

Research Article

Embedded Microprocessor Wireless Communication Data Collection Aids in Early Warning of Default Risk for Internet Finance Bank Customers

Hui Li 

School of Economics and Management, Northwest University, Xi'an 710127, China

Correspondence should be addressed to Hui Li; lihui@nwu.edu.cn

Received 16 September 2021; Accepted 30 September 2021; Published 19 October 2021

Academic Editor: Guolong Shi

Copyright © 2021 Hui Li. 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.

Thanks to the maturity and innovation of embedded technology, products based on embedded platforms continue to penetrate people's lives and play a pivotal role in various fields of society, while the development of data acquisition systems based on wireless communication of embedded microprocessors is one of the frontier directions of embedded development. Modern commercial bank risk management system includes risk control organization, risk measurement techniques (including risk warning), risk avoidance techniques, and total risk management model. In this paper, a multichannel data acquisition system combined with a wireless sensor platform is designed for early warning of default risk of Internet financial bank customers, which can realize real-time monitoring, acquisition, display, and data storage of DC signals and indoor environment information output from sensor platforms or electronic devices, and for frequent transactions, fund splitting, and mortgage pledging related to complex customers in today's e-commerce platform, which triggers when the associated enterprise credit risk presents complexity, continuity, wholeness, multiplicity, volatility, severity, etc., it is used to improve risk judgment and risk warning ability, enhance investment risk response ability, and reduce the related losses caused by marketing risk.

1. Introduction

With the iterative update of embedded microprocessor platforms, data acquisition devices based on embedded microprocessor platforms have been fully used in a wide range of fields. Data acquisition using data acquisition systems with integrated multiple sensors is a common way for people to understand the external world and themselves, and it converts continuous analog signals from the outside world into digital signals for computer processing for user observation and analysis [1]. With the development of the modern financial industry, large customers of commercial banks are becoming more and more prominent and significant. Increasing the screening services of these customer groups and consideration of their risks are of far-reaching significance for strengthening the control of the entire financial risk within the bank. The boom of the consumer side also makes the credit business in Internet

finance hot, and various credit products start to be launched, which also brings more challenges to credit risk control. These customer businesses, with the development of the market economy and continuous restructuring within the financial industry, have surfaced more financial services needs that have an important impact on the risks of the financial industry. In this process of the bank's business itself and management system, there are also many complications. Recently, the development of information technology in the financial market and the advancement of technology have led to radical changes in the previous traditional business model of commercial banks. The continuous pioneering innovation and perfect development of the commercial bank's marketing model and business philosophy have prompted a major turnaround of the bank's traditional service concept and information exchange platform, information medium, and carrier. With the rapid development of the market economy of the

financial industry and the trend of complexity, diversification, and multipolarity of investment subjects, the risk investment management of commercial banks has not only become a pillar of modern microfinance but also one of the core service functions of modern commercial banks [2]. These risk managements in banks include the early warning of risks in the early stage, mechanisms of risk disposal programs, and specific implementation methods of risk treatment. Therefore, risk management, including risk early warning, has an extremely important position in the management of commercial banks. Specifically, the risk management system of modern commercial banks includes risk control organization, risk measurement techniques (including risk early warning), risk avoidance techniques, and a comprehensive risk management model [3]. The content included in the early warning is all-encompassing and extremely rich, which includes factors of human resources, internal high-tech early warning identification and assessment scoring systems, and partly from other sources such as the expected judgment of historical, current, and future social development, economic, and industry trends.

Through embedded microprocessors, combined with wireless communication acquisition technology, we have taken the initiative to introduce and actively apply advanced risk warning, risk measurement methods, and risk management techniques. The focus is to improve risk judgment and risk early warning capability, enhance investment risk response capability, and reduce losses associated with marketing risks by targeting frequent transactions, fund splitting, and collateral pledging related to complex customers in today's e-commerce platforms, which leads to the complexity, continuity, integrity, multiplicity, volatility, and severity of associated enterprise credit risks, and else, and then promote the banking industry in the financial and commercial aspects. In this way, we will promote the healthy development of the banking industry in the areas of finance, business, and corporate investment.

2. Related Work

The application of embedded microprocessor wireless communication data collection systems in the field of credit risk management is mainly at two levels: various types of structured as well as unstructured customer data processing capabilities and evaluation models based on data mining, artificial intelligence, and other technologies.

The literature [4] argues that under the Basel II regulation, the competition of commercial banks is the competition of banks' risk management capability, which includes risk identification, measurement, evaluation, control, and risk culture. The literature [5] uses the theory of information asymmetry to analyze the causes of credit risk in commercial banks, manifested both internally and externally to the bank, and argues that credit risk management in commercial banks is whole process management. The literature [6] introduced the characteristics of credit risk that distinguish it from other financial risks, including probability distribution thick-tailedness, non-systematic, and information asymmetry, and argued that there are problems in commercial banks' credit risk management such as insufficient rating system, internal control system, legal system, and relatively backward quantitative manage-

ment level. The literature [7] discusses the techniques and practices of Chinese commercial banks to implement internal rating method (IRB) to enhance risk measurement and refinement management by introducing the Basel Accord, where credit risk measurement is discussed in the sections on nonretail exposure rating, credit risk measurement model, credit scoring for retail credit business risk, internal rating for retail exposure, and credit rating model, respectively. The literature [8] discusses the application of quantitative credit risk models from the analysis of the current situation of the construction of customer credit rating systems of commercial banks at home and abroad. The literature [9] used the Bayesian discriminant method, logistic regression model, and BP neural network model to identify credit risk and found that the accuracy of the three to improve, but the BP neural network model lacks the explanation ability than the previous two. The literature [10] analyzed the general idea of establishing a commercial bank management information system and concluded that the commercial bank management information system should be divided into a business processing system, management information system, and president's decision system in terms of architecture and eight modules including risk control management system in terms of module structure. The randomness, short-time reciprocity, time-varying, and uniqueness of wireless channel characteristics are the basis of the physical layer key generation techniques.

The literature [11] considers that the overall structure of the risk management system of commercial banks is a checkerboard-like grid system formed by the intersection of horizontal and vertical, with the horizontal structure divided according to the risk management procedures in sequence and the vertical structure divided according to different businesses. The literature [12] proposed a dual-threshold voltage-based D flip-flop based on CMOS circuit design and improved the clock response circuit of the flip-flop. The literature [13] proposes a checkpoint fault-tolerant low-power scheduling technique for real-time embedded systems. The literature [14] proposes a system power optimization idea and design flow based on hardware-software codesign and looks forward to the development direction of SOC low-power design. The literature [15] gives a method to represent the program execution process using AOV nets and, on this basis, analyzes the critical path of the program operation, finds out the part of the program that can be frequency adjusted and the adjustment magnitude, and minimizes the overall energy consumption of the program while keeping the program performance unchanged. The literature [16] improves the low-power scheduling algorithm based on the DVS technique and proposes a strategy based on average idle time allocation. The algorithm calculates the static running speed and recovers static idle time based on the utilization in the offline phase, recovers dynamic idle time to regulate the running speed of the processor in the online phase, and shuts down the processor using dynamic power management techniques at the right time to reduce the processor energy consumption. Further in-depth research in literature [17] analyzes the current situation of methods, index system, and control of domestic commercial banks' risk evaluation and points out how to build a framework system for banks' internal credit rating models and how to target to avoid the many existing risk defects.

3. Principle Model of Embedded Microprocessor Wireless Communication Data Acquisition-Assisted Internet Financial Bank Customer Default Risk Early Warning

3.1. *Embedded Microprocessor Wireless Communication Data Acquisition Principle.* Most signals in nature exist as nonelectric signals, such as temperature, humidity, and pressure, and other nonelectric signals; to collect and analyze such signals, it is necessary to convert these signals into electric signals through sensors and other devices to input computer equipment for data processing, analysis, and storage; the process can be summarized as the workflow of the data acquisition system [18]. The data acquisition system is divided into two major parts, as shown in Figure 1, the hardware part of the data acquisition card and the software part of the upper computer.

There are many types of wireless communication methods, and several common wireless transmission methods are GPRS, WI-FI, ZigBee, and Bluetooth. The ZigBee network includes three types of nodes: aggregation node (i.e., ZigBee coordinator), routing node (i.e., ZigBee router), and terminal node. The aggregation node acts as the center of the network, and this node is responsible for the establishment and maintenance of the network. The routing node provides the transmission path for data transmission. ZigBee network can support three types of topologies: star, tree, and mesh, as shown in Figure 2 for ZigBee network topology. In star topology, the aggregation node is responsible for the establishment of the network, while the end nodes (routing nodes or end nodes) communicate directly with the aggregation node [19]. In tree topology, the aggregation node functions in the same way as the star topology. And other nodes communicate indirectly through routing nodes or directly with host nodes. In a mesh topology, the aggregation nodes function as in star topology. Through these features of the wireless channel, the confidentiality, privacy, and security of user data in the process of default risk warning for Internet financial banks aided by wireless communication data collection systems can be ensured. Other nodes can communicate with the aggregation node by selecting the transmission path between them.

The establishment and maintenance of the network depend on the aggregation node. The aggregation node establishes the network by listening at a specified range of channels, obtaining the channel energy level and then performing a sequential scan, and the aggregation node selects the channel containing the least number of networks. For example, the application can also establish the network by specifying a working channel. After determining the communication channel, the device can set the remaining network parameters (including PAN identifier, network address, and extended PAN identifier) to establish the network. The principle is as follows:

$$\min K(x) = \frac{pb\sqrt{b^2 + 7ac}}{3a} \int_{\Omega} (1 + \rho) \nabla y dy dx, \quad (1)$$

where the PAN identifier is a 16-bit identifier. For the aggregation node, this identifier is 0x0000 and the broadcast

PAN identifier is 0xFFFF. The ZigBee network can accommodate up to 65536 devices.

$$F(\rho) = \int k^{-i2\pi\omega x} m(z) f(x) dx. \quad (2)$$

The aggregation node establishes the network, and other nodes communicate with the nodes by joining the network established by the aggregation node. The ways of joining the network can be classified as joining by connection, rejoining, joining by configuration, joining by isolated nodes, and joining by allowed nodes. The joining by connection method is where other nodes send a request command by searching the surrounding network and determining the network to join and the identity of the join.

$$KL(x, y) = \frac{\delta y}{\delta x} \cdot \sum_{i=1}^n X_i Y_i. \quad (3)$$

Upon receiving a connection request, the aggregation node determines to determine whether to allow other nodes to join the network. The rejoining method is like joining via connection; the difference is that the rejoin network process interacts with the rejoin request and response commands at the network layer. Configure the joining method in the aggregation node to directly configure the information of other nodes.

$$\gamma_d = \int \beta^{\text{Dirichlet}(m)} \cdot \eta^r dm. \quad (4)$$

The isolated node join method is used after the aggregation node is configured with other nodes and the other nodes join the network.

$$\tau_{ij}(t+1) = (1 - \rho) \cdot \tau_{ij}(t-1) - \prod_{k=0}^{n-1} \Delta \tau_k. \quad (5)$$

The allow node joining approach is used to cope with the coexistence of multiple networks, and the network uses this approach to control the joining of nodes. The departure of nodes is simpler than the joining of nodes. The departure of a node is divided into active and passive departure. The node active leaving is done by sending a leave command. The passive departure of a node on the other hand requires the aggregating node to send a leave command to the leaving node. With the popularity of multicore processor architectures, more and more cores are being integrated into a single chip [20]. Using off-the-shelf cores as redundancy becomes a viable solution. Redundant multithreading (RMT) is an effective approach, with the main principle of creating a thread on each core, each running the same copy of the program and comparing the resultant values computed in different cores at the point of error detection. Since errors that occur on one core are not propagated to the other core, if a resulting mismatch occurs, it can be determined that a soft error occurred.

A large part of the embedded systems is the real-time systems, which have strong real-time requirements. Embedded real-time operating systems are low-cost, low-power,

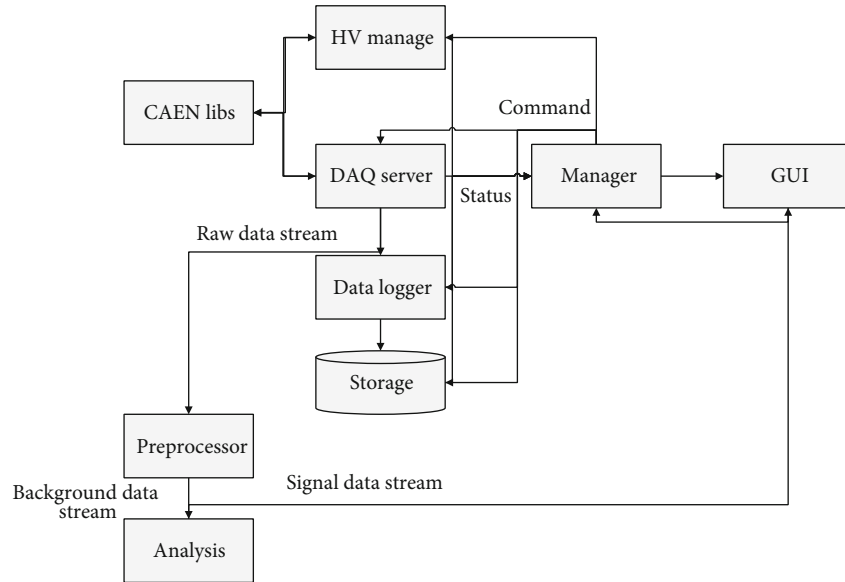


FIGURE 1: Data acquisition system architecture.

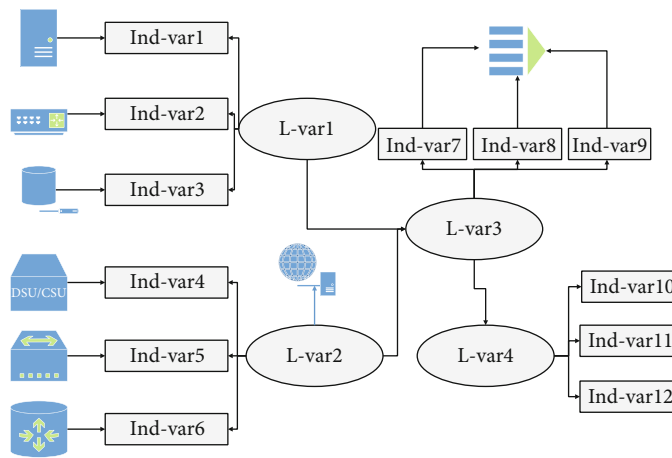


FIGURE 2: ZigBee network topology.

and small computers that are embedded in other mechanical or electrical systems. Typically, they include a processor, power supply, memory, and communication ports. Embedded systems use communication ports and communication protocols to transfer data between the processor and peripheral devices (usually other embedded systems). The processor interprets this data using software code stored in memory. The software is usually specifically designed and optimized for the functions undertaken by the embedded system. The processor may be a microprocessor or a microcontroller. A microcontroller is simply a microprocessor with a peripheral interface and integrated memory. Microprocessors use separate integrated circuits for storage and peripherals, rather than including them on the chip. Both can be used, but because microprocessors have less integrated circuitry, they typically require more support circuitry than microcontrollers. A system on a chip (SoC) is a complete system with multiple processors and interfaces

integrated on a single chip. They are often used in high-capacity embedded systems. Some examples of SoC types are application-specific integrated circuits (ASICs) and field-programmable gate arrays (FPGAs). An embedded real-time operating system is designed for real-time applications. It is intended to serve the real-time application requests of embedded systems, industrial robots, research equipment, and other systems. Although a single-core processor can run only one thread at any given point in time, the operating system allows multiple threads to execute simultaneously. Real-time systems take advantage of the concurrency of the processor precisely by making each thread appear to be executed exclusively on the processor. For example, if two tasks are executed simultaneously on a single-core processor, they may appear to be running at the same time from the perspective of both tasks. However, the processor is only executing one task at the same moment. Reliability and fault tolerance techniques are

designed to prevent as a large amount of system failures as possible when the system suffers transient errors (soft errors). Transient errors, while not permanently damaging hardware components, can cause transient bit flips (from 0 to 1 or 1 to 0) that can cause run failures or incorrect run results. Researchers in academia and industry typically take advantage of the redundancy inherent in modern systems and the availability of redundant resources (such as extra cores, time, or threads) to provide effective solutions to tolerate soft errors [21]. Error detection, error containment, and error recovery are often the basic steps in providing reliability against soft errors. Error detection techniques only need to recognize the occurrence of soft errors. To prevent the use of erroneous data by the system, the propagation of the error needs to be isolated and the errant unit rewritten using the correct data. After detection and rewriting, the system can be restored to an error-free state, which is called error recovery.

Randomness, short-time reciprocity, time-varying, and uniqueness of wireless channel characteristics are the basis of physical layer key generation techniques.

(1) Randomness: the random nature of changes in the wireless communication environment makes the wireless signal after passing through a multipath fading channel produce a random fading, and the wireless channel characteristics it experiences are random and unpredictable

(2) Short-time reciprocity: the received signals at both ends of the cochannel simultaneous half-duplex upstream and downstream links will experience the same channel fading characteristics in theory during the channel coherence time. If Alice and Bob extract the characteristics of the wireless channel between them at the same moment, the extracted channel characteristics are the same, i.e., $h_{AB} = h_{BA}$. If the nodes communicate in TDD mode, the channel characteristics measured by both legitimate parties in a channel coherence time satisfy $h_{AB} = h_{BA}$ due to the time-varying and random nature of the wireless channel characteristics. This provides the possibility of using the wireless channel characteristics to generate shared key which provides the possibility to generate shared keys

(3) Time-varying: due to the diversity of network topologies, the mobility of wireless terminals, and the variability of wireless communication environments, the wireless channel is always in dynamic change. Therefore, the wireless channel characteristics are randomly time-varying at any moment. This provides the possibility to prevent eavesdroppers from obtaining valid key information and to achieve the ideal confidential communication of “one secret at a time” between legitimate parties

(4) Uniqueness: it means that the channel characteristics experienced by receivers in different frequency bands, different locations, and different times are not the same; then, the channel characteristics are unique. This characteristic provides a guarantee for the security of physical layer key generation

The development of quantitative criteria for default identification of group customers by commercial banks, qualitative criteria for default identification of group customers by commercial banks, data on which the classifica-

tion in risk early warning is based, criteria for selecting sample data in the risk early warning system, establishing the definition of default for group customers, determination of financial data and valid information selected by commercial banks for rating group companies, and indicators for a rating of corporate group customers by bank selection. These features of the wireless channel ensure the confidentiality, privacy, and security of user data during the wireless communication data collection system-assisted Internet finance bank customer default risk warning.

3.2. Early Warning System for Default Risk of Internet Finance Bank Customers. The establishment of the model of bank risk early warning as shown in Figure 3 requires a huge number of different data conditions as a basis for judgment support. These different data come from all kinds of information that commercial banks may obtain, and the comprehensive posting and judgment of these different information sources and information conditions are also a vital part of its risk early warning, moreover a prerequisite part. All data need to be analyzed and analyzed before they can become valid factors for judging the authenticity of the situation [22]. However, these factors are not the only ones, but also contain a lot of additional information to determine the general situation of the group's customers in a comprehensive manner: quantitative criteria for identifying the default of the group's customers by commercial banks, qualitative criteria for identifying the default of the group's customers by commercial banks, data on which the classification in the risk warning is based, criteria for selecting the sample data in the risk warning system, establishing the definition of default of the group's customers, determination of financial data and valid information selected by commercial banks for rating group companies, and selection of rating indicators for banks' rating of corporate group customers.

The risk early warning mechanism of a commercial bank contains two main aspects of information. The establishment of early warning models must also consider the requirements of these aspects. In addition to risk warning for group customers, commercial banks also need to consider the construction of a system for precontrol of crises, so it is a comprehensive operational system for early warning and precontrol. During the construction of this system, commercial banks integrate various risk data as the conditions for early warning judgment and precontrol measures under the situation of information asymmetry with group customers and uncertainty of various external environment and changes, make system construction for the main research and analysis in the early warning system, and explore its corresponding in monitoring and preventing commercial bank risks and preventing group customers' business risks, [23]. Commercial bank risk early warning machine includes the following functional requirements.

Early warning modeling needs to reflect the identification function of commercial bank risk early warning. The identification function of the commercial bank risk early warning mechanism is to identify commercial bank risk factors on the premise and basis of the effective judgment and identification of possible risks and to control and manage them.

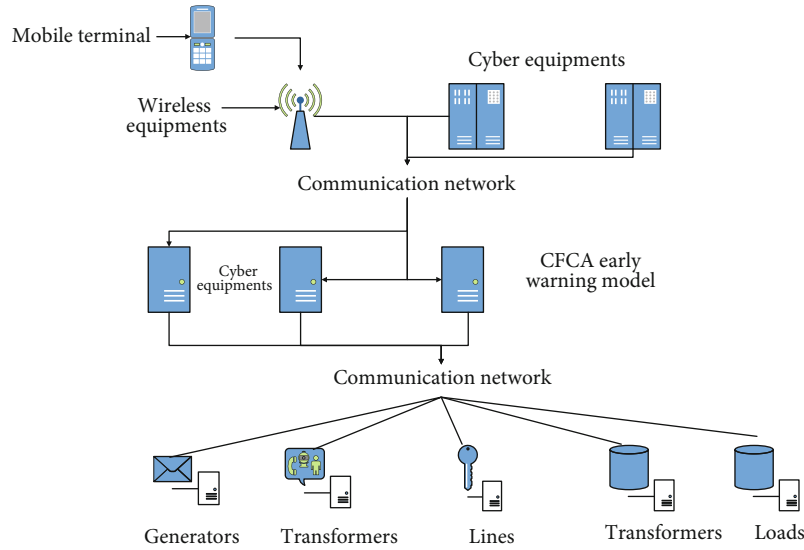


FIGURE 3: The early warning model for bank default risk.

Classification algorithm analysis classifies the risk warning approaches of commercial banks. One of the more common approaches is the qualitative calculation method. In the process of using this calculation method, the accuracy of the parameters is a very heavy issue, because although this model reflects a risk probability, it requires that the parameters must be accurate. Early warning modeling needs to reflect the identification function of commercial bank risk early warning. The identification function of the commercial bank risk early warning mechanism is to identify the premise and basis of the commercial bank risk factors, make effective judgments and identification of the possible risks, and control and manage them. This is because the certainty of these parameters can directly affect the final ratio results. This is because it directly affects the risk judgment of the business group by the commercial banks in conducting the risk assessment process. But this aspect is influenced by the quantity and quality of data and can be continuously improved. The main idea is that the first step is to use the concept of “default” in the new rating methodology as a basis for risk prediction. The second step is to use the probability of default, i.e., the likelihood of default, as a measure of the customer’s creditworthiness and to predict the level of risk that may exist. In the third step, the probability of default is measured by a credit rating model during the commercial bank’s operations. In the fourth step, the probability of default is generally divided into thousands of intervals, each representing a credit rating, to make the commercial bank’s risk warning model easier to use in practice and to grade the customer’s specific credit situation. The fifth step is to use the credit rating as a visual representation model of the level of credit risk: a higher credit rating of a customer indicates a lower level of credit risk; conversely, a lower credit rating of a customer indicates a higher level of credit risk. In these steps, the bank uses this step-by-step model to calculate data on the probability of multiple possible risks for corporate and group customers, which results in a final rating display for specific business operations with cus-

tomers. This single risk warning model helps to implement in a context where the management system, as well as the company’s cultural system and cultural traditions, is very strict, but in the current situation where the bank’s management system is not yet sound for the external environment and internal management personnel changes and system changes in the operation of the enterprise group cannot be effectively calculated, this method ultimately cannot predict the full risk situation. Many factors and political factors can also influence the final value judgment. Such data are a direct reflection of the specific business methods, operating conditions, and business models of the corporate group’s clients about the large loans granted by the banks. In turn, these different scenarios transmit the specifics of the conglomerate to the bank’s supervisory authorities.

4. Experimental Verification and Conclusions

Personal credit desensitization data from a real business scenario of an Internet bank is used as an example for personal risk assessment modeling to predict the future credit risk of customers. Most of the variables selected for constructing credit assessment models in the past are related to customers’ application information (cross-sectional data) and lack data related to customers’ loan behavior information (panel data). The data used in this paper contains information about customers’ loan behavior, which can make the model construction more comprehensive and effective. The original data includes 94 desensitized features of customers, including 18 discrete features and 71 continuous features, with a total of 15625 research samples. The customer basic information table is the desensitized customer demographic and basic situation information, and the variables mainly include gender, age, educational background, and other instances of customer authorization to collect each profile label at the time of application. The customer behavior information table is the customer’s postloan disbursement behavior data information up to a certain observation point

in time, and the variables mainly contain the repayment withdrawal behavior, late behavior, and other various routines for the profile tags at various points in time after the loan. The customer performance label information is the risk performance after the observation time point; the label variable name is “TARGET”; 0 means the customer has no default during the observation period, and 1 means the customer default.

Missing data can interfere with subsequent data analysis as well as model construction and can bias the model prediction results. Missing values are presented in various forms; the common ones are NULL values in the database and special values used to represent missing values. The special values for missing values often refer to values such as -99 and -999. Outliers are unreasonable points in the dataset. These outliers, if applied directly to the model construction without removing or correcting them, will have a very bad impact on the final empirical results and lead to biased results. In this paper, the missing data are all in the form of the value -99. Missing data can be analyzed in terms of missing records and missing field information, which can be handled in three ways: deletion, interpolation, and no processing.

Outliers are unreasonable points in the dataset, and if these outliers are not removed or corrected and applied directly to the model construction, they will have a very bad impact on the final empirical results and lead to biased results. Therefore, it is especially important to pay attention to the appearance of outliers, analyze their causes, and remove or correct them. Since the original data in this paper is desensitized data and there are no obvious outliers, the main test in the following is whether there are outliers. In this paper, outliers are identified as follows: the top 3 features in terms of importance are filtered out using the XGBoost algorithm (Figure 4), and for each record, if more than half of the top 3 features are missing, it is considered an outlier and needs to be removed. Combined with the data in this paper, 3 features contain missing data, and only two features contain missing data and ranked in the top 3 in importance, so there are no outliers and no need to delete the sample records.

To better compare the performance between the grid-XGBoost model and the GA-XGBoost model, the test set is divided with the same ratio for both models. The GA-adjusted XGBoost model successfully predicted 26476 defaulters and 31347 nondefaulters in the test set, with a recall rate of 0.9998 and a precision rate of 0.9472. Compared with the accurate index of the grid-XGBoost model, the accuracy of 0.0075 is improved, and the number of honest users misclassified as defaulters has increased slightly from 3 to 7. As shown in Figure 5, the classification accuracy and loss function of both the training and test datasets show a decreasing trend and level off during the model fitting process, which can be intuitively seen that the model shows a more stable phenomenon.

In Figure 6, the value of precision for the GA-XGBoost model has increased compared to the grid search-adjusted model. The value represents the proportion of users who are predicted not to default who do not default, and a larger value of this value also represents a smaller proportion of

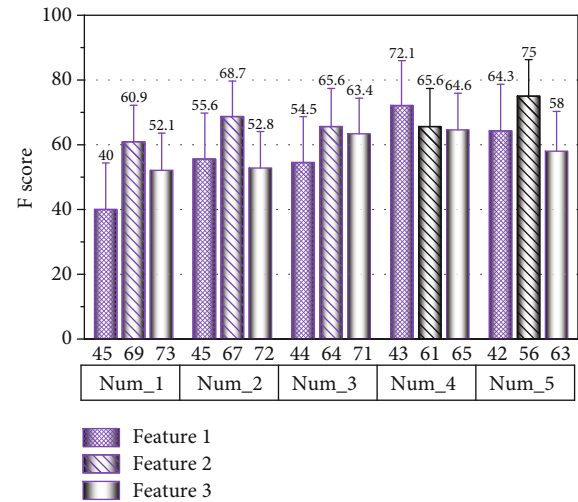


FIGURE 4: Ranking the importance of key features.

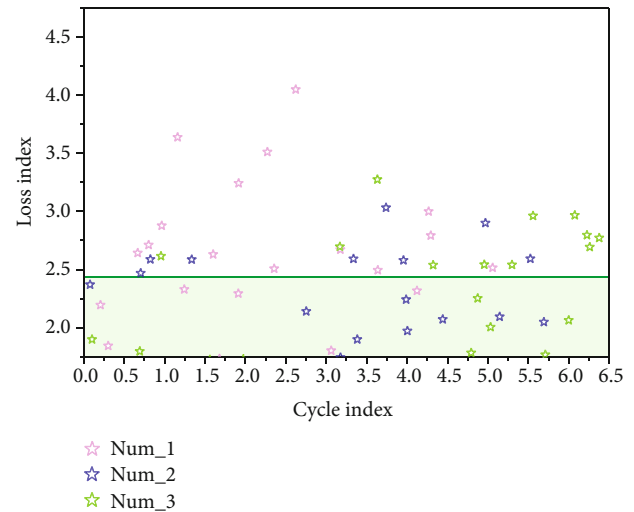


FIGURE 5: Model loss function.

users who are incorrectly judged to be nondefaulters. The GA-XGBoost model does show better classification results than the lattice-XGBoost model when compared with the other model. The GA algorithm is better than the grid search because firstly the GA algorithm adjusts seven parameters at the same time, which is faster than the grid search; secondly, the selection probability of the GA algorithm is based on the fitness value during the parameter adjustment process, which is the $F1$ -score in this training process, and the probability of selecting the parameter value with a large $F1$ -score is also larger. “The principle of “superiority and inferiority” is used to select the parameters, and the crossover is to exchange information between the chromosomes, using the principle of uniform crossover, so that the parameters that perform well can be retained.” Approach allows him to perform so well. In addition to the GA algorithm performing very well in the process of adjusting the parameters, the ranking of the importance of the features in both models

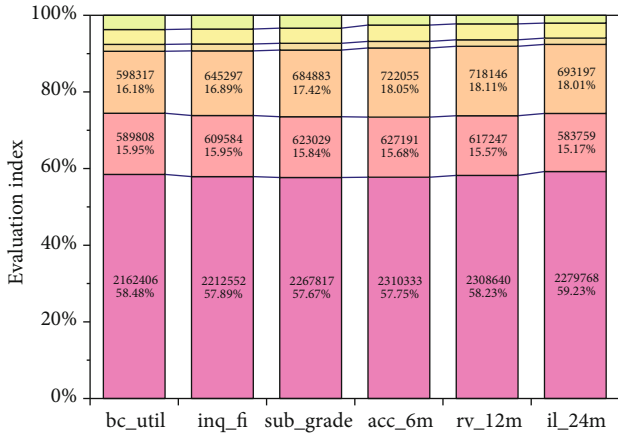


FIGURE 6: Evaluation metrics for different tuning algorithm models.

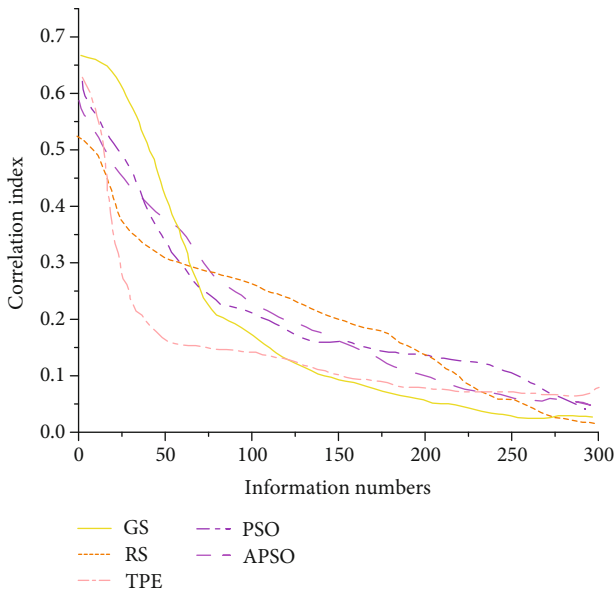


FIGURE 7: Correlation of algorithm evaluation metrics with soft information as well as hard information.

likewise reflects that the lender’s financial status and the lender’s past defaults carry a very large weight in determining whether a user is in default, in particular, the variables bc_util (ratio of total current balance to high credit limit for all bank card accounts), inq_fi (personal finance amount), sub_grade (borrowing subgrade), open_acc_6m (amount of open transactions in the past 6 months), open_rv_12m (number of revolving transaction options in the past 12 months), and open_il_24m (number of installment accounts in the last 24 months).

Soft information is numerical data that cannot be measured by a clear scale or specific metrics; in contrast, information that can be expressed in terms of precise values with definite measures is called hard information. In the field of finance and economics, there are two types of information that one is often exposed to, one is nonfinancial information such as personal reputation and relationships;

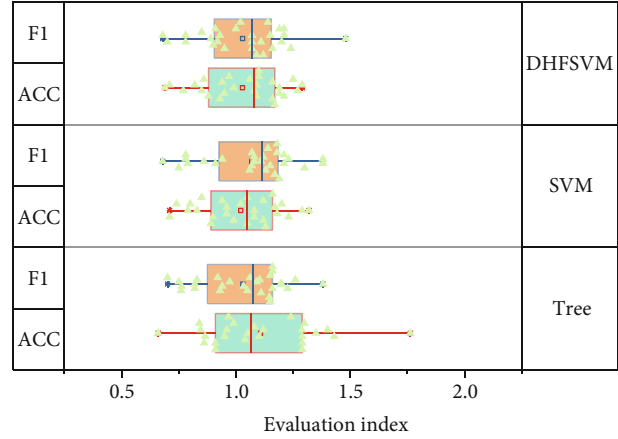


FIGURE 8: Comparison of the two types of classification error rates for multiple models.

the other is financial information such as age, gender, and annual income, which are soft and hard, respectively. Simply understood, soft information is numerical data that cannot be measured by a clear scale or a specific indicator; in contrast, information that can be represented by an accurate numerical value with a definite measure is called hard information. Soft information is specific to loan-related information and responds to a range of trends. Hard information acts more like a skeleton, while soft information explores more possibilities beyond the skeleton. Soft information is the overall scope, and the scope includes a variety of factors, both subjective and objective, that can affect loan defaults. Hard information is objective, specific, and accurate, metrics that can be measured and data that can be processed. The field of credit forecasting specifically refers to the lender’s loan data. Hard information is all information that reflects objective facts. Soft information, on the other hand, refers to the textual description of the application in the credit application. The difference between hard information and soft information is that hard information is recorded objective facts and answers are not questions, while soft information contains personal judgment touching on the subject’s values and knowledge structure system, reflecting more subjective preferences as well as quality personality. Figure 7 shows the correlation between algorithmic evaluation indexes and soft information as well as hard information.

With 144 correctly predicted samples and 10 incorrectly predicted samples out of 160 nondefaulted samples (i.e., negative class samples), this yields a type I error rate of approximately 10.0% from equation (4). With 24 correctly predicted samples and 16 incorrectly predicted samples out of 40 defaulted samples (i.e., positive class samples), the second type of error rate can be obtained as approximately 40.0%, which indicates that the error rate of logistic regression defaulted borrowers is high relative to the prediction error rate of nondefaulted borrowers. The more accurate the credit risk assessment results for borrowers in the personal loan business, the less likely they are to default and the lower their credit risk. Comparing the first classification error rate with the second classification error, the first classification error rate, as well as the second classification error

rate based on the three classification models mentioned above, is plotted, as shown in Figure 8. It is easy to find that among decision tree, support vector machine, and embedded microprocessor-based wireless communication data collection aid model, the proposed embedded microprocessor-based wireless communication data collection aid model in this paper has the lowest classification error rate in both categories with values of 7.5% and 25.00%, respectively, which shows the effectiveness and reliability of embedded microprocessor-based wireless communication data collection aid model.

5. Conclusions

The boom of the consumer side also makes the credit business in Internet finance hot, and various credit products start to be launched, which also brings more challenges to credit risk control. The research of this paper is mainly based on the research of embedded microprocessor wireless communication data acquisition auxiliary system, through the design of signal conditioning circuit, the minimum system circuit to achieve multichannel data acquisition, upload of the data to the upper computer through two transmission methods, and realize the data measurement, display, storage, and other functions through the review of literature and information to understand the current development status and application background of data acquisition system at home and abroad. In the general environment of the whole financial market, there are various types of data acquisition systems with different functions according to specific occasions or user needs. After understanding the acquisition principle and structural composition of different acquisition systems, a design scheme of an embedded microprocessor-based wireless communication data acquisition-assisted Internet financial bank customer default risk warning system is proposed. Suitable samples, index system, and SMOTE sampling method are selected, combined with parameter tuning to optimize the enterprise credit rating prediction model, construct a low-cost prediction model with strong generalization ability, and improve the enterprise credit rating prediction scheme. The sample data in this paper are selected from the Wind database, and the original data are processed based on the SMOTE sampling technique. The XGBoost algorithm is innovatively introduced into the field of enterprise credit rating prediction, and the regular term is added to the model and the second-order Taylor expansion is applied. In addition, the Bayesian optimization algorithm is applied to perform parameter search, and the Gaussian process is used to continuously update the posterior distribution of the objective function to find the global optimal solution, which solves the problem that the grid tune is easy to get local optimal for nonconvex problems and increases the prediction accuracy of the default risk rating of Internet financial bank customers.

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 that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] H. Farrell and A. L. Newman, "Weaponized interdependence: how global economic networks shape state coercion," *International Security*, vol. 44, no. 1, pp. 42–79, 2019.
- [2] D. B. Avancini, J. J. P. C. Rodrigues, S. G. B. Martins, R. A. L. Rabêlo, J. al-Muhtadi, and P. Solic, "Energy meters evolution in smart grids: a review," *Journal of Cleaner Production*, vol. 217, pp. 702–715, 2019.
- [3] J. Kaur Chahal, A. Bhandari, and S. Behal, "Distributed denial of service attacks: a threat or challenge," *New Review of Information Networking*, vol. 24, no. 1, pp. 31–103, 2019.
- [4] B. R. Gaines, "From facilitating interactivity to managing hyperconnectivity: 50 years of human-computer studies," *International Journal of Human-Computer Studies*, vol. 131, pp. 4–22, 2019.
- [5] H. C. Chou and C. Y. Huang, "Development of a walking aid with remote electrocardiogram measurement and global positioning system function," *Journal of the Chinese Institute of Engineers*, vol. 44, no. 3, pp. 261–276, 2021.
- [6] C. D. Warren, E. Nel, and P. J. Boyd, "Controlled comparative clinical trial of hearing benefit outcomes for users of the Cochlear™ Nucleus®7 sound processor with mobile connectivity," *Cochlear Implants International*, vol. 20, no. 3, pp. 116–126, 2019.
- [7] A. L. Imoize, S. D. Odeyemi, and J. A. Adebisi, "Development of a low-cost wireless bee-hive temperature and sound monitoring system," *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*, vol. 8, no. 3, pp. 476–485, 2020.
- [8] Y. Wei, J. Zhou, Y. Wang et al., "A review of algorithm & hardware design for AI-based biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 2, pp. 145–163, 2020.
- [9] X. Li, N. Zhao, R. Jin et al., "Internet of Things to network smart devices for ecosystem monitoring," *Science Bulletin*, vol. 64, no. 17, pp. 1234–1245, 2019.
- [10] R. T. Dwyer, J. Roberts, and R. H. Gifford, "Effect of microphone configuration and sound source location on speech recognition for adult cochlear implant users with current-generation sound processors," *Journal of the American Academy of Audiology*, vol. 31, no. 8, pp. 578–589, 2020.
- [11] R. K. Lomotey, S. Sriramoju, and R. Orji, "Machine-to-infrastructure middleware platform for data management in IoT," *International Journal of Business Process Integration and Management*, vol. 9, no. 2, pp. 90–106, 2019.
- [12] T. Wesarg, Y. Stelzig, D. Hilgert-Becker et al., "Application of digital remote wireless microphone technology in single-sided deaf cochlear implant recipients," *Journal of the American Academy of Audiology*, vol. 31, no. 4, pp. 246–256, 2020.
- [13] M. D. Fletcher, "Using haptic stimulation to enhance auditory perception in hearing-impaired listeners," *Expert Review of Medical Devices*, vol. 18, no. 1, pp. 63–74, 2021.
- [14] W. Xu and J. Leng, "Data transmission and real-time information acquisition based on fine-grained algorithm and wireless

- network,” *Peer-to-Peer Networking and Applications*, vol. 13, no. 4, pp. 1120–1130, 2020.
- [15] L. M. Thibodeau, “Benefits in speech recognition in noise with remote wireless microphones in group settings,” *Journal of the American Academy of Audiology*, vol. 31, no. 6, pp. 404–411, 2020.
- [16] S. Roy, N. Mazumdar, and R. Pamula, “An energy and coverage sensitive approach to hierarchical data collection for mobile sink based wireless sensor networks,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 1267–1291, 2021.
- [17] U. J. L. dos Santos, G. Pessin, C. A. da Costa, and R. da Rosa Righi, “AgriPrediction: a proactive Internet of things model to anticipate problems and improve production in agricultural crops,” *Computers and Electronics in Agriculture*, vol. 161, pp. 202–213, 2019.
- [18] D. A. Lilien, B. H. Hills, J. Driscoll, R. Jacobel, and K. Christianson, “ImpDAR: an open-source impulse radar processor,” *Annals of Glaciology*, vol. 61, no. 81, pp. 114–123, 2020.
- [19] S. Ivanov, J. Renneson, J. Fontaine et al., “Evaluating IoT based passive water catchment monitoring system data acquisition and analysis [J],” *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 4, pp. 1373–1382, 2019.
- [20] F. Al-Turjman, M. H. Nawaz, and U. D. Ulusar, “Intelligence in the Internet of Medical Things era: a systematic review of current and future trends,” *Computer Communications*, vol. 150, pp. 644–660, 2020.
- [21] M. A. Khan, P. Paul, M. Rashid, M. Hossain, and M. A. R. Ahad, “An AI-based visual aid with integrated reading assistant for the completely blind,” *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 6, pp. 507–517, 2020.
- [22] A. S. Dasgupta, B. M. S. Zambare, C. N. Kulkarni, and A. D. Shaligram, “Real-time water quality monitoring system for rivers [j],” *Acta Technica Napocensis*, vol. 60, no. 2, pp. 17–22, 2019.
- [23] A. Heidary, H. Radmanesh, S. H. Naghibi, S. Samandarpour, K. Rouzbehi, and N. Shariati, “Distribution system protection by coordinated fault current limiters,” *IET Energy Systems Integration*, vol. 2, no. 1, pp. 59–65, 2020.