

Retraction

Retracted: A Relevant Customer Identification Algorithm Based on the Internet Financial Platform

Journal of Mathematics

Received 23 January 2024; Accepted 23 January 2024; Published 24 January 2024

Copyright © 2024 Journal of Mathematics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] G. Yangyudongnanxin, "A Relevant Customer Identification Algorithm Based on the Internet Financial Platform," *Journal of Mathematics*, vol. 2021, Article ID 9770471, 9 pages, 2021.

Research Article

A Relevant Customer Identification Algorithm Based on the Internet Financial Platform

Guo Yangyudongnanxin 

Business School, Hunan International Economics University, Hunan, Changsha 410205, China

Correspondence should be addressed to Guo Yangyudongnanxin; 201914040036@zknueu.edu.cn

Received 1 October 2021; Accepted 29 October 2021; Published 23 November 2021

Academic Editor: Lazim Abdullah

Copyright © 2021 Guo Yangyudongnanxin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the intelligent search capabilities of Internet financial customers, this paper proposes a search algorithm for Internet financial data. The proposed algorithm calculates the customers corresponding to the two selected financial platforms based on the candidate customer set selected from the seed dataset and combined with the restored social relationship. Moreover, it also calculates the similarity of each field between the pairs. Furthermore, this article proposes an entity customer classification model based on logistic regression. Through the SNC model, threshold propagation, and random propagation, the model is transformed into an algorithm that identifies the associated customers, eliminates redundant customers, and realizes associated user identification. Experimental results verify that pruning increases the accuracy of identifying related customers by 8.44%. The average sampling accuracy of the entire customer association model is 79%, the lowest accuracy is 40%, and the highest is 1. From the sampling results, the overall recognition effect of the model reaches the expected goal.

1. Introduction

In recent years, IT has benefited from the increasing development of the Internet and the in-depth development of financial disarmed. The Internet finance has rapidly and finally evolved into a global financial phenomenon and become one of the most popular financial topics. Internet finance is a new financial business model for traditional financial institutions and Internet enterprises to realize funds, payments, investments, and information intermediary services. On the one hand, the Internet finance has played a role of traditional financial institution to replace the active role in promoting direct financing, improving the efficiency of financial services and resident financial services, and fully excavating and catering the market demand. It is called “squid effect” and contributing positive energy [1]. Though there will be some drawbacks, due to the relative lack of supervision norms, the industry has led to its development, which is characterized by “good, slime,” highlighting in some market participation, the main motive is not pure, high interest, etc. For the scorpion, the scam is

implemented, and the scandal of misappropriation of customer funds or even the curved running road is constantly occurring [2].

Basically, the term “Internet finance” is used for the funding achieved by traditional financial institutions and Internet companies through using the Internet and information communication technologies. The sellers on the platform characterize the features of the Internet financial platform. That means, the higher number of buyers attract the sellers to that very platform. Henceforth, the buyers and the sellers are attracted to the platform by following the operational models, specifically the pure platform mode and guarantee mode. Furthermore, the information regulations are essential parts of financial markets. The selection and configuration of the information regulatory tool is obligatory to be emphasized on the crucial links of Internet financial platform information regulation. There is a direct regulation and indirect regulation, mandatory regulation, and excitation regulations, and the information tools hidden by different types of information regulations are naturally different. In our experiments, we filtered data. 838 customers

were chosen for manual matching and labeling in a financial platform; thus, the training and testing of logical regression models took place. Furthermore, the data were processed to get similarity feature vector which is actually the input data of the model. The total data were divided into the training and test sets. Then, the remaining data were decomposed into 4 training sets and a testing set. In 5 groups of data, 1760 customers were chosen which were randomly paired to find the positive and negative cases for acquiring the initial classification model. Next, the semisupervised learning method was adopted for the calculation process. Finally, the results are used for superposition training of the model.

First, this article is designed with a cloud protocol based on a cookie single sign-on, used to implement cross-platform, cross-domain storage of customers. According to the design of the storage protocol, the personalized recommendation system proposed herein needs to meet the requirements of efficient return recommendations. Conventional content-based personalized recommendation algorithm first maintained all customers' configuration files. Each recommended traversing all products or pages to get the best recommendation. This is a cross-platform cloud personalized recommendation and related clients proposed in this article. The identification system is not applicable.

Finally, this paper applies the recommended algorithm to the Internet actual system, performing data mining and verifying personalized recommendation algorithms and customer identification from accuracy and efficiency angles. This paper puts forward a cloud storage identification designed from the customer's personal demand and associated customer identification and designs the associated data processing algorithm that matches it. While considering the actual system requirements, the algorithm also reflects the characteristics of the customer's class and realizes the integration of accuracy. It has a certain value for enhancing the Internet customer experience, identifying associated customer information, and implementing the Internet business integration.

2. Literature Review

Recently, a continuous development is being made in the fields of information and technology. Moreover, the information technology also gained attention due to the growing use of Internet and financial disarmed. As a result, the Internet finance has come up as a global financial phenomenon. It is actually an innovative business model for financial institutions as described earlier. Lu et al. [3] proposed a new and rapid SimRank algorithm, He and Yao [4] came with a structural regularization in a quadratic logistic regression model, Yao et al. [5] studied the resources trading in blockchain based on industrial IoT, Zhuang and Zhang [6] worked on the legal and risk prevention of third-party payment in Internet finance, and Wv and Dh [2] discussed the characteristics and consensus of blockchain in modern business processes. Guan et al. [7] worked on the sharing of demand information in competing supply chains with manufacturer provided service, and Pang and Yang [8] presented a loss model with its application in the Internet

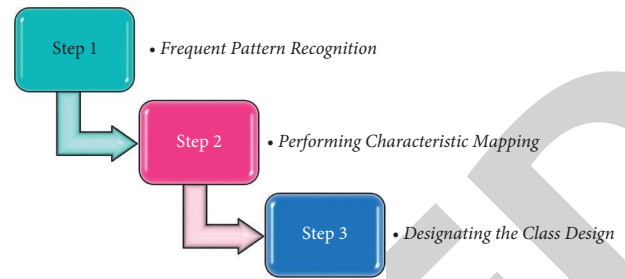


FIGURE 1: Methodology.

financial platform. Later in 2020, Peng [9] wrote an article on the Internet financial platform based on 5G network. Ju et al. [10] explained the effective fault localization of evolution software which is based on the multivariate logistic regression model, Wang et al. [11] adopted the forward local push with its parallelization for accurate and fast SimRank computation, Chang et al. [12] worked for the person re-identification and proposed a transductive semisupervised metric learning. Deng et al. [13, 14] proposed some enhanced and evolutionary algorithms for optimization problems.

With the implication of technical methods for financing, the threats and risks became a challenge. Thus, the researchers also came forward to solve the issue by proposing the security techniques and risk analysis. Han [15] worked on the legal regulations of the price war of the security company's brokerage business under the Internet finance, Yang et al. [16] proposed an identification algorithm of a high-breaching-risk member of the Internet financial platform, and Zhang et al. [17] researched the risk management of the Internet financial platform. Furthermore, Qu et al. [1] gave the idea of two-factor cross-domain authentication schemes such as biometric and password based on the technology of blockchain. Han et al. [18] described a method for dynamically assessing credit risk of the Internet financial platform, and Yu et al. [19] proposed privacy protection as a base for the multiparty secure computing financial shared platform.

3. Methodology

For the efficient return requirements, this paper proposes a personalized recommendation algorithm based on associated rules, using three steps to meet the real-time requirements of the cloud personalized recommendation system. The first step is using a frequent pattern recognition algorithm for non-real-time update, establishing a frequent mode library based on the customer's previous cloud data. In the second step, in order to reduce the difficulty of personalized recommendation in real time, the characteristic mapping based on frequent mode libraries is performed, and the clustering algorithm is designed and analyzed. In the third step, according to the result of clustering, the specific criteria for designing the class are designated, and the frequent mode of each type of customer is obtained, thereby completing the design of personalized recommendation and customer identification algorithm. Figure 1 illustrates the methodology of the proposed algorithm.

4. Internet Financial Platform

The historical process in the Internet era has been advanced from the depth, and people not only no longer satisfy the breakthroughs of Internet technology but also pay more attention to the application of Internet thinking. Due to Weibo, WeChat, social networking sites, and the “Taobao” representative of C2C (personal and personal) e-commerce websites, individuals in all sides are free to exchange information on the network platform, for low-cost trading of goods. With the development of the times, people exchange the content or will not be limited to information, and people trading should no longer stop in the goods. In recent years, Internet thinking has gradually penetrated into the field of funding and capital markets, and as a sign in the emerging Internet finance industry, it is a “detriment” nature of the Internet financial platform [7].

My country’s official document has defined “Internet Finance.” On July 14, 2015, the Top Ten Committee of the State Council’s Guidance Opinions on Promoting Internet Financial Development (hereinafter referred to as “Guiding Opinions”) pointed out that Internet finance refers to “traditional financial institutions and Internet companies use Internet technology and information communication technology to achieve funding.” New financial business models for financing, payment, investment, and information intermediary services are introduced. “In this analysis, it was not difficult to find that “Guidance Opinions” did not remove traditional financial institutions to the Internet financial industry, and the business main body of the Internet financial industry can be both traditional financial institutions or an Internet buccaneering. Obviously, this is a quite ambient definition. It is also because “Guidance Opinions” has adopted a broad sense of Internet finance. Subsequently, there are six kinds of activities of Internet finance: Internet payment, network borrowing, equity crowdfunding financing, Internet fund sales, Internet insurance, Internet trust, and Internet consumption finance. However, the author believes that Internet fund sales, Internet insurance, Internet trust, and Internet consumption finance, although carried out on the Internet platform, are still “one-to-many” online sales models of traditional financial institutions, and the end of the field belongs to the Internalization of traditional financial institutions or is the upgraded version of traditional finance [8]. Although “Guidance Opinions” incorporate it into a scope of generalized Internet finance, it is not an Internet finance in the true sense. Insufficient production capacity and high requirements on personnel bring inevitably high costs [19]. Due to the inability to get rid of basic business such as accounting, enterprises are gradually reforming their financial strategies and organizational structures and starting to build Financial Shared Service Centers. Bryan Bergeron first summarized the concept of financial sharing in Essentials of Shared Services as a semiautonomous approach that integrates operational functions of existing business units into a centralized new business unit [5]. The Internet payment is a business of Internet finance, but it does not have a distinct platform attribute. The essence of the Internet financial platform is “It is to build a trading platform, let all needs and

materials search and match on this platform by themselves, and turn centralized matching into a distributed “point-to-point” transaction state.” Go to center, low-cost trading ecology [17].

4.1. Operation Mode. The operation mode of the Internet financial platform fundamentally determines the legal attributes and risk levels of the platform. “In the platform, both sides (or more) interact on a platform. The characteristics of this model are characterized by the seller on the platform. The greater the attraction of the buyer, the more the seller is considering whether to use this platform, the more the buyers on the platform, and the more attractive it is to the seller. “In order to attract the buyer and the seller to the platform, my country’s Internet financial platform is constantly evolving out the following operational models [9].

4.1.1. Pure Platform Mode. In the pure platform mode, the platform strictly positions itself as a pure information intermediary, does not interfere in the essence of the transaction, but only provides investors and financiers with technical services such as information release, information matching, credit rating, “one-to-one” matching, and capital settlement. Equity crowdfunding financing platforms belong to the pure platform model because in equity crowdfunding financing activities, financiers share benefits and risks with investors and financiers do not promise to repay principal and interest. Therefore, the platform does not play any form of credit intermediary role, but pure information intermediary [16]. However, in contrast, in the field of P2P, the pure platform mode is very small. In 2007, the “shooting loans” established in Shanghai is China’s first information intermediary service platform for China’s information providing P2P-free network borrowing and is also a P2P platform that is not much more purely charged for a profit model. “Loan” is not guaranteed, does not set up a fund pool, always standalone, and the role of third-party platforms, and the risk is borne by the borrowing transactions. The borrowing process of “shooting loans” is as follows: borrowers release borrowings, borrowers compete for bidding, borrowers are successful, borrowers get borrowings, and borrowers are repayable on time. In order to control the borrowing risk on the platform to attract more investors, “shooting” is based on the wind control model of the big data, and a risk score is given to each borrowing to reflect the forecast of overdue rates. Each score interval will be displayed to the borrower and the lender in the form of a letter rating, such as from AAA to F, and rising in turn. In order to protect the interests of investors, “shooting” is implemented, and when the “Overdue payment” is dealt, the platform extraction is placed in the “risk spare gold account.” Once a borrowing is overdue for more than 30 days, the platform will pay the remaining borrowing of this borrow from the investors through the risk spare gold.

4.1.2. Guarantee Model. Since the equity crowdfunding financing platform does not have a guaranteed space, the Internet financial platform for the warranty mode is also

limited to the P2P platform. In recent years, there have been many guaranteed P2P platforms in China, and it can be divided into two types of self-guarantees and third-party guarantees in accordance with the guarantee main body. Platform self-warrants means that the properties of the platform have changed fundamentally, and the platform evolved into a guarantee mechanism, not a simple information intermediary organization, and credit risk is borne by the platform. This type of nature transform has been banned by the current regulatory policy. As early as the disposal of illegal fundraising, the relevant person in charge of the China Banking Regulatory Commission clearly pointed out four boundaries of the P2P network lending platform: first, it is necessary to clarify the mediation properties of the platform; second, it is necessary to clarify the platform itself must not provide guarantees; the third is to return the funds; and the four is to not illegally absorb public funds [18].

4.2. Information Regulation. The financial market is a typical information market. The analysis paradigm of information is an analysis of financial institutional analysis, which is of course applicable to the legal regulation of Internet financial platforms. From the state of view, Internet technology helps to simplify financial transaction processes and improve financial regulation efficiency and plays a unique role in creating money flow and information disclosure so that information is asymmetrical in the Internet finance market and credit risk issues. In this context, the introduction of the information regulation concept and the necessity of the system are highlighted [6]. At present, the concept of information regulation is widely used in the field of environmental protection, food safety, consumer rights relief, legal control, and legal governance of shared economy, but its connotation is different in different contexts. According to the authority of Anthony Ogs, the information regulation mainly includes two aspects. The first is the information disclosure; that is, the supplier is obliged to provide information about the price, identity, ingredient, and quantity or quality of the goods, and the other is to control errors or distortion information. The descriptive definition is clearly defined from the perspective of consumer rights protection and cannot be the basis for this article. Some scholars believe that the information regulation refers to the detailed and accurate information of the goods or services to the information or service to provide information or service to the information of the government's announcement of the government announcement information or the provisions of the government, in order to reduce the negative impact of information bias. There are also scholars who believe that the connotation of information regulation includes "regulation of information" and "regulation of information." The former is to adjust the information in accordance with the legal means to achieve the orderly flow of information and ensure the accuracy and effectiveness of information collection and reduce the overall operation cost of the society, and the latter

is to adjust the social relationship with information as a regulatory tool to provide support for the country to moderate and effectively intervene and achieve effective implementation of social political economic decision making [15]. Information regulations are collectively referred to using various information tools to boot, specification, and governance of the Internet financial platform, pointing to "regulation of information rather than" regulation of information. About this concept, it is necessary to emphasize that the key link of Internet financial platform information regulation is the selection and configuration of the information regulatory tool. There is a direct regulation and indirect regulation, mandatory regulation, and excitation regulations, and the information tools hidden by different types of information regulations are naturally different. For example, the tool for mandatory information includes mandatory information disclosure, forced information storage, and forced information sharing, and the tools such as the excitation information include information exchange and information protection. Different types of information regulatory tools should be properly configured to promote timely replacement and benign interactions, thereby increasing the effectiveness of Internet financial platform legal regulation.

5. Related Customer Information Identification

This section presents the discussion about the customer information identification.

5.1. Introduction to the Cross-Network Customer Relationship Model

5.1.1. Logic Regression. The candidate customer set according to the seed dataset is used, combined with the reduced social relationship; first, the similarity of the two financial platforms corresponding to each field between the two financial platforms is calculated, combining the similarity of the field into similarity feature vectors as follows:

$$\vec{S} = \langle V_0, V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9 \rangle, \quad (1)$$

where V_0 represents the similarity of the customer's nickname, V_1 represents the similarity of customer gender, etc., and different dimensions represent the similarity between different attributes. When conducting customer confirmation, the confirmation classification is considered to be a two-point issue; that is, it is confirmed that the same entity customers are a two-point issue. The logistic regression (LR) [10] model is a commonly used classification model. Binomial logistic regression is commonly used in logistic regression, and there are only two classifications. Therefore, in this paper, the customer association confirmation selects the logistic regression model for classification. The results belong to positive examples (represented by 1) and negative examples (represented by 0), and the conditional probability of binomial logistic regression is

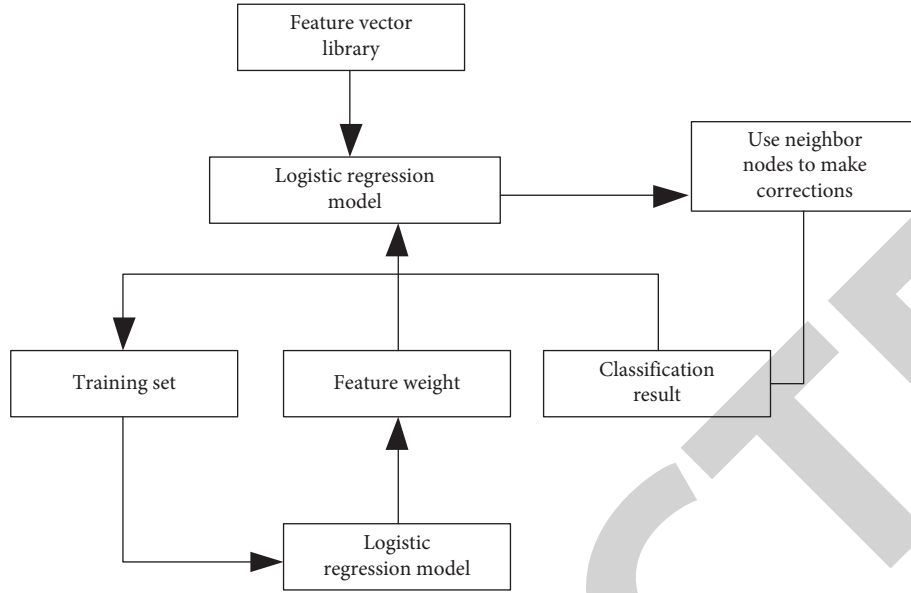


FIGURE 2: Classification model flow.

$$P(Y = 1|x, w) = \frac{e^{w^T + b}}{1 + e^{w^T + b}} = \frac{1}{1 + e^{-(w^T + b)}} \quad (2)$$

$$P(Y = 0|x, w) = \frac{1}{1 + e^{w^T + b}} \quad (3)$$

In formulae (2) and (4), xRN is an input, representing the characteristics of an example; $Y \{0, 1\}$ indicates the output, which is of only two types, simple representation or no. xRN and PR represent parameters, where w represents the weight vector, and its corresponding value represents the weight of the input feature; b expresses the offset. During classification, according to formulae (2) and (4), $P(Y = 1|x, w)$ and $P(Y = 0|x, w)$ can be obtained, respectively, and logic regression compares the size of these two conditional probability and divides the entered instance into the kind of probability value relatively large. The classification model flow is shown in Figure 2.

To use the model to classify the predictive determination requires training to train the model, obtain the feature weight parameters, and then, calculate the input feature vector according to the feature weight and compare the comparative determination according to the calculated results [4]. When the logic regression calculation result is greater than the threshold, the result is divided into a correct case, indicating that the client of the two platforms belongs to the same entity customer; otherwise, the result is divided into a negative, indicating that it is not to point to the same entity customer.

5.1.2. SNC Algorithm. Simrank [11], originally proposed by the MIT Lab Glen Jeh and Jennifer Widom in 2002, is a model of the topology information that uses the map to measure the similarity of the two objects. The core idea is if the two objects are referenced (in the social network is expressed as a similar neighbor), the two objects are similar.

There is a strong homogeneity in the social network, which is gathered by a class, and people in a population who have the same symbol or interest become friends. During the customer discovery, if most of the friends have similarity, then you can think that this is the same customer, and this law and Simrank are consistent, so it is widely used in the study of social networks. To eliminate these excess customers, Simrank's [3] neighbor similarity ideas are used to learn from, and the Simneighbor-Cut (SNC) algorithm of the neighbor node is proposed to make the model.

The algorithm involves the following formula:

$$s = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{m} \sum_{m=1}^m u_m \right), \quad (4)$$

$$I_i = \otimes s_u - s \otimes. \quad (5)$$

In formula (5), n represents the number of direct neighbor nodes that the customer matches and m means that if a direct neighbor is also present in set A , there is a plurality of correct classification results, and then, taking the result, the adjacent variance is taken in the formula as a neighbor similarity value. In formula (5), s_u represents the predicted value calculated by the logic regression function, I_i is an absolute value of the predicted value and the neighbor similarity, and the final output customer pair is the lower value of the predicted value and the neighbor similarity difference.

6. Experimental Research

In this section, the experimental research is discussed. Moreover, the tests and their results are also illustrated.

6.1. Training Model and Test. When filtering data, 838 customers in a financial platform are selected for manual labeling and matching, for the training and testing of logical

Input: Determination of the predictive result set A, user pairing set B, seed user relationship matrix M_r , candidate set matrix M_h ;

Output: Best Paired Combination $U_r - U_k$ in Collection B;

The specific steps are as follows:

- (1) Extract the pairing user in B, divide users into u_r and u_h ;
- (2) Search all direct neighbors of u_r in M_r , labeled, deposits set S_r ;
- (3) Search all direct neighbors of u_h in M_h , labeled, deposits set S_h ;
- (4) Use the user in S_r and S_h to combine the collection A, find a common neighbor of u_r and u_h , and label;
- (5) The similarity s of u_r and u_h is calculated by the following formula;
- (6) The absolute difference l_i is calculated by the following formula;
- (7) Repeat the above steps until all the pairing users in the collection B are complete;
- (8) Take the l_i value to the user output, the pairing user $U_r - U_h$ corresponding to $l = \text{Min}(l_i)$.

ALGORITHM 1: SNC algorithm.

regression models. Processing data get the similarity feature vector, namely, the input data of the model. The 838 customers' data are divided into training sets and test sets, and 38 isolated nodes are removed to reduce the impact on the relationship between friends. The remaining 800 customers are then divided into 5 groups in the proportion of 4 to 1, of which 4 are used for training and 1 set is used for testing. In 5 groups of data, 1760 customers were selected from which 1760 customers were selected and then randomly paired with customers in the seed dataset as a negative example in the training. The proportion of positive cases and negative cases in each group is different, which is used for training to obtain the initial classification model. After obtaining the initial model, the semisupervised learning method [12] is adopted in the subsequent calculation process, and the final customer association results are used for superposition training of the model. After completing the data packet, it is trained with the number of iterations, and packet training, incremental training, and combined training, respectively. Grouping training is to train models by four packets, producing four models, corresponding to training sets, numbered M1 (training set 1), M2 (training set 2), M3 (training set 3), and M4 (training set 4); incremental training uses four groups to enter the model in turn, producing a model, numbered M5; during combined training, all four groups are synthesized to train, and a model is generated, numbered M6. Table 1 is a comparison of the determination result of the input test set after the completion of the training (proportion of training concentration and negatives).

From the test results of M1, M2, M3, and M4 in Table 2, the number of negatives added during the training process has little effect on the prediction results; the results of comparison M1 to M4 and M5 and M6 have found that the number of original training concentrations will affect the final result; the more accurate it is, the higher will be the final result. The accuracy of model M5 and M6 is higher than that of the other models, and the accuracy of the two is different, but the recall rate of the model M6 is relatively high, so model M6 of combined training is used to classify the associated customers. At the time of statistics on the final test results, there is a pair of multiphenomena discussed in the classification result, and the SNC algorithm is used to score, and the results after the twig are shown in Table 2.

As can be seen from the results of the correction in Table 2, the number of positive numbers in the posttwig decrease is 31, the number of negatives increases by 31, the number of correct classifications has increased by 27, the correct number of correct classifications decreased to 4, there is an error in the twig process, a small amount of correct classification is determined as an error classification, and the correct result is deleted, resulting in a slight decrease in the correct number of normal number and the recall rate. The result of pruning is 8.44% higher than that of the original logistic regression model. The purpose of pruning is to delete the wrong classification in the positive example and keep the correct classification as far as possible. The more matched the customers, the more obvious the pruning effect. The larger the number of positive examples of correct classification retained in the positive example, the higher the final accuracy of the model.

6.2. Customer Association Experiment. There are a total of 100 in the seed dataset in this paper, which is recorded as a seed dataset. The two target datasets used to screen are a certain financial platform customer dataset (recorded as So) and the second financial platform customer dataset (recorded as SW). The SW has 5,448,509 client data, and all customers are combined with data fusion. The customer's behavior data are used to fill the partial fields, and the SW will retain 5,428,959 customers after processing. The So contains 24,950,474 customer data. After completing the data, the joint filtering vector and threshold in the client filter module are selected, the 73-tested amount of the 2.2 knit is used, whose value is (0.4, 0.3, 0.2, 0.1), and the screening threshold is set to 0.36. The logic regression model selects the logical regression model M6 of combined training. This experiment mainly uses cross-network customers to associate an overall model, from customers in seed data, and uses the proliferation inspiration to discover and identify customers. The entire customers' associated model is running 11 rounds, and its output is shown in Figure 3.

The number of matching customers in Figure 3 represents the total number of associated customers that run each round; each round increases the number of clients representing the number of related clients in each round; the original customer

TABLE 1: Judgment result.

Results	M1(1:1)	M2(1:1)	M3(1:2)	M4(1:3)	M5(2:5)	M6(2:5)
Number of positive cases	183	188	185	177	182	199
Correct number of classifications	115	126	121	117	156	151
Correct number of positive examples	69	77	73	67	89	95
Accuracy	0.3592	0.3935	0.3745	0.3652	0.4851	0.4756
Recall rate	0.4311	0.4874	0.4512	0.4156	0.5563	0.5968

TABLE 2: Decision results of the test set.

Results	Original result set	The result set after pruning
Number of positive cases	199	168
Correct number of classifications	121	152
Correct number of positive examples	151	178
Accuracy	95	91
Recall rate	0.4752	0.5569
Number of positive cases	0.5938	0.5688

represents the number of customers in the seed customer concentrated in the model. In addition to the first round, each round of input is a direct friend of the previous round of identification results. According to the “six-degree space” theory, everyone can connect with 6 people to establish a friend relationship with 6 people. However, after the study of social networks is 13, that is, after connecting with 13 people, customer relationships can form a loop. At this time, the number of friends will not increase. After the model is running 11 rounds, the number of friends (i.e., the original customer) increases; that is, the social network restriction has been reached, and the operation is stopped and output. Since the number of initial customers is small, after the first two rounds of slow start, the model starts from the third round, and the new customer volume starts to increase rapidly, but after running 3 rounds, the number of new customers starting from the 6th round is rapidly reduced, because the average path of the social network is 3.6, there are many shorter paths than the actual path length, and there is a high coarability. After 3 rounds, many customers have formed a loop, resulting in a decrease in new customers. After the model runs 11 rounds, I finally discovered 53,351 to the associated customers. From the number of customers from the entire dataset, the dataset included 5,428,959 customers, the identification ratio was 1%, the proportion of customers identified is very low, and the integrity of primary and customer personal data is very big. Many fields in a large number of customers are blank, and they can be complete and “serious” to fill in the number of customers, which causes this inevitable randomness leading to a low ratio. In order to evaluate the associated customers in the 53,351 discoveries, from these 53,351, 1,000 to the customer randomly, they are divided randomly into 100 groups, 10 customers per group, using artificial manual determination.

6.3. Determination Results Experiment. When the manual determination is used to determine the character of the customer avatar, the customer behavior information and the determination results are shown in Figure 4.

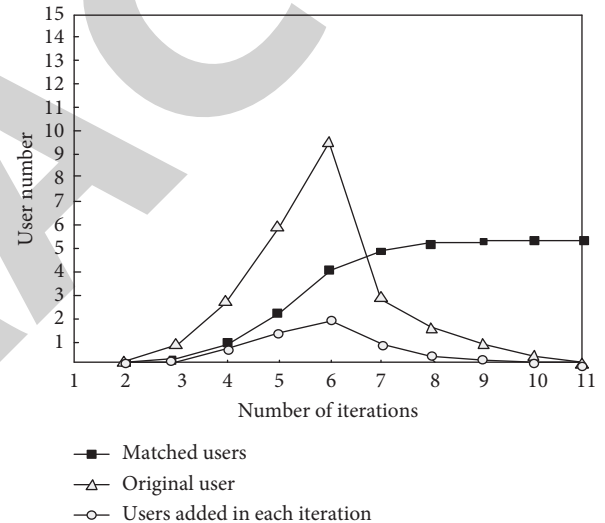


FIGURE 3: User association model operation results.

As can be seen from the accuracy distribution in Figure 4, the average accuracy of the entire customer’s association model is 79%, the minimum accuracy is 40%, the highest is 1, and from the sample result, the overall identification effect of the model can be achieved the expected goal. In order to evaluate the customer’s recognition ability, the customers in the final result set are sorted in similar degree, and the customers who are ranked 1000 are manually determined.

6.4. Judgment Results Experiment. They are also randomly divided into 100 groups, each set having 10 customers, and their judgment results are shown in Figure 5.

As can be seen from the accuracy distribution in Figure 5, the correct rate is 0.5, the highest is 1, the average correct rate is 0.855, and the correct rate of most of the packets is higher than the average value; note that the recognition capacity of the association model is strong; i.e., the probability of customers having similar characteristics in the dataset is high.

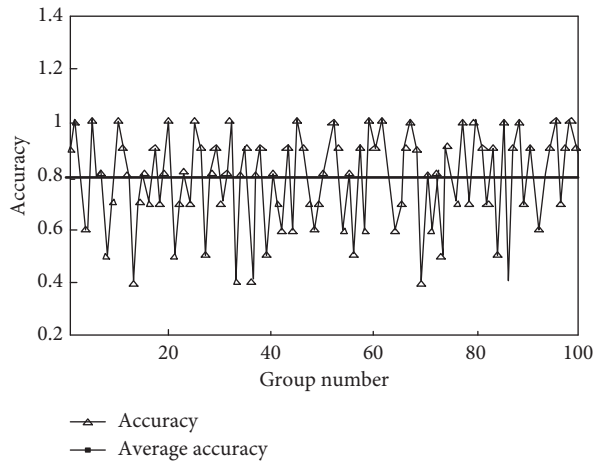


FIGURE 4: Distribution of user-associated random sampling accuracy.

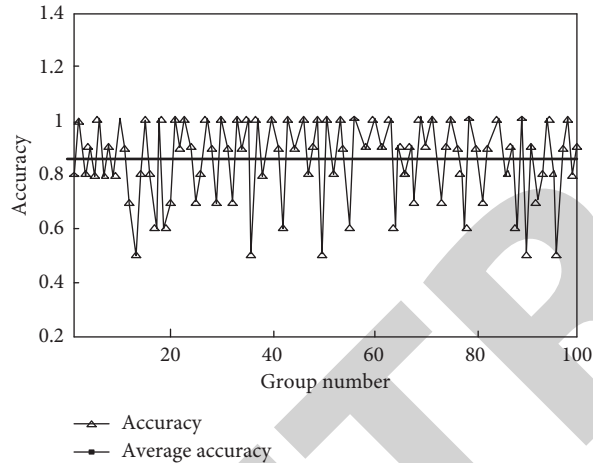


FIGURE 5: Distribution of user recognition accuracy in the top 1000.

7. Conclusions

This paper selects two Internet financial platform networks that have not been studied in China. Henceforth, this paper studies the client technologies related to Internet finance through these selected networks. Moreover, it proposes a cross-network client correlation model. The model uses the relationship between the associated customer network and the customer's candidate set and restores the customer relationship network. Furthermore, it uses the logical regression model to determine whether the customer belongs to the same entity. Finally, the SNC algorithm is used to make a result reduction correction, delete the repeated classification result, and thus, improve the accuracy of the model. Ultimately, this paper discovered 53351 related customers through experiments and proved that the model can effectively associate customers between Internet financial platforms across networks, thereby facilitating the associated certification and behavioral difference analysis of online customer entities. It is helpful to streamline customer information and improve the efficiency of public opinion

supervision. However, we still need to design a more realistic model for practical application.

Data Availability

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

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This paper was supported by the Provincial Soft Science Key Project (2013ZK2024 and 2014GK3147) of the Science and Technology Department of Hunan Province.

References

- [1] Q. Qu, C. Liu, and X. Bao, "E-commerce enterprise supply chain financing risk assessment based on linked data mining and edge computing," *Mobile Information Systems*, vol. 2021, Article ID 9938325, 9 pages, 2021.
- [2] L. Wang, M. Gao, and Z. Liang, "Application of data envelopment and internet of things technology for asset value evaluation," *Mobile Information Systems*, vol. 2021, Article ID 9934090, 8 pages, 2021.
- [3] J. Lu, Z. Gong, and X. Lin, "A novel and fast SimRank algorithm," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 3, pp. 572–585, 2017.
- [4] J. He and D. Yao, "Structural regularization in quadratic logistic regression model," *Knowledge-Based Systems*, vol. 163, pp. 842–857, 2019.
- [5] H. Yao, T. Mai, J. Wang, Z. Ji, C. Jiang, and Y. Qian, "Resource trading in blockchain-based industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3602–3609, 2019.
- [6] W. Zhuang and X. Zhang, "Legal risk and prevention of third-party payment in internet finance," *Journal of Anhui Radio & TV University*, 2019.
- [7] Z. Guan, X. Zhang, and M. Zhou, "Demand information sharing in competing supply chains with manufacturer-provided service," *International Journal of Production Economics*, vol. 220, 2020.
- [8] X. Yu, J. Yang, and Z. Xie, "Training SVMs on a bound vectors set based on Fisher projection," *Frontiers of Computer Science*, vol. 8, no. 5, pp. 793–806, 2014.
- [9] H. Peng, "Research on credit evaluation of financial enterprises based on the genetic backpropagation neural network," *Scientific Programming*, vol. 2021, Article ID 7745920, 8 pages, 2021.
- [10] X. Ju, J. Qian, Z. Chen, C. Zhao, and J. Qian, "Mulr4FL: effective Fault localization of evolution software based on multivariate logistic regression model," *IEEE Access*, vol. 8, pp. 207858–207870, 2020.
- [11] Y. Wang, Y. Che, and X. Lian, "Fast and accurate SimRank computation via forward local push and its parallelization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 99, p. 1, 2020.
- [12] T. Ouyang and X. Lu, "Clustering analysis of risk divergence of China government's debts," *Scientific Programming*, vol. 2021, Article ID 7033597, 9 pages, 2021.

- [13] W. Deng, J. Xu, X. Z. Gao, and H. Zhao, "An enhanced MSIQDE algorithm with novel multiple strategies for global optimization problems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020.
- [14] W. Deng, S. Shang, X. Cai, H. Zhao, Y. Song, and J. Xu, "An improved differential evolution algorithm and its application in optimization problem," *Soft Computing*, vol. 25, no. 7, pp. 5277–5298, 2021.
- [15] X. Han, "Legal regulation of the "price war" of the securities company's brokerage business under the background of internet finance," *Securities Market Herald*, vol. 6, 2019.
- [16] X. Yang, B. Guo, and Z. Mi, "Identification algorithm of high breaching risk member for Internet financial platform," *Application Research of Computers*, vol. 5, 2019.
- [17] C. Zhang, L. I. Xiao-Feng, and Y. Y. Ren, "Research on risk management of internet financial platform based on data mining," *Value Engineering*, vol. 1, 2019.
- [18] H. Han, Y. Yang, and R. Zhang, "A method for dynamically assessing credit risk of internet financial service platform," *DEStech Transactions on Engineering and Technology Research*, vol. 4, 2020.
- [19] Z. Yu, L. Li, W. Zhang, H. Lv, Y. Liu, and U. Khaliq, "A model-based collaborate filtering algorithm based on stacked autoencoder," *Neural Computing and Applications*, vol. 8, pp. 1–12, 2021.