ISSN:2790-1688 Volume-10-(2024)

A Novel Financial Anti-fraud Method based on Machine Learning Algorithms

Daokang Jiang *

New York University Shanghai, Shanghai, China,

* daokangjiang@gmail.com

Abstract. Digital finance is booming, financial technology is maturing, and the development of information technology has had a massively positive impact on society. However, such progress also introduces a new type of risk: the underground network industry is experiencing explosive growth, and telecommunications network fraud has caused enormous financial losses. In the age of digital finance, although commercial banks have ushered in new possibilities and created momentum, they must also confront new obstacles and needs for digital transformation. In this context, online financial services have emerged as the new primary battleground. In this study, based on the RFM high-dimensional derived features and machine learning techniques, a high-dimensional transaction behavior portraits-based anti-fraud machine learning model is created. Using big data, stream computing, and other technologies, as well as systematic deployment, application strategy, and iterative model optimization, we developed a set of machine learning-based in-event risk control solutions. Confirmed to have an AUC of 0.972, this model provides key insight into fraud risk, identifies fraudulent transactions in milliseconds, and has value and relevance for enhancing the in-transaction risk management capabilities of online digital financial organizations.

Keywords: digital finance; machine learning; anti-fraud; smart finance; fintech.

1. Introduction

1.1 Background and significance

In the context of the new wave of technological revolution and industrial change, digitalization in the banking industry has flourished and new technologies, such as big data, artificial intelligence, and cloud computing, have become deeply integrated with financial business. These tools have served as a new engine to promote financial transformation and upgrading, boost the real economy, and prevent and resolve financial risks; thus, digital transformation has become a popular choice for commercial banks to improve their service quality and competitiveness [1-3].

With the acceleration of commercial banks' innovative transformation utilizing new technologies, online banking, mobile banking, mobile payments, and other online digital financial services have provided more room for innovation and more efficient and high-quality financial services for the customers of these banks, while also posing a significant challenge in regard to anti-fraud measures [4-7]. Driven by the massive economic interests, numerous instances exist of unscrupulous agents stealing and defrauding consumers' monies through a variety of ways, including phishing links, trojan viruses, and telecom theft, have occurred. Fraudulent internet transfers and payments have evolved into a well-organized and well-defined black industrial chain, resulting in substantial losses for both clients and banks [8-9].

The purpose and objective of banks is to safeguard their clients' funds. The need for delivery of safer Internet services to consumers and protect them from fraud, skimming, and other forms of Internet blackmail has raised the bar for banks' anti-fraud capabilities and become a basic competency that banks must develop in the digital financial era [10].

1.2 Difficulties in anti-fraud transaction measures

The peculiarities of digital online financial transactions have made it increasingly difficult for banks to authenticate user identities throughout the transaction process [11]. After years of

development, major banks have gradually established a supporting anti-fraud system in accordance with the growth of online business. This system is predominantly based on the rule model designed by risk-control business experts, effectively protecting the security of customer funds and combating criminal activity [12-13].

With the assistance of Internet and mobile Internet technology, criminals' fraudulent practices are expanding and evolving, and the conflict between banks and criminals continues to intensify. The relatively static expert rule model has certain limitations in the accurate identification of fraud cases and the ability to adapt to new cases, primarily due to three factors: first, the expert rules are formed by summarizing and generalizing fraud cases, rather than based on the behavioral differences between fraudulent transactions and normal transactions, and the lack of experience makes it difficult to avoid the " benighted " limitation; second, the expert rules are rigid, and a set of standards applies to all customers. Therefore, it is impossible to identify fraudulent transactions on a "person-by-person" basis. While missing fraudulent transactions, the model will mistakenly determine normal transactions of normal customers as fraudulent transactions, affecting customer experience; third, in a quick confrontation with network blackmail, the expert rules for the prevention and control of new kinds of cases have a certain lag, necessitating that risk control experts continually summarize the case's laws in order to improve the rule model.

Moreover, online digital finance businesses will generate high concurrent, massive, and multi-dimensional data. How to effectively use the data generated by online businesses, the new technical architecture to solve the problem of timeliness, and improve the fraud prevention capability of online transactions based on big-data methods are urgent issues in the field of transaction anti-fraud.

1.3 New direction of risk control based on big data and artificial intelligence

In recent years, innovative technologies, such as big data and artificial intelligence, have offered commercial banks new avenues for enhancing their risk management skills. Based on the "streaming" technology of big data, the Spark distributed cluster computing engine can quickly and efficiently process the massive data generated by real-time transactions [2]; based on the historical data storage and computing capability of the big data platform, it can flexibly process long-term, multi-dimensional historical transaction behavior characteristics, laying a solid technical foundation for using massive data to conduct real-time ex post analysis. This has established a robust technological basis for using enormous amounts of data for real-time risk control during events.

Some commercial banks have actively investigated the application of new technologies, such as big data, stream computing, and artificial intelligence, and have established a new type of intelligent anti-fraud risk control system by adopting an advanced distributed big data system architecture and stream processing engine. The system has high concurrency, low latency, high accuracy, and high reliability, can determine risk in milliseconds, and can enable real-time machine learning models and intelligent risk judgments.

Based on the intelligent anti-fraud system's stream processing capability, big data computing capability, artificial intelligence support capability, and real-time decision-making capability, this paper conducts model research and practical application of anti-fraud machine learning algorithms for the transaction dimension of online scenarios. We model fraudulent transactions using machine learning and other artificial intelligence algorithms, adopt high-dimensional features to profile each transaction, and recognize transactions that do not match the historical profile through a detailed depiction of each transaction and a comprehensive "understanding" of historical transaction behavior. This provides a flexible security service to each customer, identifies fraudulent transactions more precisely and thoroughly, detects new types of fraud in a timely manner, prevents fraud to the greatest extent possible, and effectively improves the risk prevention and control capability of online digital finance businesses.

2. Modelling Programmes Based on Transactional Behavioural Profiling

2.1 Algorithm selection

A modified version of the extreme gradient boosting (XGBoost) decision tree machine learning technique is adopted for modeling based on considerable study and practical validation. The method employs the tree model as its primary classifier. It is generally acknowledged by the industry in terms of classification effectiveness, business interpretation, and modeling efficiency, and is currently one of the most popular machine learning models.

XGBoost is an improvement on the traditional gradient boosted decision tree (GBDT) model, which mainly optimizes the processing of the objective function and the missing value samples.

The XGBoost objective function is [3-4]:

$$\lambda^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{1}$$

The objective function is rewritten through a second-order Taylor expansion as

$$\lambda^{(t)} \cong \sum_{i=1}^{n} [l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_i(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \tag{2}$$

Substituting the traditional GBDT training process, we obtain the XGBoost objective function:

$$Obj^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_i(x_i) + \frac{1}{2} h_i f_t^2(x_i) + \Omega(f_t) + constant \tag{3}$$

The first-order and second-order terms are derived by Taylor expansions are as follows:

$$g_i = \partial_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right) \tag{4}$$

$$h_i = \partial^2_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right) \tag{5}$$

From the point of view of the objective function, XGBoost executes a second-order Taylor expansion on the objective function in conjunction with a second-order Taylor expansion of the loss term, which maintains more information about the goal function. Simultaneously, XGBoost introduces a new regularization term of branch node weights, which reduces the model's variance [5] and enhances its performance and efficacy.

From the perspective of missing value processing, XGBoost adds the strategy of automatically processing missing values by calculating the objective function of missing value samples in different branching states and automatically classifying missing value samples based on the merits of the objective function [6], which can effectively solve the problem of missing information of devices, locations, IP addresses, and other elements in anti-fraud scenarios.

In-depth comparison with logistic regression (LR), random forest (RF), support vector machines (SVM), artificial neural networks, and other artificial intelligence algorithms reveals that XGBoost has superior classification impact, effective business interpretation, and rapid online prediction. The comparison of prevalent AI algorithms is shown in Table 1.

Based on the characteristics of high transaction volume and sparse positive samples in online digital finance businesses, as well as the requirement for high response speed for real-time decision making, especially when fraud patterns or samples change, the speed of prevention and generalization ability of the XGBoost algorithm make it more suited to real-time decision making. Consequently, XGBoost is chosen as the machine learning method for this application.

Table 1. Comparison of mainstream artificial intelligence model algorithms

Algorithm	Advantages	Disadvantages
XGBoost	Excellent classification results; good operational interpretation; Fast prediction speed	High model complexity, easy to overfit
LR	simplicity of execution; rapidity of foresight; business-savvy interpretation	High workload of the data processing process of logistic regression when the dimensionality of the features is high; Inability to consider the link between characteristics; Simple to under- and over-size;
RF	The capacity to efficiently process massive data sets; high capacity for generalization of the model; Typically requires no additional trimming	tendency to over-fit; features that need more value divisions tend to have a bigger influence on the choice, which in turn influences the effect of the fitted model.
SVM	; capable of handling nonlinear issues; able to prevent "dimensional disasters"； Contains a certain amount of durability	Implementing big training samples is challenging; sensitive to the absence of data; sensitive to parameter and kernel function selection
artificial neural networks	excellent classification precision; Noise sensitivity with high durability and fault tolerance; Associational memory	; need to adjust a large number of parameters, such as network topology, weights, and thresholds; difficulty in interpreting the output results; long learning time, which may not even achieve the learning objective

2.2 Feature construction solution

Feature engineering is extremely important for the construction of machine learning models, and the quality of features directly determines the upper limit of the performance of machine learning models [7]. This paper designs a feature derivation scheme based on RFM (recency, frequency, monetary) and constructs over 100,000-dimensional transaction behavior features by feature explosion; the aim is to precisely characterize each transaction in order to discover the latent fraud pattern and accurately perceive the risk. On this basis, a set of feature screening schemes is explored to select the best ones from among the many features, control the final scale of the input features, and ensure the feasibility of real-time prediction without reducing model performance.

1) RFM feature derivation scheme: RFM is a prevalent solution in customer relationship management settings that may accomplish customer value segmentation by depicting consumer consumption behavior. In specific application practice, the basic RFM framework is extended according to the business scenario of real-time transaction interception and the characteristics of online transaction flow, and the feature-derived subjects (statistical objects) with customers, accounts, and devices as the core are defined. Additionally, R is extended to time windows, F is extended to

aggregation functions, and M is extended to data variables, and the specific derivation methods are illustrated in Table 2.

2) By extending the RFM solution from the statistical variables of transaction flow, using flexible statistical windows and a variety of aggregation functions, the independent transaction flows are derived from the time dimension, value dimension, and spatial dimension to construct spatio-temporal information that satisfies the multi-faceted, fine-grained, and deep-field portrait of trading behavior and deep depth-of-field photos, giving the model adequate samples of its features.

Table 2: RFM feature derivation method

RFM Expansion Solutions	**Description**	**Elements**
Statistical entities	Subjects for profiling multidimensional trading behavior	Customers, accounts, trading devices
Statistics Window	On the basis of the range of previous transactions (including time windows and frequency windows) of future transactions, the long-term and short-term behavioral profiles of the present trading subjects may be characterized independently.	Time frame: short-term window in hours, including One hour, 24 hours, etc., up to 48 hours; long-term window In months, including one month,three months, up to six months;frequency window: three times, five times, 10 times.
Aggregation functions	Statistical analysis functions for mining of statistical variables' deep features	continuous functions: Maximum value, lowest value, and, mean value, etc.; discrete functions: count, frequency, ratio, etc.
Statistical variables	According to field properties, the processing object of the aggregation function—that is, various types of original transaction fields and derived feature fields—can be partitioned into continuous and discrete variables.	Continuous variables are variables that can take on any value within a given interval, such as transaction amount; discrete variables, on the other hand, have limited fixed values and can be enumerated, such as transaction type.

3) Feature screening program. Feature screening is the process of selecting beneficial features from the derived sample space and discarding irrelevant or redundant features. The core purpose is to improve the prediction efficiency of the model and the generalization ability of the model with unknown data. To ensure that the features after dimensionality reduction have good classification ability and meet the strict requirements of timeliness in risk control scenarios, a "four-step feature screening scheme" set was explored through application practice: the first step is single feature coarse-grained screening based on AUC; the second step is based on XGB feature importance screening; the third step is iterative feature screening based on recursive feature elimination (RFE); and the fourth step is based on feature screening for deep business understanding. With the assistance of algorithms and business professionals, the curse of dimensionality caused by the explosive derivation of features is resolved, redundant information and noisy features can be effectively eliminated, and incoming features with fast real-time calculation speed, good business interpretability, strong model generalizability, and stable prediction results are screened.

3. Feature Construction and Feature Selection

3.1 Data preparation

In order to deeply explore the difference between fraudulent transactions and normal transactions, billions of online transaction flows including customer information, account information, transaction time, transaction amount, transaction place, transaction time spent, were extracted during the data preparation process for 14 months, including transactions of moving accounts, login, setup, and loans.. Through data quality evaluation and business understanding of data fields, based on missing data filling and field coding, fields with a high rate of missing data, unique values, and poor business value are deleted, and a wide table of basic data is created.

Lastly, based on the accumulation of fraud samples, fraudulent transactions are marked line by line to create a wide table of modeling data with positive or negative sample markers.

3.2 Data down-sampling and feature derivation

Because of the massive number of online banking transactions, fraudulent transactions appear to be exceedingly rare when compared to the daily flow data of tens of millions of transactions, and the ratio of positive to negative samples is highly uneven. To accurately perceive the difference between fraudulent and normal transactions and to prevent the features of fraudulent transactions from being overshadowed by those of normal transactions, as well as to improve the modeling efficiency and reduce the computational and storage resource overheads of the feature derivation link, the negative samples are down-sampled proportionally to the number of positive samples, and the ratio of positive to negative samples is 1:100.

After sampling, a balanced analysis of gender, age, quantity, time, and activity is conducted to ensure that the samples accurately represent the distribution features of the entire data and to prevent the problem of model bias due to sampling bias.

Finally, based on the RFM derivation scheme, the feature derivation calculation is performed on the sampled samples, and 198,000 dimensional features are obtained for each transaction.

3.3 Feature selection

The selection of features follows a "four-step feature selection approach."

The first step is single-feature coarse-grained screening based on AUC. The single-feature AUC was calculated based on the decision tree algorithm, and the features with no and poor distinguishing ability were eliminated. Finally, 30,000-dimensional features with strong distinguishing ability of positive/negative samples were selected.

The second is screening based on XGB feature significance. The 30,000-dimensional features are sent to the XGBoost model, and the algorithm performs rough screening according to the feature importance (weight), and selects 1,303-dimensional features.

The third is iterative feature screening based on RFE. Based on the XGBoost model, recursive feature deletion is performed, and 10% of the features with the lowest contribution are deleted in each iteration. After 29 iterations, the 61-dimensional features are finely screened.

The fourth is feature screening, based on deep business understanding. Combining data stability, engineering implementation difficulty, and business interpretability of the features, risk control experts conduct business interpretability analysis and classification result evaluation on the features. Finally, 55-dimensional entry features with good business interpretability, good differentiation ability and strong model generalization ability are selected. Figure 1 depicts the process of selecting features.

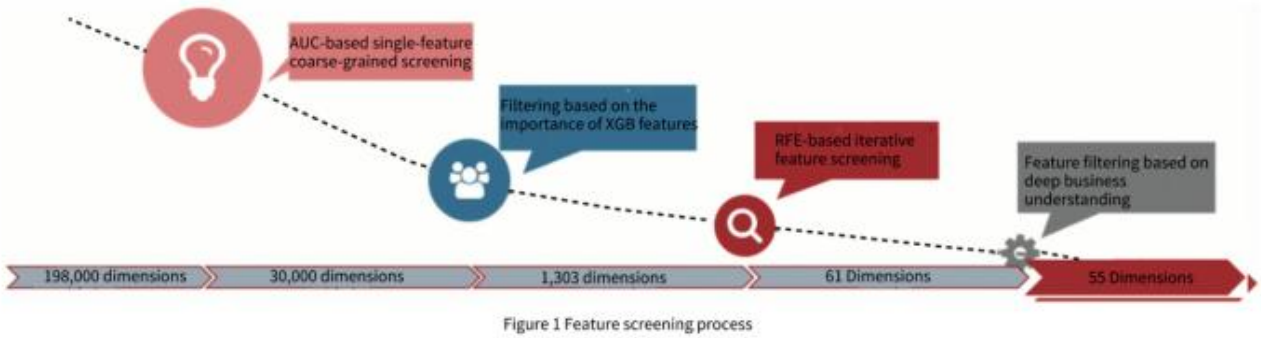

Fig. 1. the process of selecting features.

In the feature selection procedure, the primary assessment metric is the AUC of the model on the validation set, and Figure 2 depicts the AUC of the validation set from 30,000-dimensional features to 55-dimensional features. In the first round of feature selection, the AUC of 30,000-dimensional features is 0.991, but the XGBoost model uses only 1,303 of these features; after 29 rounds of RFE iteration, in addition to eliminating redundant features, it also reduces the risk of overfitting. The AUC of the 61-dimensional feature on the validation set increased slightly to 0.994 7; in the 31st round, the business-understood AUC of the 55-dimensional feature was 0.994 8, which was essentially the same level as that of the 61-dimensional feature.

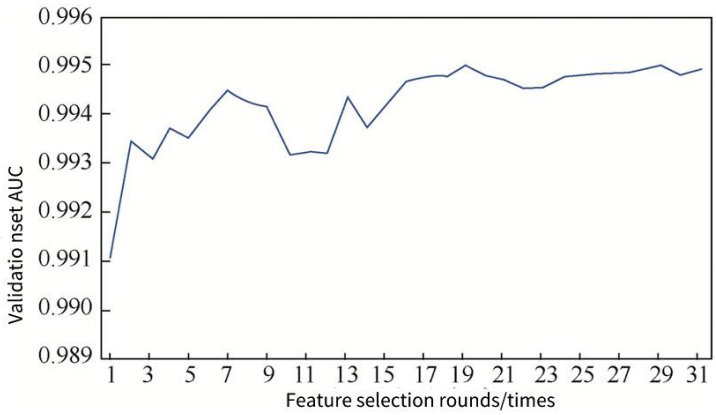

Fig. 2. Feature selection verification set AUC curve

From 198,000 to 55 dimensions, the feature compression ratio is as high as 3,600:1. With the help of the "algorithm+business" of the "four-step feature filtering scheme", the AUC of the model on the validation set does not decrease, and it demonstrates strong generalization ability on the evaluation set: from the AUC indicator, the new evaluation set the data AUC is only 0.025 lower than that of the validation set, and there is no significant decrease; from the business point of view, the model recall rate and precision rate are better than the expert rules, and reach the best level in the industry.

3.4 Evaluation of entry characteristics

After completing the feature screening, further evaluation and analysis were carried out from the perspectives of feature correlation and feature importance. The 55×55 feature correlation matrix is shown in Figure 3. Each small square in this figure represents the correlation degree of one pair of features. The white square represents a linear positive correlation, and the black square represents a linear negative correlation. The fewer the black and white squares, the smaller the correlation between features. From Figure 3, it can be seen that the correlation between the model features is low and the independence is high, which reflects the "diversity" of the featuresagainst the transaction dimension.

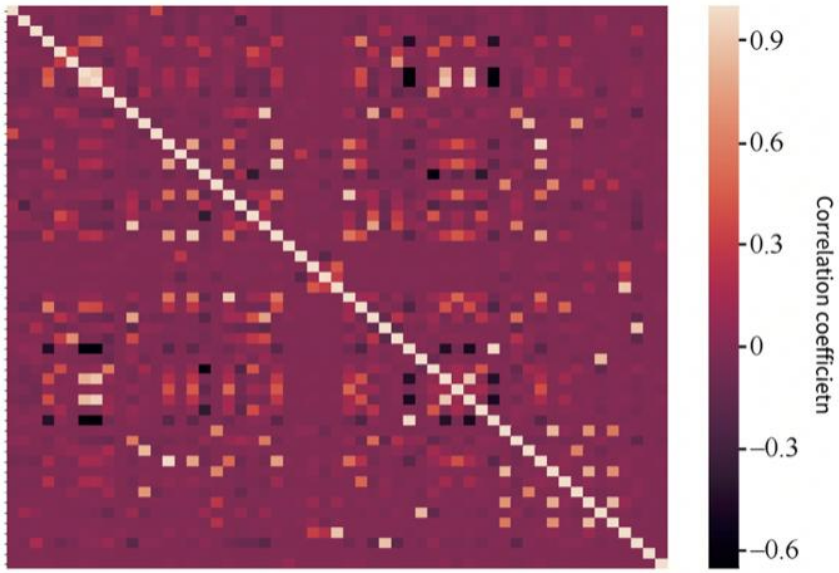

Fig. 3. 55-dimensional feature correlation matrix

From the perspective of the importance of model entry features, model entry features include time, quantity, location, frequency, distance, short-term, long-term, aggregated, weighted, etc., which indicates the "richness" of the depiction viewpoint. The feature importance ranking is shown in Figure 4.

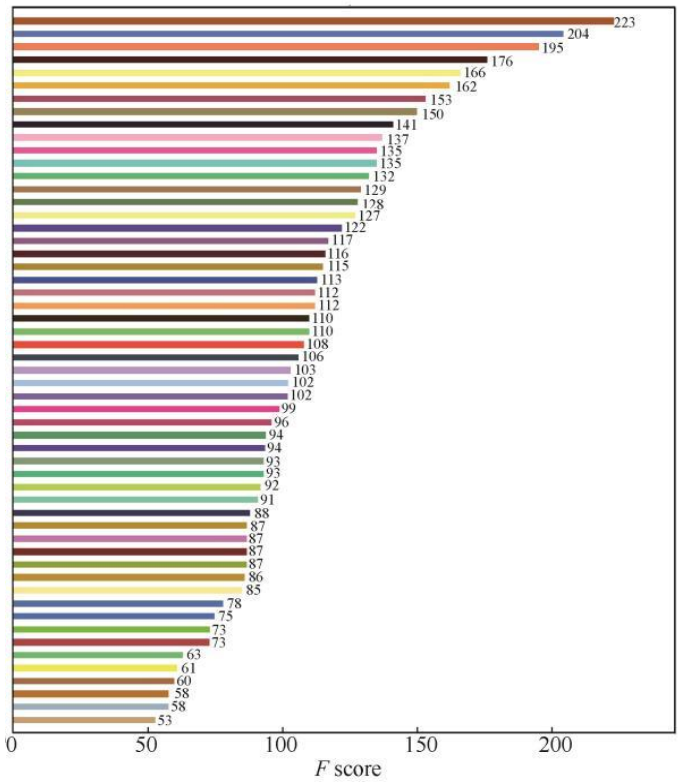

Fig. 4. Feature importance ranking

4. Experiments and Results

We performed comprehensive feature correlation and importance evaluation. The 55-dimensional model feature can accurately describe each transaction based on multiple dimensions and multiple perspectives, realizing the "thousands of people and thousands of faces" of transaction fraud risk identification.

4.1 Data set division

The 14-month data set was divided into a training set, validation set, and evaluation set in the ratio of 14:3:3. The training set is used to train models, the validation set is used to optimize models, and the evaluation set is used to validate models. In order to accurately evaluate the model's ability to predict the "future", the data set is split into natural months, with 10 months for model training, two months for model validation, and two months for model assessment.

The training and validation sets for the first 12 months were randomly downsampled based on the number of positive samples, and, after sampling, the data size was decreased to 818,000. The full data for the 13th and 14th months was reserved as the evaluation set, which avoids the impact of data traversal on the generalization ability of the model, and enables more accurate evaluation of the real performance of the model.

4.2 Model training and superparameter tuning

In accordance with the business goals of "double improvement" of fraud transaction recognition accuracy and recall rate, and engineering achievability requirements for real-time decision-making, 21 iterations of model training and optimization were carried out. According to the model effect of each iteration, a large number of parameters such as positive and negative sample weights, learning rate, maximum tree depth, and maximum number of trees were adjusted and optimized. Finally, the 55-dimensional model feature was used to train a decision tree consisting of 949 trees(each tree has a depth of 3) the XGBoost machine learning model. Technical evaluation indicators.

For the characteristics of extreme imbalance between positive and negative samples in transaction anti-fraud scenarios, AUC and KS, which are less sensitive to data imbalance, are used as technical evaluation metrics. Both evaluation metrics can be calculated by the true positive rate (TPR) and false positive rate (FPR) (Table 3).

Table 3. Confusion matrix of classification results

Actual Results	Predicted Results	
	Positive	Negative
True	true positive (TP)	false negative (FN)
False	false positive (FP)	true negative (TN)

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{6}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP+TN}} \tag{7}$$

The AUC value is the area under the ROC curve created by TPR and FPR at various thresholds, which reflects the model's ability to discriminate between positive and negative samples. The closer the AUC number to 1, the better the model's classification skill; the KS value represents the greatest disparity between TPR and FPR, which can be used to quantify the optimal discriminating effect of the model, and can also be used to guide the formulation of the optimal threshold segmentation point for the output probability value of the model.

4.3 Business evaluation metrics

Starting from the business purpose of enhancing risk control capabilities, the recall rate and false alarm ratio are employed as the primary indicators to evaluate the model's efficacy in specific application practice.

The recall rate, also known as the detection rate, indicates the percentage of fraudulent transactions identified by the model to all fraudulent transactions under a specified probability threshold. The higher the recall rate, the higher the accuracy of the model identification.

The false alarm ratio is the reciprocal of the precision ratio, which represents the number of judgments the model needs to perform to correctly identify a fraudulent transaction. The lower the value, the lower the disturbance to normal customers.

$$\text{Recall Rate} = \frac{\text{Number of samples predicted to be positive and actually positive}}{\text{Actual number of positive samples}} \quad (8)$$

$$\text{False alarm ratio} = \frac{\text{Number of samples with positive predictions}}{\text{Number of samples predicted to be positive and actually positive}} \quad (9)$$

The recall rate shows the ability to identify fraudulent transactions and the false alarm ratio represents the accuracy of fraudulent transaction identification, which is the most direct evaluation metric of risk control capabilities.

4.4 Model performance

In this experiment, after multiple iterations of optimization, all aspects of the model reached a good standard, and the AUC of the validation set model was 0.997. The model generalization ability was evaluated by applying the complete number of transactions in the 13th and 14th months; the model AUC reached 0.972 and the KS value reached the maximum value (0.83). As depicted in Figure 5, the AUC of the model on the evaluation set decreases by a smaller amount than the AUC of the training set, indicating that the model has strong generalization ability and still shows good performance with new data.

The KS curve is shown in Figure 6, and the KS value for this model is 0.83, at which point TPR=0.87, FPR=0.05, and the scoring threshold for determining positive and negative is 2.25.

The false alarm ratio curve is a measure of the operational level. As the false alarm ratio (cost) increases, the model recall rate (benefit) gradually increases, as shown in Figure 7. The steeper the curve, the better the ability of the model to identify fraudulent transactions. Concurrently, the curve can be referred to, and the threshold can be altered flexibly based on the system's operating capability, business objectives, and the fraud situation in order to construct a flexible fraud prevention system.

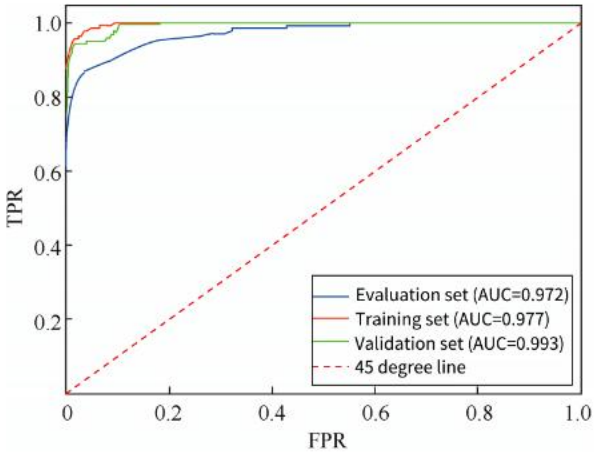

Fig. 5. ROC curve

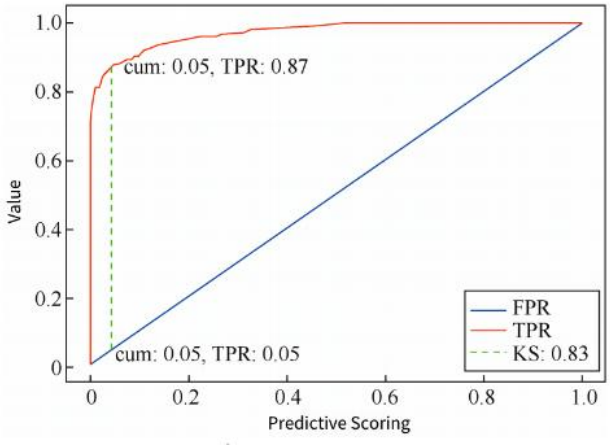

Fig. 6. KS Curve

Fig. 7. Recall-False Alarm Curve

5. Model Application Scheme

5.1 System deployment scheme combining streaming and batching

The systematic deployment of machine learning models mainly involves five links: real-time feature processing, historical feature processing, model calculation and scoring, risk control system decision-making, and business system disposal. In order to improve the operating efficiency of the model and meet the millisecond-level real-time decision-making response requirements, in the systematic deployment of the model, this article proposes "streaming + batch" and "rule + model" based on an intelligent risk control system and machine learning platform. The application scheme of the system takes advantage of enterprise-level platform resources. Figure 8 depicts the deployment strategy for the system.

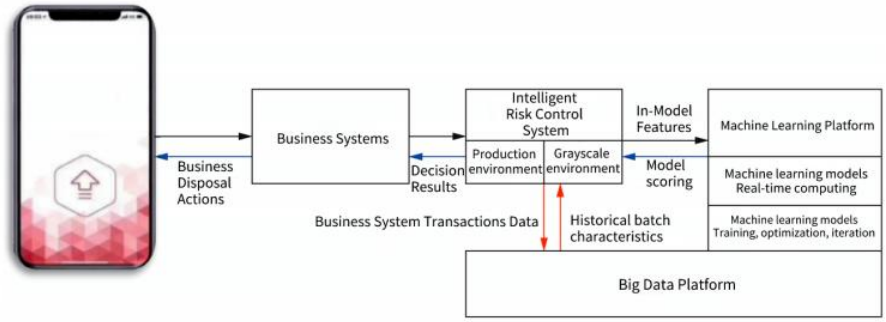

Fig. 8. System Deployment Scheme

Through the intelligent risk control system, functions such as real-time feature processing based on streaming data, historical feature merging, decision-making result generation based on model scoring, and decision-making result push business system can be realized.

Through the big-data platform, using the advantages of massive data storage and distributed computing resources, the historical behavior data is processed and calculated in batches every day to generate historical features.

Using the platform for machine learning, the model score of each transaction is computed in real time based on the computational power of the model run.

Through the business system, actions such as release, interception, challenge, and challenge enhancement are performed according to the decision results of the intelligent risk control system.

5.2 Business application scheme of model and rule combination

In order to effectively optimize model utility, through a comprehensive analysis of the rule model early warning data and machine learning model early warning data, we found that the combined application of the machine learning model and the existing expert rule model to fully exploit their respective strengths enables risk inspection and control to be effectively improved. As illustrated in Figure 9, based on the fraud score calculated by the machine learning model for each transaction, all transactions are divided into three categories: high-risk head region, medium-risk middle region, and low-risk tail region. Different types of transactions and rules are applied in various combinations. When the transaction belongs to the head region, regardless of the expert rule determination result, enhanced authentication is performed; when the transaction belongs to the tail region, regardless of the expert rule determination result, it is considered a normal transaction and released directly; when the transaction belongs to the middle region, the method of expert rule judgment as the main method and model scoring as a supplement is adopted, and selective enhanced certification is carried out according to the real-time availability of risk control operations during the event.

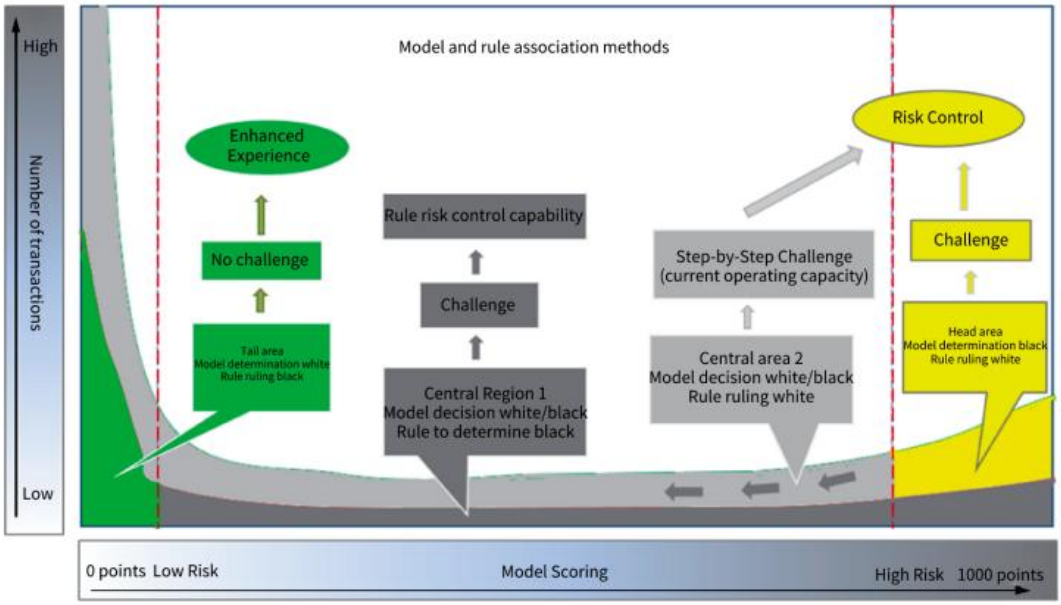

Fig.9. Distribution of rule/model alert transactions

Using the combination approach outlined in Table 4, it is possible to achieve two major benefits: first, for the high-risk head region, fraud patterns outside the expert rules can be effectively explored to improve the recall rate and achieve the goal of "risk control"; second, for the low-risk tail region, the disturbance to normal customers can be significantly reduced to improve the interception accuracy rate and achieve the effect of "enhanced experience."

Table 4. Combined rule/model application strategy

	Machine learning model probability scores		
Expert Rule Determination	Head area	Middle area	Tail area
High Risk	Enhanced Certification	Enhanced Certification	Release
Low Risk	Enhanced Certification	Selective Enhanced Certification	Release

5.3 Model operation safeguards

In the application scenario of a machine learning model embedded with high-frequency trading and real-time decision disposal, two sets of emergency plans of timeout and disaster recovery are designed to ensure efficient operation of the model, control model risk, and safeguard customer experience, and establish a model abnormality monitoring and automatic processing mechanism to enable safe, stable, and controllable model operation.

First, a system for timed-out disposal is devised. If the model score has not been calculated for more than 30 ms, the model is automatically deactivated and the rule judgment result will be used as the transaction challenge basis.

Second, an anomalous meltdown mechanism is implemented. Considering the dimensions of transaction types, customer groups, special dates, the daily challenge transaction threshold is set in groups. If the threshold is exceeded, the model score will no longer trigger transaction decisions; set a daily threshold for the number of high-risk trades, if the threshold is exceeded, the model score will no longer trigger a trading decision.

6. Model Application Effect and Optimization Iteration

6.1 Model to rule optimization and enhancement

With the expert rule model becoming increasingly complex and the room for improvement being limited, the introduction of machine learning models can provide acute insights into anomalous user behavior. From the standpoint of high-dimensional behavioral characteristics, consumers' transaction habits are relatively stable; therefore, when the mode of operation of the Internet black market changes, it is possible to identify fraudulent transactions that are different from those operated by the customers themselves. Based on application of the model to the production data and the monitoring effect, after combining the application of the rules and the machine learning model, the recall rate increases by 20% while the false alarm ratio decreases by 40%, thus achieving the goal of "double improvement" in the recall rate and accuracy rate. This effectively improves the system identification capability and reduces operating costs.

6.2 Optimization iteration of the model

6.2.1 Monitoring and iteration mechanism.

The effectiveness of the model cannot be improved overnight: with the online operation of the model system and changes in the fraud mode, a professional data analyst team is required to continuously iterate and optimize, and improve the effect, so that the model can play a greater role and have more value. In order to accomplish this, we have established a mechanism for daily monitoring and regular optimization and iteration of the model in specific application practices, thereby establishing a lifecycle management system of model monitoring-optimization-iteration-grayscale operation-release production.

1) Model monitoring.

The model monitoring report is integrated into the system, focusing on model stability monitoring indicators such as key segment missing rate and characteristic PSI indicator, model performance

indicators such as AUC value and KS value, and business evaluation indicators such as recall rate and false alarm ratio; through periodically rechecking the model operation effect, the system generates the model monitoring report.

2) Model optimization and iteration

- "Small" iteration. The model compares the operation monitoring effect with the benchmark when it first comes online, and when the model effect continues to weaken or runs abnormally (e.g., triggering fuse mechanism, or if the alarm volume is too large or too small), a "small" iteration is performed on the model, i.e., the feature processing scheme is not modified and only the model file is updated, so that the model can be rapidly improved. Without modifying the system, the model may be iterated and improved.
- "Large" repetition. To maintain the long-term stability of the model, every year or every two, a significant evolution of the complete risk control model is undertaken, which may involve changes in the processing scheme of the model entry features.

3) Optimization of iterative solutions.

On the basis of the operational effect of the model once it is placed online, the accumulation of fraud data, and the development of artificial intelligence technology, the model's optimization iteration scheme is classified into following four categories.

- Introducing fresh data samples: The effect and prediction ability of the model are highly dependent on the training samples, and new fraud samples are continuously introduced to iterate the model. This allows the model to learn the latest fraudulent transaction patterns and the latest normal transaction patterns, which is a crucial method for ensuring the model's capacity to generalize.
- Model tuning: The XGBoost machine learning model has dozens of hyperparameters. When the model effect is not ideal or cannot achieve the expected business goals, the key parameters such as learning rate and positive and negative sample weights can be adjusted in combination with the characteristics of fraudulent transactions to improve the model effect.
- Introduction of new features: With the development of business and the escalation of anti-fraud and Internet black market confrontation, the introduction of new features will improve the model's ability to depict transactions, which is an efficient method for enhancing the model's effectiveness after long-term operation.
- Introduction of new algorithmic procedures: The development of artificial intelligence technology is changing on a continuous basis. Monitoring and studying artificial intelligence technology, as well as actively investigating and implementing the use of other artificial intelligence technologies in the anti-fraud sector, is a critical path for the model's continuous improvement.

7. Conclusion

Driven by criminal interests, the confrontation between banks and fraudsters is never-ending. External transaction fraud will remain the greatest danger to banks performing online financial transactions. This paper conducts a comprehensive and in-depth study that explores subjects ranging from feature engineering schemes and model training optimization to systematic deployment, model application strategy, and iterative optimization. The results represent useful exploration and practice of using big data and artificial intelligence to improve fraud risk control capability, can resolve the risk control problem of fast-changing and difficult identification of transaction fraud patterns. The findings have value and should serve as a reference for the construction of risk prevention and control systems in commercial bank transactions in the digital financial era.

References

[1] Shaohui Zhan, Keqian Tang, Kaixuan Chang, Liang Yuan, Shuo Liu, and Zhaoming Li. 2021. Credit Anti Fraud Identification Method Based on Power Big Data. In 2021 International Conference on

Aviation Safety and Information Technology (ICASIT 2021). Association for Computing Machinery, New York, NY, USA, 735–741. https://doi.org/10.1145/3510858.3511373

[2] Yingjun Mo. 2020. Research on Theory and Practice of Financial Frauds. In 2020 The 4th International Conference on E-Business and Internet (ICEBI 2020). Association for Computing Machinery, New York, NY, USA, 61–65. https://doi.org/10.1145/3436209.3436380

[3] Mahmut Öğrek, Eyüp Öğrek, and Şerif Bahtiyar. 2019. A deep learning method for fraud detection in financial systems: poster. In Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '19). Association for Computing Machinery, New York, NY, USA, 298–299. https://doi.org/10.1145/3317549.3326299

[4] Bambang Leo Handoko and Dessy Tandean. 2021. An Analysis of Fraud Hexagon in Detecting Financial Statement Fraud (Empirical Study of Listed Banking Companies on Indonesia Stock Exchange for Period 2017 – 2019). In 2021 7th International Conference on E-Business and Applications (ICEBA 2021). Association for Computing Machinery, New York, NY, USA, 93–100. https://doi.org/10.1145/3457640.3457657

[5] Zeinab Rouhollahi, Amin Beheshti, Salman Mousaeirad, and Srinivasa Reddy Goluguri. 2021. Towards Proactive Financial Crime and Fraud Detectionthrough Artificial Intelligence and RegTech Technologies. The 23rd International Conference on Information Integration and Web Intelligence. Association for Computing Machinery, New York, NY, USA, 538–546. https://doi.org/10.1145/3487664.3487740

[6] Chung Min Tae and Phan Duy Hung. 2019. Comparing ML Algorithms on Financial Fraud Detection. In Proceedings of the 2019 2nd International Conference on Data Science and Information Technology (DSIT 2019). Association for Computing Machinery, New York, NY, USA, 25–29. https://doi.org/10.1145/3352411.3352416

[7] Shelly Xiaonan Wu and Wolfgang Banzhaf. 2008. Combatting financial fraud: a coevolutionary anomaly detection approach. In Proceedings of the 10th annual conference on Genetic and evolutionary computation (GECCO '08). Association for Computing Machinery, New York, NY, USA, 1673–1680. https://doi.org/10.1145/1389095.1389408

[8] Meiryani Meiryani, Marcellinus Anggito Darmawan, Lusianah Lusianah, Ridho Bramulya Ikhsan, and Nugroho Juli Setiadi. 2021. The Effect of Fraud Detection and Prevention on Financial Performance Study on Trading Company. In The 2021 7th International Conference on Industrial and Business Engineering (ICIBE 2021). Association for Computing Machinery, New York, NY, USA, 219–227. https://doi.org/10.1145/3494583.3494634

[9] Yi Liu, Jiawen Peng, and Zhihao Yu. 2018. Big Data Platform Architecture under The Background of Financial Technology: In The Insurance Industry As An Example. In Proceedings of the 2018 International Conference on Big Data Engineering and Technology (BDET 2018). Association for Computing Machinery, New York, NY, USA, 31–35. https://doi.org/10.1145/3297730.3297743

[10] Bambang Leo Handoko and Ameliya Rosita. 2022. The Effect of Skepticism, Big Data Analytics to Financial Fraud Detection Moderated by Forensic Accounting. In Proceedings of the 6th International Conference on E-Commerce, E-Business and E-Government (ICEEG '22). Association for Computing Machinery, New York, NY, USA, 59–66. https://doi.org/10.1145/3537693.3537703

[11] Bambang Leo Handoko, Hery Harjono Muljo, and Adelia Yulma Budiarto. 2021. The Connection and Underlying Cause between Financial Statement Fraud Cases. In 2021 5th International Conference on E-Business and Internet (ICEBI 2021). Association for Computing Machinery, New York, NY, USA, 162–167. https://doi.org/10.1145/3497701.3497727

[12] Yurou Wang, Ruixue Li, and Yanfang Niu. 2021. A Deep Neural Network Based Financial Statement Fraud Detection Model: Evidence from China. In 2021 4th Artificial Intelligence and Cloud Computing Conference (AICCC '21). Association for Computing Machinery, New York, NY, USA, 145–149. https://doi.org/10.1145/3508259.3508280

[13] Shunyu Yao. 2021. Application of Data Mining Technology in Financial Fraud Identification. 2021 4th International Conference on Information Systems and Computer Aided Education. Association for Computing Machinery, New York, NY, USA, 2919–2922. https://doi.org/10.1145/3482632.348754