International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6970
***Article Received**: 30 March 2023 **Revised**: 19 May 2023 **Accepted**: 05 June 2023*

Fraud Detection Using Machine Learning and Blockchain

Dr. Vaishali Gaikwad(Mohite)[1], Dr. Kunal Meher[2], Ryan Dass[3], Athisaya Sarah Jonista[4], Jeston D'Souza[5], Raymun Victor[6]

[1,2,3,4,5,6]Computer Engineering,
Xavier Institute of Engineering, University of Mumbai
Mumbai, India
vaishali.g, kunal.m@xavier.ac.in
201901006.ryandam, 201901001.athisayasss, 201901011.jestondgc, 201901061.raymunvvj @student.xavier.ac.in

Abstract—In the 21st century financial fraud is on the rise in many institutions. Newly released Federal Trade Commission data shows that consumers reported losing nearly $8.8 billion to fraud in 2022, an increase of more than 30 percent over the previous year. The main goal for us is to develop an efficient fraud detection system and utilize blockchain to create a decentralized banking application. Our team has collaborated on various cutting-edge technologies, such as machine learning and blockchain, to create a sophisticated fraud detection system. We have implemented three machine learning algorithms, namely Logistic Regression, Decision Tree, and Random Forest, which have been used to improve the accuracy of the model for detecting fraudulent activities. As for fraud aversion, we have used the Blockchain technology which provides a tamper-proof system that can securely record and track financial transactions, ensuring transparency and security. This feature makes it an ideal solution for fraud prevention, as it guarantees that all transactions are legitimate and free from any manipulations. By combining the power of machine learning and blockchain technology, our team is confident in providing an innovative solution that will benefit all stakeholders involved

Keywords- MachineLearning, Blockchain, Algorithms, FraudDetection, Transactions.

I. INTRODUCTION

The digitization of the world has brought about a plethora of benefits to society, such as greater connectivity, convenience, and access to information. However, it also comes with certain drawbacks that we cannot ignore. One of the most pressing issues is the rising number of fraudulent activities in various sectors such as banking, healthcare, retail, education, and charity. These sectors rely on online transactions and data storage methods, making them vulnerable to fraudulent activities. According to Juniper Research, online payment fraud losses are expected to exceed $206 billion by 2025. The effects of fraud go beyond financial losses and can have serious repercussions on an individual's personal and professional life. It can erode trust in institutions and systems and undermine the integrity of the entire financial system. To tackle this problem, our team has developed a comprehensive fraud detection system that leverages the power of machine learning and blockchain technology.

Our primary objective is to build an efficient fraud detection system and use blockchain technology to create a decentralized banking application that is secure and transparent. Our team has collaborated on several state-of-the-art technologies, including machine learning and blockchain, to design a sophisticated fraud detection system. We have incorporated three machine learning algorithms, namely Logistic Regression, Decision Tree, and Random Forest, to enhance the accuracy of the model for detecting fraudulent activities. These algorithms are used to identify patterns and anomalies in data that could indicate fraudulent activities. Blockchain provides a tamper-proof system that can securely record and track financial transactions, ensuring transparency and security. This feature makes it an ideal solution for fraud prevention, as it guarantees that all transactions are legitimate and free from any manipulations. Blockchain technology is designed to create a decentralized, transparent, and secure digital environment, where all financial transactions can be monitored in real-time.

Our project aims to address this issue by developing a sophisticated fraud detection system that leverages the power of machine learning. It is time to take proactive measures to safeguard ourselves and our organizations from fraudulent activities.

II. AIMS AND OBJECTIVES

The main aim of this project is to detect fraudulent transactions which are happening lot these days using Machine Learning and to help avert future fraudulent operations with the help of Blockchain.

The objectives we focused on can be explained in four points:

i. To perform an initial survey of the existing work to know its limitations.
ii. Create user-friendly interface for users to perform banking operations on the blockchain easily.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6970
Article Received: 30 March 2023 Revised: 19 May 2023 Accepted: 05 June 2023

iii. Making an accurate and efficient machine learning model to predict frauds in financial transactions.
iv. Optimizing the system for more accurate detections.

These are the major objectives we worked on during this project implementation. This is a feature rich project which will bring privacy, security and efficiency to a traditional banking sector and help solve many existing problems.

III. LIMITATIONS OF EXISTING SYSTEM

Although impressive, the existing systems had multiple limitations which was understood by us while working on the literature survey where we analysed several research papers from reputed journals. We covered an in-depth explanation of these in our survey paper titled "Survey on Fraud Detection using Blockchain". Given below is the Literature survey table to give one an overall understanding about the different systems out there.

Table: Comparison of present mechanisms with proposed methodology

PAPERS	SECURITY	SCALABILITY	SPEED	COMPLEXITY
Machine Learning and Blockchain for Fraud Detection: Employing Artificial Intelligence in the Banking Sector	✓		✓	
A study on blockchain technology in banking sector	-	-	-	-
Detecting Fraudulent Accounts on Blockchain: A Supervised Approach			✓	✓
Role of Smart Contract Technology Blockchain Services in Finance and Banking Systems	-	-	-	-
A Fraud detection system using Machine Learning	✓	✓		✓
Blockchain Revolution in Banking Industry	-	-	-	-
Blockchain and Machine Learning for Fraud Detection: A Privacy-Preserving and Adaptive Incentive Based Approach	✓			✓
A Review on Blockchain Technology and the Impact on Finance Sector by Blockchain Technology	-	-	-	-
Credit Card Fraud Detection using Machine Learning and Data Science		✓	✓	
Fraud Detection using Machine Learning and Deep Learning	✓		✓	
Fraud detections for online businesses: a perspective from blockchain technology	✓	✓		✓
Avoiding Insurance Fraud: A Blockchain-bas ed Solution for the Vehicle Sector	✓		✓	✓
Blockchain for Fraud Prevention: A Work-History Fraud Prevention System	✓			✓
The Influence of Blockchain Technology on Fraud and Fake Protection.		✓		✓
Are blockchains immune to all malicious attacks	✓			✓
A Blockchain-Bas ed Framework for Fraud Detection	✓		✓	✓
Counterfeit Detection of Documents using Blockchain	✓		✓	✓
A Multifaceted Approach to Bitcoin Fraud Detection: Global and Local Outliers	✓	✓		✓
Fraud Detection Systems using Blockchain	✓	✓	✓	✓

The above table gives us an overall understanding about the different pros and cons of the existing systems. As you can see that only the proposed system which we have mentioned at the end has all the pros, meaning it has overcome all the limitations of the existing systems given users the upmost experiences. The major limitations or concerns for these existing systems would be:

i. Security Issues
ii. Service Charges, Middle-man costs

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6970
Article Received: 30 March 2023 Revised: 19 May 2023 Accepted: 05 June 2023

iii. Geographical Barriers
iv. Improved data quality

Security Issues aroused with the existing systems where one can say that Internet banking is completely insecure as there are many chances of data leakage and hackings which can lead to financial losses to the users. Even though many organizations or systems does not charge the users some do, and this might be seen as a limitation, but these charges are quite nominal and sometimes even zero.

IV. FRAUD DETECTION SYSTEM

We propose a system that addresses the limitations of existing fraud prevention technologies by making a system where fraud detection is done using machine learning and fraud aversion is done using blockchain. Existing fraud prevention technologies have limitations, such as being rule-based, expensive, and generating false positives. We propose a system that addresses these limitations by using machine learning algorithms to detect fraud. Machine learning algorithms can be trained on historical data of fraudulent transactions to learn what to look for.

Once the algorithms are trained, they can be used to scan new transactions for signs of fraud. This system is more effective at detecting fraud, less expensive to implement and maintain, and generates fewer false positives. Here are some of the machine learning algorithms that can be used for fraud detection:

Logistic regression: Logistic regression is a supervised learning algorithm that can be used to predict the probability of an event occurring. In the context of fraud detection, logistic regression can be used to predict the probability of a transaction being fraudulent.
Decision trees: Decision trees are a supervised learning algorithm that can be used to classify data. In the context of fraud detection, decision trees can be used to classify transactions as fraudulent or non-fraudulent.
Random forests: Random forests are an ensemble learning algorithm that combines multiple decision trees to make a prediction. In the context of fraud detection, random forests can be used to improve the accuracy of fraud detection by combining the predictions of multiple decision trees.

These algorithms help identify patterns and anomalies in data that could indicate fraudulent activities. Our proposed system provides a more robust solution for fraud prevention that addresses the limitations of existing systems and ensures the security and integrity of financial transactions

V. FRAUD AVERSION SYSTEM

We propose a system where we have developed a web3 decentralized banking application where users will be able to send and receive money using their Metamask wallet.
The application will be built on top of the Ethereum blockchain and will use smart contracts to store and manage user funds. Users will be able to create accounts, deposit funds, and send and receive money to other users.
The application will be secure and transparent, as all transactions will be recorded on the blockchain. Users will also have complete control over their funds, as they will be able to access their wallets directly without the need for a third party.
We believe that this application will revolutionize the way people bank. It will provide a secure, transparent, and user-friendly alternative to traditional banking.
Here are some of the benefits of using our decentralized banking application:

- Security: All transactions are recorded on the blockchain, which is a secure and transparent ledger. This makes it very difficult for fraudsters to operate undetected.
- Transparency: All transactions are visible to everyone on the network. This makes it difficult for fraudsters to tamper with financial records.
- Control: Users have complete control over their funds. They can access their wallets directly without the need for a third party.
- Efficiency: Transactions are processed quickly and efficiently. There is no need to wait for approvals from banks or other intermediaries.
- Cost-effectiveness: Transactions are very cost-effective. There are no fees charged by banks or other intermediaries.

We believe that our decentralized banking application is a superior alternative to traditional banking. It is more secure, transparent, efficient, and cost-effective. We are confident that it will revolutionize the way people bank.
How Users Send and Receive Money Using Metamask:
Metamask is a cryptocurrency wallet that allows users to interact with decentralized applications (DApps). To send or receive money using Metamask, users first need to create an account and add funds to their wallet. Once they have funds in their wallet, they can use Metamask to connect to DApps and send or receive money.
To send money, users simply need to enter the recipient's address and the amount of money they want to send. Metamask will then encrypt the transaction and broadcast it to the network. The recipient will then be able to see the transaction and claim the funds.
To receive money, users simply need to share their address with the sender. The sender can then use Metamask to send the funds to the address. The recipient will then be notified of the incoming funds and will be able to claim them.

Overall, the decentralized banking system is a more secure and fraud-resistant alternative to the traditional banking system.

The figure below shows the flow of how our system processes and analysis data for prediction purposes. It explains or highlights the overall working of the machine learning part, the diagram is quite self-explanatory

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6970
***Article Received**: 30 March 2023 **Revised**: 19 May 2023 **Accepted**: 05 June 2023*

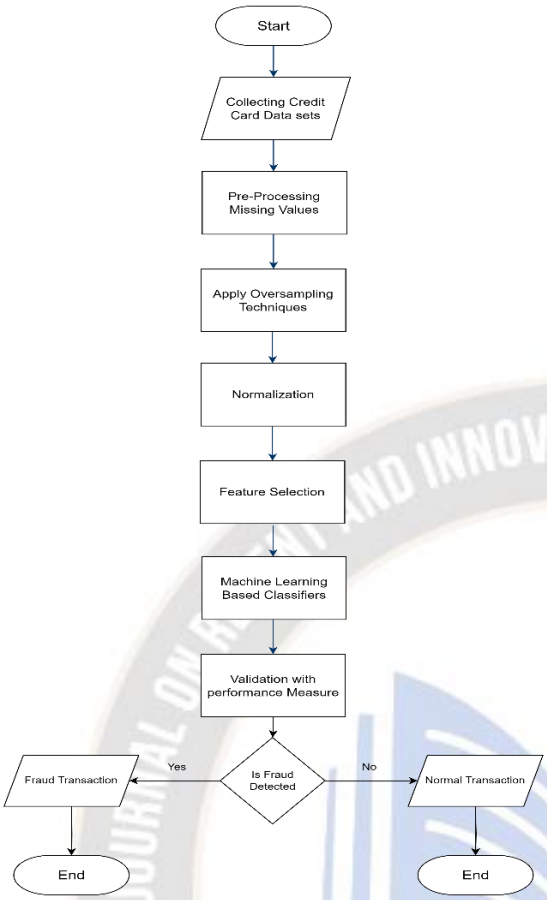


Figure 1: Process Flow Chart

The Block diagram below explains the stepwise procedure of how the entire system works.

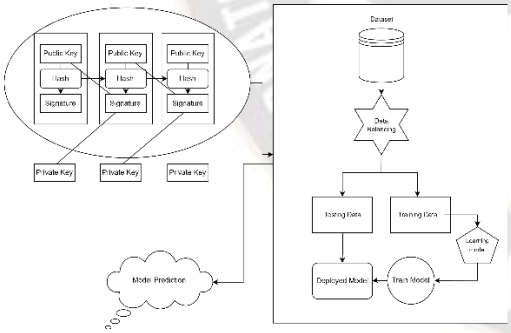


Figure 2: Block Diagram of entire model

VI. RESULTS

Below are the screenshots of our Model's Application:

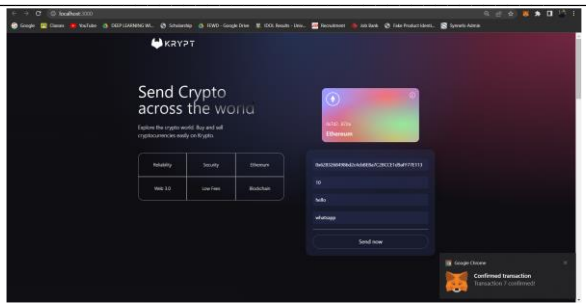


Figure 3.1: Transaction Completion

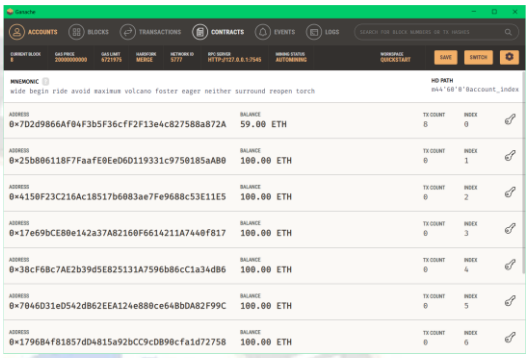


Figure 3.2: Ganache Workspace

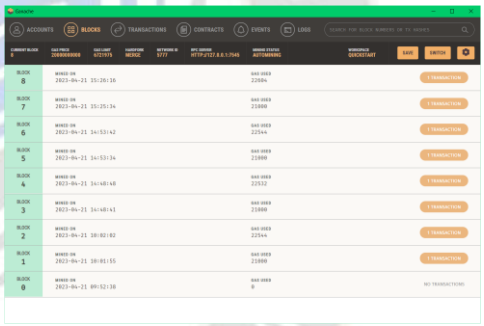
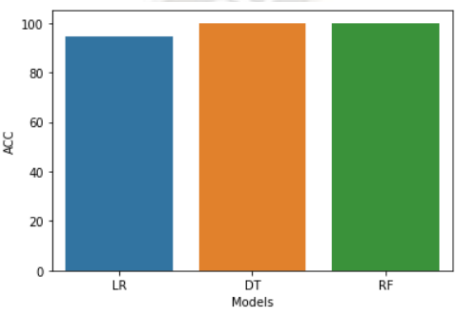


Figure 3.3: Blocks of Blockchain



	Models	ACC
0	LR	94.343908
1	DT	99.818307
2	RF	99.991824

Figure 3.4: Comparison of Algorithms

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6970
Article Received: 30 March 2023 Revised: 19 May 2023 Accepted: 05 June 2023

Figure 3.5: Output of Fraud value

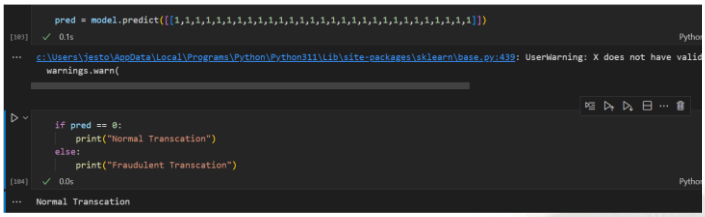


Figure 3.6: Output of Normal value

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper.

COMPARISON OF ALGORITHMS:

We used three different machine learning algorithms to detect fraud in credit card data: random forest, decision tree, and logistic regression. You found that random forest had the best performance, with an accuracy of 99.99%. Decision tree had an accuracy of 99.81%, and logistic regression had an accuracy of 94.87%.

There are several reasons why random forest performed better than the other two algorithms. First, random forest is more robust to overfitting. Overfitting occurs when a model learns the training data too well, and as a result, it does not generalize well to new data. Random forest is less likely to overfit because it is made up of a collection of decision trees, each of which is trained on a different subset of the data.

Second, random forest can handle imbalanced data better than decision tree and logistic regression. Imbalanced data is data where there are more instances of one class than another. In the case of credit card fraud, there are far more instances of non-fraud transactions than fraud transactions. Random forest is better at handling imbalanced data because it uses a technique called bagging, which helps to reduce the impact of the majority class.

Finally, random forest can capture non-linear relationships between features better than decision tree and logistic regression. Decision trees and logistic regression can only capture linear relationships between features, while random forest can capture non-linear relationships as well. This is important because many of the relationships between features in credit card data are non-linear.

Overall, random forest is a more effective algorithm for detecting fraud in credit card data than decision tree and logistic regression. It is more robust to overfitting, it can handle imbalanced data better, and it can capture non-linear relationships between features.

IMPLEMENTATION EXAMPLES:

There are multiple big companies working on the same principle to make an efficient fraud detection model. Some are successful but others are still trying to figure out their work. These companies' are: CipherTrace [21], Chainalysis [22], Elliptic [23], AnChain.AI [24], Scorechain [25], Nuggets [26], MetaCert [27], BigPanda [28], Unbound Tech [29], Forte [30].

These Companies act as an example or can provide motivating guidance for new comers to work on their projects efficiently.

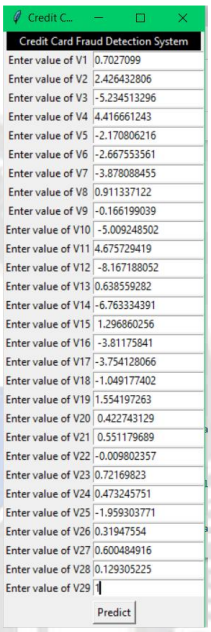


Figure 3.7: GUI of ML Model

VII. CONCLUSION

In conclusion, the benefits of digitization are evident, but the increasing frequency of fraudulent activities is a significant challenge that requires a comprehensive solution. Fraudulent activities can have far-reaching effects beyond financial losses, undermining trust in institutions and systems, and eroding the integrity of the entire financial system. Our team has responded to this challenge by developing a fraud detection system that leverages machine learning and blockchain technology. By incorporating sophisticated algorithms and tamper-proof systems, our solution aims to enhance the accuracy of fraudulent activity detection, improve efficiency, transparency, and trust in financial transactions, and create a more secure digital environment for all stakeholders involved. We believe that our project has the potential to have a significant impact on the financial sector and beyond. It is time for proactive measures to prevent fraudulent activities and safeguard ourselves and our organizations from financial losses and reputational damage. Our team's solution is one step towards this goal, and we are committed to ensuring its success and promoting a more secure and trustworthy digital environment for all.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6970
***Article Received**: 30 March 2023 **Revised**: 19 May 2023 **Accepted**: 05 June 2023*

REFERENCES

[1] V. Silaparasetty, "Machine Learning and blockchain for fraud detection: Employing artificial intelligence in the banking sector," GCU INTERNATIONAL KNOWLEDGE TRANSFER CONCLAVE - (ISBN 978-93-86516- 46-6), 2018.

[2] C.Mallesha& S.Haripriya,A STUDY ON BLOCKCHAIN TECHNOLOGY IN BANKING SECTOR, International Journal of Advanced Research in Commerce, Management & Social Science,, Volume 02, No. 03, July - September, 2019, pp 123-132.

[3] Mr. Dharmesh Dhabliya, Mr. Rahul Sharma. (2012). Efficient Cluster Formation Protocol in WSN. International Journal of New Practices in Management and Engineering, 1(03), 08 - 17. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/7

[4] M. Ostapowicz and K. Z' bikowski, "Detecting fraudulent accounts on blockchain: A supervised approach," in Web Information Systems Engineering – WISE 2019, Cham: Springer International Publishing, 2019, pp. 18–31.

[5] Dr.Ahmed Muayad Younus, Mohanad Abumandil, "Role of Smart Contract Technology Blockchain Services in Finance and Banking Systems: Concept and Core Values" in Advanced Engineering Informatics - January 2022, 1013465.

[6] Dr. Dhananjay Kalbande, Pulin Prabhu, Anisha Gharat, Tania Rajabally, A Fraud detection system using Machine Learning, Department of Computer Engineering Sardar Patel Institute of Technology.

[7] Deshpande, V. (2021). Layered Intrusion Detection System Model for The Attack Detection with The Multi-Class Ensemble Classifier . Machine Learning Applications in Engineering Education and Management, 1(2), 01–06. Retrieved from http://yashikajournals.com/index.php/mlaeem/article/view/10

[8] Thulya Palihapitiya, Blockchain Revolution in Banking Industry-October 2020.

[9] Pemith Waidyaratne,"A Review on Blockchain Technology and the Impact on Finance Sector by Blockchain Technology", General Sir John Kotelawala Defence University: July 2022.

[10] T. H. Pranto, K. T. A. M. Hasib, T. Rahman, A. B. Haque, A. K. M. N. Islam, and R. M. Rahman, "Blockchain and machine learning for fraud detection: A privacy-preserving and adaptive incentive-based approach," IEEE Access, vol. 10, pp. 87115–87134, 2022.

[9] S P Maniraj, Aditya Saini, Shadab Ahmed, Swarna Deep Sarkar, "Credit Card Fraud Detection using Machine Learning and Data Science"- IJERT-September 2019.

[10] Pradheepan Raghavan, Neamat El Gayar, Fraud Detection using Machine Learning and Deep Learning, December 2019.

[11] Cai, Y., & Zhu, D. (2016). Fraud detections for online businesses: a perspective from blockchain technology. Financial Innovation, 2(1), 1-10.

[12] Rui Roriz[a], José LuisPereira, "Avoiding Insurance Fraud: A Blockchain-based Solution for the Vehicle Sector" ScienceDirect-Procedia Computer Science 164 (2019) 211–218.

[13] Paul Sarda, Mohammad Jabed Morshed Chowdhury, Alan Colman, Muhammad Ashad Kabir, Jun Han, "Blockchain for Fraud Prevention: A Work-History Fraud Prevention System"-IEEE-2018.

[14] Joshi, P., Kumar, S., Kumar, D., & Singh, A. K. (2019, September). A blockchain based framework for fraud detection. In 2019 Conference on Next Generation Computing Applications (NextComp) (pp. 1-5). IEEE.

[15] Dhiran, A., Kumar, D., & Arora, A. (2020, July). Video Fraud Detection using Blockchain. In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 102-107). IEEE.

[16] Pathak, D. G. ., Angurala, D. M. ., & Bala, D. M. . (2020). Nervous System Based Gliomas Detection Based on Deep Learning Architecture in Segmentation. Research Journal of Computer Systems and Engineering, 1(2), 01:06. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/3

[17] Nerurkar P, Bhirud S, Patel D, Ludinard R, Busnel Y, Kumari S. Supervised learning model for identifying illegal activities in Bitcoin. Appl Intell. 2020;209(1):120.

[18] Ostapowicz, M., & Z' bikowski, K. (2020, January). Detecting fraudulent accounts on blockchain: a supervised approach. In International Conference on Web Information Systems Engineering (pp. 18-31). Springer, Cham.

[19] Raikwar, M., Mazumdar, S., Ruj, S., Gupta, S. S., Chattopadhyay, A., & Lam, K. Y. (2018, February). A blockchain framework for insurance processes. In 2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS) (pp. 1-4). IEEE.

[20] B. K., P., Naikodi, C. ., & Suresh L. (2023). Hybrid Meta-Heuristic Technique Load Balancing for Cloud-Based Virtual Machines. International Journal of Intelligent Systems and Applications in Engineering, 11(1), 132–139. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2451.

[21] Rodriguez, L., Rodríguez, D., Martinez, J., Perez, A., & Ólafur, J. Leveraging Machine Learning for Adaptive Learning Systems in Engineering Education. Kuwait Journal of Machine Learning, 1(1). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/103

[22] Vaishali Mohite, Lata Ragha, "Cooperative Security Agents", World Congress on Information and Communication Technologies (WICT-2012), 30th Oct - 2nd Nov 2012, Trivandrum, India, Publication- IEEE Xplore and indexed by EI Compendex and ISTP, 978-1-4673-4805-8/12/$31.00© 2012 IEEE.

[23] Vaishali Gaikwad (Mohite), Lata Ragha, "Security Agents for Detecting and Avoiding Cooperative Blackhole Attacks in MANET", Applied & Theoretical Computing & Communication Technology, from October 29-31, 2015, SCI, IEEE Part Number: CFP15D66-USB IEEE ISBN: 978-14673-9222-8-$31.00© 2015 IEEE held in BIT, Karnataka, India.

[24] Isabella Rossi, Reinforcement Learning for Resource Allocation in Cloud Computing , Machine Learning Applications Conference Proceedings, Vol 1 2021.

[25] CipherTrace.[Online].Available:https://ciphertrace. com/. [Accessed: May. 13, 2023].

[26] Chainalysis.[Online].Available:https://chainalysis.com/. [Accessed: May. 13, 2023].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.6970
***Article Received**: 30 March 2023 **Revised**: 19 May 2023 **Accepted**: 05 June 2023*

[27] Elliptic.[Online].Available:https://www.elliptic.co/. [Accessed: May. 13, 2023].
[28] AnChain.AI.[Online].Available:https://www.anchain.ai/. [Accessed: May. 13, 2023].
[29] Scorechain.[Online].Available:https://www.scorechain.com/. [Accessed: May. 13, 2023].
[30] Nuggets.[Online].Available:https://nuggets.life/. [Accessed: May. 13, 2023].
[31] MetaCert.[Online].Available:https://metacert.com/. [Accessed: May. 13, 2023].
[32] BigPanda.[Online].Available:https://www.bigpanda.io/. [Accessed: May. 13, 2023].
[33] UnboundTech.[Online].Available:https://www.unboundtech.com/. [Accessed: May. 13, 2023].
[34] Forte.[Online].Available:https://www.forte.io.[Accessed: May. 13, 2023]..