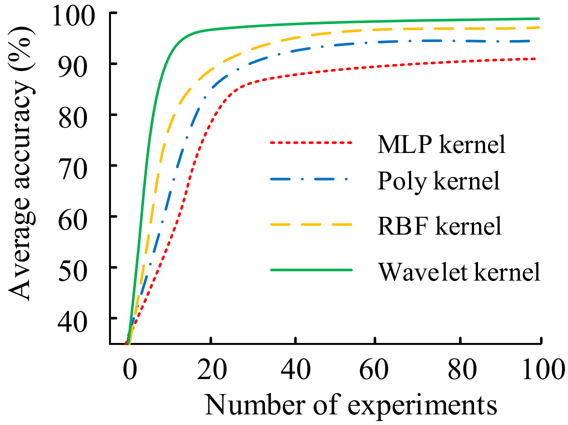


Fig. 10. Average classification accuracy of FLS-SVM (Min) for archive data sets under different kernel functions.

Table 4
Statistical test results.

Data set	Index	C-SVM	V-SVM	LS-SVM	FLS-SVM
Teaching assistant evaluation data set	Accuracy	88.45%	90.12%	96.24%	98.70%^{abc}
	Loss rate	11.55%	9.88%	3.76%	1.30%^{abc}
Census-income (KDD) data set	Accuracy	88.48%	89.28%	95.89%	98.16%^{abc}
	Loss rate	11.52%	10.72%	4.11%	1.84%^{abc}
Restaurant & consumer data set	Accuracy	88.29%	91.02%	95.84%	98.12%^{abc}
	Loss rate	11.71%	8.98%	4.16%	1.88%^{abc}
Turkiye student evaluation data set	Accuracy	87.98%	89.89%	96.25%	97.89%^{abc}
	Loss rate	12.02%	10.11%	3.75%	2.11%^{abc}

Note: “a" indicates that FLS-SVM is $P < 0.05$ compared to C-SVM; “b" represents a $P < 0.05$ between FLS-SVM and V-SVM; “c" indicates a $P < 0.05$ between FLS-SVM and LS-SVM.

classification accuracy of managed files reached 98.38%. The results showed that FLS-SVM classifier based on wavelet kernel function can improve the speed and shorten the training time without affecting the classification accuracy. The main contribution of this study is the application and optimization of least squares fitting methods in support vector machines. By optimizing LS-SVM with a new wavelet kernel function, the classification training speed has been improved and the training time has been shortened. The scientific value of this study is reflected in the following three aspects. Firstly, by optimizing the kernel function in LS-SVM and using the least squares fitting method, the performance of the classifier was verified through a large number of experiments. This provides a more feasible and effective solution to the classification problem in archive management. The second aspect is to introduce wavelet kernel functions as the kernel functions of FLS-SVM classifiers, which improves classification accuracy and reduces training time. The improved FLS-SVM classifier in the third aspect improves the situation where data is not separable in multi classification problems in archive management. This indicates the feasibility and effectiveness of FLS-SVM classifier in dealing with complex classification problems. However, there are still shortcomings in the research. The CV method is only used to optimize the parameters of several common kernel functions. The selection of kernel functions and the optimization of hyperparameters are still hot topics in LS-SVM research. However, most studies have neither explored other effective kernel functions or combined kernel functions, nor utilized further optimization algorithms to optimize them. Therefore, in future research, it is possible to consider exploring other kernel functions and adopting other optimization algorithms to further improve the performance and robustness of LS-SVM. The classification effect of this study can also be applied to the classification problems of other projects, and has certain applicability in the field of classification.

Data availability statement

The data used to support the findings of this study are available from the corresponding author upon request.

Ethics approval

Not applicable.

Informed consent

Not applicable.

Funding statement

This paper was supported by Scientific Research Projects of Department of Education of Hubei Province (Grant No. Q20202705).

Author contribution statement

Caichang Ding: conceived and designed the experiments; wrote the paper.
Hui Liang: Conceived and designed the experiments.
Na Lin: Contributed reagents, materials, analysis tools or data.
Zhimin Li: Contributed reagents, materials, analysis tools or data.
Zenggang Xiong: Contributed reagents, materials, analysis tools or data.
Peilong Xu: Analyzed and interpreted the data.

Additional information

No additional information is available for this paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M.T. Tran, L.Q. Trieu, H.Q. Tran, Document representation and classification with Twitter-based document embedding, adversarial domain-adaptation, and query expansion, J. Heuristics 28 (2) (2022) 211–233.
[2] W. Zhang, Z. Wu, D.W. Bunn, Optimal hybrid framework for carbon price forecasting using time series analysis and least squares support vector machine, J. Forecast. 41 (3) (2022) 615–632.
[3] Rostum M A e A E S, H.M.M. Moustafa, I.E.S. Ziedan, A.A. Zamel, A combined effective time series model based on clustering and whale optimization algorithm for forecasting smart meters electricity consumption, COMPEL 41 (1) (2022) 209–237.
[4] V. Kalra, I. Kashyap, H. Kaur, Generation of domain-specific vocabulary set and classification of documents: weight-inclusion approach, Int. J. Inf. Technol. 14 (1) (2022) 75–285.
[5] M. Gayathri, R.J. Kannan, Ontology based concept extraction and classification of ayurvedic documents, Procedia Comput. Sci. 172 (2020) 511–516.
[6] Y.L. Chen, L.C. Cheng, Y.J. Zhang, Building a training dataset for classification under a cost limitation, Electron. Libr. 39 (1) (2021) 77–96.
[7] S. Seifollahi, M. Piccardi, A. Jolfaei, An embedding-based topic model for document classification, ACM Transactions on Asian and Low-resource Language Information Processing 20 (3) (2021) 52.
[8] H. Azarbonyad, M. Dehghani, M. Marx, J. Kamps, Learning to rank for multi-label text classification: combining different sources of information, Nat. Lang. Eng. 27 (1) (2020) 89–111.
[9] M.C. Wijaya, The classification of documents in Malay and Indonesian using the naive Bayesian method uses words and phrases as a training set, Mendel 26 (2) (2020) 23–28.
[10] S. Solikin, H. Manik, S. Pujiyati, S. Susilohadi, Support vector machine classification method for predicting Jakarta bay bottom sediment type using multibeam echosounder data, Pertanika Journal of Science and Technology 48 (2) (2020) 477–491.
[11] Y. Chen, K. Ma, P.L. Xu, H.Z. Si, Y.B. Duan, H. Zhai, Design and Screening of New Lead Compounds for Autism Based on QSAR Model and Molecular Docking Studies, Molecules 27 (2022), 7285.
[12] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, A. Lopez, A comprehensive survey on support vector machine classification: applications, challenges and trends, Neurocomputing 408 (2020) 189–215.
[13] H.D. Praveena, C. Subhas, K.R. Naidu, Classification and discrimination of focal and non-focal EEG signals using hybrid features and support vector machine, Int. J. Adv. Intell. Paradigms 18 (3) (2021) 417–437.
[14] J. Zhang, G. Ma, J. Cheng, S. Song, Y. Zhang, L. Shi, Diagnostic classification of solitary pulmonary nodules using support vector machine model based on 2-[18F]fluoro-2-deoxy-D-glucose PET/computed tomography texture features, Nucl. Med. Commun. 41 (6) (2020) 560–566.
[15] S.K. Mishra, V.H. Deepthi, Brain image classification by the combination of different wavelet transforms and support vector machine classification, J. Ambient Intell. Hum. Comput. 12 (6) (2021) 6741–6749.
[16] Z. Zhang, J. He, H. Zheng, J. Cao, G. Wang, Y. Shi, Alternating minimization-based sparse least-squares classifier for accuracy and interpretability improvement of credit risk assessment, Int. J. Inf. Technol. Decis. Making 22 (1) (2023) 537–567.
[17] W. Zhang, Z. Wu, D.W. Bunn, Optimal hybrid framework for carbon price forecasting using time series analysis and least squares support vector machine, J. Forecast. 41 (3) (2022) 615–632.
[18] H. Chen, M. Zeng, H. Zhang, Prediction of carbon dioxide solubility in polymers based on adaptive particle swarm optimization and least squares support vector machine, ChemistrySelect 7 (3) (2022) 1–8.
[19] W. Lyu, Z. Wang, Global classical solutions for a class of reaction-diffusion system with density-suppressed motility, Electronic Research Archive 30 (3) (2022) 995–1015.
[20] X. Li, Y. Sun, Stock intelligent investment strategy based on support vector machine parameter optimization algorithm, Neural Comput. Appl. 32 (6) (2020) 1765–1775.
[21] X. Dai, Z. Xiao, H. Jiang, M. Alazab, J.C.S. Lui, G. Min, S. Dustdar, J. Liu, Task offloading for cloud-assisted fog computing with dynamic service caching in enterprise management systems, IEEE Trans. Ind. Inf. 19 (1) (2023) 662–672.
[22] G. Premalatha, V.T. Bai, Wireless IoT and cyber-physical system for health monitoring using honey badger optimized least-squares support-vector machine, Wireless Pers. Commun. 124 (4) (2022) 3013–3034.

[23] J.R. Soble, D.W. Sharp, D.A. Carter, K.J. Jennette, Z.J. Resch, G.P. Ovsiew, E.A. Critchfield, Cross-validation of a forced-choice validity indicator to enhance the clinical utility of the Rey Auditory Verbal Learning Test, Psychol. Assess. 33 (6) (2021) 568–573.
[24] I. Tougui, A. Jilbab, J.E. Mhamdi, Impact of the choice of cross-validation techniques on the results of machine learning-based diagnostic applications, Healthcare Informatics Research 27 (3) (2021) 189–199.
[25] X. Liu, G. Zhou, M. Kong, Z. Yin, X. Li, L. Yin, W. Zheng, Developing multi-labelled corpus of twitter short texts: a semi-automatic method, Systems 11 (8) (2023) 390.
[26] X. Liu, T. Shi, G. Zhou, M. Liu, Z. Yin, L. Yin, W. Zheng, Emotion classification for short texts: an improved multi-label method, Humanities and Social Sciences Communications 10 (1) (2023) 306.
[27] J. Martínez, C. Iglesias, J.M. Matías, J. Taboada, M. Araújo, Solving the slate tile classification problem using a DAGSVM multiclassification algorithm based on SVM binary classifiers with a one-versus-all approach, Appl. Math. Comput. 230 (2014) 464–472.
[28] S. Zhang, L. Zhang, A.G. Hauptmann, Fuzzy least squares support vector machine with adaptive membership for object tracking, IEEE Trans. Multimed. 22 (8) (2020) 1998–2011.
[29] S. Chen, J. Cao, F. Chen, B. Liu, Entropy-based fuzzy least squares twin support vector machine for pattern classification, Neural Process. Lett. 51 (1) (2020) 41–66.
[30] U. Gupta, D. Gupta, Least squares structural twin bounded support vector machine on class scatter, Appl. Intell. 53 (12) (2023) 15321–15351.
[31] U. Gupta, D. Gupta, Bipolar fuzzy based least squares twin bounded support vector machine, Fuzzy Set Syst. 449 (2022) 120–161.
[32] U. Gupta, D. Gupta, Least squares large margin distribution machine for regression, Appl. Intell. 51 (2021) 7058–7093.
[33] P. Borah, D. Gupta, Affinity and transformed class probability-based fuzzy least squares support vector machines, Fuzzy Set Syst. 443 (2022) 203–235.