

An Analysis of the Characteristics and Detection Techniques for Credit Card Fraudulent Transactions

Haojun Shi^a

Lee Shau Kee School of Business and Administration, Hong Kong Metropolitan University, Hong Kong, China

Keywords: Credit Card Fraud, Machine Learning, Deep Learning, Transaction Security, Data Imbalance.

Abstract: With the rapid growth of the modern economy, credit cards have become more and more essential for daily transactions. At the same time, the rise in fraudulent activities poses significant financial risks, making it of critical importance for cardholders and merchants to implement effective detection measures. This paper reviews various machine learning and deep learning techniques aimed at accurately identifying fraudulent credit card transactions. It covers key aspects such as technology analysis, detailed feature extraction, algorithm selection, and the overall effectiveness of these methods. The discussion includes the strengths and weaknesses of different algorithms, particularly in handling imbalanced data and complex fraud patterns. Ultimately, this research offers insights for institutions seeking to enhance security, mitigate financial risks, and advance detection technology with innovative solutions. In the future, the research should focus on improving scalability and real-time detection capabilities to better address evolving fraud strategies, ultimately contributing to more secure financial ecosystems.

1 INTRODUCTION


As a non-cash credit tool provided by financial institutions, a credit card plays an indispensable role in the modern economic life. It allows cardholders to make purchases without having to pay cash immediately and to complete repayments within subsequent billing cycles, promoting convenience and flexibility. Unlike debit cards, which are charged directly from the user's account, credit card transactions are limited by a preset credit limit, determined by the card issuer based on an assessment of the cardholder's credit status (Bertaut & Haliassos, 2006).

Credit card transactions are a complex and sophisticated ecosystem typically involving five core entities: cardholders, merchants, card issuers, credit card organizations, and acquiring agencies. The card issuer is responsible for issuing credit cards to cardholders and providing related services. The merchant processes the payment made by the customer using the credit card through the acquiring agency. Credit card organizations act as clearing centers for transactions, ensuring accurate fund transfers between parties. After the cardholder

completes the consumption, the merchant submits the transaction details to the bank for approval, and once the approval is passed, the bank will pay the merchant on behalf of the cardholder while reducing the credit limit of the cardholder.

Given the central role of credit card transactions in the modern economy, transaction detection is particularly important (Chu et al., 2023). Effective transaction detection mechanisms detect and prevent fraud promptly, protecting the property of cardholders, merchants, and financial institutions. Through advanced technologies such as big data analysis and artificial intelligence, transaction patterns can be deeply mined and abnormal transaction characteristics can be identified to effectively prevent risks such as credit card theft and fake transactions (Varmedja et al., 2019). In addition, transaction detection helps improve the quality of customer service and provides cardholders with a more convenient and secure payment experience by optimizing the transaction process and improving processing efficiency.

This paper starts with the characteristic analysis of credit card fraudulent transactions and discusses the detection methods based on machine learning and

^a <https://orcid.org/0009-0003-2765-4696>

deep learning in detail. Through a systematic review of the application of these technologies in credit card fraud detection, the purpose is to reveal their effectiveness and existing technical limitations and provide references for future research directions. It is hoped that this analysis will serve as a useful resource for researchers and practitioners, encouraging further exploration and refinement of fraud detection methods in future studies.

2 FRAUD CHARACTERISTICS

The development of credit card transactions began in the mid-20th century. In 1950, Diners Club issued the first modern credit card for restaurant and travel purchases. In 1958, American Express and banks issued credit cards, which gradually became popular in the retail industry. In the 1960s, advances in computer processing technology led to the widespread use of credit cards. Subsequently, the development of magnetic strips, chip technology, and online payments made credit card transactions more secure and convenient.

Fraudulent credit card transactions often exhibit multiple significant characteristics, including but not limited to abnormal transaction patterns, abnormal transaction time, abnormal transaction location, abnormal transaction type, and abnormal login information. The abnormal trading pattern is embodied in two aspects. The first is the unusually high amount of spending, which far exceeds the cardholder's previous transaction records and is highly unusual. This is followed by frequent small test transactions, which may be used as a prelude to test the validity of the card and then gradually increase the transaction amount for fraudulent purposes. A transaction time anomaly refers to the period in which a transaction occurs that does not match the cardholder's regular activity pattern and can be replaced by a discrete period (year, month, day, minute, and second) with a more easily analyzed period label (such as morning, noon, or evening) using machine learning technology. If the transaction occurs during a very long period, such as late at night or during holidays, and frequent trading activities occur in a short period, such as a large number of transactions in a few minutes or hours, it may be a fraudulent transaction (Lu & Wung, 2015). The transaction location anomaly is reflected in the fact that the cardholder of the transaction goes to a place that is infrequent or far away from his residence, especially the transaction that spans multiple

countries or cities in a short period, which is contrary to normal consumption habits.

Abnormal transaction type refers to the transaction behavior that is inconsistent with the cardholder's daily consumption habits, such as sudden large purchases of virtual products, high-risk products, luxury goods, or repeated purchases of the same goods in a short period, which are not in line with the normal consumption logic. Finally, abnormal login information, which includes multiple attempts to log in with the wrong password, using false personal information to make transactions, and initiating transactions from unknown or high-risk new devices and IP addresses. These behavior patterns are highly abnormal, and there is a high probability that they are fraudulent transactions.

3 APPLICATION OF MACHINE LEARNING

Through the use of data and algorithms, machine learning is a branch of computer science that enables computers to perform better and learn continuously from their data in order to enhance data processing accuracy. The creation and study of statistical algorithms that can learn from data and generalize to invisible data to carry out tasks without explicit instructions is known as machine learning, a branch of artificial intelligence research. Chen et al. (2018) used the machine learning technique of integrated trees and neural networks to synthesize the literature on credit card fraud and demonstrate the validity of the machine learning model for credit card fraud. Support vector machines (SVM), K-nearest neighbor (KNN), and artificial neural networks (ANNs) were the methods Asha and KR (2021) employed to detect fraud; the artificial neural network's accuracy rate was an astounding 100%. It is evident that machine learning techniques are quite beneficial when it comes to credit card fraud detection.

Because there are significantly fewer fraudulent transactions than there are legitimate transactions, the distribution of data in this standard binary classification problem of credit card transactions is imbalanced, making it more challenging for machine learning to identify outliers. The machine learning model's accuracy won't be impacted even if it misinterprets fraudulent transactions as legitimate ones. Therefore, how to ensure the stability of data and construct a stable model is the direction of research.

In a recent study, the training set was balanced and the ratio of positive to negative samples was adjusted to 1:1 using six oversampling techniques, including Synthetic Minority Over-sampling Technique (SMOTE) (Huang, 2023). Awoyemi et al. (2017) used downsampling and oversampling methods and found that K-proximity performed better than naive Bayes and logistic regression algorithms on unbalanced data problems. Duman (2013) conducted experiments with different proportions of data to downsample transaction data. Yang (2021) used a generative adversarial network (GAN) to alleviate the imbalance of data and combined GAN and convolutional neural network (CNN) algorithms to demonstrate the superiority of model checking. Liao (2022) used a random forest (RF) classifier to build a detection model with low sensitivity to unbalanced data sets. These techniques can significantly enhance the capacity to identify credit card transactions.

4 APPLICATIONS OF DEEP LEARNING

Deep learning is a branch of machine learning that processes complex data by simulating artificial neural networks. This process involves layers of sampling, layers of representation, and learning algorithms for data representation. Compared with other machine learning algorithms, it is more inclined to the goal of artificial intelligence, and common models include CNN and recurrent neural networks (RNNs). Inspired by real neurons, Warren McCulloch and Walter Pitts created the first artificial neurons in the 1940s and 1960s. Frank Rosenblatt created Perceptron, an early neural network model that could classify binary data. It lays a foundation for the future fraud detection class binary classification problem. In the 1970s-1990s, Ronald J. Williams popularized backpropagation, a method for training multi-layer neural networks that made it possible to train more complex networks; however, neural networks declined due to computational limitations and the rise of SVMs and other algorithms. New algorithms and improvements in processing power have resulted from this. As a result, deep learning was used in many domains in the 2010s, and methods like CNN, RNN, and GAN became indispensable instruments.

According to the research, the combination of RFE and oversampling, such as SMOTE, is better than the traditional method in dealing with the high imbalance of data, while the traditional method often treats the transaction as an isolated event, to lack of

mining the interactive information, which largely satisfies the imbalance of credit card fraud data (Liao, 2022). In the aspect of model introduction, a deep belief network (DBN) model is used, whose function is to extract and classify fraudulent transaction features. The model is trained using a variety of classifiers, such as gradient lift trees (GBDT), extreme gradient boosting (XGBoost), and mesh search methods, to optimize the model parameters. Huang (2023) may be able to help us further decompress the concept of common P-values by controlling marginal P-values and calibrating P-values to increase discrimination since marginal P-values have the advantage of error discovery rate control. It is concluded that the Adaboost model is the best; light, GBDT, and XGBoost are second. Xue (2023) believes that compared with the traditional model, the stochastic integrated deep learning model has more advantages in data processing and recognition ability, especially in identifying rare fraudulent transaction samples, and it is not suitable for big data samples such as credit card detection. She proposed the innovative and improved Bagging and Borderline-SMOTE algorithms. Can improve the accuracy of the model. Deep learning techniques like ANN, CNN, Gated Recurrent Units (GRU), and Long Short-Term Memory (LSTM) each have their own advantages in terms of algorithms. LSTM is frequently utilized for anomaly detection in the dynamic changes of complicated networks. Wang (2019) proposed an anomaly detection method for a dynamic graph model based on LSTM to effectively characterize the changes in dynamic graph structure. It is precisely because LSTM is capable of processing time series data to capture the timing of transactions in credit card transaction detection. By combining LSTM with the feature fusion method, transaction information of neighborhood nodes can be obtained through random walks and dynamic features can be added to the LSTM model to better capture the interrelationship between transactions and potential abnormal transaction patterns. GRU is similar to LSTM, but its structure is simpler and the operation speed is faster. The researchs can use GRU to model and learn the historical transactions of users to find abnormal transactions that deviate from the normal pattern. In addition, CNN is a neural network model that can extract spatial correlations of data. CNN is less effective in the field of credit card detection than LSTM and GRU because it is better suited for processing static image data and has comparatively poor time dependence and sequence history processing capabilities. ANN's results are also relatively simple, and it is suitable for dealing with

large-scale nonlinear problems. However, like CNN, Ann's processing of time dependence and sequence features is unstable, and it is easily affected by gradient disappearance or explosion, so its effect is far less than CNN's.

5 LIMITATION AND PROSPECT

Credit card fraud detection has become significantly more effective due to the rapid advancements in machine learning. There are still a lot of restrictions, though. The percentage of fraudulent transactions generally makes up very little of all transactions. A significant disparity between positive and negative samples might impact the model's training effect, leading to the majority of learning models exhibiting a tendency to forecast typical transactions. Secondly, fraud changes too quickly and is too diversified. Fraudsters constantly update their fraud methods, which increases the difficulty of detection. Traditional rule-based methods make it difficult to deal with new fraud behaviors, and the rules need to be updated and adjusted constantly. To increase the precision and resilience of the detection model, future research should incorporate multi-source data, which can include information from several sources like social media, bank transaction records, and geographic location data. Algorithms that are adaptive to detect fresh fraud patterns and improve the ability to recognize new fraud behaviors can also be developed. Additionally, real-time detection efficiency can be increased by optimizing the decision-making process and detection method using reinforcement learning.

6 CONCLUSIONS

Finding and stopping fraudulent credit card transactions is crucial to lowering financial risks. This study examines machine learning and deep learning-based detection techniques and evaluates the capabilities and drawbacks of several algorithms for handling fraudulent transactions. Although these techniques have significantly increased the effectiveness of detection, there is still much work to be done to address the issues of data imbalance and fraud diversity. In order to improve the model's capacity to identify novel fraud patterns, integrate data from multiple sources, and improve the system's real-time performance, future research should concentrate on these areas. This will enable financial

institutions to implement a more dependable risk prevention strategy. Generally speaking, credit card fraud cannot be prevented, and the transaction detection model must be updated on a regular basis. In order to identify irregularities and stop credit card fraud, cardholders must cultivate a sense of security and routinely organize the transaction flow.

REFERENCES

- Asha, R. B., KR, S. K., 2021. Credit card fraud detection using artificial neural network. *Global Transitions Proceedings*, 2(1), 35-41.
- Awoyemi, J. O., Adetunmbi, A. O., Oluwadare, S. A., 2017. Credit card fraud detection using machine learning techniques: A comparative analysis. In 2017 international conference on computing networking and informatics (ICCNI) (pp. 1-9). IEEE.
- Bertaut, C. C., Haliassos, M. (2006). *Credit cards: facts and theories*.
- Chu, Y. B., Lim, Z. M., Keane, B., Kong, P. H., Elkilany, A. R., Abusetta, O. H., 2023. Credit Card Fraud Detection on Original European Credit Card Holder Dataset Using Ensemble Machine Learning Technique. *Journal of Cyber Security*, 5, 33-46.
- Duman, E., Buyukkaya, A., Elikucuk, I., 2013. A novel and successful credit card fraud detection system implemented in a turkish bank. In 2013 IEEE 13th International Conference on Data Mining Workshops (pp. 162-171). IEEE.
- Huang, A., 2023. Research on credit card fraud detection based on machine learning and conformal p-values (Master's thesis, Huazhong University of Science and Technology).
- Lu, H. P., Wung, Y. S., 2021. Applying transaction cost theory and push-pull-mooring model to investigate mobile payment switching behaviors with well-established traditional financial infrastructure. *Journal of theoretical and applied electronic commerce research*, 16(2), 1-21.
- Liao, Q., 2022. Research on anomaly detection in Bitcoin transactions (Master's thesis, People's Public Security University of China).
- Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M., Anderla, A., 2019. Credit card fraud detection-machine learning methods. In 2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH) (pp. 1-5). IEEE.
- Wang, K., Chen, D., 2019. Research on anomaly detection algorithm for dynamic graph model based on LSTM. *Computer Engineering and Applications*, (05), 76-82.
- Xue, X., 2023. Design of random ensemble learning algorithm and credit card fraud detection (Master's thesis, Yunnan University of Finance and Economics).
- Yang, H., 2021. Research on credit card detection and recognition based on GAN-CNN imbalanced classification algorithm (Master's thesis, Shantou University).