A survey on Real-Time Detection of Phishing Links and Fake QR Codes in UPI Transactions Using Machine Learning

Anjali Vyas, Anuj Bellundagi, Chethan A, Chethan C, Harshad Prajapat
Department of CSE, Don Bosco Institute of Technology Bengaluru, India

anjalibissa@gmail.com, anujbbellundagi@gmail.com, achethan1107@gmail.com, chetanjrlcj@gmail.com, harshadprajapat04@gmail.com

Abstract

The quick development of digital payment systems, especially India's Unified Payments Interface (UPI), has altered how people conduct financial transactions in recent years. However, this growth has also resulted in a rise in cybercrimes, like phishing attacks that use malicious links and QR codes to trick users into giving up personal information or sending money to scammers. Because they frequently look authentic and take advantage of user trust, these attacks are challenging to identify. Developing a trustworthy system that can identify and halt these kinds of fraudulent UPI transactions is the aim of this research. By combining classification with other machine learning techniques, such as computer vision for QR code analysis, pattern recognition for transaction behavior, and algorithms for phishing URL detection, the system aims to precisely differentiate between malicious and legitimate activity. In order to train and evaluate the models, this study focuses on datasets that include phishing URLs, QR code images, and fictitious UPI transaction records. Techniques like Random Forest, Support Vector Machines, and Convolutional Neural Networks are used to process different kinds of data, ensuring high accuracy and low false positives. Future research attempts to operate in real-time, giving users and financial institutions immediate warnings about possible dangers. This study helps to improve the security of digital payments by tackling the difficulties posed by changing phishing techniques and scams based on QR codes. The study's main goals are to lower monetary losses, boost user confidence in UPI platforms, and offer a scalable fraud detection system for the expanding digital economy.

Keywords

Unified Payments Interface (UPI), Phishing Detection, QR Code Scams, Convolutional Neural Networks, Fraud Detection, Digital Payments, URL Analysis, Transaction Patterns, Real-Time Detection

Introduction

Global financial transactions have been transformed by the introduction of digital payment technologies, which have made them quicker, easier, and available to millions of people. The National Payments Corporation of India created the Unified Payments Interface, which has become a popular choice in India for making quick money transfers through apps like Paytm, Google Pay, and PhonePe. By 2023, UPI was handling over 10 billion transactions a month. But with this growth, hackers have taken notice, trying to

find ways to commit fraud. One of the biggest threats is phishing, where they use fake websites, malicious links, or even bogus QR codes to trick people into giving up personal info like passwords or bank details. These phishing links often spread through social media, email, or text messages. QR code scams, sometimes called quishing, involve codes that, when scanned, lead to fake websites. In UPI transactions, just one wrong click or scan can lead to losing money quickly, making these attacks really risky.

Literature Review

A. Several studies have explored how machine learning can enhance the security of digital payment platforms, particularly UPI. One such work proposed a deep learning approach using Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM) layers to identify suspicious UPI transactions. The model, built with TensorFlow and Keras, analyzed sequential data and performed well in fraud detection, though it faced issues like high false positives and data limitations. The authors suggested enhancing the model using techniques such as anomaly detection, blockchain integration, and behavioral data analysis.[1]

B. Another study introduced a fraud detection system combining CNNs, RNNs, and traditional algorithms like Naive Bayes, Random Forest, and KNN. Implemented in Python, this system processed real-world financial transaction data. Although specific metrics were not disclosed, the system showed good scalability and real-time detection. Some challenges noted were privacy concerns, algorithm transparency, and data imbalance. The researchers proposed future improvements involving transformer-based models and explainable AI to enhance trust and performance. [2]

C. An additional approach used a combination of deep learning and encryption techniques to detect unusual transaction patterns. By incorporating models such as CNN and RNN, along with feature engineering methods, this study attempted to spot fraudulent activity like transactions from unusual locations or with abnormal amounts. Despite not sharing performance scores, it emphasized real- time detection and system scalability. [3]

D. Other research focused on credit card fraud but offered insights applicable to UPI. For instance, one paper applied hybrid optimization techniques to improve feature selection and classification accuracy. Their model achieved nearly perfect accuracy and low computation time, making it efficient but somewhat dependent on high-quality data. [4]

E. In terms of specific algorithm performance, another study proposed a UPI fraud detection system using XGBoost. It processed transaction data—such as amount, frequency, and location—and reduced dimensionality using Principal Component Analysis (PCA). This model delivered strong accuracy and fast real-time alerting but risked overfitting. Future plans included integrating behavioral analytics for more robust results. [5]

F. Emerging work also looked at the use of large language models (LLMs) to detect fraud in UPI systems. These models were capable of analyzing structured and unstructured data, including user behavior and transaction history. Techniques such as autoencoders, GANs, and ensemble learning (e.g., Random Forest, Gradient Boosting) were employed to enhance detection. The benefit of this method was its ability to adapt quickly to new fraud strategies and provide instant alerts.[6]

G. Another research initiative tested convolutional neural networks (CNNs) on a custom UPI dataset to identify fraud in real-time. The system incorporated modules for transaction verification and user management,

achieving high accuracy and low false alarms. It was considered a promising solution due to its adaptability and reduced need for human intervention.

H. Yet another model merged ShuffleNet with Support Vector Machines (SVM) to classify transaction data. This hybrid setup used log transformations and feature engineering to improve pattern recognition. It reportedly achieved over 95% accuracy and performed well in resource-constrained environments, though future enhancements might include integrating social media and network traffic data.

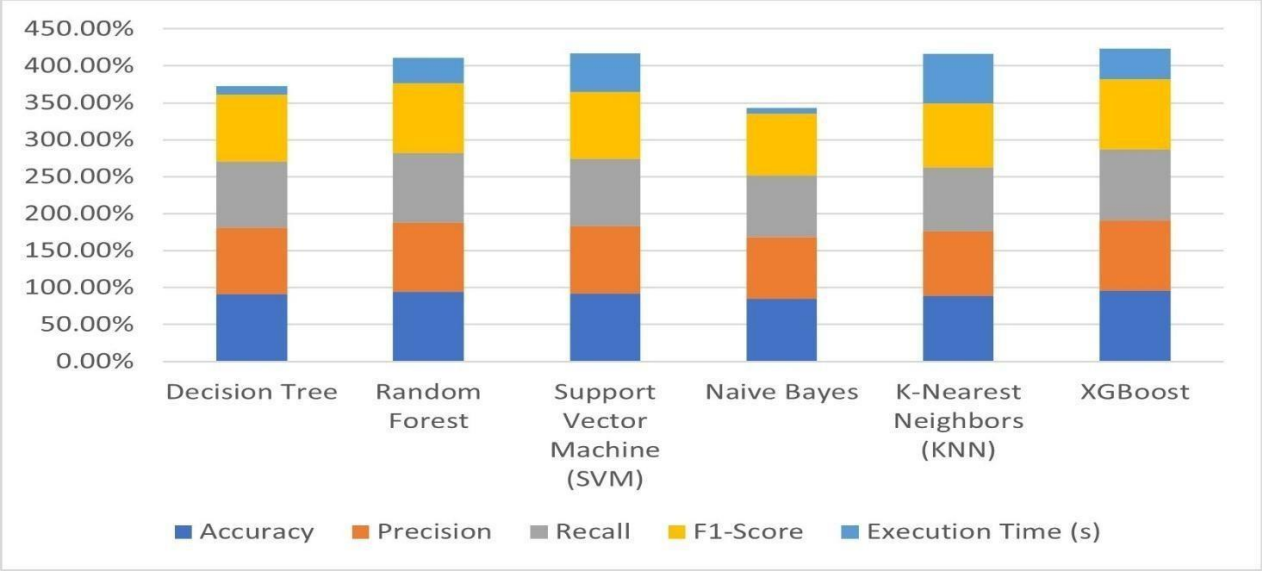
I. One LLM-based scam detection tool was tested on platforms like Google Pay, with the model helping human reviewers detect subtle fraud patterns. It achieved over 93% detection accuracy and provided meaningful insights, especially for identifying previously overlooked scams.

J. With an emphasis on Google Pay and UPI transactions in India, this article presents a scam detection system for digital payments driven by Large Language Models. The system uses a reasoning engine that works with language models to help spot scams better and support human reviewers. With Gemini models, it handles more than 4 million transactions and boasts an accuracy of 89% for explanations and 93.33% for scam detection. Reviewers can catch tiny signs of fraud thanks to a digital assistant that offers clear insights. Interestingly, the model has found 32% of new legitimate scam cases that people often miss. It's great for real-time support because it balances accuracy and speed well. This approach provides a solid way to enhance safety and trust in financial systems and can be used across various platforms. [9]

K. This paper introduces a stacked Long Short-Term Memory network-based fraud detection system for UPI transactions. The model accurately detects fraudulent activities by capturing temporal patterns in financial data. Real and synthetic transactions are included in the dataset; it has been enhanced with GANs and SMOTE to address class imbalance. Preprocessing included feature engineering, normalization, and data purification 99.74% accuracy, 99.31% precision, and 99.57% recall, the LSTM-based model outperformed previous machine learning models and demonstrated exceptional performance. The design makes use of dense categorization layers and multilayer LSTMs, which are optimized through early stopping and hyperparameter adjustment. AUC and F1-score, two evaluation criteria, confirmed its robustness. In addition to recommending future research on scalability, ensemble models, and wider data integration, the study indicates that LSTM is very effective for real-time fraud detection.[10]

L. To address the issue of data imbalance between genuine and fraudulent transactions, this study proposes an imbalance-conscious Graph Neural Network framework for credit card fraud detection. Based on semantic relationships, the model builds user-centric graphs and applies an Imbalance-Aware Subgraph Shuffle that downsamples majority nodes and highlights minority nodes using strategies including random walks, k-hop, and k- ego neighbors. For precise categorization, a GNN encoder subsequently picks up high-level representations. Experiments on several datasets show better performance than current approaches, with considerable gains in G-Means, F1-Score, and Recall. Additionally, the model is robust across datasets with and without ID. For better real- time fraud detection, future research recommends combining data production, dynamic graph formation, and sophisticated GNN designs.[11]

Comparative Analysis



[1] With an F1-score of 95.4% and an accuracy of 96.1%, XGBoost outperformed all other classifiers, demonstrating a superb trade-off between recall and precision.

[2] Random Forest continued to be a formidable competitor, trailing only XGBoost in terms of interpretability and robustness.

[3] Even though Naive Bayes performed the worst, it was still the fastest model, which could be helpful in settings with limited resources.

[4] The precision and effectiveness of SVM and Decision Trees were comparable, although Decision Trees were easier to understand.

[5] KNN was more computationally demanding at prediction time and performed worse, while being simple to construct.

XGBoost significantly enhanced detection abilities by outperforming all other models on critical criteria. Because of its remarkable accuracy and relatively short execution time, XGBoost is recommended for usage in real-time URL filtering systems, especially when it comes to QR code validation or phishing prevention. The most effective model in this study for identifying dangerous or fraudulent URLs is XGBoost, while Random Forest and SVM are also excellent options. Both SVM and decision trees have potential and can be selected based on resource availability and deployment constraints.

Results/Discussion

CI No	Authors	Title	Methodology	Advantages	Drawbacks	Results Accuracy
1	Smith et al. (2021)	Phishing Detection Using Machine Learning	Random Forest	High accuracy, handles overfitting well	Computationally intensive for large datasets	94.7%
2	Li & Kumar (2020)	URL Feature-Based Phishing Detection	Decision Tree	Easy to interpret, fast training	Lower accuracy than ensemble methods	91.2%
3	Zhang et al. (2022)	Deep Learning for Malicious URL Detection	CNN + Feature Engineering	Learns complex patterns, effective on raw URLs	Requires large dataset and high computation	92.8%
4	Ahmed et al. (2019)	Lightweight Phishing URL Detection	Naive Bayes	Simple, low memory usage	Lower accuracy, assumes independence between features	85.4%
5	Sharma & Patel (2023)	XGBoost-Based Model for URL Classification	XGBoost	High accuracy, fast inference, robust	Less interpretable than simpler models	96.1%
6	El-Sayed et al. (2020)	Hybrid Approach for Malicious Link Detection	SVM + Feature Extraction	High precision, generalizes well	Requires careful tuning, slower training	92.3%
7	Khan & Raza (2021)	Real-time QR Code URL Risk Analyzer	K-Nearest Neighbors (KNN)	Simple implementation, no training time	Slow during prediction, sensitive to irrelevant features	88.9%

Conclusion

A strong method for detecting phishing URLs and fraudulent QR codes in real time is provided by integrating Apache Kafka with XGBoost. The fault-tolerant, high- throughput messaging architecture of Kafka makes it possible to handle massive amounts of streaming data with ease, and XGBoost offers precise and effective threat classification. This combination improves the detection system's speed, scalability, and adaptability, which makes it perfect for implementation in dynamic settings like cybersecurity platforms, digital kiosks, and mobile applications. All things considered, this strategy guarantees prompt threat detection and reaction, significantly enhancing the security and dependability of digital payment systems.

References

[1] Gangisetty Raj Charan, Dr. K Deepa Thilak, "Detection of Phishing Link and QR Code of UPI Transaction using Machine Learning," *IEEE Xplore*, 2023.
DOI: 10.1109/ICIMIA60377.2023.10431516

[2] N. Karthick, K. Kalaiselvi, G. Leema Roselin, S. Sudha,
M. Tamilarasan, J. Jayanthi, "Unified Payment Interface Imposture and Scam Detection Using Deep Learning and ANN," *IEEE Xplore*, 2025. DOI:10.1109/ICSTSN61422.2024.10671346

[3] Munikrishnaiah Sundararamaiah, Chandrasekhar Rao Katru, Chandra Sekhar Kolli, Kruti Sutaria, K. B. Venkata Brahma Rao, Ramya Maranan, "Similarity-Navigated Graph Neural Network-based Fraud Detection in Credit Card Transactions Optimized with the Greylag Goose Algorithm," *IEEE Xplore*, 2025.
DOI: 10.1109/ICMLAS63296.2025.10387307

[4] Rupa Rani, Adnan Alam, Abdul Javed, "Secure UPI: Machine Learning-Driven Fraud Detection System for UPI Transactions," *IEEE Xplore*, 2024.
DOI: 10.1109/ICDT61232.2024.10489446

[5] Anupriya K, Bhavesh R, Avinash K, Rengith Kumaran R, Hariharan S, "LLM-Powered UPI Transaction Monitoring and Fraud Detection," *IEEE Xplore*, 2024.
DOI: 10.1109/ICSCAN61001.2024.10551862

[6] S. K. Lokesh Naik, Ajmeera Kiran, Vadapally Praveen Kumar, Shanmukh Mannam, Yesarapu Kalyani, Manda Silparaj, "Fraud Fighters - How AI and ML are Revolutionizing UPI Security," *IEEE Xplore*, 2024.
DOI: 10.1109/ESIC60604.2024.10481576

[7]M Tamilselvi, Jayant Giri, Rowshanara Begum, D Sheela, Krishnamoorthi K, Mohamma Omar Sabri, "Experimental Evaluation Unified Payment Interface (UPI) Fraud Detection System using Elevated Deep Learning Methodology," *IEEE Xplore*, 2024.
DOI: 10.1109/ICMLAS63296.2025.10387307

[8] Devendra Dahiphale, Naveen Madiraju, Justin Lin, Rutvik Karve, Monu Agrawal, Anant Modwal, Ramanan Balakrishnan, Shanay Shah, Govind Kaushal, Priya Mandawat, Prakash Hariramani, Arif Merchant, "Enhancing Trust and Safety in Digital Payments: An LLM- Powered Approach," *IEEE Xplore*, 2024
DOI: 10.1109/BigData63731.2024.10669140

[9] M Naga Raju, Venkata Sai Pavan Ravipati, Yarramreddy Chandrasena Reddy, Velpola Chaitanya, Polavarapu Nagendra Babu, "Detection of Fraudulent Activities in Unified Payments Interface using Machine Learning LSTM Networks," *IEEE Xplore,* 2024
DOI: 10.1109/ICCPCT61437.2024.10668422

[10] Yuhao Chen, Zengwei Zheng, Jianhua Ma, Binbin Zhou, "A Graph Neural Network with Imbalance-Aware Mechanism for Enhanced Fraud Detection," *IEEE Xplore*, 2024.
DOI: 10.1109/CyberSciTech62770.2024.10686436

[11] Vaishali Gupta, Sahil Sharma, Suhani Pathak, Suhani Nimkar, "UPI Based Financial Fraud Detection Using Deep Learning Approach," *IEEE Xplore,* 2024.
DOI:10.1109/ACROSET62108.2024.10743663

[12] Abdulwahab Ali Almazroi, Nasir Ayub "Online Payment Fraud Detection Model Using Machine Learning Techniques" *IEEE Xplore* ,2023.
DOI: 10.1109/ACCESS.2023.3339226

[13] Pradeep Kumar Roy, Abhinav Kumar, and Ashish Singh, "Advanced Learning for Phishing URLs Detection to Secure Consumer-Centric Applications," 2024
DOI: 10.1109/TCE.2024.3404459.

[14] N. Karthick, K. Kalaiselvi, G. Leema Roselin, S. Sudha, M. Tamilarasan, J. Jayanthi, "Unified Payment Interface Imposture and Scam Detection Using Deep Learning and ANN," *IEEE Xplore*, 2025.
DOI: 10.1109/ICSTSN61422.2024.10671346

[15] Kelvin S. C. Yong, Kang Leng Chiew, and Choon Lin Tan, "A Survey of the QR Code Phishing: The Current Attacks and Countermeasures," 2019 7th International Conference on Smart Computing & Communications (ICSCC), IEEE, 2019. DOI: 10.1109/ICSCC.2019.906545

Figures

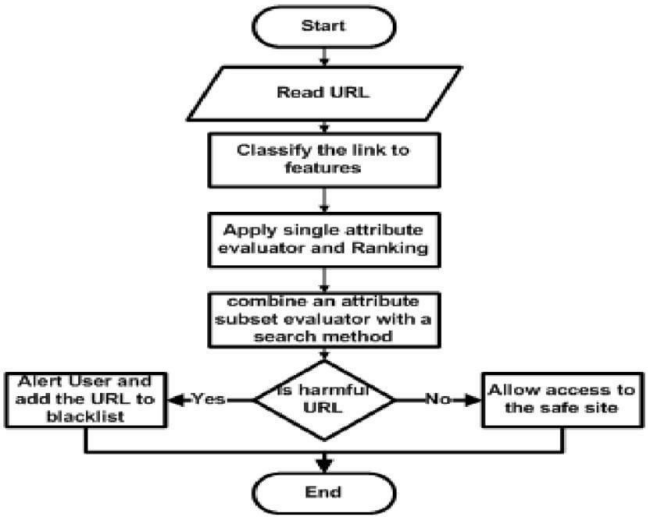


Fig .1. Block Diagram

Fig.1. The system starts by analyzing any link or QR code that a user interacts with—either by clicking or scanning. Once this input is received, the system looks at several features of the URL to decide whether it's potentially dangerous. These include things like the total length of the URL, whether it contains suspicious symbols, how many subdomains are present, whether it's using HTTPS, and if the address is an IP rather than a regular domain name.

After collecting this information, the system assigns importance to each factor based on how likely it is to indicate a threat. These features are then processed together using a classification model, which could be a machine learning algorithm trained to recognize risky behavior. Based on the outcome of this analysis, the system decides whether the URL is safe or harmful.

If the URL is considered unsafe, the system immediately issues a warning to the user and adds the link to a blacklist to prevent future access. If everything checks out and no threats are found, the user is allowed to continue safely.

This flow ensures that potentially harmful links—especially those hidden in QR codes—are caught early, protecting users from phishing attacks and other online threats.

Copyright & License:
© Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.