Hindawi

Research Article

Optimization of the Economic and Trade Management Legal Model Based on the Support Vector Machine Algorithm and Logistic Regression Algorithm

Zhihai Nie [ID],[1] Xue Bai [ID],[2] Lihai Nie [ID],[3] and Jin Wu [ID][4]

[1]*Law School, Dalian Maritime University, Dalian 116026, China*
[2]*Law School, Dongbei University of Finance and Economics, Dalian 116025, China*
[3]*College of Intelligence and Computing, Tianjin University, Tianjin 300350, China*
[4]*Amazon Payment Product, Amazon.com Services LLC, San Francisco 94105, CA, USA*

Correspondence should be addressed to Xue Bai; baixue@dufe.edu.cn

Nowadays, various algorithms are widely used in the field of economy and trade, and economic and trade management laws also need to introduce scientific and effective data models for optimization. In this paper, support vector machine algorithm and logistic regression algorithm are used to analyze and process the actual economic and trade case data and bank loan user data, and a hybrid model of support vector machine and logistic regression is established. This study first introduces the basic definitions and contents of the support vector machine algorithm and logistic regression algorithm, and then constructs a hybrid model by randomly dividing the data, first using the support vector machine algorithm to calculate the results, and then inputting them into the logistic regression algorithm. The first mock exam is that the efficiency of the hybrid model is much higher than that of the single model. This study mainly optimizes and upgrades the legal system of economic and trade management from two aspects. In the prediction of economic and trade legal cases, the hybrid model is significantly better than FastText and LSTM models in accuracy and macro recall performance. In terms of credit risk prediction of economic and trade loan users, the subset most likely to default in the loan customer set is obtained.

1. Introduction

As mankind enters the digital age, from the perspective of the Internet, we are in an era of big data. Big data means huge amount. Usually, the amount of big data is more than 10 TB to 1 PB. The scale of these data is too large to obtain the desired information directly, so it needs to process and analyze big data, which is collectively referred to as data mining. Traditional data analysis often adopts sampling survey. Although this method can quickly draw conclusions, there are great errors. Data mining is to analyze and process all data, which has the advantages of comprehensiveness and accuracy. To duplicate data mining, we need to know more about big data. According to the distribution of data, big data can be divided into structured big data, semistructured big data, and unstructured big data. Among companies and

enterprises, unstructured big data accounts for the highest proportion, which can reach 83.5%, and this proportion will continue to increase with the continuous improvement of data collection and storage. This study first understands big data from the theoretical level. American scholars have summarized that big data has the following five characteristics: huge volume, high speed, diversity, low value density, and authenticity [1]. Big data is a new resource and tool for everyone. It has three levels: application, technology, and resources [2]. Economic and trade management laws need to use big data resources and realize the application of big data and attention to personnel through various algorithms, so as to distinguish them from old technologies.

With the exponential growth of data volume, its core has been transferred from storage to mining and application. The tentacles of the network extend to every corner

of the world, which not only brings convenience, but also brings new threats to the data security of companies, enterprises, and public sectors. In the traditional context of data security, economy, and trade have formed a relatively perfect data security protection system in the face of illegal intrusion [3]. However, in the context of big data economy, data security is not only the security of stored data, but also potential security risks in the process of data transmission, application, and analysis. Moreover, with the increase of the number of network nodes, the difficulty of security protection of economic and trade data also increases exponentially [4]. The big data of economic and trade enterprises usually contains a large amount of user identity and behavior information. Different channels cause the possibility of cross inspection, which is very easy to cause the leakage of user privacy. Taking this as the background, the data crime under the background of big data economy is no longer just adding, deleting, and modifying the information stored in the computer, but has transformed into a crime that takes big data as the object and spreads horizontally to all walks of life, including individuals, society, politics, economy, military, and personal fields. Vertically penetrate into big data, especially in economic and trade industries, including taking account data as the core and spreading to daily life behaviors such as e-mail, shopping records, and transfer records to commit crimes [5]. In the face of big data crime, it poses a severe test for legislative and judicial organs all over the world. In the face of increasing data security threats, the ability of legal sanctions and crackdown have lagged behind. Only by making full use of big data can we effectively carry out defense and counterattack under the background of big data economy.

2. Related Work

Banks are the basis of economy and trade in various countries. It can be said that the building of human modern finance is built on the solid foundation of banks. Under the economic background of big data, the business structure of banks has also changed, gradually shifting from offline to online. With the comprehensive popularization of e-money and the change of bank business environment, data security has become the top priority in security defense. At present, most banks in the world have implemented the New Capital Accord [6]. After 2016, the CBRC, together with many of the above-mentioned banks, established a customer big data risk statistics system to conduct real-time supervision and month by month tracking of the risk data of large loans and retail credit of each bank [7]. After years of accumulation, in the big data of existing customers in economy and trade, relevant management legal models can be established for in-depth mining and analysis [8].

With the arrival of the big data era, the economic and trade fields are facing new challenges. The development of science and technology also leads to more means of economic and trade crimes, and the related cases are becoming more and more complex. At the same time, the

detection means are becoming more and more perfect, and there will be an excessive number of cases in the end. In the face of increasing economic and trade cases, relevant criminal investigators are facing more and more severe challenges. First of all, the number and data involved in the case have increased significantly, which makes the case complex, and it is difficult for judges to achieve unity in their judgment [9]. Secondly, the information related to the case has increased from a few pages in the past to dozens of pages today, and there are even documents in different formats, which makes it more difficult for staff to sort out the case [10]. The prediction and judgment technology of the economic and trade management law based on data algorithm can solve the above problems to some extent. It is an important embodiment of the intellectualization of the legal assistant model of economic and trade management. On the one hand, this algorithm model can provide a reasonable reference for the judges of economic and trade cases and improve the efficiency of handling cases; on the other hand, this model can horizontally compare similar economic and trade cases, provide judgment reference for different judges, and avoid the artificial deviation of different judgment results in the same case [11–13]. The prediction and judgment technology of the economic and trade management law refers to a new technology in which the computer predicts the judgment result according to the detailed case description. It is one of the most important technologies in the economic and trade legal circle, especially in the civil law system, in the era of big data. The legal and Internet circles have studied this technology for many years. The early research work mainly focused on the use of algorithms to summarize and classify specific economic and trade cases [14]. With the advent of the era of big data and the rise of artificial intelligence technology, the legal prediction and judgment technology of economic and trade management have also developed rapidly. Using computer language to prejudge economic and trade legal cases through black box machine learning has become the mainstream. However, because machine learning is a black box operation, it is difficult to explain the prediction results [15]. The corresponding technologies are highly transparent codes and models such as the linear regression or decision tree algorithm. Their prediction results can be well explained, but the accuracy of the results is not high because of their limited ability [16].

In view of the above problems, in order to carry out the credit risk management of commercial banks and improve the accuracy of economic and trade legal cases, this paper selects the hybrid model constructed by the logistic regression (LR) algorithm and support vector machine (SVM) algorithm. This research is referred to as the regression vector model [17]. This experiment takes customer information, enterprise financial data, economic and trade case details, and other data as the research object, constructs the regression vector model by using the combination of the logical regression algorithm and support vector machine algorithm, deeply studies the real-time of customer risk early

warning and the correlation of risk transmission, and puts forward an analysis method of influencing factors of economic and trade legal judgment prediction.

3. Principle and Introduction of the Logistic Regression Algorithm and Support Vector Machine Algorithm

3.1. Logistic Regression Algorithm. The logistic regression algorithm, also known as the logarithmic probability model, is a classical algorithm of machine learning, which is often used to solve classification and regression problems. The dependent variables of the economic and trade measurement model are usually continuous, but there is a selection problem in the actual economic and trade, that is, companies, enterprises, and individual merchants must choose from the options provided by several banks [18]. Such a choice can be represented by discrete data in the logistic regression algorithm. It is easy to realize and widely used in industrial problems. When classifying, the amount of calculation is very small, the speed is very fast, and the storage resources are low. Convenient observation sample probability score. For logistic regression, multicollinearity is not a problem. It can be solved by combining L2 regularization. The computational cost is not high, and it is easy to understand and implement. This experiment evaluates the behavior of bank loan customers in economy and trade. The default of loan customers is expressed by 0, and the repayment on time is expressed by 1. According to the results of such selection as independent variables, the logistic regression algorithm is used for modeling. This kind of model is called the dynamic discrete model. In this paper, the information of bank customers and their repayment characteristics are predicted through the following formula:

$$Y_a = X_a\beta + \varepsilon_a, a = 1, 2, \ldots, n, \tag{1}$$

where a represents the customer, n is the sample size of the customer, and Y_a represents the repayment behavior of the customer a, which is a potential variable that cannot be observed. X_a represents the observable characteristics of customers and is a k-dimensional explanatory independent variable. In this experiment, it represents the behavioral characteristics of customers, such as loan amount, number of loan banks, degree of loan risk, and loan concentration. The model is dynamic discrete, so the customer repayment behavior is defined by the following formula:

$$Y_a = \begin{cases} 1, & Y_a^* > 0, \\ 0, & Y_a^* \le 0, \end{cases} \tag{2}$$

when Y_a^* is greater than 0, the customer will repay the loan on time; when Y_a^* is less than or equal to 0, the customer defaults on the loan. The critical value in the logistic regression model is 0, but in practice, as long as the constant term is selected in X_a, the critical value can be any integer. However, if the critical value is set to 0 in this study, the following formula can be obtained:

$$P(Y_a = 1|X_a) = P(Y_a^* > 0) = P(\varepsilon_a > -X_a\beta) = 1 - F(X_a\beta),$$
$$P(Y_a = 0|X_a) = P(Y_a^* \le 0) = P(\varepsilon_a \le -X_a\beta) = F(X_a\beta). \tag{3}$$

P in the above formula is probability, and both formulas are monotonically increasing functions.

3.2. Support Vector Machine Algorithm. Support vector machine is a linear binary classifier. Its main advantage interval is to identify high latitude models and solve small sample data and linear classification. Among them, having the largest interval in the feature space is the most important feature that distinguishes it from other perceptrons [19]. Although it is a linear classifier, it can also be used as a nonlinear classifier because it includes the kernel technique. The essence of the support vector machine algorithm is to find the optimal solution of convex quadratic programming and can accurately divide the training data set [20]. Replacing the optimal solution of convex quadratic programming with geometric description is to find the separation hyperplane with the largest geometric interval. It can also be represented in Figure 1.

In Figure 1, there are countless separation hyperplanes for a linear data set that can be divided, but there is only one separation hyperplane with the largest geometric spacing.

The previous data set of economic and trade space is defined as

$$T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}, x_i \in R^n, \tag{4}$$
$$y_i \in \{+1, -1\}, i = 1, 2, \ldots, N,$$

where x_i refers to the *ith* eigenvector and y_i refers to the dynamic class mark. The value is +1 when it is a positive example and −1 when it is a negative example. Then, through the set economic and trade management data set and hyperplane formula, take the sample point (x_i, y_i), and several intervals can be obtained through the following formula:

$$\gamma_i = y_i \left(\frac{w}{\|w\|} \cdot x_i + \frac{b}{\|w\|} \right). \tag{5}$$

The maximum γ can be obtained, and then the Lagrange multiplier method is used to obtain the newly constructed objective function:

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i (y_i (w \cdot x_i + b) - 1). \tag{6}$$

α_i in the formula is a Lagrange multiplier, and its value range is greater than or equal to 0. When the sampling point x is in the feasible region, α_i is set to infinity, and the constraint conditions are met. When the sampling point x is in the infeasible area, the value can be expressed by the following formula:
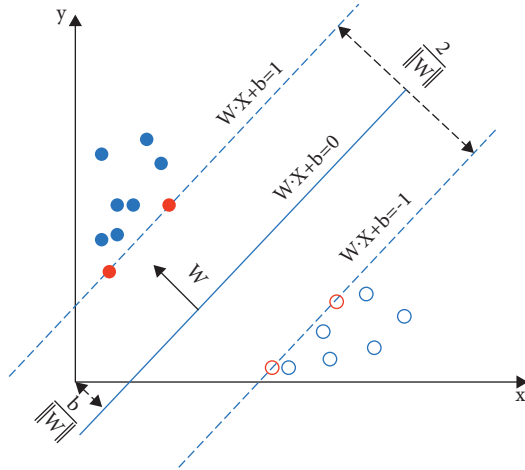
FIGURE 1: Schematic diagram of separating hyperplanes for support vector machines.

$$\theta(w) = \begin{cases} \dfrac{1}{2}\|w\|^2, \\ \\ +\infty. \end{cases} \tag{7}$$

After obtaining this new objective function, the original constrained economic and trade management problem is equivalent to

$$\min_{w,b} \theta(w) = \min_{w,b} \max_{\alpha_i \geq 0} L(w,b,\alpha) = p^*. \tag{8}$$

It will be quite complicated if the above formula is calculated first and then the minimum value. Therefore, Lagrange duality is used to exchange the positions of the maximum value and the minimum value to facilitate the solution of the two. The transformed formula is

$$\max_{\alpha_i \geq 0} \min_{w,b} L(w,b,\alpha) = d^*. \tag{9}$$

To make $d^* = p^*$, the economic and trade management problem needs to be a convex optimization problem and

meet the KKT condition, which satisfies the following inequality equations:

$$\begin{cases} \alpha_i \geq 0, \\ y_i(w_i \cdot x_i + b) - 1 \geq 0, \\ \alpha_i(y_i(x_i \cdot x_i + b) - 1) = 0. \end{cases} \tag{10}$$

The maximum hyperplane obtained by separation is

$$w^* \cdot x + b^* = 0. \tag{11}$$

Finally, the classification decision function of economic and trade management of the classification support vector machine can be obtained as follows:

$$f(x) = \text{sign}(w^* \cdot x + b^*). \tag{12}$$

4. Establishment of the Mixed Model of the Logistic Regression Algorithm and Support Vector Machine

A dynamic method combining the logistic regression algorithm and support vector machine is usually used for algorithm discrimination. The main idea is to provide the results of the support vector machine algorithm to logistic regression algorithm for analysis. In this experiment, all economic and trade management data are divided into eight regions. In each region, the classification accuracy of support vector machine is evaluated, and then the evaluation results are input into logistic regression for calculation to improve its classification accuracy:

$$\overline{p} = \text{mean}(p_i \mid y_i = \pm 1, i = 1, 2, \ldots, N). \tag{13}$$

Firstly, the partition probability mean of economic and trade management data is defined, which is referred to by $\overline{p}$ in the above formula. Then, different random sample values to calculate the accuracy of 8 regions are taken. The calculation formula of accuracy is shown in the following formula:

$$f_i = \begin{cases} \dfrac{\text{num}\{x_i \in I_j \mid y_i = -1, i = 1, 2, \ldots, N\}}{\text{num}\{x_i \in I_j, i = 1, 2, \ldots, N\}}, x_i \in I_j, j = 1, 2, 3, 4, \\ \\ \dfrac{\text{num}\{x_i \in I_j \mid y_i = 1, i = 1, 2, \ldots, N\}}{\text{num}\{x_i \in I_j, i = 1, 2, \ldots, N\}}, x_i \in I_j, j = 5, 6, 7, 8. \end{cases} \tag{14}$$

In the formula, I_j represents the region, with values from 1 to 8.

The eight regions are classified by the support vector machine algorithm and logistic regression algorithm, and the accuracy is obtained. Then, the results obtained by the support vector machine algorithm are input into the logistic regression algorithm for hybrid calculation to obtain the third group of accuracy. According to this method, the

average accuracy of various methods is calculated every 8 times, and the results shown in Figure 2 can be obtained.

As can be seen from Figure 2, the average accuracy obtained by multiple calculations is close to that of the single algorithm and hybrid algorithm, indicating that the data have a certain reliability. The two average accuracy rates of the support vector machine and logistic regression hybrid model are higher than that of the single algorithm, which
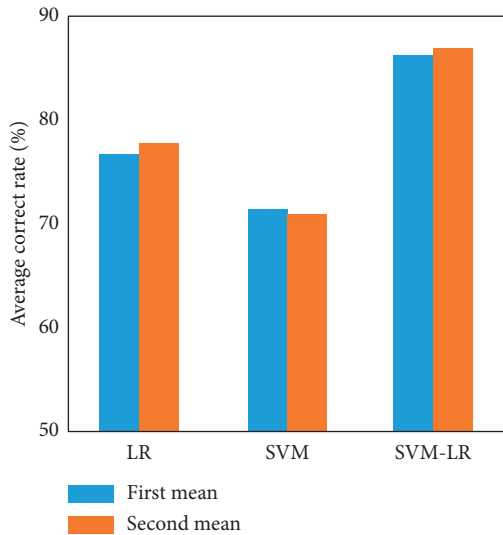
Figure 2: Comparison histogram of the accuracy rate of the classification results of the single algorithm and the hybrid algorithm.

shows that the hybrid algorithm model used in this experiment is scientific and effective.

5. Practical Application of the Mixed Model of Logistic Regression and Support Vector Machine

5.1. Establishment of the Mixed Model for Prediction of Economic and Trade Legal Judgment. This experiment selects the actual cases in the field of economy and trade for analysis and then uses the database provided by the "China legal research cup" judicial artificial intelligence challenge (cail2018). Compared with other databases, the cail2018 database has the advantages of larger scale and real cases. Moreover, the case in the cail2018 database is composed of case description and judgment results, which is equivalent to a certain preprocessing of the original data. A technology company to steal information for analysis has to be selected. The specific case description is as follows: from April 17 to 18, 2018, a technology company crawled a total of 220552 shopping software orders, including the victim Li, who lives in the Yuecheng District of Shaoxing City, through the 116.63 IP address rented by the defendants Zhou and Huang (a technology network Co., Ltd., in Zhejiang actually output 10000). A technology company maliciously added 137093 friends to a designated shopping software account (a technology network Co., Ltd., in Zhejiang actually output 20000). Defendant: a technology Co., Ltd; charges and punishment: for the crime of illegally obtaining computer information system data, a fine of 10 million yuan shall be imposed. Defendant: Zhou Moumou; charges and term of imprisonment: for the crime of illegally obtaining computer information system data, he shall be sentenced to fixed-term imprisonment of three years and six months and fined 100000 yuan. Relevant criminal law: Article 285.

Preprocessing the original data can improve the utilization rate of the data, make the data more suitable for the mixed model of logistic regression and support vector machine, and better match the needs of the model, so as to help improve the accuracy and ability of the conclusion. This paper takes the case description of economic and trade management legal cases as the input. Firstly, the input case description text is segmented by word segmentation software, and then the obtained word segmentation is transformed into word vector for hybrid model processing. Chinese word segmentation Python Jieba tool is used to process the text of case description, eliminate special characters, and extract key words. The directed acyclic process diagram of maximum efficiency using dynamic programming method is as follows.

Figure 3 uses the dictionary provided by the python Jieba Chinese word segmentation tool to generate a prefix tree, which is used for fast vocabulary scanning of economic and trade cases based on the prefix tree structure. Then, the search path is used to find the maximum probability, and finally, the word segmentation result is obtained. For words not in Jieba dictionary, the hidden Markov model is used to add this new word.

After the word segmentation preprocessing of the economic and trade management data, the input case description related word vector is obtained. Then, the experiment carries out machine learning on the obtained word vector, and finally obtains an interpretable economic and trade legal judgment prediction model. The mixed support vector machine is used to learn and manage the data.

5.2. Establishment of the Mixed Model for Economic and Trade Credit Risk Prediction. This study selects the data of loan customers in the economic and trade field for a total of 36 months from January 2019 to December 2021. Firstly, customers without nonperforming loans, interest arrears, and overdue loans are divided into the normal customer set, and customers with at least one nonperforming loan, interest arrears, or overdue loans are divided into the default customer set. Using the hybrid model of the support vector machine and logistic regression algorithm established in the previous section, on the basis of the economic and trade credit risk index system, 126 subindexes are comprehensively considered, and the significance is tested. 56 candidate early warning indicators were selected and included in the analysis scope of the hybrid model. With the addition of 4 indicators in the basic system, a total of 60 early warning indicators are called economic and trade credit dependent variables in the hybrid model. Then, the "up-down" screening method is used to improve the accuracy of prediction. The flowchart is shown in Figure 4.

Considering the nonlinear classification problem of the support vector machine and logistic regression hybrid algorithm, the selection of kernel function is the key to whether the output performance of the hybrid model is good. If the selection is inappropriate, it will lead to
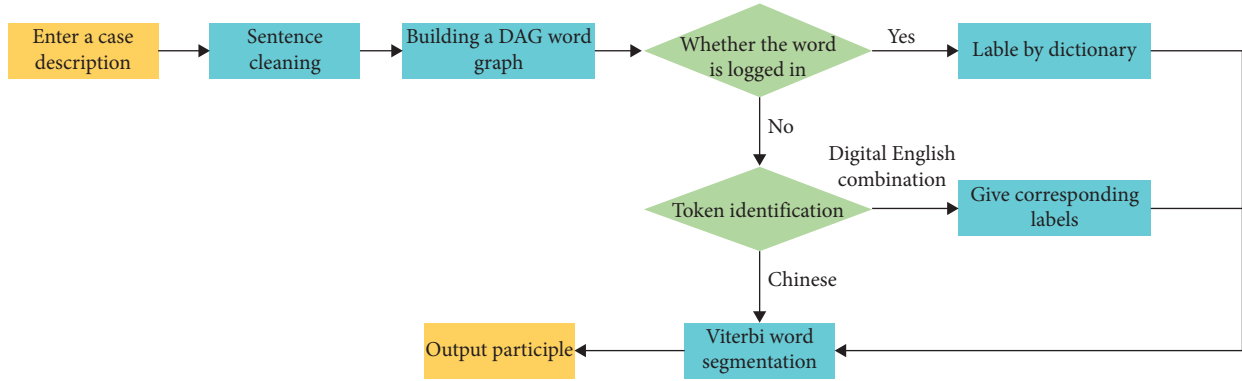
FIGURE 3: A directed acyclic graph composed of all possible word formations of Chinese characters in fact descriptions.
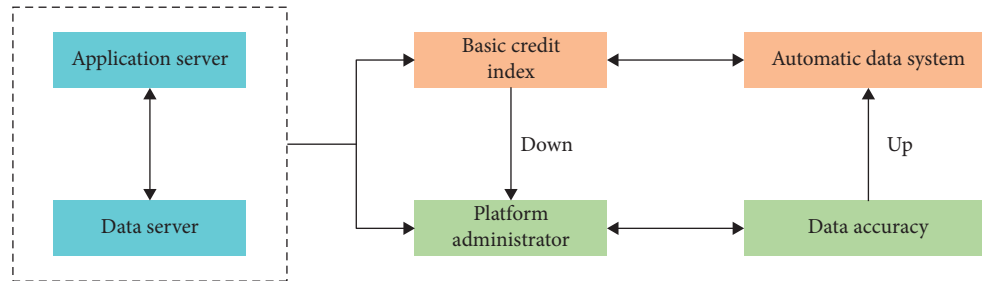

FIGURE 4: Flowchart of improving data accuracy using up-down screening.

overfitting or insufficient fitting. In this experiment, RBF kernel function with strong nonlinear mapping ability is selected. In order to achieve the best performance effect of the hybrid model, the data are standardized by the pruning method. The core problem of application pruning optimization is to design the pruning judgment method. This is to determine which branches should be abandoned and which branches should be retained. Three principles of pruning optimization are as follows: correct, accurate, and efficient. In principle, most search algorithms need pruning, but not all branches can be pruned. This requires designing a reasonable judgment method to determine the choice of a branch. In the hybrid model, the output result is 1 or 0 as the shutdown condition, and the dynamic discrete cross inspection method is used to optimize the economic and trade management data. For the evaluation of the hybrid model, four parameters are selected: the overall classification accuracy, the accuracy of default samples, the accuracy of normal samples, and the false positive rate. The overall classification accuracy rate is the ratio of the number of correctly classified samples to the number of overall samples, which reflects the performance of the hybrid model in the overall data test; the accuracy rate of default samples is the ratio of the number of correctly classified samples to the number of default samples, which further considers the accuracy of the samples used; the correct rate of normal samples is the ratio of the number of normal samples to the number of normal samples with the correct classification of normal samples; the false positive rate is the ratio of the number of samples classified as default samples but actually normal

samples to the number classified as default samples. It reflects the severity of miscalculation in the hybrid model.

6. Analysis of Experimental Results

6.1. Case Analysis of Legal Judgment Prediction. The performance of the mixed model of economic and trade legal judgment prediction is compared with the baseline, and the influencing factors of the output results of the mixed model are analyzed. Three classical metrics, accuracy, macro recall, and macro precision, are also used to evaluate the performance of the hybrid model. In order to test the efficiency of the hybrid model of the support vector machine and logistic regression algorithm in predicting economic and trade legal decisions, FastText and LSTM are also selected for the comparative test.

The comparison results of the research in verifying the effectiveness of the hybrid model of support vector machine and logistic regression algorithm are shown in the Figure 5. Based on the existing economic and trade management data set, a hybrid model of the support vector machine and logistic regression algorithm is established. It is obviously superior to FastText and LSTM models in accuracy and macro recall performance, which proves the effectiveness and reliability of the hybrid model established in this experiment.

Figure 6 shows the comparison of the macro accuracy of the hybrid model with FastText and LSTM. It can be seen that the macro accuracy of the hybrid model is higher than that of FastText and LSTM when the number of samples is 200, 400, and 600, respectively. Although the macro accuracy decreases with the increase of the number of economic and
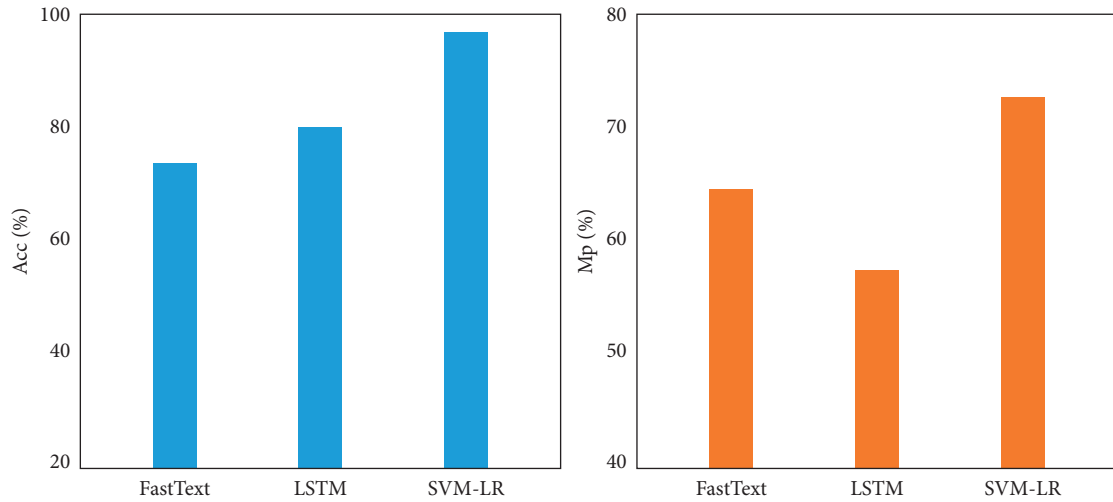
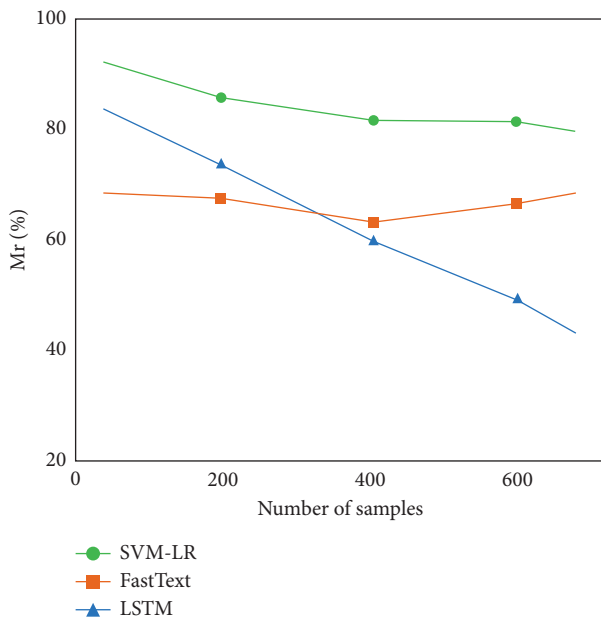FIGURE 5: Comparison histogram of accuracy, macro recall of the hybrid model, and FastText and LSTM.

FIGURE 6: Discounted plot comparing macro accuracy of hybrid models with FastText and LSTM.

FIGURE 7: Schematic diagram of the P-R curve and equilibrium point obtained by hybrid algorithm analysis data.

trade data samples, it can be maintained at more than 80%, so it is an effective and reliable model.

6.2. Example Analysis of Credit Risk Prediction.

The mixed model of the support vector machine and logistic regression algorithm is used to analyze the economic and trade management data. Firstly, the subset most likely to default in the loan customer set is obtained, and the results are represented by the P-R curve in Figure 7.

The P-R curve intuitively shows the recall and precision of the mixed model after the analysis of economic and trade samples. The three equilibrium points in Figure 7 are the measurement of the prediction accuracy of each sample subset. The larger the value, the greater the default probability of loan customers in this subset. Taking the loan
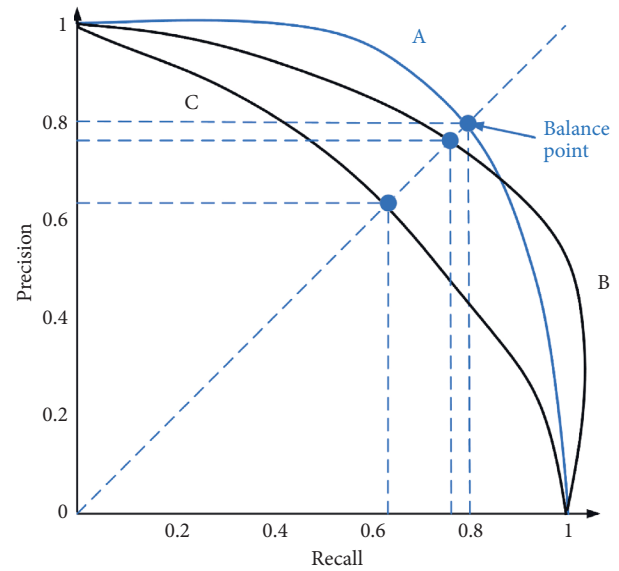
customers in subset a, we can use the hybrid model to count their common characteristics. The characteristic proportion distribution of the subset of loan customers with the largest default probability within 6 months is shown in Figure 8.

As can be seen from the above result chart, loan users with high economic and trade credit risk have the following characteristics. Among them, the highest proportion of overdue records reached 28%; secondly, 22% of customers without stable jobs have no ability to repay on schedule; thirdly, affected by the epidemic, entrepreneurs of board games and other projects also have a high proportion of overdue loans; the male in the client's portrait who is about 20 years old and has a high school education or below also has low economic and trade credit.

The mixed model of the support vector machine and logistic regression algorithm cannot predict the dishonest customers of economic and trade loans with 100% accuracy.
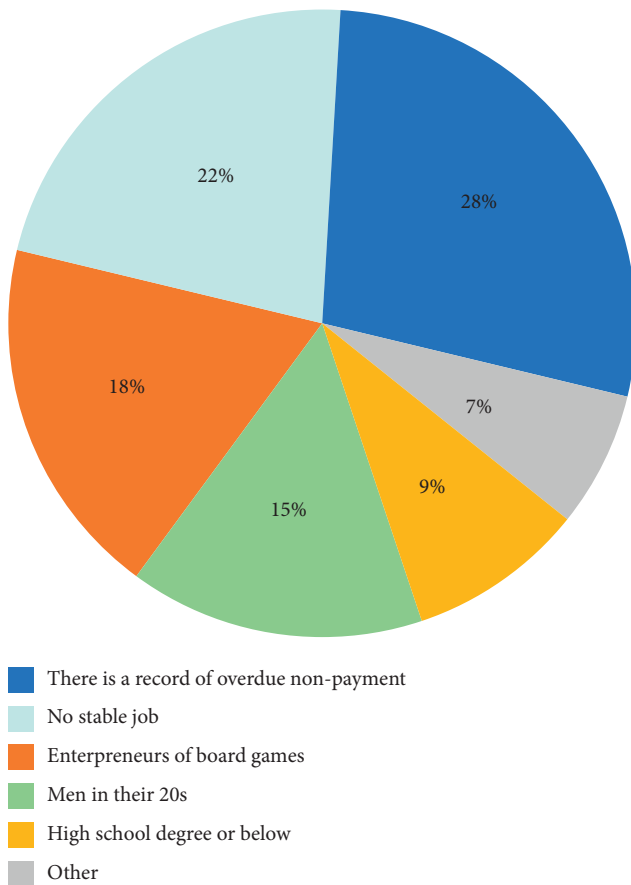
FIGURE 8: The distribution map of the characteristic proportion of the subset of loan customers with the largest default probability.
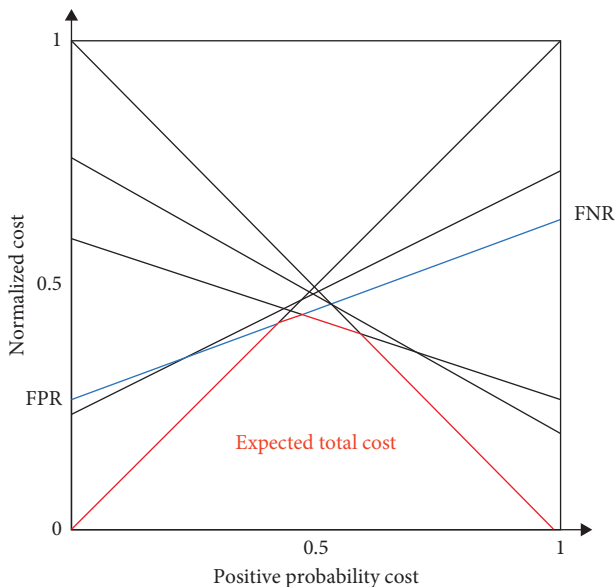

FIGURE 9: Plot of the cost curve and expected total cost.

In practical application, when a customer is classified into a subset of possible overdue according to the model, the expected total cost of this subset needs to be calculated. Only when the current expected total cost is lower than a certain

value, the algorithm model is available. In this experiment, the subset obtained above is normalized, and then the expected overall cost is obtained. The results are shown in Figure 9.

The red curve in Figure 9 is the cost curve of the experimental sample subset customers, and the blue rate is the curve value of the false positive rate (FPR) and false negative rate (FNR) of the test sample. And the area enclosed by it and the horizontal axis is the expected overall cost of the mixed mode analysis results. The final expected overall cost is 0.21, which is less than the established upper limit of 0.25, so we can get the conclusion that the results of this customer subset are available.

7. Conclusion

In this paper, a hybrid model of the support vector machine and logistic regression is established, which provides a scientific research method for this complex system. After the establishment of the mixed model, this study mainly optimizes and upgrades the legal system of economic and trade management from two aspects. In the prediction of economic and trade legal cases, on the existing economic and trade management data set, the hybrid model of the support vector machine and logistic regression algorithm is significantly better than the FastText and LSTM model in accuracy and macro recall performance. When facing a large number of samples, the macro accuracy of the hybrid model is higher than that of FastText and LSTM models. In terms of credit risk prediction of economic and trade loan users, we get the subset of loan customers who are most likely to default. Customers with the highest risk subset are characterized by overdue payment records, customers without stable jobs, entrepreneurs in projects such as board games, men around the age of 20, and people with high school education or below. Finally, using the cost curve, it is calculated that the expected total cost is 0.21, which is lower than the set upper limit of 0.25. The result of this customer subset is scientific and effective.

However, the research has some limitations. The hybrid model combining the support vector machine and logistic regression algorithm also has shortcomings in mining efficiency, which needs to be further improved. At the same time, the risk of data leakage in the economic and trade management industry is also greater and needs to be strengthened. Both the prediction of the judgment results of economic and trade management cases and the evaluation and prediction of the risk of loan customers can only provide some scientific references. In practical application, we need to be cautious and reasonably apply the analysis results given by big data.

Data Availability

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

Conflicts of 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.

Acknowledgments

References

[1] P. Matri, Y. Alforov, Á. Brandon et al., "Mission possible: unify HPC and Big Data stacks towards application-defined blobs at the storage layer," *Future Generation Computer Systems*, vol. 109, pp. 668–677, 2020.

[2] V. Papakonstantinou and P. de Hert, "Big data analytics in electronic communications: a reality in need of granular regulation (even if this includes an interim period of no regulation at all)," *Computer Law & Security Review*, vol. 36, p. 105397, 2020.

[3] J. Minou, J. Mantas, F. Malamateniou, and D. Kaitelidou, "Health professionals' perception about big data technology in Greece," *Acta Informatica Medica*, vol. 28, no. 1, p. 48, 2020.

[4] L. Gossec, J. Kedra, H. Servy et al., "EULAR points to consider for the use of big data in rheumatic and musculoskeletal diseases," *Annals of the Rheumatic Diseases*, vol. 79, no. 1, pp. 69–76, 2020.

[5] E. W. Deutsch, N. Bandeira, V. Sharma et al., "The ProteomeXchange consortium in 2020: enabling 'big data' approaches in proteomics," *Nucleic acids research*, vol. 48, no. D1, pp. D1145–D1152, 2020.

[6] K. Y. Wang, Y. Wen, T. L. Yip, and M. Luo, "How does a bank's involvement interplay with a firm's capacity investment? An analysis and comparison of different consortium structures," *International Transactions in Operational Research*, vol. 27, no. 5, pp. 2658–2682, 2020.

[7] J. Valdes, "Participation, equity and access in global energy security provision: towards a comprehensive perspective," *Energy Research & Social Science*, vol. 78, p. 102090, 2021.

[8] S. Sivanenthira and S. D. Kumar, "A study of performance appraisal of people'S bank & Bank of Ceylon in Vavuniya in Sri Lanka," *Asian Journal of Multidimensional Research (AJMR)*, vol. 9, no. 4, p. 55, 2020.

[9] S. Ghosh, P. Wronski, J. Koetsenruijter, W. Mueller, and M. Wensing, "Watching people decide: decision prediction using heatmaps of reading of a decision-support document," *Journal of Vision*, vol. 21, no. 9, p. 2631, 2021.

[10] M. Elizabeth, A. Justo Julie, K. Joseph, P. B. Bookstaver, R. Winders Hana, and M. Al-hasan, "313Using a clinical decision prediction tool to improve empirical antimicrobial therapy in ceftriaxone-resistant enterobacterales bloodstream infections," *Open Forum Infectious Diseases*, vol. 7, p. S153, 2020.

[11] F. Yao, X. Sun, H. Yu, W. Zhang, and K. Fu, "Commonalities-, specificities-, and dependencies-enhanced multi-task learning network for judicial decision prediction," *Neurocomputing*, vol. 433, pp. 169–180, 2020.

[12] E. Filtz, S. Kirrane, and A. Polleres, "The linked legal data landscape: linking legal data across different countries," *Artificial Intelligence and Law*, vol. 29, no. 4, pp. 485–539, 2021.

[13] F. F. Sukini, P. Lestari, S. N. Purwaningsih, and S. R. Ontran, "Legal protection of dental and oral therapists for oral delegation by dentists[J]," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 2, pp. 1756–1761, 2021.

[14] T. E. Meskela, Y. K. Afework, N. A. Afework, M. W. Ayele, T. B. Teferi, and T. B. Mengist, "Designing time series crime prediction model using long short-term memory recurrent neural network," *International Journal of Recent Technology and Engineering*, vol. 9, no. 4, pp. 402–405, 2020.

[15] B. Yang, L. Liu, M. Lan, Z. Wang, H. Zhou, and H. Yu, "A spatio-temporal method for crime prediction using historical crime data and transitional zones identified from nightlight imagery," *International Journal of Geographical Information Science*, vol. 34, no. 9, pp. 1740–1764, 2020.

[16] H. I. Jo and J. Y. Jeon, "Overall environmental assessment in urban parks: modelling audio-visual interaction with a structural equation model based on soundscape and landscape indices," *Building and Environment*, vol. 204, p. 108166, 2021.

[17] R. Sutradhar and L. Barbera, "Comparing an artificial neural network to logistic regression for predicting ed visit risk among patients with cancer: a population-based cohort study," *Journal of Pain and Symptom Management*, vol. 60, no. 1, pp. 1–9, 2020.

[18] X. Liu, M. Lu, Y. Chai, J. Tang, and J. Gao, "A comprehensive framework for HSPF hydrological parameter sensitivity, optimization and uncertainty evaluation based on SVM surrogate model- A case study in Qinglong River watershed, China," *Environmental Modelling & Software*, vol. 143, p. 105126, 2021.

[19] W. Huang, H. Liu, Y. Zhang et al., "Railway dangerous goods transportation system risk identification: comparisons among SVM, PSO-SVM, GA-SVM and GS-SVM," *Applied Soft Computing*, vol. 109, p. 107541, 2021.

[20] Y. Lao, V. Y. Yu, A. Pham et al., "Voxel-wise GBM recurrence prediction based on post-operative multiparametric MR images using multidimensional SVM coupling with stem cell niches proximity estimation," *International Journal of Radiation Oncology, Biology, Physics*, vol. 108, no. 3, p. e771, 2020.