Fraud Fighters - How AI and ML are Revolutionizing UPI Security

S. K. Lokesh Naik[1], Ajmeera Kiran[2], Vadapally Praveen Kumar[3], Shanmukh Mannam[4], Yesarapu Kalyani[5], Manda Silparaj[6]

[1]Department of Computer Science and Engineering, MLR Institute of Technology, Hyderabad, India
[2]Department of Computer Science and Engineering, MLR Institute of Technology, Hyderabad
[3]Department of Computer science Engineering (Data science), CVR College of Engineering, Hyderabad-501510.
[4]Computer Science and Engineering, MLR Institute of Technology, Hyderabad, India
[5]Computer Science and Engineering, MLR Institute of Technology, Hyderabad, India
[6]Department of Computer Science and Engineering (CS), CVR College of Engineering, Hyderabad-501510

E-mail : lokesh.naik@mlrinstitutions.ac.in, kiranphd.jntuh@gmail.com, micro091983@gmail.com, shanmukhmannam@gmail.com, yesarapukalyani@gmail.com, silparajm@gmail.com,

2024 International Conference on Science Technology Engineering and Management (ICSTEM) | 979-8-3503-7691-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICSTEM61137.2024.10560740

Abstract- Unified Payments Interface (UPI) is a highly efficient and securely encrypted monetary mechanism extensively deployed across India for the immediate transfer of funds between various bank accounts of the involved parties. However, with the rise of digital payments, UPI fraud has become a serious problem. Since more and more transactions are taking place online, scammers have discovered new ways to take advantage of holes in the system and steal money from blind people. The consequences of UPI fraud can be devastating for individuals and businesses alike. Not only does it result in financial losses, but it can also damage reputations and erode trust in the system. That's where AI and ML come in. By leveraging the power of artificial intelligence and machine learning, we can effectively detect potentially fraudulent Unified Payments Interface transactions. By using advanced algorithms and machine learning strategies, we can detect as well as prevent fraud more effectively than ever before. The goal of this project is to identify frauds by analyzing public data, managing imbalanced data, adapting to changes in fraud approaches, and minimizing false alarms. The primary goal is to utilize Machine Learning algorithms and Artificial Intelligence that have been recently developed for this objective.

Keywords— Unified Payment Interface (UPI), Encrypted Monetary Mechanism, Artificial Intelligence, Machine learning.

I. INTRODUCTION

In an era that uses digital finance and cashless transactions, the way individuals exchange money in India has been completely transformed by Unified Payments Interface (UPI), which has emerged as another game-changer[1]. However, with increase of UPI's popularity, so too has the sophistication of fraud along with security threats. It is now essential to integrate AI and ML technologies in order to protect users and ensure the security of this critical financial infrastructure[2].

This project explores the transformative role that AI and ML play in bolstering UPI security. It delves into the mechanisms behind UPI, elucidates the evolving security challenges, and, most importantly, examines how AI and ML are leveraged to detect and prevent fraudulent activities[3]. Through a comprehensive exploration of cutting-edge techniques, real-world case studies, and the future prospects of this dynamic field, we uncover the fraud fighters of the digital age and their pivotal role in securing the financial transactions that underpin modern India[4].

We have included a dataset of fictitious mobile-based payment transactions in this project[5]. We classify this data according to the various kinds of transactions it includes in order to analyze it. The dataset includes transaction categories labeled with the following information: UPI number, UPI holder name, date of birth, state, pin code, date and time of transaction, transaction amount, seller name, and merchant category.

II. LITERATURE SURVEY

Detecting card fraud with Deep Convolution Neural Network Model: The research focuses on fraud detection using ML algorithms that leverage data preprocessing. It employs a large quantity of data as a dataset. The algorithms have an accuracy of up to 75% and offer less efficient outcomes. The security is reduced since it must train numerous times to attain the requisite precision. The financial industry and its stakeholders face far-reaching consequences as a result of financial fraud. Several obstacles have risen in recent years, notwithstanding rising technology advancement and dependability. In the age of big data, old strategies remain ineffective. One of the models utilizes card fraud data to detect financial fraud using DCNN. Despite the large number of data, the model outperforms previous methods in terms of detection. Deep learning provides high accuracy and rapid pattern recognition for recognizing complicated and unfamiliar patterns. This model solves the inefficiencies of the previous models. To evaluate performance, current

machine learning models, auto-encoder models, and other deep learning models are compared to the proposed approach, which incorporates a real-time credit card fraud dataset. In future work, fraud location, time calculation, and other aspects might be combined into a single algorithm [6-8].

Prudent Fraud Detection in Internet Banking: This study concluded that the data mining approach extracts all relevant information from a vast quantity of knowledge and alters all high-level decision making in the banking and retail sectors. It involves mixing varied data from different databases and storing it in an acceptable format for analysis. This data is then used to target and exploit new clients, detect fraud in real-time, manage risk, and analyze customer data. Data mining is crucial for preventing fraud activities in banks and enhancing decision-making power. It extracts essential patterns from huge databases, hence enhancing database quality. This research paper discusses numerous concerns linked to financial information security and suggests strategies to overcome these issues using the supplied methodologies [9-10].

In this article, electronic payment systems are those that allow transactions to be completed without the need for either checks or cash. Machine Learning algorithms are used in these systems, with an accuracy rate of about 90%. Artificial Intelligence (AI) tools and algorithms are being developed to create disruptive technology for consumers. Major AI payment industry firms like PayPal, Amazon's Kindle, MasterCard, and Google are focusing on customer satisfaction in payment processing. This article provides an integrated view of e-payments, discussing their types, returns, obstacles, risks, and future prospects, aiming to provide a comprehensive understanding for those interested in electronic payments [11].

The paper discusses fraud detection and prevention solutions using Machine Learning technology. It analyzes existing methods and their impact on data protection legislation and respondents' views on privacy in fraud identification. The paper also discusses adversarial Machine Learning techniques in banking fraud detection, identifying challenges and designing novel evasion attacks. The models are validated by simulating attackers against state-of-the-art systems. The Hidden Markov Model technique is used in this article to detect and prevent fraud in online banking [22]. It also promises that legitimate transactions are not denied by sending individuals a one-time password created by the bank server via SMS [12-14].

This study proposes a sequence classification model for online banking fraud detection, focusing on features and transactions without recurrent or convolutional connections. The model, based on South American bank data, outperforms LSTM and identifies relevant features and transactions for final judgment [15].

Artificial intelligence (AI) is transforming financial institutions by enhancing customer experience through payments and digital transactions. Companies like PayPal, Amazon's Kindle, MasterCard, and Google prioritize customer satisfaction in this field. A study proposes integrating Artificial Neural Networks (ANN) and Human-Sensing Algorithms (HSA) to create a fraud-detection model called NNHS [24]. The system accurately identifies hidden algorithms and predicts the ideal ANN structure, with the best accuracy achieved from a German dataset. Financial fraud is a significant issue that undermines customer confidence and causes economic losses for banks and financial institutions. To combat this, financial institutions are exploring methods to detect fraud using Artificial Neural Network techniques and Harmony Search Algorithm. However, the accuracy of these methods is only 72%. The 2019 paper, published in 2019, highlights the importance of strong, effective, and efficient banking in an economy. The Fuzzy Darwinian System and Markov model are used to detect fraud, but the accuracy is only 74%. The banking sector is a cornerstone of any economy, and its strong, effective, and efficient nature is crucial for economic growth. [16-17].

The article analyzes current fraud banking systems, their problems, and available machine learning strategies for increasing detection rates and lowering fraud rates. It examines and contrasts various methods to emphasize their assumptions and goals. Big data difficulties impede the development and efficacy of Fraud Detection Systems (FDSs), and numerous academics offer a variety of machine learning strategies to improve detection accuracy, speed, and decrease false warnings. The article's accuracy is just 74%. Online banking activities are rapidly expanding, resulting in a drastic increase of attempting a fraud. However, there is little research on internet banking fraud detection. The research presents a block diagram used for categorizing bank transactions as fraud or legitimate by solving a sequence classification issue without recurring or convolution connections. The approach can be considered basic but is also effective, as illustrated using instantaneous statistics through bank located in South America. The article also covers credit card fraud's influence on the financial industry, including theoretical elements as well as comparative assessments of existing machine learning algorithms used to combat credit card fraud. It indicates that supervised algorithms are more widely employed than other approaches like SVM. Research focuses on relevant machine learning algorithms and data engineering approaches to provide a contribution for dealing with unbalanced data [18-20].

Credit card fraud is defined as using a defrauder's credit card information or committing theft for financial gain[23]. Economic prosperity has led to greater card expenditure, that encountered resulted in a leap of

misleading activity. This problem has resulted in substantial losses for both businesses and individuals. Various other models turned out to be used for detection of card fraud. Decision Tree classifier is more suitable for fraud detection, while LGBM Model outperforms AdaBoost and XGBM in boosting classifiers. Bagging ensemble techniques have the best results, making them suitable for credit card fraud detection [21].

III. Existing System

The existing system gathers information from a number of sources, such as past fraud incidents, user profiles, and UPI transaction records. Its reliance on historical data raises the possibility of a lag in fraud detection because it might not be indicative of changing fraud practices. It can be challenging to modify current systems to accommodate growing transaction volumes, doing so may have an influence on system performance and necessitate ongoing modifications. According to the information at hand, machine learning (ML) approaches have been used by a number of banks and financial institutions that are using UPI (Unified Payments Interface) fraud detection systems in order to improve security and stop fraudulent activity. Organizations are improving their UPI fraud detection systems as technology and security procedures change.

Problem Identified

The analysis and justification of the problem that was found for the project "Fraud Fighters: How AI and ML are Revolutionizing UPI Security" are summarized as follows:

Identification of the Problem: Due to the prevalence of its use in India, Payments Interface has emerged as a prominent target for security concerns. Transaction fraud, unauthorized access, identity theft, and data breaches are just a few of the security issues that UPI must deal with. These flaws have the potential to cause monetary losses and erode user confidence in the UPI system that must be resolved.

Security Repercussions: In the UPI ecosystem, security lapses can have disastrous effects, including monetary losses for institutions and consumers, a decline in confidence, and damage to UPI's brand. For security to be maintained, regulatory framework compliance is essential.

The Role of AI and ML: AI and ML technologies are increasingly vital in addressing these security challenges. They offer real-time monitoring, predictive analysis, and adaptive security measures to detect patterns indicative of fraud, adapt to evolving threats, and enhance user authentication processes.

IV. Proposed System

With the use of AI and ML algorithms, the proposed system can swiftly and reliably analyze enormous volumes of data, lowering false positives and raising fraud detection rates. Over time, AI and ML systems can become cost-effective, as they reduce the need for manual intervention and investigation. AI and ML systems can scale to handle increasing transaction volumes, ensuring that they remain effective as the UPI platform grows. The proposed method may constantly acquire knowledge from new information, enhancing its fraud detection skills over time.

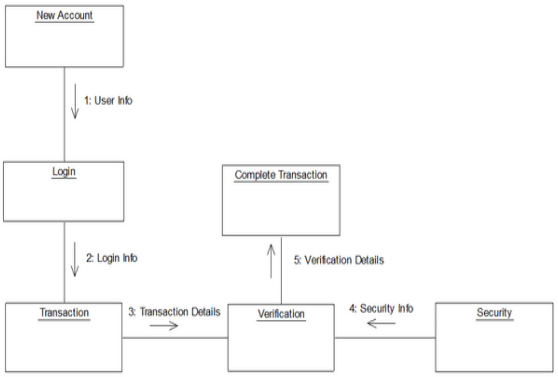

Fig 1. Architecture Diagram

Fig1 depicts the suggested system architecture for fraud detection, which includes an initial account, logging in, transactions, authentication, security, and completion of transactions. A new account is created for a user with a username and password, and the user logs in after verifying their details. After successful login, security information is checked for authentication, and details are verified. Verification is necessary for figuring out whether the transaction is legitimate or counterfeit. If fraud is identified, the financial transaction will proceed. Transaction completion is the final stage.

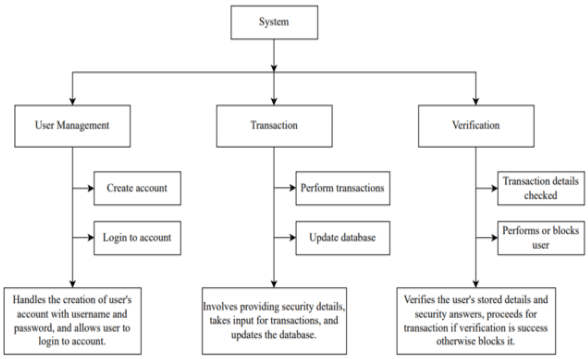

Fig 2. Module Diagram

User Management Module

User management plays a significant role in UPI (Unified Payments Interface) fraud detection by implementing

various security measures and controls to mitigate potentially fraudulent activities as represented in fig2. effective user management practices within UPI services involve a combination of security measures, behavioral analysis, real-time monitoring, risk assessment, and user education to detect, prevent, and mitigate fraud effectively.

Transaction Module

The Transaction module consists of providing security details, takes input for transactions, and updates the database. UPI users are each given a communications device as part of the pre-authorization technique and apparatus. A UPI transaction is initiated by the UPI owner contacting a UPI number and recording a unique piece of information that defines an activity that will be carried out by an authorized UPI user at a later time. Only when an accurate unique identification code (PIC) is applied during the transmission is the information recognized as internet network data in the database. Subsequently, internet network data will be utilized to authorize the particular purchase. Only UPI owner or another approved user will thereafter be able to use the UPI to execute the transaction.

Verification Module

Unified Payments Interface (UPI) fraud detection often involves various security measures, and a verification module is a crucial component in this context. UPI transactions involve digital payments that can be susceptible to fraudulent activities. A robust verification module for UPI transactions incorporates multiple layers of security measures, authentication methods, and fraud detection techniques to enhance the security of digital payments and prevent fraudulent activities.

Proposed Algorithm

The proposed system introduces UPI fraud detection using Convolutional Neural Network (CNN). It is detected based on the cardholder's spending history. The FDS (Fraud Detection System) in the bank monitors the cardholder's normal expenditure. The technology monitors the user's whole spending history. When the situation becomes odd, the technique stops the transaction on the card and notifies the bank. It happens without your intervention. There is no need for any personnel. The fig3 illustrates the proposed system model steps.

UPI Fraud Detection System uses the CNN algorithm and consists of the following working procedure:

- Data Collection: Labeled dataset of UPI transactions, including both genuine and fraudulent transaction details.
- Data Preprocessing: Transform raw data into a format suitable for training a model, and ensuring that it can effectively learn from the given dataset.
- Feature Extraction: Involves extraction of features like transaction amount, user details, and transaction history.
- CNN Model: Tailoring the model for fraud detection using training and testing data.
- Transaction Verification: Evaluation of the transaction where it is legitimate or fraudulent

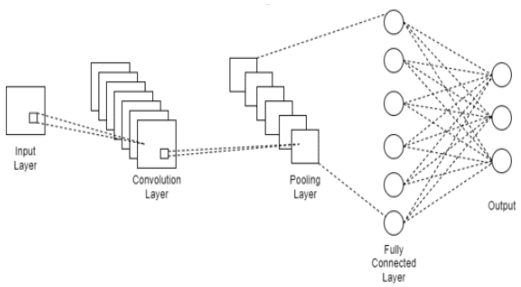

Fig 3. Working of CNN

The fig 3 illustrates the architecture of the Convolutional

V. Results And Discussion

Dataset

The necessary dataset, which includes transactional data, user data, authentication data, and project-specific models, has been gathered in order to train the model. The dataset was designed specifically for this model with sample data for testing and training. The dataset used is not taken from any sources since the obtaining of genuine data from the bank is not mandatory and is made for sample purposes only. The dataset contains numerous attributes, including identification numbers, payment hour, payment day, payment month, payment year, class, UPI number as well, age, payment amount, the state, area code, and possibility of fraud value.

Table 1. Sample values of the dataset

Id	trans_hour	trans_day	trans_month	trans_year	category
0	0	1	1	2022	12
1	1	1	1	2022	3
2	3	1	1	2022	8
3	6	1	1	2022	4
4	6	1	1	2022	0
5	8	1	1	2022	3
6	13	1	1	2022	11
7	19	1	1	2022	10
8	19	1	1	2022	10

upi_number	age	trans_amount	state	zip	fraud_risk
9957000001	54	66.21	22	49879	0
9957000002	15	55.81	14	62668	0
9957000003	60	8.68	4	96037	0
9957000004	44	89.52	40	29911	0
9957000005	72	1.9	38	16421	0
9957000006	24	61.74	15	46765	0
9957000007	41	23.25	18	70808	0
9957000008	75	81.94	35	45860	0
8753000004	48	71.86	49	24927	0

Table 1 illustrates the information that was utilized for training the model, which includes identity numbers, payment hour, payment day, payment month, payment year, class, UPI number, age, payment amount, state, area code, and the likelihood of fraud value.

Working

UPI Fraud Detection System uses the CNN algorithm and consists of the following working procedure:

Data Collection: Labeled dataset of UPI transactions, including both genuine and fraudulent transaction details. Data Preprocessing: Transform raw data into a format suitable for training a model, and ensuring that it can effectively learn from the given dataset. Feature Extraction: Involves extraction of features like transaction amount, user details, and transaction history. CNN Model: Tailoring the model for fraud detection using training and testing data. Transaction Verification: Evaluation of the transaction where it is legitimate or fraudulent transaction and blocks it if it is fraudulent.

Fig 4. Input for Valid Transaction

In fig4, the input given is the data that coincides with the valid transaction in the dataset given to the model.

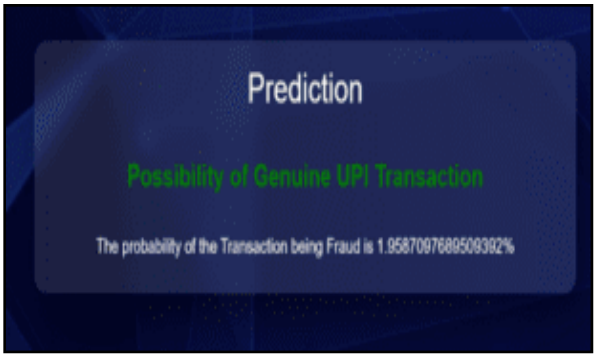

Fig 5. Results from valid test data

The fig 5 represents the verification that the transaction is a valid one.

Fig 6. Input for Fraud Transaction

In fig 6, the given data input is the one that matches with the fraudulent data.

Table 2. Table of several algorithm metric values

Algorithm	Accuracy	Precision	Recall	F1- score
Convolutional Neural Network	97.68%	0.96	96.2%	0.97
Random Forest	95.50%	0.93	86.7%	0.30
Decision Tree	94.75%	0.70	93.6%	0.92
K- Nearest Neighbors	83.00%	0.79	46.2%	0.88
Support Vector Machine (SVM)	81.50%	0.89	89.7%	0.27
Naïve Bayes	81.25%	0.16	11.5%	0.86
Logistic Regression	80.25%	0.76	53.8%	0.94

The table 2 represents the table that includes several algorithms accuracy, precision, recall, f1-score respectively. The table shows seven algorithms: Convolutional Neural Network, Random Forest, the Decision Tree, KNN, Support Vector Machine, Naïve Bayes, as well as Logistic Regression. The CNN algorithm used in the model for UPI fraud detection gave the highest accuracy amongst all the algorithms.

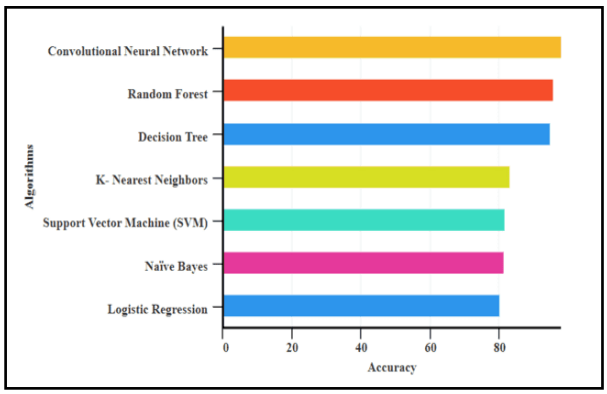

Fig 7. Graph with accuracy of algorithms

The fig 7 illustrates the graph with various algorithms accuracy. The CNN model represented the highest accuracy with 97.68%

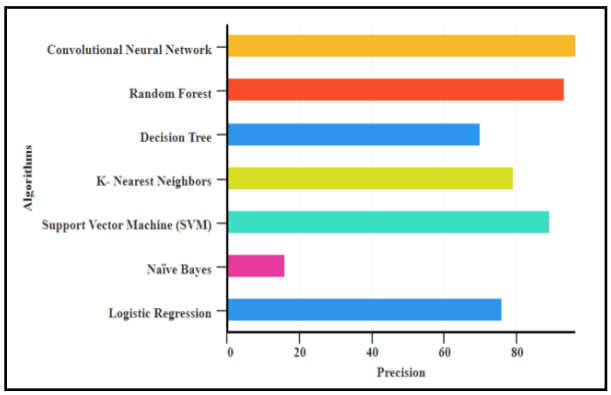

Fig 8. Graph with precision of algorithms

The fig 8 illustrates the graph with various algorithms with precision values. Precision is a criterion that was utilized for evaluating the performance of the model, and it measures preciseness of true positives provided by the CNN model for UPI fraud detection.

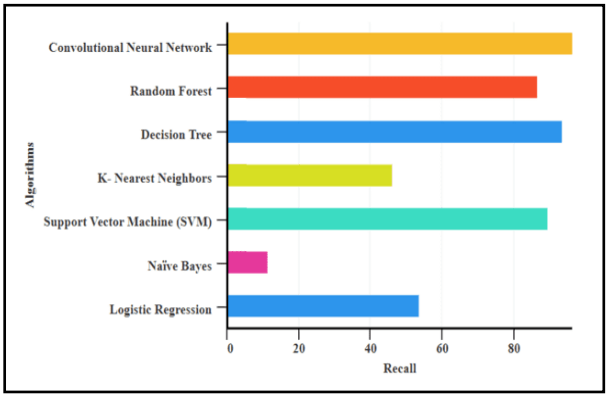

Fig 9. Graph with recall of algorithms

The fig 9 illustrates the graph with various algorithms with precision values. Recall is another parameter used to assess the effectiveness of the CNN and other fraud detection algorithms. Recall is the measure that specifies the completeness of the positives predictions made by the CNN model and other models.

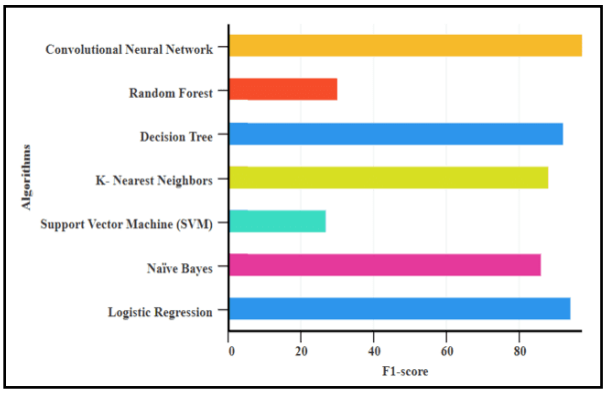

Fig 10. Graph with F1-score of algorithms

The fig 10 illustrates the graph with various algorithms with F-measure values. The model presented the exceptional F-measure value of 0.97%

VI. Conclusion

The present paper proposed a technique for detecting UPI transaction fraud using machine learning techniques. The system, which combines machine learning and artificial intelligence algorithms, can detect fraudulent UPI transactions with high accuracy and few false positives. The paper provides a large dataset of transactions as input, offering valuable insights from previous studies on online transaction fraud detection. Future research may use machine learning algorithms to detect fraudulent transactions by varying input and output considerations. This approach can help address common fraud detection and prevention issues.

References

[1] Lakshmi, K. Krithiga; Gupta, Himanshu; Ranjan, Jayanthi, "UPI Based Mobile Banking Applications – Security Analysis and Enhancements",2019 Amity International Conference on Artificial Intelligence (AICAI),pp1–6, doi:10.1109/AICAI.2019.8701396 , 2019.

[2] S. Mason and N. Bohm, "Banking and fraud," Comput. Law Secur. Rev., vol. 33, no. 2, pp. 237–241, Apr. 2017.

[3] Roshna Thomas, Dr. Abhijeet Chatterjee, "Unified Payment Interface (UPI): A Catalyst Tool Supporting Digitalization – Utility, Prospects & Issues", International Journal of Innovative Research and Advanced Studies, vol.4, pp. 192-195, February. 2017.

[4] Kanika, Singla, Jimmy, A Survey of Deep Learning based Online Transactions Fraud Detection Systems, 2020 International Conference on Intelligent Engineering and Management (ICIEM), pp 130–136. doi:10.1109/ICIEM48762.2020.9160200, 2020.

[5] Rahul Gochhwal, "Unified Payment Interface—An Advancement in Payment Systems", American Journal of Industrial and Business Management, vol.7, pp. 1174-1191, Oct. 2017.

[6] A. Kiran, C. Vimalarani, L. Ashwini, G. Gayithri, J. Supriya and T. Vinod, "Secure Reversible Image Data Hiding (SRIDH) Using LSB Prediction Method," 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-4, doi: 10.1109/ICCCI56745.2023.10128232.

[7] A. Kiran, S. W. Prakash, B. A. Kumar, Likhitha, T. Sameeratmaja and U. S. S. R. Charan, "Intrusion Detection System Using Machine Learning," 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-4, doi: 10.1109/ICCCI56745.2023.10128363.

[8] A. Kiran, G. R. Sakthidharan, D. D. Priya, B. U. Mahesh, K. P. Kumar and K. P. Kumar, "Voice Controlled Home Automation System using Google Assistants," 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-5, doi: 10.1109/ICCCI56745.2023.10128217.

[9] Allam Balaram, T. Sakthivel, Radha Raman Chandan, "A context-aware improved POR protocol for Delay Tolerant networks," Automatika, 2023, Vol. 64, Issue no.1, pp. 22–33.DOI: 10.1080/00051144.2022.2095830.

[10] Allam Balaram, Manda Silparaj, Rajender Gajula, " Detection of malaria parasite in thick blood smears using deep learning," Materials Today: Proceedings, Vol, 64, Part 1, pp. 511–516, 2022, DOI: 10.1016/j.matpr.2022.04.1012.

[11] Allam Balaram, Shaik Abdul Nab" A hybrid approach of segmentation in brain tumor data using fuzzy c-means and level set," Journal of Advanced Research in Dynamical and Control Systems, Vol. 12, Issue No. 6, pp. 47–56, 2020. DOI: 10.5373/JARDCS/V12I6/S20201006.

[12] Nilay YILDIRIM and Asaf VAROL "A Study on Security Vulnerabilities in Online and Mobile Banking Systems,", IEEE 2019.

[13] Murad Obaid, Mahmoud Obaid, and Musbah Aqel, "Blockchain Security for Mobile Payments," 2021.

[14] Raphael Olufemi Akinyedea and Odoseiye Aidohelen "Development of a Secure Mobile E-Banking System," Esese, IJCRT 2019.

[15] S. Sanjana, V. R. Shriya, Gururaj Vaishnavi and K. Ashwini, "A review on various methodologies used for vehicle classification helmet detection and number plate recognition", Evolutionary Intelligence, vol. 14, no. 2, pp. 979-987, 2021.

[16] M. Young, The Technical Writer's Handbook. Mill Valley, CA:

University Science, 1989.
[17] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
[18] O. Adepoju, J. Wosowei, S. Lawte and H. Jaiman, "Comparative evaluation of credit card fraud detection using machine learning techniques", 2019 Global Conference for Advancement in Technology (GCAT), pp. 1-6, 2019.
[19] V. Jain, M. Agrawal and A. Kumar, "Performance analysis of machine learning algorithms in credit cards fraud detection", 2020 8th International Conference on Reliability Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), pp. 86-88, 2020.
[20] Noor Saleh Alfaiz and Suliman Mohamed Fati, "Enhanced Credit Card Fraud Detection Model Using Machine Learning", Electronics, vol. 11, no. 4, pp. 662, 2022.
[21] Z. Kazemi and H. Zarrabi, “Using deep networks for fraud detection in the credit card transactions,” 2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), 2017.
[22] A. Pumsirirat and L. Yan, “Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine,” International Journal of Advanced Computer Science and Applications, vol. 9, no. 1, 2018.
[23] Y. Heryadi and H. L. H. S. Warnars, “Learning temporal representation of transaction amount for fraudulent transaction recognition using CNN, Stacked LSTM, and CNN-LSTM,” 2017 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), 2017.
[24] Gupta, Himanshu, Sharma, Vinod Kumar, “Role of Multiple Encryption in Secure Electronic Transaction”, International Journal of Network Security & its Applications, vol. 3, pp. 89-96, Nov. 2011.