The Classification of Fraudulent Bank Transactions

Subjects: Mathematics, Interdisciplinary Applications

Contributor: Alexey Ruchay , Elena Feldman , Dmitriy Cherbadzhi , Alexander Sokolov

The modern financial sector requires high security standards due to rapid technological progress. Banks utilize many security measures for transactions, including artificial intelligence (AI), neural networks, and machine learning. The use of these techniques allows advanced technological problems to be solved. The most common types of financial frauds are money laundering, identity fraud, and credit and debit card frauds.

bank transactions Fraudulent AI

1. Introduction

The modern financial sector requires high security standards due to rapid technological progress [1][2]. Banks utilize many security measures for transactions, including artificial intelligence (AI), neural networks, and machine learning [3][4][5][6]. The use of these techniques allows advanced technological problems to be solved. The most common types of financial frauds are money laundering, identity fraud, and credit and debit card frauds [1].

The key role in securing financial systems is anti-money-laundering (AML), while identifying illegal transactions [7][8][9]. The launch of Bitcoin (BTC) has created a paradox: the anonymity allows criminals to remain unknown; however, forensic or AML analysis can be performed with access to the BTC transaction database. The goal of AML analytics is to identify fraudulent transactions in massive, ever-growing datasets. Semi-automatic or manual transaction analysis produces a large number of errors while the success of machine learning methods shows great potential in AML analysis [10][11][12].

Current AI methods are not sufficient to process the online transaction data stream. To analyze different security methods, a set of transactions filtered by specially created rules needs to be used. All transactions passing through the filter are considered legitimate and do not get analyzed further. The filter parameters include the direct characteristics of the transaction or its parties with the exclusion of any indirect characteristics. The results of this process can be invalid; thus, the most important obstacles in AML are [1] the inability to make real-time online selections of parameters that are unique to specific financial activities, the lack of a centralized transaction analysis tool, a large volume of false positives for fraudulent transactions, and the high cost of manual analysis.

The current task is to develop a system to identify suspicious transactions. Such a system would help speed up and simplify data processing, decrease the risks of transactions, store information and data about criminals, provide credit score verification, and trace the users involved in money laundering. Due to the high cost of these

systems, only large financial institutions can afford them. In 24 h, large companies process up to 20 million transactions [13]. Unfortunately, manual analysis identifying suspicious activities often produces false results and requires extra time and resources in order for it to be performed correctly. Therefore, the development and implementation of an antifraud system is essential [14][15][16][17].

2. The Classification of Fraudulent Bank Transactions

It is important for banks to identify suspicious transactions before fraud happens. As an example the Nilson report states that cumulative damages from credit card fraud were USD 22.8 billion in 2016 and USD 28.7 in 2019 [18].

All international credit card payment systems use their own fraud detection methods. Often, all of the transactions are checked using antifraud rules, lists, and filters. In Russia these systems are also used to identify cyber criminals; however, the first priority is to find the illegal exchange of goods and services.

Bank transaction security has been improved due to the addition of a built-in EMV chip to credit and debit cards and the client authentication process, by means of two-step authentication. Despite this, the number of fraudulent bank transactions has not reduced: credit card fraud totals EUR 1.5 billion a year [14]. Experts agree that to reduce credit card fraud using authentication methods—the development and implementation of the fraud identification systems based on data analysis—is required.

Experts classify credit card fraud into different forms, for example, fraud with false identity or using behavioral biometrics. In the first case, a criminal applies for a credit card with false identification, while in the case of behavioral biometrics, a criminal obtains and uses the cardholder’s existing credit card credentials. Fraudulent bank transactions are divided into six categories [14]: lost or stolen card fraud, fake card fraud, online fraud, bankruptcy fraud, retailer fraud, and stolen in transit card fraud. In addition, bank transaction fraud can be classified by three categories: bankcard fraud, retail fraud, and Internet fraud.

Recently, it has been suggested to distinguish two categories of fraud: fraud with a card (face-to-face) and fraud without a card (e-commerce fraud). Bank transaction fraud can be categorized into five different types [14]:

- Lost and stolen cards (<1% of all fraudulent transactions). Most of the time elderly people are the victims in this category; fraudsters get the PIN by “shoulder-reading” and subsequently steal the card. In this case, the fraudster is the thief and the credit card does not go through an organized crime resale network.
- Cards that do not reach their destination (<1% of all fraudulent transactions). In this case the credit card may have been stolen during production or upon delivery by mail. To avoid this type of fraud, banks ask customers to obtain and activate the card at their office.
- Identity theft. A card obtained by using fake or stolen documents.

- Counterfeit cards (>10% of all fraudulent transactions). The card is copied using a genuine card or during database hacking, and then reproduced on a counterfeit plastic card by international organized crime groups. The criminal obtains and reproduces the magnetic stripe data of the cards. This type of fraud had been prevalent in the past but has been partially solved by EMV technology as magnetic stripe terminals are no longer used in the EU; however, they remain in Asia and America. This kind of non-contact payment is not appealing for criminals, since only small value payments are processed.

- Fraud without a card makes up over 90% of all fraudulent transactions. Most credit card fraud occurs during electronic transactions. The card number, expiration date, and CVC are extracted during database hacks, then these credentials are sold. The value of the credentials depends on the re-sale capabilities (the first digits of the card numbers represent the bank and, accordingly, the blocking policies they have). Many retailers (90%) use 3-D Secure. This technology protects the cardholder through two-step authentication; however, some large retailers, such as Ebay or Amazon, do not protect their users' transactions with 3-D Secure. Another problem hindering the fight against cardless fraud is that companies do not report attacks that result in a data breach since it can cause bad publicity, which might lead to financial losses.

Typical fraud detection systems have several control levels. Each level can be automated or managed manually. Part of the automated approach consists of machine learning algorithms. These algorithms are used to create predictive models based on annotated bank transactions. Over the past ten years regular credit card fraud research has made it possible to develop models based on supervised or unsupervised machine learning, and partial machine learning [17].

The authors analyze methods of money laundering detection based on Big Data [19]. Big Data is represented in complex systems that perform tasks to prevent legalization and laundering of money that has been obtained illegally (the SAS Anti-Money Laundering System, SAS AML).

Machine learning methods for binary classification to predict fraudulent transactions using multilayer perceptrons, random forest, logistic regression, and convoluted network graphs are used to detect anomalous BTC transactions [10]. Convoluted network diagrams have emerged as a potential tool for AML analysis, and they are particularly attractive as a new way to capture and analyze transactions. The results reflect the advantage of the random forest method of classification, though the F1 score of 0.796 is not sufficiently reliable for classifying abnormal BTC transactions.

Recent studies built deep neural networks for the detection of irregularities either by discriminatively mapping normal samples and abnormal samples to different regions of the feature space, or by fitting multiple different distributions [20][21][22][23][24].

The effect of high cardinality attributes on credit card fraud detection has been investigated, and the Value Clustering for Categorical Attributes algorithm was used to reduce the cardinality of their domains while preserving fraud-detection capabilities [25].

Using machine learning to predict fraudulent financial transactions without balancing data, as in real life there are usually many more honest than fraudulent ones, an accurate and reliable method is proposed. Additional fraudulent transaction samples were generated using a conditional generative adversarial network for tabular data (CTGAN) to make the classifier more robust and accurate [26].

By combining a nature-inspired hyperparameter setting with several supervised classifier models implemented in a modified version of the XGBoost algorithm, a unique hybrid technique for financial payment fraud detection is presented [27]. Using machine learning and domain data to identify fraudulent payments, a modified XGBoost model can be created and tested by tuning. Experiments showed that the model successfully classified with 99.64% accuracy.

An AE-PRF fraud detection method has been proposed which uses AE to reduce the dimensionality of the data and extract the features of the data [28]. In addition, the method uses RF with a probabilistic classification to classify the data as fraudulent together with the corresponding probability. AE-PRF outputs a final classification as fraudulent if the corresponding probability exceeds a predefined probability threshold.

An AED-LGB algorithm has been proposed to solve bank credit card fraud [29]. The AED-LGB algorithm first extracts the feature data using an autoencoder. The features are then fed into the LightGBM algorithm for classification and prediction.

References

1. Khrestina, M.P.; Dorofeev, D.I.; Kachurina, P.A.; Usubaliev, T.R.; Dobrotvorskiy, A.S. Development of Algorithms for Searching, Analyzing and Detecting Fraudulent Activities in the Financial Sphere. Eur. Res. Stud. J. 2017, 20, 484–498.
2. Alsuwailem, A.; Saudagar, A. Anti-money laundering systems: A systematic literature review. J. Money Laund. Control. 2020, 23, 833–848.
3. Stojanović, B.; Božić, J. Robust Financial Fraud Alerting System Based in the Cloud Environment. Sensors 2022, 22, 9461.
4. Srokosz, M.; Bobyk, A.; Ksiezopolski, B.; Wydra, M. Machine-Learning-Based Scoring System for Antifraud CISIRTs in Banking Environment. Electronics 2023, 12, 251.
5. Razaque, A.; Frej, M.B.H.; Bektemyssova, G.; Amsaad, F.; Almiani, M.; Alotaibi, A.; Jhanjhi, N.Z.; Amanzholova, S.; Alshammari, M. Credit Card-Not-Present Fraud Detection and Prevention Using Big Data Analytics Algorithms. Appl. Sci. 2023, 13, 57.
6. Bakumenko, A.; Elragal, A. Detecting Anomalies in Financial Data Using Machine Learning Algorithms. Systems 2022, 10, 130.

7. Jullum, M.; Løl, A.; Huseby, R.B.; Ånonsen, G.; Lorentzen, J. Detecting money laundering transactions with machine learning. J. Money Laund. Control. 2020, 23, 173–186.
8. Weber, M.; Chen, J.; Suzumura, T.; Pareja, A.; Ma, T.; Kanezashi, H.; Kaler, T.; Leiserso, C.E.; Schardl, T.B. Scalable graph learning for anti-money laundering: A first look. arXiv 2018, arXiv:1812.00076.
9. Singh, K.; Best, P. Anti-money laundering: Using data visualization to identify suspicious activity. Int. J. Account. Inf. Syst. 2019, 34, 100418.
10. Weber, M.; Domeniconi, G.; Chen, J.; Weidele, D.; Bellei, C.; Robinson, T.; Leiserson, C. Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks for Financial Forensics. arXiv 2019, arXiv:1908.02591.
11. Feldman, E.V.; Ruchay, A.N.; Matveeva, V.K.; Samsonova, V.D. Bitcoin abnormal transaction detection model based on machine learning. Chelyabinsk Phys. Math. J. 2021, 6, 119–132.
12. Feldman, E.V.; Ruchay, A.N.; Matveeva, V.K.; Samsonova, V.D. Bitcoin Abnormal Transaction Detection Based on Machine Learning. Recent Trends in Analysis of Images, Social Networks and Texts (AIST 2020). Commun. Comput. Inf. Sci. 2021, 1357, 205–215.
13. Deng, W.; Huang, T.; Wang, H. A Review of the Key Technology in a Blockchain Building Decentralized Trust Platform. Mathematics 2023, 11, 101.
14. Lucas, Y. Credit Card Fraud Detection Using Machine Learning with Integration of Contextual Knowledge; Artificial Intelligence; Universite de Lyon: Lyon, France; Universitat Passau: Passau, Germany, 2019.
15. Maniraj, S.P.; Aditya, S.; Shadab, A.; Swarna, S. Credit Card Fraud Detection using Machine Learning and Data Science. Int. J. Eng. Res. Technol. 2019, 8.
16. Lebichot, B.; Le Borgne, Y.A.; He-Guelton, L.; Oble, F.; Bontempi, G. Deep-Learning Domain Adaptation Techniques for Credit Cards Fraud Detection. In Recent Advances in Big Data and Deep Learning: Proceedings of the International Neural Networks Society (INNSBDDL 2019); Springer: Berlin/Heidelberg, Germany, 2020.
17. Carcillo, F.; Borgne, Y.L.; Caelen, O.; Kessaci, Y.; Oble, F.; Bontempi, G. Combining unsupervised and supervised learning in credit card fraud detection. Inf. Sci. 2019, 557, 317–331.
18. HSN Consultants, Inc. Card Fraud Losses Reach 22.80 Billion; Technical Report 1118; The Nilson Report: Oxnard, CA, USA, 2017.
19. Plaksiy, K.; Nikiforov, A.; Miloslavskaya, N. Applying Big Data Technologies to Detect Cases of Money Laundering and Counter Financing of Terrorism. In Proceedings of the 6th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), Barcelona, Spain, 6–8 August 2018; pp. 70–77.

20. Zong, W.; Zhou, F.; Pavlovski, M.; Qian, W. Peripheral Instance Augmentation for End-to-End Anomaly Detection Using Weighted Adversarial Learning. In Database Systems for Advanced Applications. DASFAA 2022; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2022; Volume 13246.
21. Pang, G.; Shen, C.; Hengel, A. Deep Anomaly Detection with Deviation Networks. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19), Anchorage, AK, USA, 4–8 August 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 353–362.
22. Huang, Z.; Zhang, B.; Hu, G.; Li, L.; Xu, Y.; Jin, Y. Enhancing unsupervised anomaly detection with score-guided network. arXiv 2021, arXiv:2109.04684.
23. Kumar, N.; Shaju, S.J.; Kayathwal, K.; Agarwal, K.; Singh, A.; Chaurasia, D.; Asthana, S.; Arora, A. Intent2vec: Representation learning of cardholder and merchant intent from temporal interaction sequences for fraud detection. In Proceedings of the IJCAI-21 Workshop on Applied Semantics Extraction and Analytics (ASEA), Virtual, 21–23 August 2021.
24. Zhou, Y.; Song, X.; Zhang, Y.; Liu, F.; Zhu, C.; Liu, L. Feature Encoding With Autoencoders for Weakly Supervised Anomaly Detection. IEEE Trans. Neural Netw. Learn. Syst. 2022, 33, 2454–2465.
25. Carneiro, E.M.; Forster, C.H.Q.; Mialaret, L.F.S.; Dias, L.A.V.; da Cunha, A.M. High-Cardinality Categorical Attributes and Credit Card Fraud Detection. Mathematics 2022, 10, 3808.
26. Alwadain, A.; Ali, R.F.; Muneer, A. Estimating Financial Fraud through Transaction-Level Features and Machine Learning. Mathematics 2023, 11, 1184.
27. Dalal, S.; Seth, B.; Radulescu, M.; Secara, C.; Tolea, C. Predicting Fraud in Financial Payment Services through Optimized Hyper-Parameter-Tuned XGBoost Model. Mathematics 2022, 10, 4679.
28. Lin, T.-H.; Jiang, J.-R. Credit Card Fraud Detection with Autoencoder and Probabilistic Random Forest. Mathematics 2021, 9, 2683.
29. Du, H.; Lv, L.; Guo, A.; Wang, H. AutoEncoder and LightGBM for Credit Card Fraud Detection Problems. Symmetry 2023, 15, 870.