RESEARCH

Research on data mining algorithm of logistics time series based on intelligent integrated network structure

Zhengyi Zhang[1*] and Jian Cui[2]

* Correspondence:
ngqiheng1966dr@163.com
[1]School of Economics and
Management, Xi'an Aeronautical
University, Xi'an 710077, People's
Republic of China
Full list of author information is
available at the end of the article

Abstract

With the continuous development of information, big data analysis has become important and dependent technical means-increasingly in various fields. By data mining through time series, the development regular of the object could be grasped, so we could predict its future development trend. Based on the intelligent integration architecture, a new algorithm of bi-weighted support vector machines (SVM) based on category weighting, and feature weighting was proposed to solve the problem of unbalanced samples in time series. In the non-balanced sample set classification, the recognition ability of the traditional classification method was low; the supported vector machine as classifier was taken in the new algorithm based on cost-sensitive learning, and different weighting coefficients to less and more samples were given, and Gauss kernel function with the weight coefficients of different features was reconstructed, thus the recognition ability of less samples was improved. In the experiment, classification accuracy, g-mean, f-measure, TP, and FP were selected as evaluation indexes, indicating that the two-weighted SVM algorithm is effective in the classification of non-balanced sample sets.

Keywords: Data mining, Time series, Metric learning, Weighted support vector machines, Intelligent integrated, network

1 Introduction

With the continuous development of information, big data analysis has become important and dependent technical means-increasingly in various fields. By data mining through time series, the development regular of the object can be grasped, so as to predict its future development trend. Time series modeling and prediction methods are generally divided into two categories: traditional method and intelligent method [1, 2]. Traditional methods include linear regression analysis, nonlinear regression analysis, auto-regressive sliding average (ARMA) modeling, partial least square method, and gray prediction. The intelligent technology such as expert system, fuzzy rules, neural network, and support vector machine is used in the intelligent method to realize the predictive modeling [3, 4]. The modeling of expert system refers to the description of the production process based on expert experience and knowledge, which is very

explanatory. However, its knowledge acquisition has bottlenecks and its learning ability is poor. Similar to the expert system, fuzzy logic also describes the production process according to the expert experience and knowledge. The difference is that the fuzzy reasoning method could deal with uncertain information well. The modeling technology based on fuzzy rules is also limited by the acquired knowledge and has the problem of low model precision [5, 6]. For complex forecasting problems, a modeling method is often unable to achieve the prediction accuracy. Therefore, it is necessary to integrate multiple modeling technologies and absorb the advantages of various modeling, so as to the purpose of accurate prediction is reached. Intelligent integration modeling is to integrate two or more modeling methods to realize the modeling of complex industrial processes in a certain way, at least one of which is the intelligent modeling method [7].

In this paper, an intelligent integration architecture is proposed, and an intelligent integration structure is given. For the time-series data of a kind of random noise disturbance, an auto-regression sliding average model of nested dual-population particle swarm algorithm is proposed by using a parallel nested modeling structure [8]. The least-square support vector machine based on probability density control is proposed for mining deterministic trend in data. The effectiveness of the proposed method was verified by a set of experiments [9].

The specific contributions of this paper include (1) a double weighted support vector machine algorithm based on category weighting and feature weighting is proposed; (2) solved the problem of imbalanced time series samples; (3) support vector machine classification based on cost-sensitive learning; (4) reconstructed Gaussian kernel functions with different feature weight coefficients.

The rest of this paper is organized as follows. Section 2 discusses intelligent integration structure, followed by the methods in Section 3. Experiment is discussed in Section 4. Section 5 concludes the paper with summary and future research directions.

2 Intelligent integration structure

Intelligent integration refers to two or more pattern mining methods that are integrated in a certain way to realize the mining of complex data rules or patterns, of which at least one is an intelligent modeling method. There are four main forms and structures of intelligent integration pattern mining [10].

2.1 Parallel complement integration structure

The parallel complement integration structure includes two sub-models, and the two models have no primary and secondary points and complement each other. The two sub-models in the structure are usually obtained by two modeling methods [11]. The single modeling method can mine part of the information in the time series data to know the corresponding law. However, due to the limitation of the method, all the information in the data cannot be obtained. The two modeling methods complement each other to fully exploit the laws or patterns implied in the data [12].

2.2 Weighted overlay integrated structure

The weighted superposition integration structure is composed of multiple sub-models weighted and superimposed, and the weight of each sub-model corresponds to its role

in the integration model. The multiple sub-models in the structure are usually obtained by a variety of modeling methods [13]. A single modeling method can mine part of the information in the time series data to know the corresponding law. However, due to the limitation of the method, all the information in the data cannot be obtained, so a variety of modeling methods complement each other to fully exploit the laws or patterns implied in the data [14].

2.3 Series integrated structure

The tandem integration structure consists of two or more sub-models. Except for the first and last models, each model is the output of the previous model and the input of the latter model. Nonlinear dynamic systems usually adopt this form. For example, the neural network is used to reflect the nonlinear characteristics of the system static, and the dynamic characteristics are characterized by NARMX (a nonlinear auto-regressive moving average with exogenous variables) [15].

2.4 Model nested integration structure

The nested integration structure includes at least two sub-models, one of which is called a base model, which is used to model the main structure of the industrial process, and the other sub-models are nested in the base model to build unknown parameters in the base model [16]. For example, the bionic algorithm such as the ant colony algorithm, particle swarm optimization algorithm, and genetic algorithm is applied to the system identification to realize the parameter estimation in the model.

The smart integrated architecture diagram is as follows (Fig. 1) [17]:

3 Methods

3.1 Data mining technology

In recent years, with the development of technologies such as data collection and storage, the data of the information society has exploded, and there has been a situation of "rich data and poor information." Massive data not only makes it difficult to distinguish useful data but also increases the complexity of data analysis. In order to solve this problem, data mining technology came into being [18, 19]. The birth of data mining aims to transform a large amount of data that could be widely used in society into useful knowledge and information for market analysis, fraud monitoring, hacking, product control, and scientific exploration.

In general, data mining can be divided into the following seven steps (Fig. 2):

(1) Data cleansing—eliminating noise and data not related to the mining theme.
(2) Data integration—integrating data from multiple data sources.
(3) Data selection—select data related to mining topics.
(4) Data transformation—using data such as normalization to transform data into a form suitable for data mining.
(5) Data mining—the core steps to mine knowledge using methods such as classification, fusion, and association rules.
(6) Mode evaluation—evaluation of the effect of the model, the commonly used indicators have accuracy, recall rate, etc.
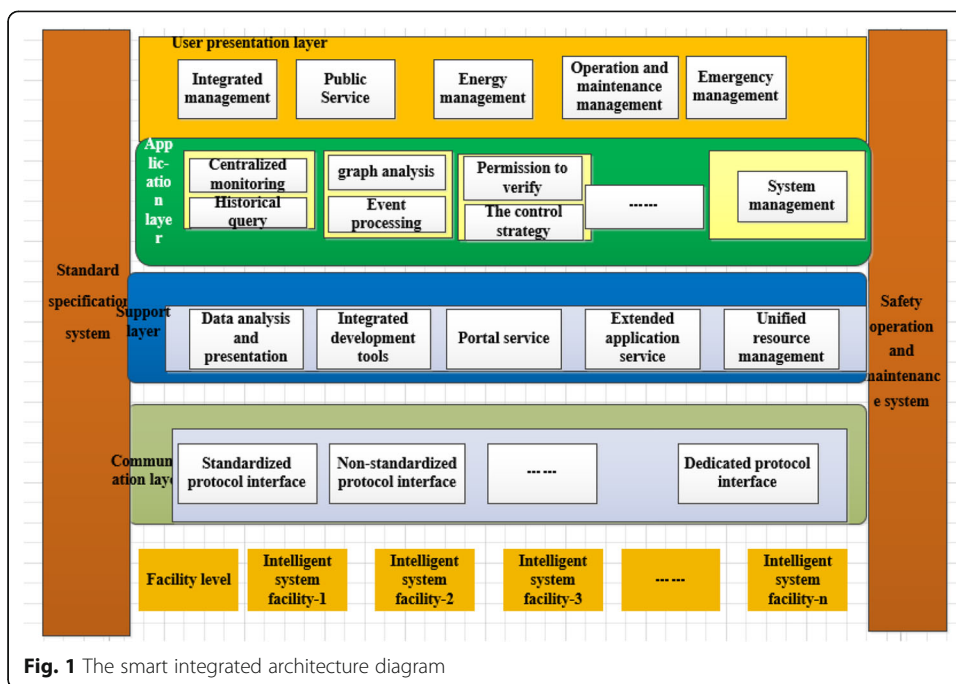
Fig. 1 The smart integrated architecture diagram

(7) Knowledge representation—the model is represented by a technically understandable model, and the knowledge obtained by the mining is presented to the user.

3.2 Time series data mining algorithm

The increasing non-equilibrium data in recent years has brought new challenges to data mining research. When a classification model is established for an unbalanced data set, the cost of misclassifying less types of data is more expensive than that of misclassified multi-class data, so the traditional classification method is not applicable to it. Nowadays, research on unbalanced data sets is concentrated [20]. There are two main types, one is to undersample multiple data, and the other is to generate some small data manually, but these two methods are for time series data which is not applicable. Ojha tried to set different classification penalty parameters for different categories of samples in the process of establishing classification model, and proposed weighted support
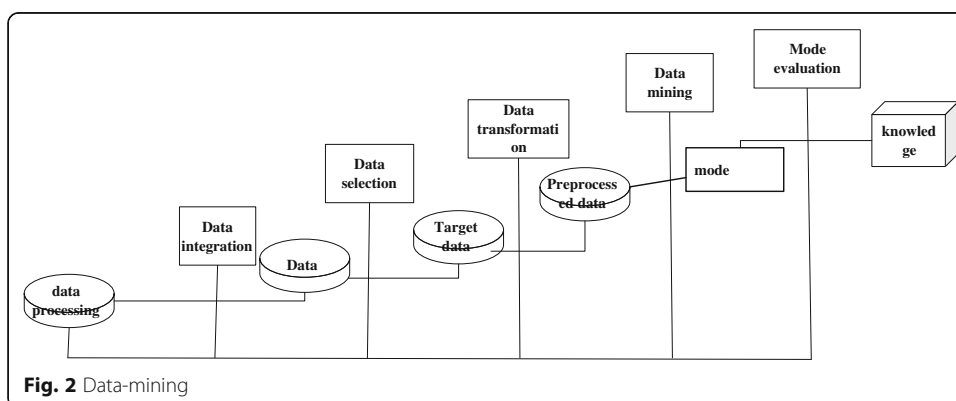

Fig. 2 Data-mining

vector machine WSVM [21]. In addition, Ram [13] et al. proposed a support vector machine FWSVM [22] which is based on feature weighting. Firstly, the information gain is used to evaluate the importance of each feature corresponding to the classification task, and the weights are respectively assigned, and then the weight coefficients are applied to the calculation of the SVM kernel function. Based on the sample weighted support vector machine and the feature weighted support vector machine, a dual-weighted support vector machine (DWSVM) is proposed, which maintains the advantages of WSVM and FWSVM.

The dual-weighted support vector machine algorithm DWSVM is proposed in this paper, first, different misclassification penalty coefficients to different sample categories when constructing the classification hyperplane is assigned, then the weights of each feature and reconstructs the kernel function is calculated, thus which makes the algorithm better, generalization and robustness.

3.3 Introducing different penalty factors

For the two-category problem, the multi-class is a negative class, and the other is a positive class. Different types of samples are given different error-discriminating coefficients:

$$
\begin{aligned}
&\min f(x) = \frac{1}{2}\omega^T\omega + C^+ \sum_{y=1}\xi_i^y + C^- \sum_{y=-1}\xi_i^y \\
&s.t. \ y\left(\omega^T\vartheta(x)\right) + b \geq 1-\xi_i \\
&\xi_i \geq 0, i = 1...T
\end{aligned}
\tag{1}
$$

In which, $C+$ is the penalty factor of the misclassified positive sample, $C-$ is the penalty factor of the misclassified small sample, and the problem of (1) is converted into the Wolfe dual problem by the convex quadratic programming method in the optimization theory. After getting it,

$$
\begin{aligned}
&\max_{\alpha} \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j y_i y_j \phi(x_i)(x_j) \\
&y_i = 1, 0 \leq \alpha_i \leq C^+ \\
&y_i = -1, 0 \leq \alpha_i \leq C^-
\end{aligned}
\tag{2}
$$

In which, $\phi(x_i)(x_j)$ is Kernel function.

3.4 Kernel function based on feature weighting

The definition of feature weighting is based on the degree of contribution of each feature of the sample to the pattern recognition, giving different weight coefficients. There are many methods for feature weighting, such as the information weight-based feature weighting method proposed by Wang Yan et al. The feature weighting method based on manifold learning is adopted here. A weighting matrix P is defined as follows:

$$
P = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix}
\tag{3}
$$

The function of the kernel function is to find the nonlinear mode by using the linear function in the eigenvector space established by the nonlinear feature map [23].

According to Theorem 1 and Theorem 2, the matrix P is introduced, the shape of the input geometric space can be scaled, and the geometry of the feature space can be scaled, thereby changing the weights assigned to different linear functions in the feature space during the modeling process.

Theorem 1: K is made a kernel function which is defined on XX, which is a mapping from input space to feature space. X-F, P is a linear transformation matrix, then

$$\| \vartheta(\chi_i) - \vartheta(\chi_j) \| \neq \| \vartheta(\chi_i) \| - \| \vartheta(\chi_j) \| \tag{4}$$

Theorem 2: If there is $1 < \_k < \_h$, then the kth eigenvector of the data set is independent of the calculation of the weighted kernel function, and has nothing to do with the output of the classifier. The smaller the ωk, the smaller the influence is on the calculation of the kernel function, and the smaller the effect is on the classification result.

Therefore, by introducing P into the Gauss radial basis kernel function, the following formula can be obtained.

$$f(\chi) = \operatorname{sgn}\left(\sum \alpha_i k_i \left(\| x_i^T P - x_j^T P \| \right) + b \right)$$
$$\sum k_i \left(\| x_i^T P - x_j^T P \| \right) = \sum \exp\left\{ -\frac{\| x_i^T P - x_j^T P \|}{\sigma^2} \right\} \tag{5}$$

Using the weighted kernel function in equation (5) for the support vector machine classifier, the classification model with weighted sample features can be obtained.

3.5 Double weighted support vector machine

According to the introduction of the previous two sections, the construction steps of the dual feature support vector machine (DWSVM) are as follows:

Step 1. Collects the data x, and the feature set in x is (fs, n), where n is the number of the feature, that is, the feature of X is represented by fs: $(f_1, f_2, ..., f_n)$.

Step 2. Calculate the weight coefficient ωi of the fi feature off by the MBFS method, and generate a linear transform weight matrix P which is based on ωi.

Step 3. Transform the Gauss kernel function with a linear transformation weighting matrix P, which is obtained a kernel function based on feature weighting (4).

Step 4. Construct a minimized structural risk function (1) which is based on the weighting of the two classification samples, and add the kernel function in equation (4) to the construction of the classification hyperplane is to establish a support vector machine classification model.

Step 5. Evaluate the obtained classifier.

3.6 Unbalanced classification evaluation index

In the process of establishing the *SVM* model, we continuously debug the parameters of the *SVM* (including the kernel function t, the penalty parameter c, the kernel function parameter g, and the weighting coefficients w0 and w1, etc.) to obtain better prediction results. Because in the unbalanced classification, classification accuracy could not be used as an evaluation index to measure classification performance, we also need to select appropriate model evaluation indicators according to actual needs. In addition

to introducing the classification correct rate Acc, the evaluation indicators such as G-mean, F-measure and auc-roc are also selected.

The classification correct rate Acc represents the proportion of the sample with the correct classification to all samples. The total number of samples is M, and the number of samples with the correct classification is TM.

Define the classifier's Acc.

$$\mathrm{Acc} = \frac{TM}{M} \tag{6}$$

The TP rate and FP rate of the classifier are given by the following definition, defining the TP rate of the classifier.

$$TP\ \mathrm{rate} = \frac{TP}{TP + FN} \tag{7}$$

Define F-prate of the classifier:

$$FP\ \mathrm{rate} = \frac{FP}{FP + TN} \tag{8}$$

First define the sensitivity and specificity of the classifier. Define the sensitivity of the classifier:

$$\mathrm{sensitivity} = TP/(TP + FN) \tag{9}$$

Define the specificity of the classifier:

$$\mathrm{specificity} = TN/(TN + FP) \tag{10}$$

By definition, the sensitivity is positive class sample accuracy, and specificity is negative class sample accuracy. Based on the above two indicators, Ku-bat et al. proposed a new metric G-means to evaluate the unbalanced classification, which is given by the following definition.

Define the classifier's G-means:

$$G-\mathrm{mean} = \sqrt{\mathrm{specificity} + \mathrm{sensitivity}} \tag{11}$$

From the definition point of view, G-means takes into account the positive and negative precision of the class, and can better reflect the comprehensive performance of the classifier. Many researchers use G-mean as a measure when evaluating unbalanced classification performance.

In some special applications, more attention is paid to the classification performance of sample positive categories, such as credit card fraud detection, customer churn in telecommunications, arrears forecast, intrusion r in intrusion detection, abnormal state, and disease monitoring in medical diagnosis. Wait F-measure is mainly used to measure the classification effect of positive samples.

First, the definitions of the classifier precision and recall are given.

Define the classifier's precision:

$$\mathrm{precision} = \frac{TP}{TP + FP} \tag{12}$$

Define the classifier's recall:

$$\text{recall} = \frac{TP}{TP + FN} \tag{13}$$

By definition, precision is the positive class sample coverage, and recall is the positive class sample accuracy. Based on the above two indicators, *F*-measure is given by the following definitions:

Define the *F*-measure of the classifier:

$$F\text{-measure} = \frac{(\beta + 1)\text{precision} \times \text{recall}}{\beta\text{precision} + \text{recall}} \tag{14}$$

Usually $\beta = 1$. It can be seen from the above definition that *F*-measure fully embodies the classification performance of positive classes and is a trade-off between positive coverage and accuracy. Many researchers use *F*-measure as a measure when evaluating unbalanced classification performance.

Define the AUC of the classifier:

$$\text{AUC}(f) = \frac{\sum_{i=1}^{n^+} \sum_{j}^{n^-} f(x^+) > f(x^{1-})}{n^+ \times n^-} \tag{15}$$

The middle $n+$ is the number of samples of all the minority classes, and $n-$ is the number of samples of all the majority classes. For any sample of a few classes, if the probability that the classification algorithm f can divide it into a minority class is greater than the probability of dividing it into a majority class, the value of $f(x+) > f(x-)$ is 1, otherwise 0, the same is true for the AUC for solving most classes. Then multiply the two and divide by the product of the minority class and the majority class to get the AUC value.

4 Experiment

In order to verify the classification performance of the dual-weighted support vector machine in time series data, this chapter uses the Libsvm toolkit and modifies it accordingly, and then uses MATLAB to perform dual-weighted support vector machine modeling on the real data set.

Table 1 Data format before dimension compression

Gender	Breakfast Frequency	Breakfast Amount	Lunch Frequency	Lunch Amount	Dinner Frequency	Dinner Amount	Total consumption Amount
1	30	120	15	335	12	200	655
1	28	178	22	400	18	260	838
1	28	298	21	253	23	300	851
2	15	265	12	200	30	310	775
2	16	153	19	210	15	280	643
2	23	256	25	310	12	223	789
2	22	123	25	260	12	245	628

Fig. 3 Data model diagram
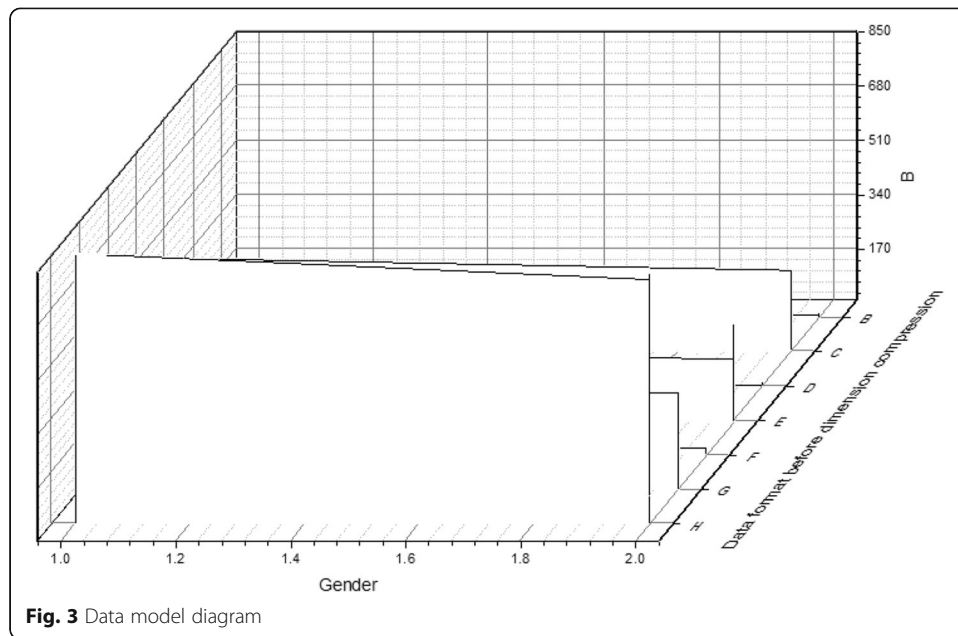
4.1 Data set introduction

This experiment used a university student card consumption data set (2011.09–2011.10). Among them, the consumption data comes from the logistics department of the university, the list of impoverished students comes from the school, the sign of the poor students is 1, and the mark of non-poor students is 0. The sample size of the consumption data in September and October after pre-processing is 8093 × 10, and the data format is shown in Table 1.

By establishing a classification model for the data set, it is possible to judge whether it is a poor student based on a student's card consumption data, to help the school to identify poor students and to see if there are poor students who have not applied for subsidies for other reasons. The data model diagram is shown in Fig. 3.

4.2 Parameter settings

The pre-processed September card consumption data set is used to establish a dual-weighted SVM model, and the G value, F value, and Acc value of the October data prediction result under different parameter settings are recorded.

Table 2 DWSVM model mesh optimization algorithm _G value

c/g	1	2	4	8	16	32	64	128
1	0.9677	0.9675	0.9673	0.9671	0.9669	0.8123	0.7923	0.6569
2	0.9677	0.9675	0.9673	0.9671	0.9669	0.8123	0.7923	0.6569
4	0.9677	0.9675	0.9673	0.9671	0.9669	0.8123	0.7923	0.6569
8	0.9677	0.9675	0.9673	0.9671	0.9669	0.8123	0.7923	0.6569
16	0.9677	0.9675	0.9673	0.9671	0.9669	0.8123	0.7923	0.6569
32	0.9075	0.9073	0.9071	0.9069	0.9067	0.7025	0.6825	0.5471
64	0.8176	0.8174	0.8172	0.817	0.8168	0.8143	0.7943	0.6589
128	0.7623	0.7621	0.7619	0.7617	0.7615	0.9088	0.8888	0.7534

Table 3 DWSVM model mesh optimization algorithm _F value

c/g	1	2	4	8	16	32	64	128
1	0.6677	0.6675	0.6673	0.6671	0.6669	0.6342	0.6142	0.4788
2	0.6677	0.6675	0.6673	0.6671	0.6669	0.6342	0.6142	0.4788
4	0.6677	0.6675	0.6673	0.6671	0.6669	0.6342	0.6142	0.4788
8	0.6677	0.6675	0.6673	0.6671	0.6669	0.6342	0.6142	0.4788
16	0.6677	0.6675	0.6673	0.6671	0.6669	0.6342	0.6142	0.4788
32	0.6075	0.6073	0.6071	0.6069	0.6067	0.6025	0.5825	0.4471
64	0.7176	0.7174	0.7172	0.717	0.7168	0.6243	0.6043	0.4689
128	0.4623	0.4621	0.4619	0.4617	0.4615	0.5088	0.4888	0.3534

1) Select the kernel function. In libsvm, t–0, 1, 2, 3, 4 represent linear kernel functions, polynomial kernel functions, RBF kernel functions, sigmoid kernel functions, and custom kernel functions, respectively. According to the above introduction, this paper uses the kernel function WRBF based on feature weighting, and the parameter is set to $t = 4$.

2) The penalty parameter c and the kernel function parameter g are selected. After selecting the kernel function, use the grid optimization algorithm to find the best c value and g value. According to experience, set the parameter range of c and g to $[2 \wedge (-5), 2 \wedge (5)]$; the step length is 1. As shown in Tables 2, 3, and 4, when c, g is between $2 \wedge (-5)$ and 1, the DWSVM model obtains the best prediction. In order to facilitate the subsequent experiments, we choose $c = 1$, $g = 1$ as the DWSVM model parameters. At this time, the F value is 0.6733, and the G value is 0.9088 and the Acc value is 89.00%.

3) Determination of weighting coefficients w0 and w1. In libsvm, the two-category samples are weighted with parameters w0 and w1. Under the premise of $s = 0$, $t = 4$, $c = 1$, $g = 1$, three experiments are carried out in this paper, as shown in Tables 5, 6, and 7: The first group, w0 = 1, w1 from 1 to 1.9 G value, F value, Acc value, TP value, and TN value are recorded in steps of 0.1; the second group, w1 = 1, w0 changes from 1 to 1.9 in steps of 0.1, and the G value and the F value are recorded. Acc value, TP value, and TN value; the third group, w1 = 1, w1 changes from 1 to 20, and records G value, F value, Acc value, TP value, and TN value. We get that

Table 4 DWSVM model mesh optimization algorithm _Acc value (%)

c/g	1	2	4	8	16	32	64	128
1	89	88.9998	88.9996	88.9994	88.9992	89.65	89.63	89.4946
2	88	87.9998	87.9996	87.9994	87.9992	89.65	89.63	89.4946
4	88	87.9998	87.9996	87.9994	87.9992	89.65	89.63	89.4946
8	88	87.9998	87.9996	87.9994	87.9992	89.65	89.63	89.4946
16	88	87.9998	87.9996	87.9994	87.9992	89.65	89.63	89.4946
32	88.06	88.0598	88.0596	88.0594	88.0592	89.65	89.63	89.4946
64	79.08	79.0798	79.0796	79.0794	79.0792	83	82.98	82.8446
128	87.47	87.4698	87.4696	87.4694	87.4692	76.05	76.03	75.8946

Table 5 DWSVM model w1 selection

w1	1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
G	0.8754	0.8478	0.8976	0.8974	0.8972	0.897	0.8968	0.8966	0.8964	0.8962
F	0.6873	0.6871	0.6869	0.6867	0.6865	0.6863	0.6861	0.6859	0.6857	0.6855
Acc	89.12	88.92	88.72	88.52	88.32	88.12	87.92	87.72	87.52	87.32
TP	975	981	987	993	999	1005	1011	1017	1023	1029
TN	6286	6268	6250	6232	6214	6196	6178	6160	6142	6124

when $w0 = 1.5$, $w1 = 1$, the DWSVM model has the best prediction effect, at this time $G = 0.8206$, $F = -0.6917$, Acc = 90.88%, TP = 828, TN = 6527.

In summary, when $t = 2$, $c = 1$, $g = 1$, $w0 = 1.5$, $w1 = 1$, the prediction result of the DWSVM model is the same and can be obtained in the unweighted support vector machine model, when $t = 2$, $c = 1$, $g = 1$, the classification effect is the best, the values of G-mean, F-measure, Acc, TP, FP are 0.8408, 0.6234, 83.78%, 1081, and 6239 respectively.

4.3 Model effect evaluation

Since the training set and test set are randomly selected in the data set in the experiment, different distribution methods have a certain impact on the experimental results. In order to make the experimental results more convincing, this experiment uses a tenfold cross-validation, taking an average of 10 experiments as the final result. For the classification of unbalanced samples, the classification accuracy of negative samples is the key to the evaluation model. In the card consumption data for October, there were 1237 positive samples and 6856 negative samples. In the classification model established in this paper, the average classification accuracy rate Acc reached 91.28%; meanwhile, the values of TP and FP were 828 and 6527, respectively, that is, the correct classification rates of positive and negative categories were 67% and 95.2%. The values of G-mean and F-measure are 0.8206 and 0.6917, respectively. Figure 4 is a sample graph for predicting October data:

Figure 5 is the ROC curve for which the October data is predicted.

It can be seen that the overall effect of the classification model is better, which meets the requirements of the non-equilibrium sample classification model. If there is no classification model assistance and complete manual guessing, the average classification accuracy of positive samples is 50%. However, with the classification model, the classification accuracy of positive samples reaches 67%, which is 34%, which has practical guiding significance.

Table 6 DWSVM model w0 selection

w0	1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
G	0.8854	0.8478	0.8976	0.8974	0.8972	0.897	0.8968	0.8966	0.8964	0.8962
F	0.6273	0.6271	0.6269	0.6267	0.6265	0.6263	0.6261	0.6259	0.6257	0.6255
Acc	89	88.8	88.6	88.4	88.2	88	87.8	87.6	87.4	87.2
TP	918	924	930	936	942	948	954	960	966	972
TN	6373	6355	6337	6319	6301	6283	6265	6247	6229	6211

Table 7 DWSVM model weighting coefficient selection

w1 (w0 = 1)	G	F	Acc	TP	TN
1	0.8408	0.6234	83.78	1081	6239
2	0.8388	0.6204	83.28	1075	6207
3	0.8368	0.6174	82.78	1069	6175
4	0.8348	0.6144	82.28	1063	6143
5	0.8328	0.6114	81.78	1057	6111
6	0.8308	0.6084	81.28	1051	6079
7	0.8288	0.6054	80.78	1045	6047
8	0.8268	0.6024	80.28	1039	6015
9	0.8248	0.5994	79.78	1033	5983
10	0.8228	0.5964	79.28	1027	5951
11	0.8208	0.5934	78.78	1021	5919
12	0.8188	0.5904	78.28	1015	5887
13	0.8168	0.5874	77.78	1009	5855
14	0.8148	0.5844	77.28	1003	5823
15	0.8128	0.5814	76.78	997	5791
16	0.8108	0.5784	76.28	991	5759
17	0.8088	0.5754	75.78	985	5727
18	0.8068	0.5724	75.28	979	5695
19	0.8048	0.5694	74.78	973	5663
20	0.8028	0.5664	74.28	967	5631

5 Results and discussion

In this paper, a dual-weighted support vector machine model DWSVM based on sample class weighting and sample feature weighting is proposed, and the verification experiment is carried out on the actual data set. The experimental results show that the DWSVM model has a good classification effect on unbalanced time series, which meets

Fig. 4 Sample plot of the DWSVM model

Fig. 5 ROC graph of DWSVM model
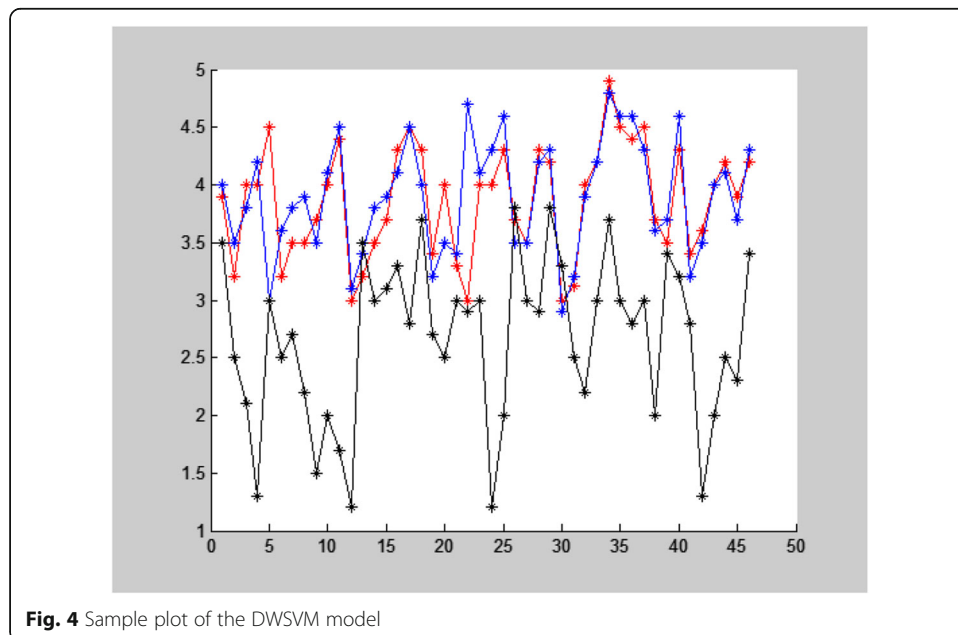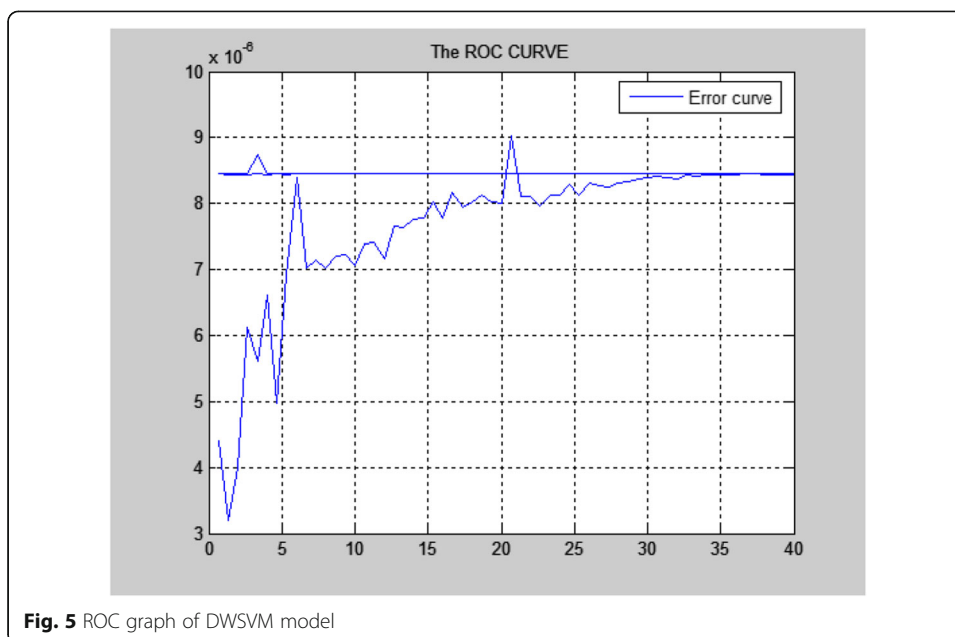
the requirements of practical applications and improves the performance of unweighted support vector machines.

In this paper, the characteristics of time series and the classification difficulties of unbalanced data sets are studied. Under the topology design of intelligent integrated architecture, based on the previous research, a double-weighted support vector machine DWSVM based on sample class weighting and sample feature weighting is proposed. In the aspect of sample category weighting, the idea of cost-sensitive learning is used. In the aspect of sample feature weighting, the feature weight calculation method in MBFS algorithm is adopted. The introduction of these two weighting methods makes the SVM more generalized and robust. The whole system has perfect functions, stable performance, strong real-time performance, good scalability, accurate data communication, and fully meets the intelligent integrated architecture. The time series dimension is large, and the sample unbalanced mining difficulty is used to verify that the practical application value of time series data mining is large and difficult, and has broad research prospects.

Abbreviations
ARMA: Agung Rai Museum of Art

Authors' contributions
Hui Xu wrote the entire article. The author(s) read and approved the final manuscript.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication
All authors agree to submit this version and claim that no part of this manuscript has been published or submitted elsewhere.

Competing interests
The authors declare that they have no conflict of interest.

Author details
[1]School of Economics and Management, Xi'an Aeronautical University, Xi'an 710077, People's Republic of China.
[2]School of Automobile, Chang'an University, Xi'an 710064, People's Republic of China.

References
1.   G. Ste, M. Gerard, D.H. Mil, Data mining over biological datasets: an integrated approach based on computational intelligence. Comput. Intell. Magazine IEEE **7**(4), 22–34 (2012)
2.   Y. Huang, Z. Zhao, X. Hua, et al., Advances in systems biology: computational algorithms and applications. Bioc. Systems Biol. **6**(3), 1–5 (2012)
3.   J. Zhang, Research on time series data mining based on intelligent integrated particle swarm optimization algorithm. J. Sichuan Univ. Sci. Eng. **3**(6), 20–28 (2015)
4.   Y. Chan, Visualization and ontology of geospatial intelligence. Sensors Actuators B Chem. **177**(1), 397–403 (2009)
5.   K.J. Kim, S.B. Cho, Meta-classifiers for high-dimensional, small sample classification for gene expression analysis. [J]. Pattern. Anal. Applic. **18**(3), 553–569 (2015)
6.   E. Fery, E. Messina, F. Arc, A p-Median approach for predicting drug response in tumour cells. BMC Bioinformatics **15**(1), 353 (2014)
7.   M. Glykas, Fuzzy cognitive strategic maps in business process performance measurement. Expert Syst. Appl. **40**(1), 1–14 (2013)
8.   W. Guo, T. Xu, Z. Lu, An integrated chaotic time series prediction model based on efficient extreme learning machine and differential evolution. Neural Comput. Applic. **27**(4), 1–16 (2016)
9.   H.J. Liao, C.H.R. Lin, Y.C. Lin, et al., Intrusion detection system: a comprehensive review. J. Netw. Comput. Appl. **36**(1), 16–24 (2013)
10. T. Ojha, S. Misra, N.S. Rag, Wireless sensor networks for agriculture: the state-of-the-art in practice and future challenges. Comput. Electron. Agric. **118**(3), 66–84 (2015)
11. G. Yu, L. Zhou, Research on an intelligent algorithm for public network attack data mining. Comp. Measur. Control **5**(3), 331–339 (2016)
12. J. Jaw, P. Kłe, Automatic classification of specific melanocytic lesions using artificial intelligence. Biomed. Res. Int. **27**, 893–902 (2016)
13. S. Ram, P.S. Rao, Collaborative concealment of spatio-temporal mobile sequential patterns. Glob. J. Comp. Sci. Technol. **26**(2), 82–99 (2012)
14. D. Xu, The intelligent information service based on the fourth paradigm. Res. Library Sci. **2**(3), 23–35 (2016)
15. T. Jing, T. Huang, G. Pan, A research based on the intelligent robot education to rehabilitation of children with autism. J. Quanzhou Normal Univ. **23**(2), 123–135 (2016)
16. H.Y. Zhao, G.X. Li, H.L. Zhang, et al., An improved algorithm for segmenting online time series with error bound guarantee. Int. J. Mach. Learn. Cybern. **7**(3), 365–374 (2016)
17. Y. Xue, M. Zhang, Z. Liao, et al., A contiguous column coherent evolution bi-clustering algorithm for time-series gene expression data. Int. J. Mach. Learn. Cybern. **9**(3), 441–453 (2018)
18. S.S. Chung, A class of semiparametric volatility models with applications to financial time series. Dissert. Theses-Gradworks **123**(3), 439–444 (2016)
19. B. Wu, T.T. Cheng, T.L. Yip, Y. Wang, Fuzzy logic based dynamic decision-making system for intelligent navigation strategy within inland traffic separation schemes. Ocean Eng. **197**, 106909 (2020)
20. Z. Huang, X. Xu, J. Ni, H. Zhu, W. Cheng, Multimodal representation learning for recommendation in internet of things. IEEE Internet Things J. **6**(6), 10675–10685 (2019)
21. L. Dong, W. Wu, Q. Guo, et al., Reliability-aware offloading and allocation in multilevel edge computing system. IEEE Trans. Reliab. (2019). https://doi.org/10.1109/TR.2019.2909279
22. W. Wei, H. Song, W. Li, P. Shen, A. Vasilakos, Gradient-driven parking navigation using a continuous information potential field based on wireless sensor network. Inf. Sci. **408**(2), 100–114 (2017)
23. Z. Huang, X. Xu, H. Zhu, M.C. Zhou, An efficient group recommendation model with multiattention-based neural networks. IEEE Transac. Neural Netw. Learn. Syst. (2020). https://doi.org/10.1109/TNNLS.2019.2955567

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.