Review Paper on Reducing Fraud Detection System using Blockchain and ML

Pooja Patel (0225CS20MT12)*

*[1](Department of CSE, Global Nature Care Sangathan's Group of Institutions, Jabalpur)

Abstract: In this paper, we address the problems of fraud and anomalies in e-banking and online transactions. However, as the financial sector evolves, so do the methods for fraud and anomalies. Moreover, blockchain technology is being introduced as the most secure method integrated into finance. Therefore, we propose a secure fraud detection model based on machine learning and blockchain. There are two machine learning algorithms—XGboost and random forest (RF)—used for transaction classification. The machine learning techniques train the dataset based on the fraudulent and integrated transaction patterns and predict the new incoming transactions. The blockchain technology is integrated with machine learning algorithms to detect fraudulent transactions in the Bitcoin network. In the proposed model, XGboost and random forest (RF) algorithms are used to classify transactions and predict transaction patterns. A security analysis of the proposed smart contract is also performed to show the robustness of our system.

Keywords: anomaly detection, blockchain, fraud detection, machine learning, random forest.

Introduction Every industry, including banking, education, health care, and others, has modernized as a result of technological growth. Moreover, with the advent of communication technology, online transactions and means of payment are also being modernized. Through this modernization, traditional currencies are being converted into digital currencies, and all financial transactions are being conducted digitally. However, these transactions are not fully secured and are vulnerable to various digital attacks, such as fraud issues, anomalies, and privacy breaches. Additionally, as the volume of transactions rises, there is an increase in fraud associated with financial transactions. As a result, billions of dollars are lost globally every year [1]. Any suspicious activity on a network that behaves abnormally is called an anomaly. In cybersecurity and digital financial exchange, anomaly detection is used to detect fraud and network invasion. The goal of anomaly detection is to protect the network from illegal and fraudulent activities. In the financial sector, anomaly detection applications have investigated suspicious activity and identified hackers and fraudulent users. However, all anomaly detection methods in traditional financial systems are designed for centralized systems. Therefore, with the development of digital currencies, such as Bitcoin, anomaly detection methods using the blockchain are improving. Despite these advances, there are still many fraud occurrences [2]. Many artificial intelligence (AI) and machine learning techniques have been proposed to detect anomalies and fraud in digital transactions; however, there is no suitable solution for centralized systems. Blockchain is the most advanced and quickly evolving technology in many fields. It first became visible with the appearance of Bitcoin in 2008, which was introduced by Satoshi Nakamoto [3]. It addresses the security issues of centralized systems and provides solutions to external threats. It is a distributed, decentralized, and immutable ledger that time stamps all records and ensures record integrity. However, some participants in the blockchain network behave maliciously [4].

Overview of Blockchain

A blockchain is a growing distributed ledger that keeps a permanent record of all transactions that have taken place in a secure, chronological, and immutable way. It wasconceptualized and first used in 2008 by an unknown person or group named Satoshi Nakamoto to create the Bitcoin cryptocurrency. The primary aim is to use a cryptosystem to encrypt the sequence of bits in electronic files so as not to be anteceded or tampered with [17,18]. When evaluating a blockchain, the notable characteristics to consider include audibility, privacy, confidentiality, consistency, decentralization, and integrity [19,20]. Blockchain technologies can be categorized into three types: Public Blockchains (anyone canjoin the network), Private Blockchains (the members are chosen based on conditions), and Consortium Blockchains (semiprivate blockchains limited to a group) [21]. All three types can additionally be classified as Permissionless (public Blockchain), permissioned (private Blockchain), or both (Consortium blockchain). A Blockchain network comprises several components and attributes, such as a distributed and immutable ledger, Peer-to-Peer (P2P) networks, a consensus mechanism, and smart contracts.

Cryptography Hash Function

A hash function is a cryptographic algorithm that is widely used in blockchain technol- ogy. A hash function returns any kind of input as a string of bytes with a fixed length andstructure. The output formed is named a hash value. A hash value formed from data usingan explicit hashing algorithm is always the same length and one-way, that is, it cannot be reversed. The SHA-256 is the most illustrious of all cryptographic hash functions, and is used widely in blockchain technology.

Immutable Ledger

Blockchain is recognized for its ability to be immutable. When people talk about Blockchain's "immutability", they are referring to the impossibility of adjusting the data after it is recorded and stored. This is an essential attribute when dealing with blockchains. Figure 1 shows how the blocks are linked and how each block contains the previous block's hash value.

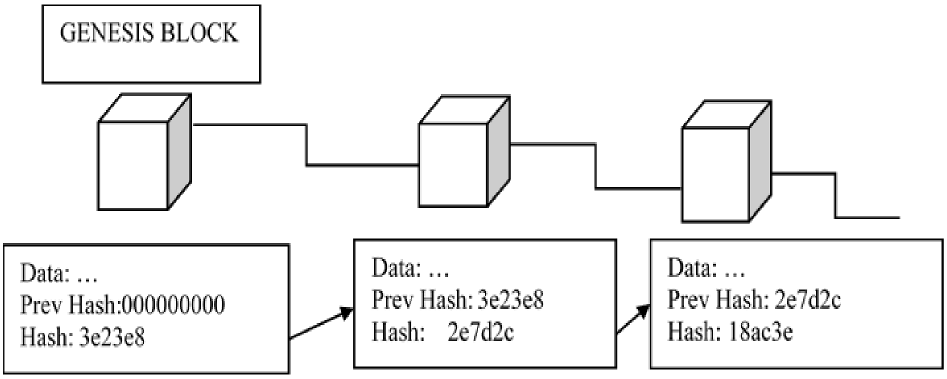

Figure 1. block links

The first block does not have previous blocks, and is named the genesis block. As can be seen, immutability emphasizes the fact that a blockchain is very secure and well designed. When the hash varies and no longer equals the previous hash in the ledger, theblockchain rejects that hash, making it invalid, similar to a bad check. A hacker would need to completely modify the next block, the block after that, and indeed the whole blockchain.

Distributed Peer-to-Peer Networks

By using a blockchain, interaction between two parties through a peer-to-peer model can be easily accomplished without any third-party requirement. Blockchain uses P2P pro- tocols, which permits all network members to hold an identical copy of contacts, allowing agreement over a consensus mechanism.

Distributed Application

To preserve an effective digital transaction platform, the blockchains used by most cryptocurrencies utilizes Distributed Applications (ÐApps). Dapps are software applications which are usually maintained and implemented on cloud services and can work on various systems at the same time. Many DApps have been built and deployed on a modelbased on Blockchain, although DApps can run on a cloud environment or other networksystems as well [22].

Consensus Protocol

A consensus protocol is an agreement between nodes in a blockchain network that submits transactional information, and is one of the most critical blockchain technology components. A blockchain network is restructured through the arrayed consensus protocol in order to certify that contacts and blocks are organized correctly, to guarantee the dis- tributed ledger's integrity and consistency, and ultimately to enhance trust between nodes. There are several consensus mechanisms used in various blockchain networks. Proof of Work (PoW) was the first consensus mechanism used in Blockchain.

Smart Contracts

Smart contracts refer to computer programs that obey a succession of previously established instructions stored on a blockchain [26]. A smartcontract allows anyone to protect an arrangement, automate payment, and eliminate the risk of scams while at the same time reducing intermediary fees. Unfortunately, accurate implementation of a smart contract's code cannot ensure its complete safety.

Public vs. Private Blockchains

Blockchains can be *public* (or *permissionless*), *private* or *consortium* (or *permissioned*). Bitcoin [20] or other cryptocurrencies (e.g., Ethereum [21]) are public. Cryptocurrencies are typically open to anyone to join the network and contribute to maintaining the integrity of transactions. However, in many other blockchain-based applications (e.g., related to company's private database), service providers may want to limit access rights to some specific groups of people. Answering the question: "who is able to join in the network, participate in the consensus algorithm and maintain the distributed ledger" allows developers to determine the suitability of a public or private blockchain. In a private or a consortium blockchain platform, as opposed to a public platform, will allow organizations to retain control and privacy while still cutting down their costs and transaction speeds. Typical examples include Hyperledger Fabric [22] and Multichain [23]. While clients are allowed to submit transactions, only pre-determined participants havepermission to execute the consensus protocol, and update the distributed ledger as well.These participants must be governed by informal arrangements, formal contracts or confidentiality agreements. The private or consortium systems will have lower costs and faster speeds than a public blockchain platform can offer.

Literature Review

Different public and private regions deploy blockchain technologies for various objectives because it is vital to protect and monitor auditing systems. These technologies help to evaluate its repositories and take care of the privacy of auditors. They allow auditors to send their queries in a reliable and accessible manner without exposing their identities to unauthorized users. In [5], consensus algorithms check the legitimacy of the performed transactions. However, it is inefficient to identify the transactions. Therefore, using blockchain as a solution for fraud detection does not completely address the problem. Because of this, new solutions are used to eliminate the vulnerabilities in the existing systems, such as machine learning algorithms. Different supervised machine learning techniques are used to detect fraudulent transactions. Furthermore, a comparative analysis of various machine learning methods is presented [6,7].

In [8], the authors proposed different supervised machine learning solutions to detect fake businesses. Moreover, they also tested over 300,000 accounts using random forest and XGBoost classifiers.

The authors in [9] also used XGboost for accurate results. In [10], the authors dealt with the problem of an imbalanced dataset. The dataset belongs to an insurance company and describes the driving patterns of individuals. They use XGboost to predict the performance of drivers along with their telematic information.

According to [11], fraudulent activities are data mining issues because the central server for credit card transactions tells whether a trading transaction is fake or legal. Fraud detection is not a new problem; yet, there are still numerous challenges. The primary reason is that researchers lack real-time data, and banks are unwilling to share their data with researchers because customer data is confidential. At the same time, it is linked to the banks' privacy policies [12].

In [13], the authors used a distributed data mining model to address the problems of slanted delivery of credit cards and non-uniform expenditures. A fraud detection algorithm aws presented in [14], which identifies fraud without relying on any fraudulent historical instances, with a proactive method capable of overcoming the well-known cold-start problem.

In [15], The authors suggested and demonstrated the application of the uncertain association law of mining to extract useful data from credit card transactions. The authors in [16] traindded a Support vector machine model to detect the improper data of credit card transactions.

In [17], the authors mixed three different techniques to decrease the wrong beeps in fraud identification. These techniques are Bayesian learning, rule-based learning and Dempster–Shafer theory.

In [18], the authors used a transaction aggregation technique to interpret the customer's behavior before any transaction is performed and then used this aggregated data to identify fake transactions. The entire analysis takes place on the behavior of the customers. The primary purpose of the work is to develop a model that can work with unknown datasets and highlight fake datasets in them. Banks give unspecified datasets due to privacy issues. Therefore, the model behaves similarly with all the participant attributes without prioritizing them. The model has also worked on the improper datasets and arranged them in two separate sections: one for legal transactions and the other for fake transactions [18].

In [19], the authors identified the issues of trust, privacy, security and verifiability in centralized-based IoT-driven smart cities. Therefore, the authors proposed a trustworthy privacy-preserving secure framework (TP2SF) for smart cities. The proposed framework comprises three modules: a module for trustworthiness, and two modules that consist of two-layered privacy modules. The trustworthiness module is a blockchain-based reputation system that ensures the system's security. Furthermore, two-layered privacy modules are based on an enhanced proof of work (ePoW) technique and principle component analysis (PCA). These modules transform the data into a reduced shape to prevent the system from poisoning attacks. However, a cloud system is used for data storage, which leads to a centralization problem.

In [20], the authors resolved the issue of privacy preservation through encryption techniques. They also used cryptographic approaches for the computation of data. The proposed system use asymmetric, symmetric and homomorphic encryption techniques to achieve privacy. However, high computational power and time are required to implement these approaches. Cyber attacks and intrusion detection are major problems that cause data privacy issues. Blockchain technology with deep learning algorithms is used to resolve the mentioned in [21]. These models provide security and privacy in virtual machines migrated to the cloud to protect IoT networks. The authors proposed a deep blockchain framework (DBF) model for intrusion detection based on bidirectional long short-term memory (BiLSTM) and blockchain.

In [22], the authors identified the issues of centralization and cyber attacks in cloud-based systems. Therefore, they proposed a mixture-of-localization-based outliers (MLO) system with a Gaussian mixture. This collaborative anomaly detection system detects insider and outsider attacks in a cloudbased system. Privacy preservation is highly important for cyber–physical systems (CPSs). In these systems, anomaly detection systems are required to protect the system from inner and outer attacks [23]. Therefore, the authors proposed a new privacy-preserving anomaly detection framework that protects the system from attacks and keeps sensitive information confidential. The proposed method is based on two modules, i.e., the pre-processing module and anomaly detection module that used a Gaussian mixture model (GMM). However, the proposed system is inefficient for tackling modern IoT attacks.

Adversarial Machine Learning Methods In adversarial machine learning, some machine learning techniques try to exploit the model's specific vulnerabilities and take advantage of the model's obtained information to generate some malicious attacks [24]. Some adversarial problems are discussed in the following papers.

In [25], the authors gave a comprehensive overview of the research conducted in the last decade, considering the pioneering research from the security of non-deep learning algorithms to the advances in this field, i.e., properties of security in deep learning algorithms.

In [26], the authors proposed unsupervised random forest algorithms to reduce the number of fraudulent transactions. Further, this proposed algorithm was used to analyze the detection of credit card fraud. Moreover, the Bayesian network assembles a coordinated non-cyclic chart, further used for the conditional probability distribution for creating a noncyclic graph. Results show that the random forest-based proposed algorithm performed better than its counterparts. Authors in [27] also proposed a random forest model for detailed feature selection, financial fraud detection, importance measurement of variables, and multidimensional and partial correction analysis. Nevertheless, the authors applied several statistical methodologies, i.e., non-parametric and parametric models, to detect accuracy. They concluded that non-parametric models have less accuracy compared with parametric models.

In [28], the authors worked on the problem of intrusion detection in cyber security. They used a dataset which has highly sensitive training data. This type of dataset is vulnerable to cyber attacks. To resolve this issue, they used a random forest algorithm that performs better in detecting cyber attacks. However, there is still room for researchers to improve the detection of cyber attacks.

In [29], the authors proposed an effective random forest classifier for anomaly detection in an IoT network. They also compared the performance of an intrusion detection system (IDS) and random forest classifier in terms of accuracy and false alarm rate. However, security is the major issue while implementing an IoT network.

In [30], the authors identified the problems of malicious data and manipulation of data by an attacker. Therefore, they implemented the evasion classifier and checked its effectiveness on a test case. The authors analyzed some potential techniques used to increase the robustness of machine learning models against the attacks of data manipulation.

In [31], the authors employed unsupervised machine learning techniques to detect the monetary anomalies.

Problem Statement With the advancement of technology, cyber crime is also increasing day by day, and the financial sector is the most affected sector by cyber crime [5]. The main reason for this problem is security vulnerabilities in financial systems. Anomalies occur in these systems, which are also known as frauds. In traditional financial systems, credit card frauds are the most common frauds, and AI techniques are used to solve these frauds. As a result, the financial industry suffers a loss of billions of dollars each year due to these frauds [1].

However, according to [32], supervised machine learning techniques are more effective for fraud detection. A large amount of learning data and labeled data is good for supervised learning. Therefore, the authors developed a complex model to learn the patterns of anomalies and fraud. However, this model is not able to provide accurate results. Moreover, blockchain innovation solves several fraud problems. It provides security and privacy to the financial sector, as it is decentralized and immutable. However, it does not address such issues as loss of privacy, Sybil attacks, and double-spending attacks. The purpose of these attacks is to discourage illegal activities and increase financial benefits. Bitcoin is a digital currency based on the concept of proof of work (PoW). In the Bitcoin network, all digital transactions are executed in a distributed manner using digital signatures and hashes via a timestamp service. Bitcoin transactions do not involve a trusted third party to verify the transactions. Therefore, a user can spend the same coin twice, which becomes a fraudulent transaction and is known as a double-spending attack [33]. The proposed model detects anomalies and thefts based on the predictive model. In the proposed work, machine learning models are trained on a dataset according to the fraud types and integrated transactions. The proposed model is linked with blockchain to overcome security and threats.

Proposed System Model The proposed system model consists of two layers: blockchain and machine learning. The blockchain model initiates transactions, and then machine learning models are used to classify these transactions as malicious or legitimate. This is a binary classification. The proposed system model is based on the integration of machine learning and blockchain for fraud and anomaly detection in the financial sector. The anomaly detection system identifies unusual suspicious events that are different from most of the data. A dataset of bitcoin transactions is used for the proposed model. We also use the random forest and XGboost classifiers to classify legitimate and malicious transactions. These classifiers are also used to predict new incoming transactions. The proposed model is trained and tested for legitimate and malicious data patterns using the given dataset. The proposed system model consists of the following steps (discussed in the below subsections).

Linkage of Blockchain with Machine Learning in the Proposed Model Blockchain technology has been used for the past few years to provide security and privacy in various networks. Despite the fascinating features of blockchain, it is still vulnerable to fraudulent activities. The malicious entities may perform invalid and fraudulent transactions using various methods, such as a double-spending attack. In the proposed system, blockchain is combined with machine learning to solve this problem. The database of bitcoin transactions is used in the underlying work, and the proposed machine learning model is trained on the dataset. The pattern of transactions stored in the database is analyzed for further use. In parallel, the transactions are performed on the Ethereum network. The pattern of these transactions is assumed to be similar to the pattern of bitcoin transactions stored in the bitcoin transaction database. Moreover, each new Ethereum transaction is made an input to the machine learning model, and the model is trained on it. The transaction pattern is analyzed and compared with the bitcoin transaction pattern. If the pattern of both transactions matches, the new transaction is classified as legitimate or malicious. To further test the robustness of the proposed system, a double-spending attack is implemented in the underlying work. In Figure 1, blockchain-based transactions are verified using a machine learning model, and the prediction result shows that the transaction is legitimate or malicious. The prediction of the machine learning model is based on the training and testing of a bitcoin transaction-based dataset.

Conclusions Nowadays, blockchain is the latest and most secure technology that covers various research areas related to security. Blockchain development is based on digital currencies and is used to secure digital financial transactions. It protects financial systems from fraudulent attacks. Therefore, a blockchain-based machine learning algorithm is proposed to secure digital transactions. The proposed model predicts whether the incoming transaction in the blockchain is fraudulent or not. The proposed machine learning algorithms are trained and tested on a bitcoin-based dataset based on bitcoin transactions and predict the behavior of the incoming transactionsTherefore, we generate synthetic malicious data points through SMOTE to achieve better results. We use XGboost and random forest to classify the model and calculate the confusion matrix. This classification allows the model to distinguish between fraudulent and real data. The simulation results show that the proposed algorithm works adequately to find transaction fraud. Moreover, two attacker models are implemented to check the efficacy of the system against bugs and attacks. The proposed system is robust against double-spending and Sybil attacks.

References

1. Staudemeyer, R.C.; Voyiatzis, A.G.; Moldovan, G.; Suppan, S.R.; Lioumpas, A.; Calvo, D. Smart cities under attack. In HumanComputer Interaction and Cybersecurity Handbook; CRC Press: Boca Raton, FL, USA, 2018.
2. Podgorelec, B.; Turkanovi´c, M.; Karakatiˇc, S. A machine learning-based method for automated blockchain transaction signing including personalized anomaly detection. Sensors 2020, 20,147
3. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 21 March 2020)
4. Farrugia, S.; Ellul, J.; Azzopardi, G. Detection of illicit accounts over the Ethereum blockchain. Expert Syst. Appl. 2020, 150, 113318
5. Ostapowicz, M.; Zbikowski, K. Detecting fraudulent accounts on blockchain: A supervised approach. In Proceedings of the ˙ International Conference on Web Information Systems Engineering, Hong Kong, China, 19–22 January 2020; Springer: Cham, Switzerland, 2020; pp. 18–31.
6. Aziz, A.S.A.; Hassanien, A.E.; Azar, A.T.; Hanafy, S.E. Genetic Algorithm with Different Feature Selection Techniques for Anomaly Detectors Generation. In Proceedings of the 2013 Federated Conference on Computer Science and Information Systems (FedCSIS), Kraków, Poland, 8–11 September 2013.
7. Hassanien, A.E.; Tolba, M.; Azar, A.T. Advanced Machine Learning Technologies and Applications: Second International Conference, AMLTA 2014, Cairo, Egypt, 28–30 November 2014. In Communications in Computer and Information Science; Springer: Berlin/Heidelberg, Germany, 2014; Volume 488, ISBN 978-3-319-13460-4.

8. Khan, H.; Asghar, M.U.; Asghar, M.Z.; Srivastava, G.; Maddikunta, P.K.R.; Gadekallu, T.R. Fake review classification using supervised machine learning. In Proceedings of the International Conference on Pattern Recognition, Virtual Event, 10–15 January 2021; Springer: Cham, Switzerland, 2021; pp. 269–288.
9. Shahbazi, Z.; Hazra, D.P.; Park, S.; Byun, Y.C. Toward Improving the Prediction Accuracy of Product Recommendation System Using Extreme Gradient Boosting and Encoding Approaches. Symmetry 2020, 12, 1566.
10. Pesantez-Narvaez, J.; Guillen, M.; Alcañiz, M. Predicting motor insurance claims using telematics data—XGBoost versus logistic regression. Risks 2019, 7, 70.
11. Li, J.; Gu, C.; Wei, F.; Chen, X. A Survey on Blockchain Anomaly Detection Using Data Mining Techniques. In Proceedings of the International Conference on Blockchain and Trustworthy Systems, Guangzhou, China, 7–8 December 2019; Springer: Singapore, 2019
12. Reid, F.; Harrigan, M. An analysis of anonymity in the bitcoin system. In Security and Privacy in Social Networks; Springer: New York, NY, USA, 2013; pp. 197–223.
13. Ngai, E.W.T.; Hu, Y.; Wong, Y.H.; Chen, Y.; Sun, X. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decis. Support Syst. 2011, 50, 559–569.
14. Saia, R.; Carta, S. Evaluating Credit Card Transactions in the Frequency Domain for a Proactive Fraud Detection Approach. In Proceedings of the 14th International Conference on Security and Cryptography (SECRYPT 2017), Madrid, Spain, 26–28 July 2017; pp. 335–342.
15. Sánchez, D.; Vila, M.A.; Cerda, L.; Serrano, J.M. Association rules applied to credit card fraud detection. Expert Syst. Appl. 2009, 36, 3630–3640.
16. Gyamfi, N.K.; Abdulai, J.D. Bank fraud detection using support vector machine. In Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 1–3 November 2018; pp. 37–41.
17. Panigrahi, S.; Kundu, A.; Sural, S.; Majumdar, A.K. Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning. Inf. Fusion 2009, 10, 354–363.
18. hi, F.B.; Sun, X.Q.; Gao, J.H.; Xu, L.; Shen, H.W.; Cheng, X.Q. Anomaly detection in Bitcoin market via price return analysis. PLoS ONE 2019, 14, e0218341.
19. Kumar, P.; Gupta, G.P.; Tripathi, R. TP2SF: A Trustworthy Privacy-Preserving Secured Framework for sustainable smart cities by leveraging blockchain and machine learning. J. Syst. Archit. 2021, 115, 101954.
20. Zhao, Y.; Tarus, S.K.; Yang, L.T.; Sun, J.; Ge, Y.; Wang, J. Privacy-preserving clustering for big data in cyber-physical-social systems: Survey and perspectives. Inf. Sci. 2020, 515, 132–155.
21. Alkadi, O.; Moustafa, N.; Turnbull, B.; Choo, K.K.R. A deep blockchain framework-enabled collaborative intrusion detection for protecting IoT and cloud networks. IEEE Internet Things J. 2020, 8, 9463–9472
22. AlKadi, O.; Moustafa, N.; Turnbull, B.; Choo, K.K.R. Mixture localization-based outliers models for securing data migration in cloud centers. IEEE Access 2019, 7, 114607–114618.

23. Keshk, M.; Sitnikova, E.; Moustafa, N.; Hu, J.; Khalil, I. An integrated framework for privacy-preserving based anomaly detection for cyber-physical systems. IEEE Trans. Sustain. Comput. 2019, 6, 66–79.
24. Kurakin, A.; Goodfellow, I.; Bengio, S. Adversarial machine learning at scale. arXiv 2016, arXiv:1611.01236.
25. Biggio, B.; Roli, F. Wild patterns: Ten years after the rise of adversarial machine learning. Pattern Recognit. 2018, 84, 317–331.
26. Xuan, S.; Liu, G.; Li, Z.; Zheng, L.; Wang, S.; Jiang, C. Random forest for credit card fraud detection. In Proceedings of the 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai, China, 27–29 March 2018; pp. 1–6.
27. Liu, C.; Chan, Y.; Alam Kazmi, S.H.; Fu, H. Financial fraud detection model: Based on random forest. Int. J. Econ. Financ. 2015, 7, 178–188.
28. Apruzzese, G.; Andreolini, M.; Colajanni, M.; Marchetti, M. Hardening random forest cyber detectors against adversarial attacks. IEEE Trans. Emerg. Top. Comput. Intell. 2020, 4, 427–439.
29. Primartha, R.; Tama, B.A. Anomaly detection using random forest: A performance revisited. In Proceedings of the 2017 International Conference on Data and Software Engineering (ICoDSE), Palembang, Indonesia, 1–2 November 2017; pp. 1–6.
30. Laskov, P. Practical evasion of a learning-based classifier: A case study. In Proceedings of the 2014 IEEE Symposium on Security and Privacy, San Jose, CA, USA, 18–21 May 2014; pp. 197–211.
31. Pham, T.; Lee, S. Anomaly detection in bitcoin network using unsupervised learning methods. arXiv 2016, arXiv:1611.03941.
32. Martin, K.; Rahouti, M.; Ayyash, M.; Alsmadi, I. Anomaly detection in blockchain using network representation and machine learning. Secur. Priv. 2022, 5, e192.
33. Pinzón, C.; Rocha, C. Double-spend attack models with time advantange for bitcoin. Electron. Notes Theor. Comput. Sci. 2016, 329, 79–103.
34. Bitcoin Network Transactional Metadata. Available online:
35. Shafiq, O. Anomaly Detection in Blockchain. Master's Thesis, Tampere University, Tampere, Finland, 2019.
36. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. J. Artif. Intell. Res. 2002, 16, 321–357
37. Sadaf, K.; Sultana, J. Intrusion detection based on autoencoder and isolation Forest in fog computing. IEEE Access 2020, 8, 167059–167068.
38. Eyal, I.; Sirer, E.G. Majority is not enough: Bitcoin Mining is vulnerable. In Proceedings of the International Conference on Financial Cryptography and Data Security, Christ Church, Barbados, 3–7 March 2014; pp. 436–454; Springer: Berlin/Heidelberg, Germany, 2014.
39. Landa, R.; Griffin, D.; Clegg, R.G.; Mykoniati, E.; Rio, M. A Sybilproof indirect reciprocity mechanism for peer-to-peer networks. In Proceedings of the IEEE INFOCOM 2009, Rio De Janeiro, Brazil, 24 April 2009; pp. 343–351.

40. Luu, L.; Chu, D.-H.; Olickel, H.; Saxena, P.; Hobor, A. Making smart contracts smarter. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016.