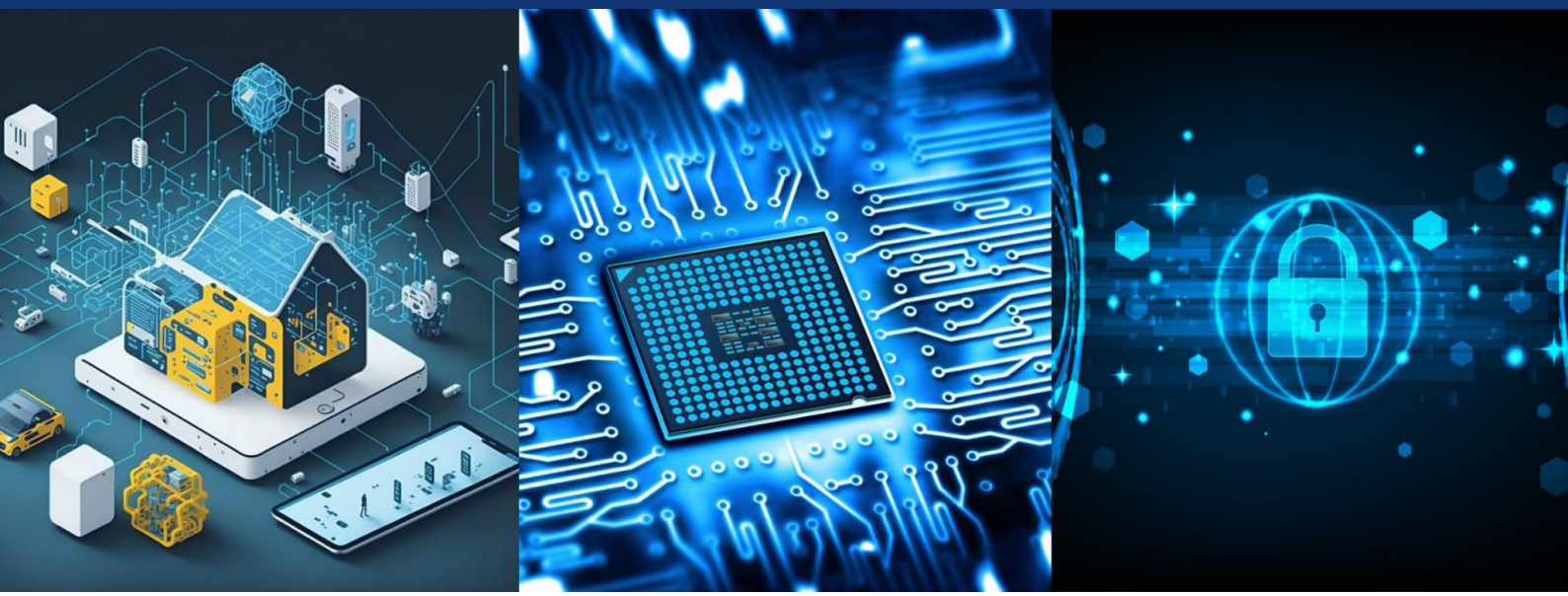
IJIRCCE

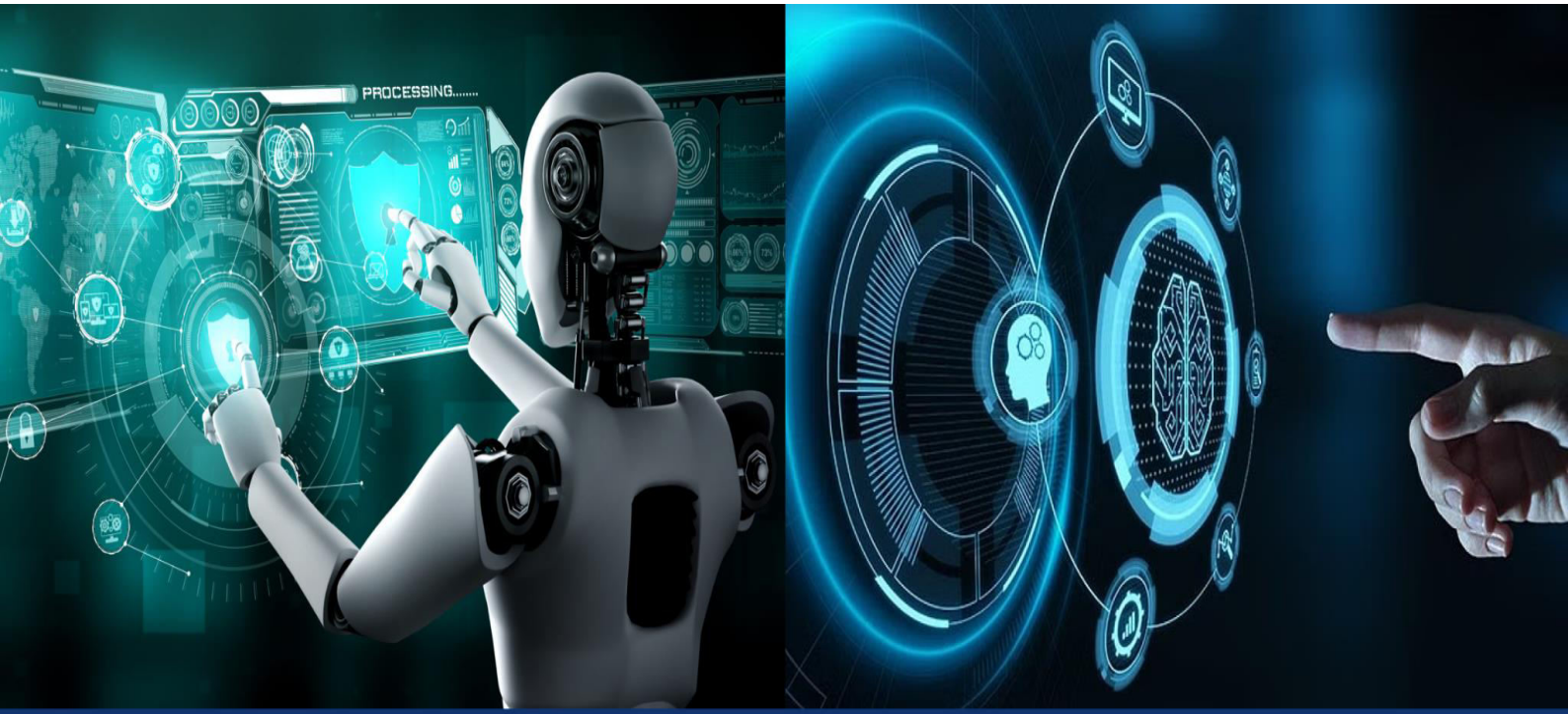
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 13, Issue 3, March 2025

© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

UPI Fraud Detection using Machine Learning

Priyanka S N, Shaharin, Namrata, Akshith, Dr T Y Satheesha

Department of Computer Science and Engineering, Reva University, Bangalore, India

ABSTRACT: The rise of digital payment systems like Unified Payments Interface (UPI) has revolutionized financial transactions, making them seamless and efficient. However, this increased convenience has also led to a surge in fraudulent activities. This project focuses on the detection of UPI fraud using Machine Learning (ML) techniques. By leveraging real- time transaction data, advanced algorithms are designed to identify and flag suspicious activities. The study evaluates various ML models, compares their performance, and highlights the importance of feature selection in fraud detection. The proposed system aims to enhance security, reduce financial losses, and improve user trust in UPI systems. achine learning- based UPI fraud detection involves analyzing transaction data to identify anomalies and patterns indicative of fraudulent behavior. Key steps include data preprocessing, feature extraction, and the application of algorithms such as Random Forest, Gradient Boosting, and Neural Networks for supervised learning, as well as clustering methods for unsupervised anomaly detection. These models are trained on historical data to detect deviations from normal transaction behavior in real time, thus ensuring security and reliability for users.

This paper discusses the importance of implementing ML for fraud detection in the UPI ecosystem. It highlights the need for dynamic, adaptive systems capable of evolving with new fraud techniques while minimizing false positives. By leveraging ML, the financial sector can enhance transaction security, build user trust, and contribute to a resilient and fraud-resistant digital payment landscape.

KEYWORDS: artificial intelligence; fraudulent banking operations; machine learning; recognition of fraudulent operations

I. INTRODUCTION

The Unified Payments Interface (UPI) is a real-time payment system facilitating inter-bank transactions in India. With its widespread adoption, fraudulent activities targeting UPI users have increased exponentially. Existing rule-based fraud detection systems are limited in their ability to adapt to evolving fraud patterns. Machine Learning offers a dynamic and scalable solution, capable of learning and predicting fraudulent behaviors based on historical transaction data. This report investigates the application of ML techniques to address UPI fraud challenges effectively.

In recent years, digital payment systems have witnessed an exponential surge in adoption, revolutionizing the way financial transactions are conducted. Among these systems, the Unified Payments Interface (UPI) has emerged as a game- changer in the Indian financial ecosystem. Launched by the National Payments Corporation of India (NPCI) in 2016, UPI facilitates real-time, interbank transactions seamlessly through mobile devices. Its simplicity, speed, and convenience have made it the preferred choice for millions of users, driving its rapid expansion across urban and rural landscapes alike.

However, with the rise in UPI transactions comes an inevitable challenge—fraud. Cybercriminals exploit vulnerabilities in the system and users' lack of awareness to conduct phishing attacks, social engineering scams, transaction spoofing, and unauthorized access. According to reports from the Reserve Bank of India (RBI), the financial losses incurred due to UPI fraud have grown alarmingly, creating an urgent need for robust fraud detection mechanisms. Traditional fraud detection systems, primarily rule-based, are becoming increasingly inadequate. These systems struggle to adapt to the sophisticated and dynamic nature of modern fraud techniques, often resulting in false positives and missed detections. Furthermore, as the volume of transactions grows exponentially, the need for automated, real-time solutions has become critical.

Machine Learning (ML) offers a promising approach to tackle these challenges. Unlike static rule-based systems, ML models are dynamic and capable of learning from historical data to predict fraudulent patterns. They can analyze

© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

large volumes of transaction data, identify anomalies, and adapt to new fraud techniques over time. Additionally, ML algorithms can significantly reduce false positives, ensuring genuine transactions are not disrupted while suspicious ones are flagged. Unified Payments Interface (UPI) has revolutionized the digital payment landscape, enabling seamless, real-time transactions across India. With its growing adoption, UPI has also become a target for fraudsters, leading to financial losses and undermining user trust. To address this challenge, machine learning (ML) has emerged as a powerful tool for fraud detection, offering innovative and efficient solutions.

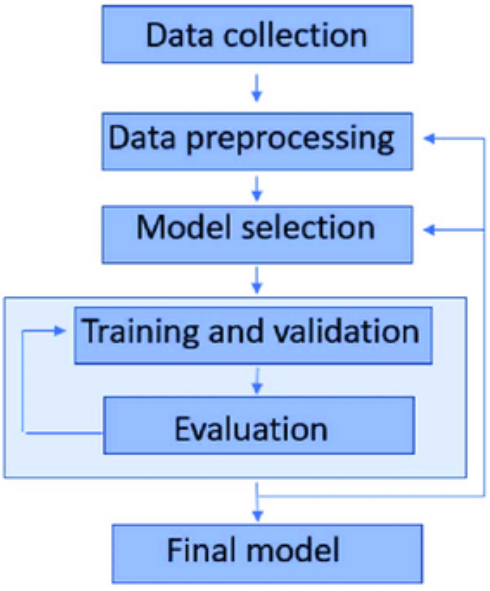
The dynamic nature of fraud tactics, such as phishing, fake payment requests, and account takeovers, makes it difficult to detect fraudulent activities using traditional rule-based methods. These approaches often fail to adapt to evolving patterns and may generate a high rate of false positives, inconveniencing genuine users. Furthermore, the massive volume of UPI transactions necessitates a scalable and robust fraud detection system.

Objectives: The objective of UPI fraud detection is to enhance the security and reliability of digital payment systems by identifying and mitigating unauthorized and fraudulent activities in real-time. With the exponential growth of Unified Payments Interface (UPI) as a fast and seamless payment mechanism, there is a parallel rise in sophisticated cyber threats, including phishing, identity theft, and transaction tampering. By employing advanced techniques such as machine learning, anomaly detection, and behavioral analysis, UPI fraud detection aims to safeguard user transactions, maintain trust in the payment ecosystem, and comply with regulatory standards. These systems focus on analyzing transaction patterns, identifying unusual activities, and promptly flagging or blocking potential threats, ensuring a secure and uninterrupted user experience.

II. METHODOLOGY

The methodology for UPI fraud detection typically involves leveraging advanced data analytics and machine learning to identify and prevent fraudulent activities in digital payment systems. Initially, the system collects and preprocesses transaction data, which includes user details, transaction amounts, timestamps, and geolocation. This data is then analyzed to extract features indicative of normal and abnormal transaction patterns.

Machine learning models such as Random Forest, Support Vector Machines (SVM), or neural networks are trained using historical data labeled with legitimate and fraudulent transactions. In real-time detection systems, algorithms like anomaly detection or clustering are used to identify deviations from normal behavior. Additionally, behavioral analysis tools monitor user interaction patterns to detect phishing attempts or account takeovers.



© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Data Collection: Gather large datasets of historical transaction data, including both legitimate and fraudulent transactions. This includes features like transaction amount, time, location, user behavior, and device details.
Feature Engineering: Extract relevant features from the raw data that can help in identifying fraudulent patterns, such as transaction frequency, unusual activity, or abnormal behavior compared to a user's typical transaction history.
Model Training: Use supervised learning algorithms, such as decision trees, random forests, or neural networks, to train models on labeled data (fraudulent vs. non-fraudulent transactions). Unsupervised learning methods, like clustering, can also be used to detect anomalies.

Anomaly Detection: Apply ML models to identify unusual or suspicious patterns in real-time transactions. This can include high transaction amounts, sudden location changes, or rapid, repeated transactions from a single account.
Risk Scoring: The model assigns a risk score to each transaction based on its likelihood of being fraudulent. High- risk transactions are flagged for further review or automatic blocking.
Model Evaluation and Optimization: Continuously evaluate the model's performance using metrics like precision, recall, F1-score, and accuracy. The model is updated and retrained to adapt to new fraud patterns.

III. LITERATURE SURVEY

In fraud detection, we often deal with highly imbalanced datasets. For the chosen dataset (Paysim), we show that our proposed approaches are able to detect fraud transactions with very high accuracy and low false positives – especially for TRANSFER transactions. Fraud detection often involves a tradeoff between correctly detecting fraudulent samples and not misclassifying many non-fraud samples. This is Often a design choice/business decision which every digital payment company needs to make. We've dealt with this problem by proposing our class weight-based approach. We can further improve our techniques by using algorithms like Decision trees to leverage categorical features associated with accounts/users in Paysim dataset. can be Very effective in improving our classification quality on this dataset [1].

Now a days Digital transactions are rapidly increasing as it results in increasing online Payment frauds too. In fact, according to the Reserve Bank of India, comparing March 2022 to March 2019, digital payments have risen in volume and value by 216% and 10%, respectively. People are starting to go all-in with digital transactions, but one can't deny the security issues that loom, and know-how when it comes to online payments. Few years ago, we could have barely seen the online payment, but today UPI payment QR code installed at doorstep.

This invited the hoaxers and attackers to develop fraudulent transactions and fool people for some amount of money. Fortunately, the online transactions are monitored and hence could be analyses using the latest tools. In this system, an attempt is made to develop a machine learning model to identify fraudulent transactions in a transaction's dataset. [2]

Fraud detection for credit/debit card, loan defaulters and similar types is achievable with the assistance of Machine Learning (ML) algorithms as they are well capable of learning from previous fraud trends or historical data and spot them in current or future transactions. Fraudulent cases are scant in the comparison of non-fraudulent observations, almost in all the datasets. In such cases detecting fraudulent transaction are quite dif?cult. The most effective way to pre-vent loan default is to identify non-performing loans as soon as possible. recall and F-1 score along with Receiver Operating Characteristic (ROC) curves [3]

Financial fraud, considered as deceptive tactics for gaining financial benefits, has recently become a widespread menace in companies and organizations. Conventional techniques such as manual verifications and inspections are imprecise, costly, and time consuming for identifying such fraudulent activities. With the advent of artificial intelligence, machine- learning-based approaches can be used intelligently to detect fraudulent transactions by analyzing a large number of financial data. Therefore, this paper attempts to present a systematic literature review (SLR) that systematically reviews and synthesizes the existing literature on machine learning (ML)-based fraud detection. Particularly, the review employed the Kitchenhand approach, which uses well- defined protocols to extract and synthesize the relevant articles; it then report the obtained results. Based on the specified search strategies from popular electronic database libraries, several studies have been gathered. After inclusion/exclusion criteria, 93 articles were chosen, synthesized, and analyzed. The review summarizes popular ML techniques used for fraud detection, the most popular fraud type, and evaluation metrics. The reviewed articles showed that support vector machine (SVM) and artificial neural network (ANN) are popular ML algorithms used for fraud detection, and credit card fraud is the most

© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

popular fraud type addressed using ML techniques. The paper finally presents main issues, gaps, and limitations in financial fraud detection areas and suggests possible areas for future research. [4]

III. SYSTEM DIAGRAM

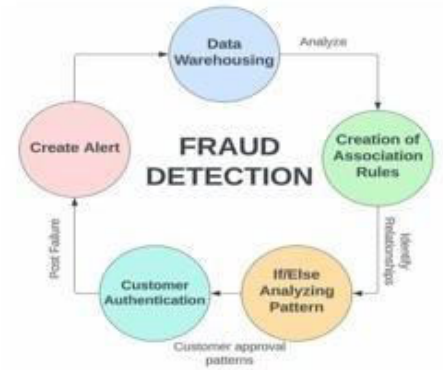


Fig: System Diagram for UPI fraud Detection using machine learning

Challenges in ML Based UPI Fraud Detection:

While ML based fraud detection offers numerous advantages, it also faces several challenges. One of the primary challenges is the complexity of fraud patterns. Fraudsters are constantly adapting and finding new ways to exploit the system, making it difficult for static models to detect novel fraud tactics. This requires the fraud detection system to be highly adaptable and continuously updated with new data.

Another challenge is the potential for false positives, where legitimate transactions are mistakenly flagged as fraudulent. This can cause inconvenience to users and lead to a loss of confidence in the system. Striking the right balance between detecting fraud and minimizing false positives is a critical aspect of any ML-based fraud detection system.

Additionally, ensuring the privacy and security of customer data is paramount. Since ML models require access to large amounts of transaction data, there must be strict protocols in place to protect sensitive information and comply with data protection regulations.

IV. CONCLUSION

The conclusion of UPI fraud detection using advanced methods like machine learning emphasizes the critical importance of leveraging predictive analytics and real-time monitoring to safeguard digital payment systems. With the increasing adoption of UPI (Unified Payments Interface) for seamless transactions, the risk of fraud has grown, necessitating robust mechanisms to detect and prevent unauthorized activities. Machine learning algorithms, through techniques such as anomaly detection, supervised learning, and clustering, can identify suspicious patterns and flag fraudulent transactions effectively.

ACKNOWLEDGEMENT

We acknowledge the collaborative efforts of researchers, financial institutions, and technology developers in advancing the field of UPI fraud detection. The development of secure and efficient detection mechanisms relies on the integration of domain expertise, cutting-edge technologies, and user-centric approaches. Contributions from data science communities, regulatory bodies, and open-access platforms have been instrumental in providing insights into fraud trends, innovative machine learning techniques, and practical applications. Special thanks to the organizations providing datasets, frameworks, and tools that enable comprehensive analysis and the creation of robust fraud prevention systems. This acknowledgment also extends to users and testers whose feedback ensures the reliability and

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

adaptability of fraud detection systems in real- world scenarios. The information provided on UPI fraud detection using Machine Learning methodologies draws from established techniques and practices in the field of fraud detection and machine learning. This includes data collection, feature engineering, model training, anomaly detection, risk scoring, and continuous optimization to ensure an effective fraud prevention system. The aim is to help create a secure and reliable UPI ecosystem, using state-of-the-art technologies to detect and prevent fraudulent activities in real time.

Implementing such systems ensures better customer trust, compliance with regulatory standards, and the overall integrity of the payment ecosystem. Continuous adaptation to emerging fraud tactics and improving dataset quality remain essential for maintaining efficiency and accuracy in fraud detection. The rise of digital payment systems has transformed the way financial transactions are conducted, with the Unified Payments Interface (UPI) emerging as one of the most popular and widely used platforms in India. However, with the growing adoption of UPI for various financial activities, there is an increasing risk of fraudulent activities that can cause significant financial losses to users and institutions. Fraudulent transactions not only undermine the trust of consumers in digital payment systems but can also affect the broader financial ecosystem.

Machine learning (ML) has become a powerful tool in combating UPI fraud by enabling real-time, adaptive, and intelligent detection systems. By leveraging vast amounts of historical transaction data and sophisticated algorithms, ML- based fraud detection systems can identify suspicious patterns and prevent fraudulent transactions more effectively than traditional methods. The application of ML in UPI fraud detection can significantly enhance security, improve user confidence, and reduce financial losses.

REFERENCES

[1] A. Oza, "Fraud Detection using Machine Learning," GitHub, 2023.
[2] M. Valavan, S. Rita, "Predictive-Analysis-based Fraud Detection," IJRASET, 2023.
[3] S. H. Projects, "Machine Learning for Fraud Detection," IEEE, 2022.
[4] N. Kumari, "UPI Fraud Detection Using Genetic Algorithms,"
[5] Middle East Journal of Scientific Research, 2024.J. Han etal., "Data Mining: Concepts and Techniques,"
[6] M. Adekunle and P. Ozoh, "Fraud detection model for illegitimate transactions", Kabale University Interdisciplinary Research Journal, vol. 2, no. 2, pp. 21-37, 2023.
[7] B. Mytnyk, O. Tkachyk, N. Shakhovska, S. Fedushko and Y. Syerov, "Application of Artificial Intelligence for Fraudulent Banking Operations Recognition", Big Data and Cognitive-Computing, vol. 7, no. 2, pp. 93, 2023.
[8] V. Chang, A. Di Stefano, Z. Sun and G. Fortino, "Digital payment fraud detection methods in digital ages and Industry 4.0", Computers and Electrical Engineering, vol. 100, pp. 107734, 2022.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details