Journal of Network Intelligence ©2024 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **9**, Number **3**, August **2024**

Financial Early Warning Model Based on Semi-supervised Learning and Transfer Learning

Jiang-Yan Cheng*

College of Business Management
Fujian Polytechnic of Information Technology, Fuzhou 350002, P. R. China
18050280088@163.com

Hong Liu

School of Information Technology and Engineering
St. Paul University Philippines, Tuguegarao 3500, Philippines
uj2860@163.com

*Corresponding author: Jiang-Yan Cheng
Received December 11, 2023, revised March 25, 2024, accepted May 30, 2024.

Abstract. *Listed companies are facing different kinds and degrees of financial risks in the actual operation process, which may lead to potential uncertainty in the expected return of the enterprise due to the influence of multiple factors, such as internal operation decisions and changes in the external market environment. For the purpose of the above issues, this paper designs a financial early warning model on the ground of semi-supervised learning and migration learning. Firstly, to address the problem of poor generalization effect of current transfer learning algorithms, Gaussian kernel is introduced to make it suitable for realistic nonlinear problems, and a semi-supervised transfer learning method with Gaussian kernel is obtained by regularization. Then, based on the optimization of the semi-supervised migration learning algorithm based on the above discriminant analysis, the financial early warning indicator system is constructed, and the common knowledge of the financial data between the cross-domain heterogeneous labeled source domain dataset and the target domain dataset is effectively acquired through the migration strategy, and based on the dynamically updated target domain, the migration and semi-supervised mechanism is processed to achieve the symbiotic fusion of the migration learning and the semi-supervised learning and to gradually improve the early warning capability of the company's finance. ability. The simulation results show that the accuracy rate, precision rate, recall rate, F1 and AUC of the model designed in this paper are 93.18%, 92.73%, 95.84%, 94.26% and 93.82%, respectively, which are better than the comparative model, and have a better early warning performance.*
Keywords: Semi-supervised learning; Transfer learning; Discriminant analysis; Symbiotic integration; Financial warning

---

1. **Introduction.** With the deepening of the globalization of the world economy, the scale of the domestic capital market is expanding. On the one hand, the expanding market attracts more shareholders to join; on the other hand, the market economic environment is complicated and changeable, and the financial crisis of listed companies appears frequently [1, 2]. In this case, if a listed company has a financial crisis, it will have a great impact on business operators, owners, investors and other stakeholders, and even affect the stability of the order of the securities market, which makes the company's financial early warning become more important. How to build a financial early warning system adapted to the advancement of China's enterprises, for the purpose of identifying and prevent financial

crises earlier, has become a concern of enterprises and the market [3]. The development of information technology provides more effective technical means for financial early warning and promotes the depth of financial early warning research. Under the continuous efforts of scholars at home and abroad, from univariate model to multivariate model to the current neural network model, the research methods related to corporate financial early warning are maturing, and the prediction accuracy is gradually improving [4, 5, 6]. How to create a more suitable financial early warning model for listed companies by quantitative and qualitative research on their financial data as well as non-financial data from enterprise operation and management data is the focus of the current research, and machine learning is the favored object of many researches in terms of many model choices.

1.1. **Related Work.** Beaver [7] was the first to incorporate cash flow indicators into the indicator system in his early warning study of financial crisis. Blum [8] found the effectiveness of cash flow indicators in early warning study of financial crisis through the comparison of model effects. Ohlson [9] pointed out that there are problems with the widely used multivariate discriminant analysis methods, such as the nature of the distribution of predictor variables has certain statistical requirements, and there is no intuitive interpretation of the output scores, etc. Korobow et al. [10] were the first to try to use neural network algorithms to classify the bond rating of enterprises. Tam and Kiang [11] chose to combine the neural network model with the banking industry to study the model's early warning of the financial risk of banks. Sun and Li [12] constructed a default risk prediction model based on logistic regression. Hong et al. [13] used five indicators used by Z-score model as input variables of neural network algorithm model for empirical study. Rosa and Gartner [14] tried to combine deep learning neural network model with stock market to study the model's prediction of stock rise and fall, and found that the deep learning model has a strong adaptability to the stock indicators. Hosaka [15] innovatively transformed the annual report data of listed companies into images and used convolutional neural networks to conduct risk warning research, and the empirical results verified the feasibility of the research method.

Subsequently, Valaskova et al. [16] applied financial data to logistic regression analysis and concluded that indicators such as cash ratio and quick ratio can best influence the future financial development of the company. Du et al. [17] combined factor analysis and logistic regression to build a model for predicting financial risk, which has a high ability to determine risk. Ouyang and Lai [18] constructed a financial risk identification model based on logistic regression, and the overall accuracy of the model reached 87.16%. Costa et al. [19] used a logistic regression model to assess the default risk of consumer loans based on the credit scoring information of financial institutions.

Support vector machine, neural network and other machine studying models have good fault tolerance for data and high prediction accuracy, but the theory is more abstract and computationally difficult. Gresnigt et al. [20] substituted the early warning indicators of the Z-Score model into the neural network model and compared it with the multivariate discriminant analysis method, and found that the forecasting ability of the artificial neural network mechanism is stronger. Nanayakkara and Azeez [21] established the Z-Score financial early warning model by multivariate discriminant analysis and stepwise regression, and the study tested the financial early warning ability of the model using data respectively. Wu et al. [22] used regression analysis to assess the financial hazard of a firm. Sdino et al. [23] highlighted that the importance of precociously warning of financial risk of a company increases significantly with the deep financialization of the market. Zhang and Luo [24] constructed an artificial neural network model to forecast the financial threat of a company. Mselmi et al. [25] used neural networks and sustain vector machines to predict the financial risk of small and medium-sized enterprises (SMEs), and concluded that firms with poor liquidity and profitability are more able to fall under financial crises. Shi and Li [26] pointed out that logistic regression and neural network models have gone to be the most expansively adopted in the field of risk prediction. Meanwhile, Shin et al. [27] analyzed the comparative effect of sustain vector machine and neural network on the prediction of corporate risk in Korean companies, and the results certified that the sustain vector machine is more effective in early warning.

1.2. **Contribution.** For the purpose of the issue of low accuracy in the current corporate financial preciously warning model, this article designs a financial preciously warning model on the ground of semi-supervised learning and migration learning. Firstly, for the problem of difficulty in estimating the distribution parameters of the current migration learning algorithm, a semi-supervised Gaussian kernel discriminant analysis method is obtained by means of regularization to migrate the data in the origin domain, so that the reusable data in the source domain can be correctly and efficiently migrated to the target task. Then, based on the optimization of the semi-supervised migration learning algorithm of the above discriminant analysis, the financial early warning indicator data are deviation normalized, and with the constructed indicator dataset as the target domain, the common scholarship of the financial data among the cross-domain heterogeneous labeled source domain dataset and the target domain dataset is effectively obtained through the migration strategy and the filtering mechanism is set up to achieve the micro-volume and effective migration of the samples of the source domain, and to gradually improve the company's financial early warning performance. The experimental outcome indicate that the model designed in this article has high accuracy, precision and recall, and is able to efficiently realize the early warning of company's finance.

2. Relevant theoretical analysis.

2.1. **Theory of semi-supervised learning algorithm.** Semi-supervised learning improves classification learning by introducing unlabeled samples into the training data, and the destination is generally to find a classification operation that not only has a small classification error on the labeled training data, but also requires the classification function to have consistent classification results based on the data distribution of the unlabeled samples. In this way, the data distribution of the data in the input space can be obtained from the unlabeled samples, and then the classification training effect of the classifier can be improved by utilizing the information of these data distributions [28]. As an abstract learning method, semi-supervision defines an optimization objective function based on interval loss Equation (1), where $y_j\mathcal{F}(x_j)$ is called the classification interval of the training samples, where N is a strictly decreasing function, and β is the training sample weights. And the objective function is solved iteratively by maximizing $- \prec \nabla C(F), f \succ$

$$-\nabla\mathcal{C}(F), f \propto D(F) = \sum_{x_j \in S_j} \beta_j N(y_j, F(x_j)) \tag{1}$$

Semi-supervised learning builds on fully supervised learning by introducing unlabeled training samples into the learning process. The interval of unlabeled training samples is defined as $|\mathcal{F}(x_j)|$, which leads to the new optimization objective Equation (2) containing unlabeled samples.

$$D(F) = \sum_{\xi_j \in S_j} \beta_j N(y_j, F(x_j)) + \sum_{x_j \in S_{cs}} \beta_j N(|F(x_j)|) \tag{2}$$

Since the absolute value function is not derivable, this is solved here by introducing the concept of pseudo-labeling. In this way, we can iteratively solve the newly obtained optimization objective Equation (2) by maximizing Equation (3).

$$- \prec \nabla D(F), f \succ = - \sum_{x_j \in S_j \cup S_H} \beta_j y_j f(x_j) N'(y_j F(x_j)) \tag{3}$$

Since M is strictly decreasing, the M' function will always be positive and greater than zero, so maximizing Equation (3) is equivalent to finding the right f to maximize Equation (4).

$$- \prec \nabla D(F), f \succ = - \sum_{y_j = f(x_j)} \beta_j N'(yF(x_j)) + \sum_{y_j \neq f(x_j)} \beta N'(y_j F(x_j)) \tag{4}$$

where the labeled sample data y_j is the real label of the sample, while the unlabeled sample data is the pseudo-label. Maximizing Equation (4) by dividing both sides by $- \sum_{x_j \in S_j \cup S_j} \beta_j N'(y_j F(x_j)))$ is equivalent to minimizing Equation (5).

$$\sum_{y_j \neq f(x_j)} C(j) - \sum_{y_j = f(x_j)} C(j) = 2 \sum_{y_j = f(x_j)} C(j) - 1 \tag{5}$$

where $C(j) = -[\beta_j N'(y_j F(x_j))]/[\sum_{x_i \in S_j \cup S_*} \beta_i N'(y_i F(x_i))]$.

2.2. **Transfer learning.** Migration learning is a type of migration that directly utilizes samples from the origin domain and samples from the object domain [28], requiring a lot of labeled data in the origin domain and only a little (or no) tagged data in the object domain, and the basic framework is shown in Figure 1. A common task of migration learning is domain adaptation, in which the sample space of the origin and object domains is the same, but the joint distribution of the data is different $p_T(x, y) \neq p_S(x, y)$. Due to $\hat{P}(x, y) = p(x|y)\hat{P}(y) = \hat{P}(y|x)p(x)$, according to the above equation, it can be obtained that the data distribution inconsistency is mainly caused by three aspects.

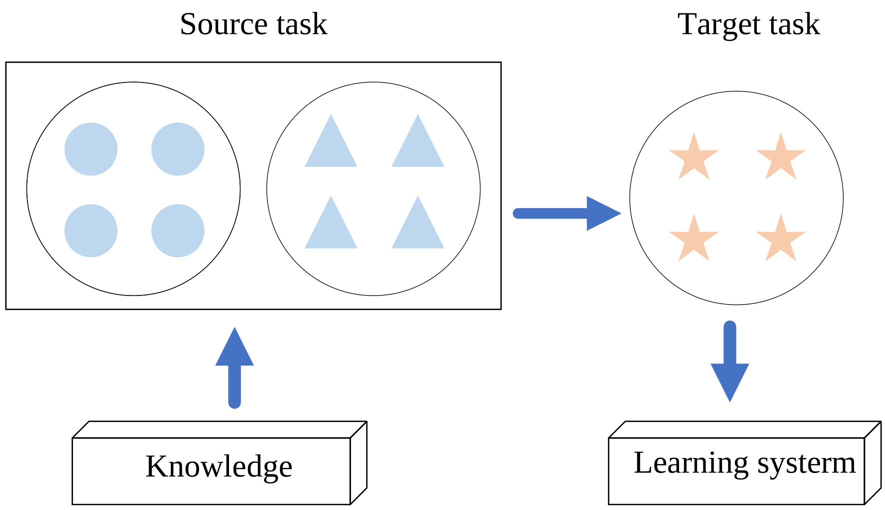

Figure 1. Basic framework of transfer learning

(1) The tasks in the origin and object areas are the same, however the input negligible dispersions $p_T(x) \neq p_S(x)$ are various, and the posterior dispersions $p_T(y|x) = p_S(y|x)$ are the same.

(2) The tasks on the origin and object areas are different, however the input marginal distributions $p_T(x) = p_S(x)$ are the same and the posterior distributions $p_T(y|x) \neq p_S(y|x)$ are different.

(3) The output dispersions $p_T(y) \neq p_S(y)$ of the origin and object areas are different, and the conditional distributions $p_T(x|y) = p_S(x|y)$ are the same. The key to solving this type of area adjustable problem is to study a domain-invariant representation. The purpose of area adjustable is to study a model $U : X \to Y$ makes:

$$\mathcal{S}_R(\delta_U) = \mathbb{E}_{(x,y)\sim p_R(x,y)}\left[L(U(\kappa, \delta_U), y)\right] \tag{6}$$

where L is the loss function and δ_U is the parameter of the model. Assuming that there is a mapping function V that maps x into a new feature space where the edge dispersions of the origin and object areas are the same, i.e., $p_R(v(x;\delta_V)) = p_R(v(x;\delta_V))$, v is the parameter of the mapping function, then the objective function is:

$$\mathcal{S}_R(\delta_U, \delta_V) = \mathbb{E}_{(x,y)\sim p_R(x,y)}\left[L(U(v(x;\delta_V);\delta_U), y)\right] + \chi_w(\mathcal{R}, \mathcal{S}) \tag{7}$$

The learning objective is to make the extracted features domain-independent and minimize the loss on the source domain. The first half of the formula is the expected risk function on the source domain, $\chi_w(\mathcal{R}, \mathcal{S})$ is a difference measure function of the distributions, which is used to calculate the distance between the sample distributions of the source domain and the object domain in the characteristic space, and χ is used to balance the importance ratio between the two parts.

3. **Semi-supervised migration learning algorithm optimization based on discriminant analysis.** Aiming at the current migration learning algorithm's difficulty in estimating distribution parameters and poor generalization effect, this paper proposes a semi-supervised migration studying algorithm based on positive discriminant analysis, and the structure of the algorithm is shown in Figure 2. Firstly, a Gaussian kernel is introduced on the basis of transfer learning, which is suitable for realistic nonlinear problems, and a semi-supervised Gaussian kernel discriminant analysis method is obtained by regularization. Second, based on the regularized discriminant analysis, the data in the origin area are migrated. The untagged data in the object domain is introduced by adding pseudo-labels in the migration, for the purpose of adopting the category message of labeled data and the distribution information of untagged data at the same time, which improves the generalization ability of the migration. In addition, a distance measure, an indicator matrix, and a new discriminant criterion function are defined to select the source data that is closest to the data in the object area. The source domain data is initially filtered using the projection matrix from the regularized discriminant analysis, and the source domain data is filtered with repeated corrections using the projection matrix obtained from the new discriminant function. In the end, the reusable data in the origin area can be correctly and efficiently transmigrated to the target task. The optimized algorithm avoids the direct estimation of domain distribution parameters and migrates part of the data in the origin domain. The process is described below.

Given N sample data with d-dimensional features $(x_1, x_2, \ldots, x_m)$, there are C categories $h_j (j = 1, 2, \ldots, C)$. The discriminant criterion function for regularized semi-supervised discriminant analysis is defined as:

$$D(H) = tr\left(\frac{H^T R_a^{\varphi} H}{H^T \left(R_h^{\varphi} + \beta\left(R_M - R_N\right)\right) H}\right) \tag{8}$$

where H is the projection matrix; H^T is the transpose of H; $\mathrm{tr}(\cdot)$ represents the trace of the matrix; R_a^{φ} represents the inter-class scatter matrix of the instance data in the

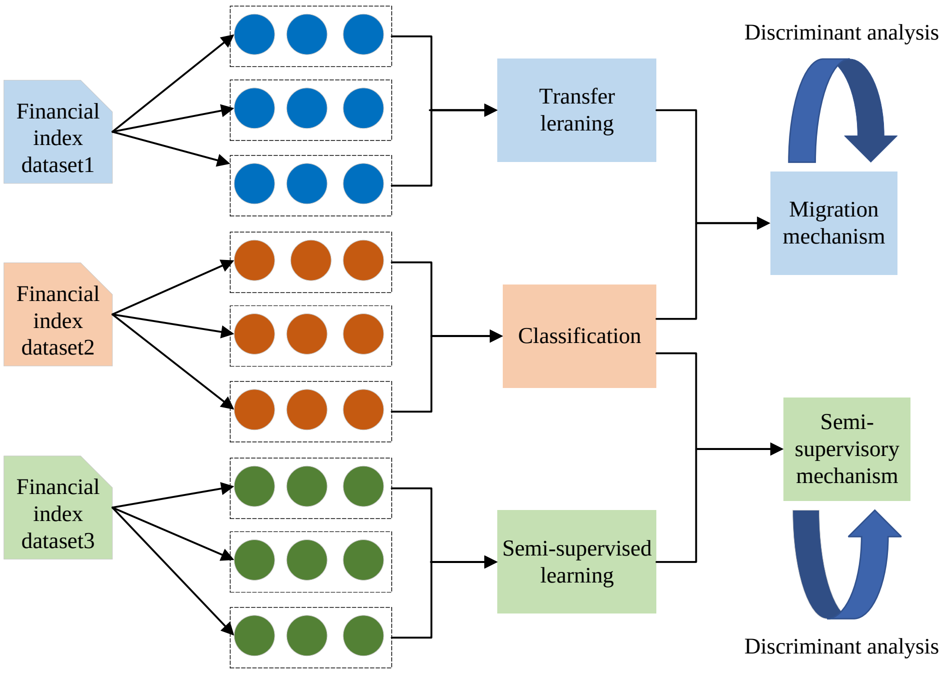

Figure 2. The structure the improved algorithm

high-dimensional characteristic space F; R_h^{φ} denotes the intra-class scatter matrix of the instance data in the high-dimensional characteristic space F; β is the adjustment coefficient of the regularity term; R_M is the extension scatter matrix; R_N is the non-local scatter matrix, and $(R_M - R_N)$ is the regularity term. The local and non-local scatter matrices are defined as:

$$R_M = \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}(x_i - x_j)(x_i - x_j)^T \tag{9}$$

$$R_N = \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} (1 - W_{ij})(x_i - x_j)(x_i - x_j)^T \tag{10}$$

In Equation (9) and Equation (10): W_{ij} is an N-dimensional adjacency matrix defined as:

$$W_{ij} = \begin{cases} 1 & x_i \in N_l(x_j) \text{ or } x_j \in N_l(x_i) \\ 0 & \text{others} \end{cases} \tag{11}$$

where $N_l(x_j)$ denotes the l-nearest neighbor of sample x_j and $N_l(x_i)$ represents the l-nearest neighbor of sample x_i.

In order to obtain the optimal projection matrix W, the discriminant criterion function for this regularized semi-supervised discriminant analysis is solved using a method similar to that used for solving the discriminant criterion function for the Gaussian kernel discriminant analysis, and the following characteristic equation is obtained.

$$\rho R_a^{\varphi} H = (R_h^{\varphi} + \beta(R_M - S_N))H \tag{12}$$

Solving Equation (12), the matrix formed by the eigenvectors in accordance with its primary $C-1$ largest eigenvalues is the desired optimal projection matrix. In order to fully utilize the local structural information and global information of the sample data, this paper combines the objective function of linear discriminant analysis and sparse local holding projection to obtain the supervisory function of sparse local discriminant analysis.

$$\arg \max_{H \in S \subseteq S^X} \left(\operatorname{tr}(H^T R_{kc} H (H^T R_{kh} H)^{-1}) \right) \tag{13}$$

where arg is the abbreviation of variable, $\arg\max(\cdot)$ denotes the value of the variable when $\operatorname{tr}(\cdot)$ reaches the maximum value. H is the projection matrix. H^T is the transpose of H. R_{kc} and R_{kh} are defined respectively:

$$R_{kc} = \mu R_c + (1-\mu) Y C Y^T \tag{14}$$

$$R_{kh} = \mu R_h + (1-\mu) Y R_{kh} Y^T \tag{15}$$

where $\mu \in [0,1]$ is the conditioning coefficient. R_c denotes the interclass scatter matrix. D is an n-dimensional solidus matrix whose elements on the virgule are the rows or columns sums corresponding to the symmetric matrix R_k of the connection weights.

Then linear mapping is used to complete the alignment between the subspaces, and in order to minimize the divergence among the distributions of the samples in the origin domain and the samples in the object domain, the coordinate transformation matrix is used to align G and G_s, and the objective function is established.

$$f(N) = \left\| G^T G N - G^T G_s \right\|_F^2 = \left\| N - G^T G_s \right\|_F^2 \tag{16}$$

where f denotes the Frobenius norm.

A closed solution can be obtained by setting the first order derivative to zero:

$$N^* = G^T G_S \tag{17}$$

Align G to G_S, i.e., use M^* to transform the source subspace to the new coordinate space $G_b(G_b = GN^*)$, and then project the instance data of the origin domain and the instance data of the object domain to G_b and G_S, respectively, which reduces the difference between the two domains. Finally, the SVM algorithm is used to train the classification model.

4. **4. Financial early warning model based on semi-supervised learning and transfer learning.**

4.1. **Construction and pre-processing of the financial early warning indicator system.**

Based on the optimization of the semi-supervised transfer learning algorithm of the above discriminant analysis, this paper will design the financial early warning model on the ground of semi-supervised learning and transfer learning from the perspective of spatial domain.

First, screen out financial and non-financial indicators based on the principle of scientificity, construct a financial preciously warning indicator system, and standardize the deviation of the indicator data; then take the constructed indicator dataset as the target domain, effectively obtain the common knowledge of financial data between the cross-domain heterogeneous labeled origin domain dataset and the object domain dataset through the migration strategy and establish the filtering mechanism, to realize the effective migration of a small amount of samples from the source domain Then, for the heterogeneous object domain containing migration and a small amount of its own labeled samples, construct a semi-supervised recognition model based on collaborative training of isomorphic comment samples, and dynamically update the target domain label set; finally, based on the dynamically updated target domain, once again carry out migration and semi-supervision mechanism to realize the symbiotic fusion of migration learning and semi-supervised learning, and the algorithm makes migration and semi-supervision mechanism synergistically supplement each other, and through the trace of the dynamic tuning, so as to gradually improve the early warning ability of the company's finance. The model structure is shown in Figure 3.

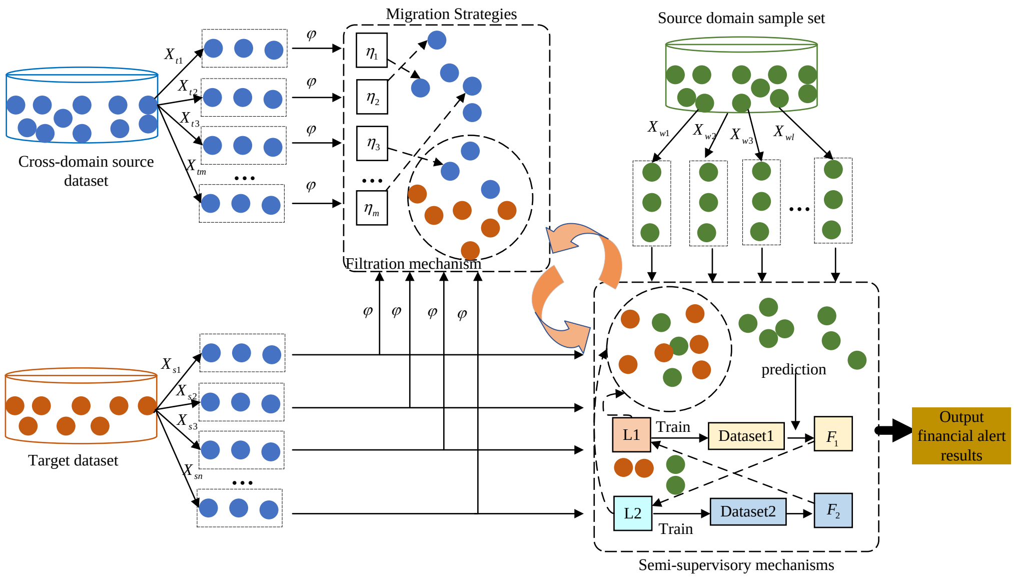

Figure 3. The overall structure

Based on the previous research and the principles of accessibility, systematicity, independence and scientificity, this paper has screened 25 financial indicators step by step, which belong to 5 level 1 indicators. In the process of selecting indicators, this paper tries to avoid eliminating indicators that may have application value in order to select as many indicators as possible. This is because the migration learning method has the ability of self-learning, and it will determine the weight of each indicator in the subsequent training process, and the selection of a large number of indicators in this paper can present the situation of the sample enterprises more comprehensively.

The system of financial early warning indicators finally adopted in this paper is shown in Table 1, where Var stands for variable, ROA (Return on Net Assets), NIROA (Net Interest Rate on Assets), GSM (Gross Sales Margin), NSM (Net Sales Margin), SER (Sales Expense Ratio), MCR (Management Cost Ratio), FCR (Financial Cost Ratio), CER (Cost-effectiveness Ratio), and NIROA (Net Interest Rate on Assets). SER (Sales Expense Ratio), MCR (Management Cost Ratio), FCR (Financial Cost Ratio), CER (Cost-effectiveness Ratio), OPM(Operating Profit Margin), CR(Current Ratio), QR (Quick Ratio), ER(Equity Ratio), GR(Gearing Ratio), ICM (Interest Coverage Multiple), RGR(Revenue Growth Rate), OPGR(Operating Profit Growth Rate), NPGR (Net Profit Growth Rate), GIMF(Growth in Money Funds), NAGR (Net Asset Growth Rate),TAGR(Total Asset Growth Rate), IT (Inventory Turnover), ARTR (Accounts Receivable Turnover Ratio), CATR (Current Asset Turnover Ratio), FAT (Fixed Asset Turnover), SCR(Sales Cash Ratio), COI(Cash Operating Index), and TACRR(Total Asset Cash Recovery Rate).

Because the data units of the 25 selected indicators are inconsistent and have different ranges of variation, the data needs to be normalized for the convenience of subsequent training. In this paper, the MinMaxScaler object of the scikit-learn library is adopted

Table 1. System of financial early warning indicators

Level 1 index	**Var**	**Level 2 index**	**Level 1 index**	**Var**	**Level 2 index**
Profitability	X1	ROA (%)	Growth capacity	X16	GIMF (times)
	X2	NIROA (%)		X17	NAGR (times)
	X3	GSM (%)		X18	TAGR (times)
	X4	NSM (%)	operating ability	X19	IT(times)
	X5	SER (%)		X20	ARTR(times)
	X6	MCR (%)		X21	CATR (times)
	X7	FCR (%)		X22	FAT(times)
	X8	CER (%)		X23	SCR(times)
	X9	OPM (%)	Cash flow	X24	COI (%)
Solvency	X10	CR (%)		X25	TACRR (%)
	X11	ER (%)			
	X12	GR (%)			
	X13	ICM (%)			
	X14	OPGR (%)			
	X15	NPGR (%)			

to process the data and normalize the sample data to 0 to 1. The specific normalization method is as follows.

$$X_{\text{scal}} = \frac{(X - X_{\min}(\text{axis} = 0))}{(X_{\max}(\text{axis} = 0) - X_{\min}(\text{axis} = 0))} \times (\max - \min) + \min \tag{18}$$

where X is the feature value to be normalized, $X_{\max}$ and $X_{\min}$ are the maximum and minimum values of the feature, max and min in the single column indicate the range of the normalized value, and $ax = 0$ means that each column is normalized. After this round of data processing, then train the model, can achieve better training results.

4.2. **Migration strategy based on financial sample data.** In the financial early warning model, the spatial dispersion of characteristics among the origin and object datasets is often inconsistent, and a migration strategy needs to be designed, the core of which is to make the spatial distribution of characteristics in the origin and object domains closer to each other by changing the weights of the origin domain sample set.

First of all, the source domain dataset has m_t entries, represented as X_t, and the object domain financial data sample set has m_s entries, denoted as X_s. Meanwhile, the labeled data set of the origin domain is denoted as Y_t and the labeled data set of the object domain is represented as Y_s. In addition to the above basic parameters, the weight value of each source domain sample relative to the object domain instance set is the key parameter of this algorithm, which is denoted as η. Here, due to the different spatial distributions of the origin and object domains, the spatial dispersion of the origin domain dataset is set as P_{st} and the spatial dispersion of the target domain dataset is set as P_{sd}. In this algorithm, the expected loss on the source domain dataset $\xi_{st} = E_{X_t,S_t \sim P_t(X_t,Y_t)}[L(f_h(X_t) \neq Y_t)]$ is somewhat different from the expected loss on the target domain $\xi_{sd} = E_{X_s,X_r \sim P_r(X_s,X_r)}[L(f_h(X_s) \neq Y_s)]$, which can be simply notated as $P_{st}(X_t, Y_t) \neq P_{sd}(X_s, Y_s)$, in view of which a set of weight values $\eta = P_{st}(X_s)/P_{sd}(X_t)$ is induced to denote the ratio of the spatial distribution of the object domain samples to that of the origin domain samples.

Define the optimization objective for minimizing the divergence in dispersion between the training set and the test set as:

$$\min \|E_{X_s,Y_s\sim P_{st}(X_s,Y_s)}[\varphi(X_s)] - E_{X_w,Y_w\sim P_{st}(X_t,Y_t)}[\eta\varphi(X_t)]\|_{\mathcal{H}} \tag{19}$$

where $\eta_j \geq 0, \quad \sum_{j=1}^{J} \eta_j = 1$.

The parameter $\varphi(x)$ refers to the mapping of the samples in Hilbert space, and the RBF kernel function is used. Through the above analysis, it can be seen that for each sample in the source domain, the closer η is to 1, the more similar it is to the spatial dispersion of the object domain. The migration filtering mechanism is set based on η in order to select the source samples with high value of micro-migration to be moved into the target domain.

4.3. **Design of Financial Early Warning Algorithm Based on Semi-supervised Learning and Migration Learning.** After the detailed discussion of the migration strategy above, this section proposes a financial early warning algorithm based on semi-supervised learning and migration learning. The proposed algorithm has the following main features: effective utilization of multi-spatial datasets, integration of multi-algorithmic mechanisms, micro-adjustment, mutual error correction between mechanisms, and multiple dynamic updates of the algorithm, as follows.

First, model optimization parameters δ^* are generally obtained by minimizing the risk function shown in Equation (20) during the migration learning phase in the multi-space domain.

$$\delta^* = \arg \underset{\delta\in o}{min}\, E_{(x,y)\in P}[l(x,y,\delta)] \tag{20}$$

where $l(x,y,\delta)$ denotes the loss function and δ denotes the corresponding parameters of the algorithm model. However, in reality, since the joint dispersion operation of x and y is unknown, the expected risk operation cannot be calculated, and for this reason, the locally optimal empirical risk function shown in Equation (21) is often used instead.

$$\delta^* = \arg \min_{\delta\in\delta_o} \sum_{(x,y)\in C_s} p(C_s) l(x,y,\delta) \tag{21}$$

To make effective use of the source domain sample set and optimize a good model for the sparsely labeled target domain dataset, a weight parameter $\eta_j = P_s(x_{s_j}, y_s)/P_r(x_{T_j}, y_{T_j})$ is introduced to make the source and target domain distributions as close as possible, except that the recognition task in the origin and object domains is the truthfulness and falsity of the comments, and thus the above equation can be simplified to $\eta_j = P_s(x_{s_j}, y_s)/P_r(x_{T_j}, y_{T_j})$. The above optimization problem can be further formulated as follows.

$$\delta^* = \arg \min_{\delta\in\delta_o} \sum_{j=1}^{m} p(x_{s_j})p(x_{T_j})p(C_s)l(x,y,\delta) \tag{22}$$

where x_{s_j} denotes a sample in the object domain and y_s is its label value, x_{T_j} denotes a sample in the origin domain and y_{T_j} is label value. The algorithm introduces $P(x_{s_j})/P(x_{T_j})$ into the objective function of Equation (23), and learns the weight parameters by optimizing the objective function in the kernel Hilbert space (RKHS).

$$\left\| \frac{1}{m_s} \sum_{n=1}^{m_s} \varphi(x_n) - \frac{1}{m_t} \sum_{j=1}^{m_t} \eta_j K_{mj} \varphi(x_j) \right\|_{\mathcal{H}} \tag{23}$$

where the kernel function K_{mj} is expressed as follows:

$$K_{mj} = k(x_m, x_j), k_m = \frac{m_t}{m_s} \sum_{i=1}^{n_h} k(x_m, x_j), K = \begin{bmatrix} K_{t,t} & K_{t,S} \\ K_{S,t} & K_{S,S} \end{bmatrix} \tag{24}$$

Finally, the newly semi-supervised labeled sample set is updated into x_j, which means that the new weight matrix η^{N+1} is thus obtained during the $M + l$-th operation, thus continuing the dynamic update for $N + 1$ times. The algorithm runs until $x_F^N = \emptyset \parallel ...\emptyset$ $|\eta_j^{N+1} - 1| \leq \delta, j = 1, 2, ..., m'$

5. **Experiment and analysis.**

5.1. **Model Training and Prediction.** Aiming at verifying the performance of the model designed in this article, this article will be the model and other existing financial early warning models for comparison experiments, all experiments are completed on the Python v3.11 environment. Using based on the dataset after comprehensive sampling and feature screening and feature dimensionality reduction processing, a financial early warning model is established for the financial status of desired companies, and its classification and prediction effect is analyzed in a horizontal comparison. To facilitate the analysis, the literature [14] is denoted as PSMT, the literature [17] is denoted as AIRE, the literature [20] is denoted as IFMC, the literature [30] is denoted as EFEW, the literature [31] is denoted as AKFT, and the algorithm in this paper is denoted as FEWM.

During the training of the FEWM model, different convergence results are achieved as the number of trainings increases. After each training an accuracy is obtained which is the accuracy of predicting the known data using the FEWM model. After the training is completed the obtained FEWM model needs to be tested, the data used for testing should not overlap with the training data, and the results obtained from testing are the valid accuracy results. The training time used in this paper is tens of hours, and the amont of training times is hundreds of times. The first 50 training samples are samples of financial crisis enterprises, of which the 1st to 25th are samples of T-5 years (enterprises five years before the crisis), the 26th to 50th are samples of T-10 years (enterprises ten years before the crisis), and the last 50 samples are samples of normal enterprises. The test results of the indicator set can be statistically analyzed to produce a table of prediction accuracy, as shown in Table 2.

As can be seen from Table 2, the accuracy of the early warning model based on supervised learning and migration learning constructed using the set of indicators containing financial-related indicators1 for the prediction of enterprises that have already experienced financial crises using the data from two years before the occurrence of their crises (T-5) is higher than that using the data from three years before the occurrence of the crises (T-10), but both of them reach more than 85%, and the overall accuracy of the judgment is up to The overall judgment accuracy rate reaches 93%, which can efficiently realize the early warning of financial crisis enterprises.

In the following, based on the balanced dataset obtained from comprehensive sampling, the early warning model based on supervised learning and migration learning is established after tuning parameterization, and the confusion matrix shown in Figure 4 is obtained after training and testing. The "0" samples indicate crisis enterprises, and the "1" samples indicate normal enterprises.

By calculating the data in Figure 4, it can be seen that the FEWM model designed in this paper has a correct classification rate of 92% for the "0" samples and 94% for the "1" samples; thus, it can be seen that the FEWM model on the ground of integrated sampling has a strong financial early warning capability for both the "0" and "1" samples, indicating that the model can accurately and effectively identify the listed companies with financial

Table 2. Accuracy statistics on the indicator set

Data sources	Tests' number	Correct predictions' number	Misjudgments' number	Accuracy
T-5 Crisis firms	25	24	1	96%
T-10 Crisis firms	25	22	3	88%
Normal firms	50	47	3	94%
Add up the total	100	93	7	93%

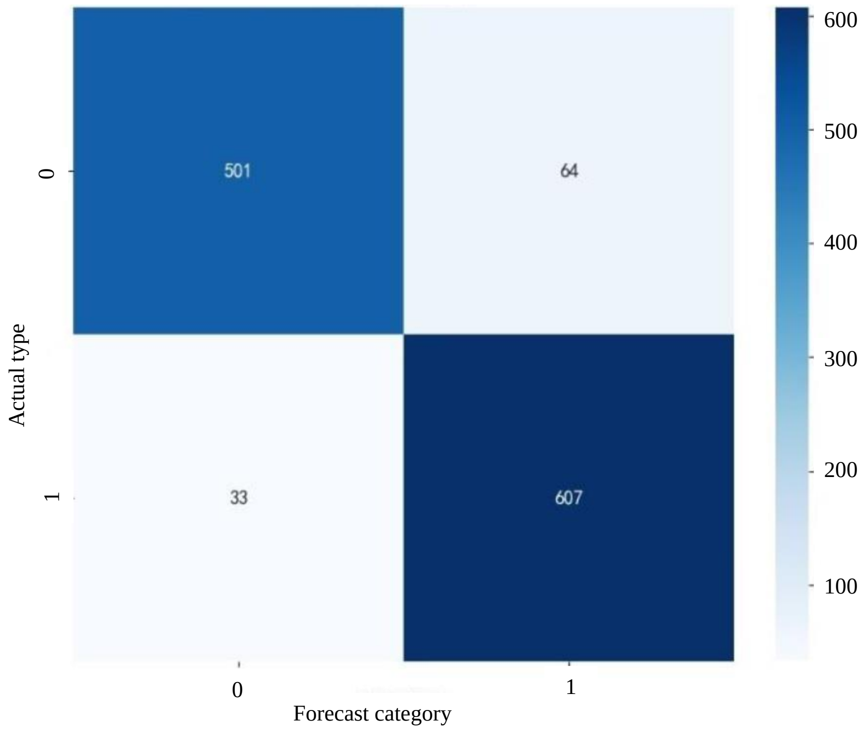

Figure 4. Confusion matrix

risks. It can be seen that the FEWM model on the ground of integrated sampling has strong financial early warning ability for both "0" and "1" samples, which indicates that the model can accurately and effectively identify listed companies with financial risks.

5.2. **Comparative performance analysis.** In order to compare the performance of each model on the final dataset obtained from the unbalanced processing of the integrated sampled data, the five evaluation indexes of each model on the test set are summarized: accuracy rate, checking accuracy rate, checking completeness rate, F1 value, and AUC value, in which the F1 value is $\left(\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}\right)\times 2$, and the AUC value denotes the area under the ROC curve, as shown in Table 3. To facilitate the analysis, the content in Table 3 is visualized to obtain Figure 5.

It can be analyzed that on the data set based on comprehensive sampling, the FEWM model designed in this paper has the highest accuracy, accuracy, completeness, F1 value and AUC value, while the PSMT model has the worst evaluation indexes, with the F1

Table 3. Indicators for evaluating models

Model	Evaluation indicators				
	Accuracy	Precision	Recall	F1 value	AUC
PSMT	83.47%	83.78%	82.31%	83.04%	88.79%
AIRE	87.14%	84.26%	86.32%	85.28%	89.02%
IFMC	84.52%	83.95%	84.91%	83.58%	87.19%
EFEW	89.37%	87.58%	86.97%	87.27%	89.14%
AKFT	89.94%	88.39%	88.16%	88.27%	90.54%
FEWM	93.18%	92.73%	95.84%	94.26%	93.82%

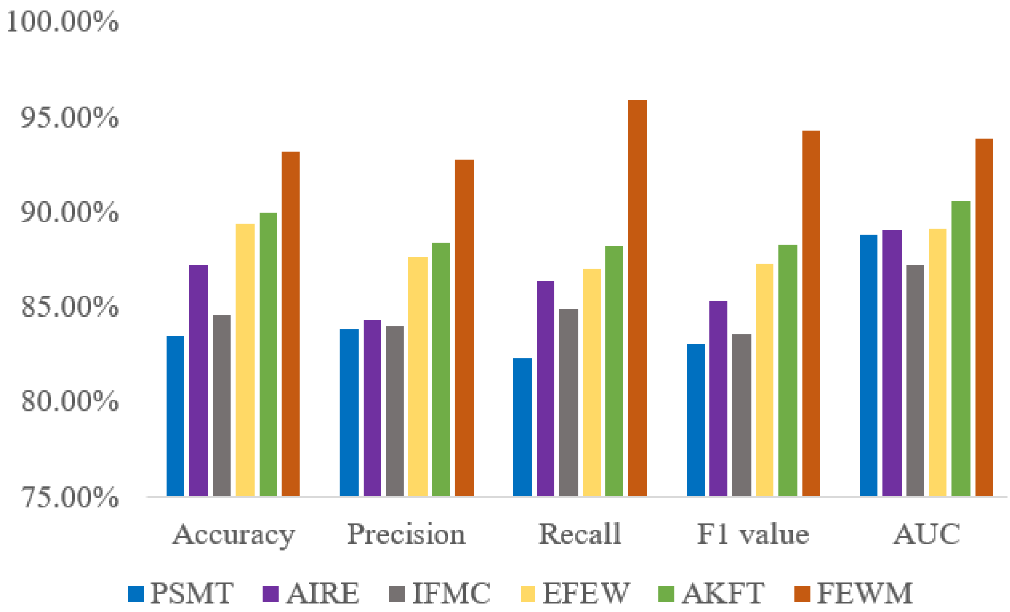

Figure 5. Evaluation result

value and AUC value of 83.04% and 88.79%, respectively. Since the F1 value is the reconciled value of the accuracy and completeness checking rate, comprehensively, on the dataset after comprehensive sampling, the model with the best performance of classification prediction ability is the FEWM model, with an accuracy rate of 93.18%, a precision rate of 92.73%, a recall rate of 95.84%, an F1 value of 94.26%, and an AUC of 93.82%; followed by the AKFT model, with an accuracy rate of 89.94%, precision rate of 88.39%, recall rate of 88.16%, F1 value of 88.27%, and AUC of 90.54%; followed by EFEW model with 89.37%, precision rate of 87.58%, recall rate of 86.97%, F1 value of 87.27%, and AUC of 89.14%.

Figure 6 visualizes that compared with other models, FEWM improves the early warning of corporate financial data at all sparse training set sizes. Compared with the PSMT algorithm, the experimental results of FEWM after semi-supervised learning and migration learning can achieve a maximum improvement of 3.5% at all sparse data sizes, and a maximum improvement of 7.3% compared with IFMC and AKFT. In addition, FEWM can get stable improvement over AIRE and EFEW algorithms in all data scales, which proves that FEWM has certain applicability to various sparse financial warning datasets.

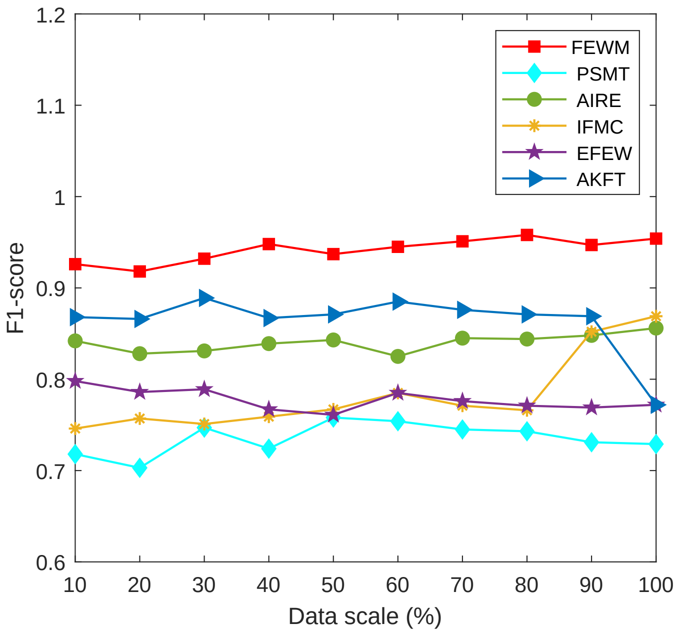

Figure 6. F1 comparisons across models

6. **Conclusion.** Aiming at the problem of low performance of existing corporate financial early warning models, this paper designs a financial early warning model based on semi-supervised learning and migration learning. Firstly, based on the regularized discriminant analysis, the data in the source domain are migrated, and the category information of the labeled data and the distribution information of the unlabeled data are used at the same time, which can improve the generalization ability of the migration. Then, based on the optimized semi-supervised migration learning algorithm, the financial early warning indicator system is constructed, and with the constructed indicator dataset as the target domain, a semi-supervised recognition model based on the isomorphic review samples of collaborative training is constructed, and the label set of the target domain is dynamically updated; secondly, based on the dynamically updated target domain, the migration and semi-supervision mechanism are processed again to achieve the symbiotic fusion of the migration learning and semi-supervision learning, and the early warning capability of company's financials is gradually improved. early warning capability of corporate finance. The experimental outcome show that the model designed in this article has high accuracy rate, precision rate and recall rate, and exhibits high early warning performance.

REFERENCES

[1] M. Bussiere, and M. Fratzscher, "Towards a new early warning system of financial crises," *Journal of International Money and Finance*, vol. 25, no. 6, pp. 953-973, 2006.
[2] A. Samitas, E. Kampouris, and D. Kenourgios, "Machine learning as an early warning system to predict financial crisis," *International Review of Financial Analysis*, vol. 71, 101507, 2020.
[3] A. S. Koyuncugil, and N. Ozgulbas, "Financial early warning system model and data mining application for risk detection," *Expert Systems with Applications*, vol. 39, no. 6, pp. 6238-6253, 2012.
[4] X. Sun, and Y. Lei, "Research on financial early warning of mining listed companies based on BP neural network model," *Resources Policy*, vol. 73, 102223, 2021.
[5] F. Zhang, T.-Y. Wu, Y. Wang, R. Xiong, G. Ding, P. Mei, and L. Liu, "Application of Quantum Genetic Optimization of LVQ Neural Network in Smart City Traffic Network Prediction," *IEEE Access*, vol. 8, pp. 104555-104564, 2020.
[6] M.-E. Wu, H.-H. Tsai, W.-H. Chung, and C.-M. Chen, "Analysis of Kelly betting on finite repeated games," *Applied Mathematics and Computation*, vol. 373, 125028, 2020.

[7] W. H. Beaver, "Financial ratios as predictors of failure," *Journal of Accounting Research*, vol. 4, no. 2, pp. 71-111, 1966.

[8] M. Blum, "Failing company discriminant analysis," *Journal of Accounting Research*, vol. 12, no. 1, pp. 1-25, 1974.

[9] J. A. Ohlson, "Financial ratios and the probabilistic prediction of bankruptcy," *Journal of Accounting Research*, vol. 18, no. 1, pp. 109-131, 1980.

[10] L. Korobow, D. P. Stuhr, and D. Martin, "A probabilistic approach to early warning of changes in bank financial condition," *Monthly Review*, vol. 7, pp. 187-194, 1976.

[11] K. Y. Tam, and M. Y. Kiang, "Managerial applications of neural networks: the case of bank failure predictions," *Management Science*, vol. 38, no. 7, pp. 926-947, 1992.

[12] J. Sun, and H. Li, "Financial distress early warning based on group decision making," *Computers & Operations Research*, vol. 36, no. 3, pp. 885-906, 2009.

[13] Y. Hong, W. Sun, B. Qianling, and X. Mu, "SOM-BP neural network-based financial early-warning for listed companies," *Journal of Computational and Theoretical Nanoscience*, vol. 13, no. 10, pp. 6860-6866, 2016.

[14] P. S. Rosa, and I. R. Gartner, "Financial distress in Brazilian banks: an early warning model," *Revista Contabilidade & Finanças*, vol. 29, pp. 312-331, 2017.

[15] T. Hosaka, "Bankruptcy prediction using imaged financial ratios and convolutional neural networks," *Expert Systems with Applications*, vol. 117, pp. 287-299, 2019.

[16] K. Valaskova, T. Kliestik, and M. Kovacova, "Management of financial risks in Slovak enterprises using regression analysis," *Oeconomia Copernicana*, vol. 9, no. 1, pp. 105-121, 2018.

[17] G. Du, Z. Liu, and H. Lu, "Application of innovative risk early warning mode under big data technology in Internet credit financial risk assessment," *Journal of Computational and Applied Mathematics*, vol. 386, 113260, 2021.

[18] Z.-S. Ouyang, and Y. Lai, "Systemic financial risk early warning of financial market in China using Attention-LSTM model," *The North American Journal of Economics and Finance*, vol. 56, 101383, 2021.

[19] E. Costa e Silva, I. C. Lopes, A. Correia, and S. Faria, "A logistic regression model for consumer default risk," *Journal of Applied Statistics*, vol. 47, no. 13-15, pp. 2879-2894, 2020.

[20] F. Gresnigt, E. Kole, and P. H. Franses, "Interpreting financial market crashes as earthquakes: A new early warning system for medium term crashes," *Journal of Banking & Finance*, vol. 56, pp. 123-139, 2015.

[21] K. Nanayakkara, and A. Azeez, "Predicting corporate financial distress in Sri Lanka: An extension to z-score model," *International Journal of Business and Social Research*, vol. 5, no. 3, pp. 41-56, 2015.

[22] W. Wu, Z. Ni, F. Jin, Y. Li, and J. Song, "Decision support model with Pythagorean fuzzy preference relations and its application in financial early warnings," *Complex & Intelligent Systems*, vol. 8, no. 1, pp. 443-466, 2022.

[23] L. Sdino, P. Rosasco, and S. Magoni, "Real estate risk analysis: The case of Caserma Garibaldi in Milan," *International Journal of Financial Studies*, vol. 6, no. 1, 7, 2018.

[24] H. Zhang, and Y. Luo, "Enterprise Financial Risk Early Warning Using BP Neural Network Under Internet of Things and Rough Set Theory," *Journal of Interconnection Networks*, vol. 22, no. 3, 2145019, 2022.

[25] N. Mselmi, A. Lahiani, and T. Hamza, "Financial distress prediction: The case of French small and medium-sized firms," *International Review of Financial Analysis*, vol. 50, pp. 67-80, 2017.

[26] Y. Shi, and X. Li, "An overview of bankruptcy prediction models for corporate firms: A systematic literature review," *Intangible Capital*, vol. 15, no. 2, pp. 114-127, 2019.

[27] K.-S. Shin, T. S. Lee, and H.-j. Kim, "An application of support vector machines in bankruptcy prediction model," *Expert Systems with Applications*, vol. 28, no. 1, pp. 127-135, 2005.

[28] J. E. Van Engelen, and H. H. Hoos, "A survey on semi-supervised learning," *Machine Learning*, vol. 109, no. 2, pp. 373-440, 2020.

[29] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big Data*, vol. 3, no. 1, pp. 1-40, 2016.

[30] X. Nie, and G. Deng, "Enterprise financial early warning based on lasso regression screening variables," *Journal of Financial Risk Management*, vol. 9, no. 04, 454, 2020.

[31] X. Huang, and F. Guo, "A kernel fuzzy twin SVM model for early warning systems of extreme financial risks," *International Journal of Finance & Economics*, vol. 26, no. 1, pp. 1459-1468, 2021.