International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Safeguarding Digital Finance from Frauds using ML Technologies in Blockchain Technology

Nimmaka Asha Sri [1], K. Venkata Rao [2]

[1] Student, Department of Computer Science and Systems Engineering, Andhra University College of Engineering (A), Andhra University, Visakhapatnam, India.
[2] Professor, Department of Computer Science and Systems Engineering, Andhra University College of Engineering (A), Andhra University, Visakhapatnam, India.

Abstract
The rapid digitization of financial services has resulted in a staggering increase in sophisticated fraud, endangering global economies and damaging public trust. The dynamic nature of current fraud is outpacing classic fraud detection systems, which frequently rely on static, rule-based methods. This study reveals a new hybrid framework that pairs distributed ledger technology for immutable transaction avoidance with Machine Learning (ML) for real-time fraud detection. The fundamental driving force is to address the inherent shortcomings of centralized systems, as well as the lack of an unchangeable audit trail in ML-only solutions. Using a range of classification algorithms, our methodology entails creating separate machine learning pathways for three important financial domains: credit card, UPI, and loan applications. A fraud verdict is subsequently produced using the top-performing model for each domain, which is determined by a thorough analysis of metrics. Through a smart contract, this decision is safely and irrevocably documented on a private blockchain. This study shows how a strong security architecture may be produced by fusing the decentralized trust and immutability of blockchain technology with the predictive performance of machine learning. The findings demonstrate that this integrated approach strengthens the integrity and dependability of digital financial transactions by achieving high performance in fraud detection as well as creating a transparent and impenetrable record.

Keywords: Machine Learning, Private Blockchain, Hybrid framework, Smart contract

1. Introduction

Digital finance refers to the acquisition, utilization, and distribution of financial resources through digital devices and technology. Because it serves as the foundation for game-changing technologies like fintech and blockchain finance, provides users with convenience and cost savings, and has the potential to expand financial inclusion, it is essential to modern finance[1]. The swift expansion of digital financial services has resulted in an increase in intricate and sophisticated fraud. The volume and speed of contemporary financial data are too much for traditional rule-based systems, which are both static and prone to high false-positive rates[2]. Financial fraud is a rising challenge, with consumer losses reaching nearly $8.8 billion in 2022 in the US alone[3]. The fast expansion of FinTech, including digital banking tools and peer-to-peer financing, has resulted in substantial security issues such as cyber attacks, fraud, and data breaches. FinTech companies were the target of two-thirds of all financial sector cyberattacks in 2020 [4].

Transaction data is analyzed using supervised, unsupervised, and reinforcement learning algorithms to detect abnormalities and patterns that may indicate fraud.[5].

Machine Learning algorithms, such as supervised models like logistic regression and neural networks, as well as unsupervised techniques like clustering and anomaly detection, are used to analyze big datasets, uncover patterns, and detect abnormalities that may indicate fraud. Two essential methods for analyzing data in real time are feature extraction and pattern recognition [6]. Blockchain's decentralized, immutable, and cryptographic capabilities can improve FinTech security by reducing fraud, data breaches, and unauthorized access. It reduces single points of failure by spreading the ledger over several nodes, strengthening the system's defenses against cyberattacks. Because it is immutable, once a transaction is recorded, it cannot be altered, which is essential for preserving the accuracy and dependability of financial records and making real-time audits easier. Additionally, blockchain offers a secure layer that guards against financial record manipulation and unauthorized access, particularly in cloud-based systems [7]. Combination of the ML and Blockchain technologies offers a more durable and flexible defense against financial fraud than traditional methods alone [8].

2. Literature Review

Machine Learning algorithms are used to analyze big datasets, uncover patterns, and detect abnormalities that may indicate fraud. Two essential methods for analyzing data in real time are feature extraction and pattern recognition [6]. Logistic Regression is still one of the most often used techniques for data-mining in practice since it is simple to use, well-understood, and a good starting point for more recent approaches. A cutting-edge data mining technique known for its exceptional performance, Random Forest is computationally effective and noise-resistant. Support Vector Machine is a sophisticated data mining methodology with a solid theoretical basis and potent generalization ability. The popularity of Random Forest can be attributed to its interpretability, adaptability, and ease of use. The instability and dependability problems of individual DT models are addressed by RF, an ensemble of DTs [9]. XGBoost is an ensemble learning technique based on decision trees; the model is used to identify the most discriminating features [10]. XGBoost has a stellar track record in numerous data mining competitions, making it an excellent candidate for fraud detection models [11]. The model is used to pick the most discriminative features [12].

Logistic Regression is a commonly used fraud detection model in conventional financial companies. It is noted for being interpretable and having relatively good performance. An empirical study also demonstrated that linear logistic networks are more effective than other machine learning models for fraud detection [11].

SVM excels at handling complex, non-linear data by transforming it to find optimal decision boundaries, making it highly effective for classifying transactions that are not linearly separable. It is efficient in high-dimensional spaces [13]. KNN is a tested algorithm in the field and is noted for being easy to implement on small datasets [14]. SVM is a powerful supervised machine learning algorithm used for classification and is widely used for binary classification tasks . SVM aims to find the best boundary that separates data points with a maximized margin, which reduces the risk of misclassification. It is particularly useful in non-linear scenarios using the kernel trick [15].

Blockchain offers a transparent, decentralized, and impenetrable record for keeping information about financial transactions. When suspicious transactions are identified, smart contracts which are self-executing agreements with code-based terms, can be configured to automatically initiate alerts or take other appropriate action [6].

Data preprocessing is crucial, as is handling categorical variables with One Hot Encoding and selecting the most predictive features using Information Value (IV) and feature engineering to produce more meaningful features from preexisting ones, like turning a Time feature into cyclic sine and cosine values [16].

There are three phases of blockchain development, where the first stage (Blockchain 1.0) includes the exchange for cryptocurrencies like Bitcoin. The second Stage 2 (Blockchain 2.0) includes the usage of s smart contracts, loans, bonds, and futures that go beyond financial transactions. and the third Stage (Blockchain 3.0) serves as a general platform for applications across a range of sectors, including government, academia, and healthcare [17].

Blockchain is a decentralized, transparent, and immutable ledger that can improve risk management, authenticity, and security in the financial industry. It offers an unchangeable record of transactions, preventing fraud and tampering, and it does so without the need for middlemen, which can lower costs, speed up transactions, and increase operational efficiency [18].

The integration of ML with blockchain takes advantage of ML algorithms' predictive analytics as well as the transparency, security, and immutability of a decentralized blockchain. . By providing highly accurate and flexible real-time fraud detection, this synergy addresses the increasing complexity of financial fraud[6].

3. Existing System

Traditional rule-based systems find it difficult to handle the additional fraud risks brought forth by the growth of Fintech. A small percentage of transactions in financial datasets are fraudulent, and "concept drift" happens when fraud strategies and consumer behavior change over time [16].

.Conventional fraud protection techniques frequently rely on centralized systems that are susceptible to manipulation, are reactive, and are prone to human error. Reputational harm and large financial losses result from this [19]. The decentralized, immutable ledger of blockchain technology provides a safe and transparent transaction platform, while traditional banking systems are susceptible to hacks. Blockchain's advantages include better data security, more auditability and transparency, increased efficiency through smart contract automation, and the capacity to speed up international transactions. It may be almost hard to change or remove recorded data due to blockchain's immutability [20].

Current credit card systems depend on a reliable third party to complete transactions, making them susceptible to frauds such as social engineering and card skimming. Modern fraud detection techniques frequently employ machine learning or security measures like OTP separately. Scalability issues, which can make systems slower and more costly, and the difficulties of protecting extremely important data, such as transaction details on a public ledger, are two significant barriers to implementing blockchain technology [21].

Machine learning would establish a baseline of regular user behavior, and blockchain smart contracts would automatically prevent transactions from "non-deviant" accounts, creating a decentralized, immutable, and secure framework for digital payments. The paper makes the case that this strategy would

improve security, foster user confidence, and increase the resilience of the digital payment ecosystem—all while necessitating a strong IT infrastructure [22].

4. Methodology

The proposed system is intended to detect fraudulent activity across numerous digital financial platforms, specifically credit card, UPI, and loan transactions, as well as to ensure the integrity and immutability of fraud determinations using a blockchain ledger. Data gathering, preprocessing, feature engineering, model training and assessment, and blockchain integration are all part of the methodology.

4.1. Data Collection and Preprocessing

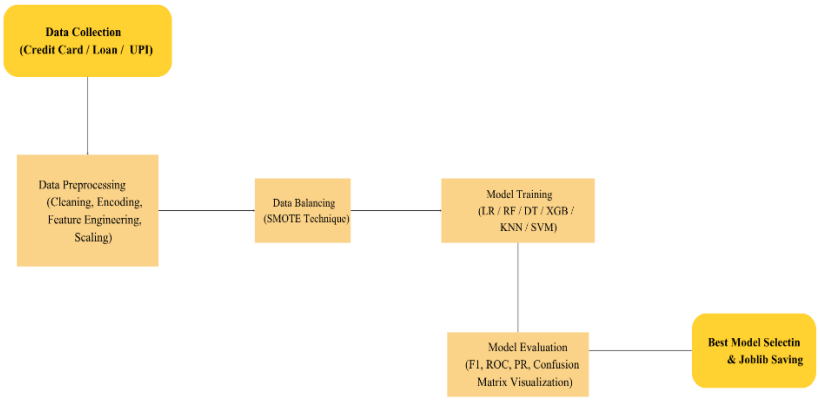
Credit card transactions data, Loan data, and UPI transaction datasets were obtained from Kaggle. For all datasets: To maintain data integrity, missing values were eliminated. Target variables and features were kept apart. Stratified by the goal label, the datasets were divided into training (80%) and testing (20%) sets. To avoid bias toward majority classes, the Synthetic Minority Oversampling Technique (SMOTE) was used to address imbalanced classes. In machine learning pipelines, standard scaling was used for numerical features.

4.2. Feature Engineering

To improve model performance, feature engineering was done: Credit Card: Encoded category features that show risk indications and transaction declines. Loan: Suspicious activity was highlighted using balance discrepancies and mismatch indicators. UPI: Label-encoded categorical characteristics and temporal features derived from transaction dates were incorporated. To ensure uniformity in the model input, all features were scaled and aligned.

4.3. Model Selection and Training

For Credit Card Fraud Detection, Four machine learning algorithms were evaluated: Logistic Regression, Random Forest, Decision Tree, and Support Vector Machine (SVM). Pipelines were constructed combining feature scaling and the classifier. Models were trained on SMOTE-resampled datasets. For Loan Fraud Detection, Two algorithms, Logistic Regression and XGBoost, were used. Pipelines included scaling, and SMOTE was applied to balance the classes. For UPI Fraud Detection, K-Nearest Neighbors (KNN) and SVM were trained. Label-encoded features ensured proper handling of categorical inputs. SMOTE was applied for class balance.



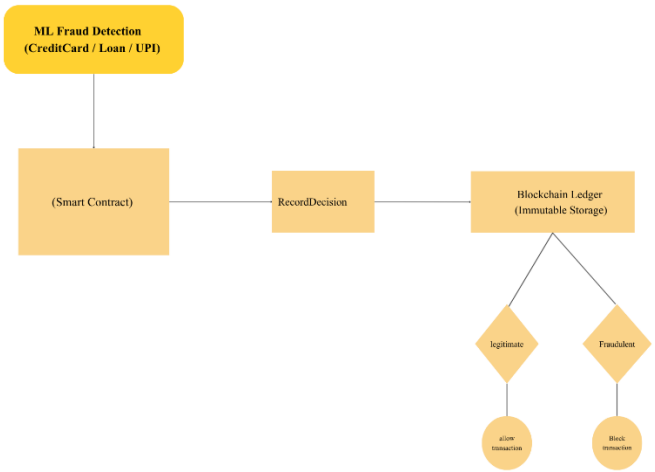
Workflow of Fraud Detections using ML algorithms

4.4. Evaluation Metrics

Model performance was assessed using: F1 Score, the Primary metric for imbalanced data. Precision-Recall Curve and Average Precision (AP), ROC Curve and AUC, Confusion Matrix
The best-performing model for each transaction type was selected based on the highest F1 Score.

4.5. Blockchain Integration

A Solidity smart contract, was deployed on a local Ethereum network (Ganache) to maintain an immutable record of all fraud decisions. Then, the contract was compiled and deployed using Python, with the contract address and ABI stored for integration with ML pipelines. Now Integrated Machine Learning Pipelines. The system pipeline, receives a new transaction and identifies its type, loads the corresponding trained ML model, performs fraud prediction, and records the prediction on the blockchain ledger. And then, a separate module queries the smart contract to verify transaction status, ensuring transparency, auditability, and tamper-proof logging.



The conceptual workflow

5. Implementation

5.1. Fraud Detection using Machine Learning Algorithms

The initial step is loading the dataset and follows with Data Preprocessing which includes Cleaning, feature engineering, encoding, scaling and then splitting the dataset where the data is split into training (80%) and testing (20%), here Stratification: stratify=y is used to ensure that the proportion of fraudulent (positive) cases is preserved in both the training and testing sets, which is crucial for imbalanced datasets.The datasets is highly imbalanced (fraud cases are rare). The Synthetic Minority Over-sampling Technique (SMOTE) is applied only to the training data to balance the class distribution.

SMOTE Algorithm Summary:

For each minority sample x_i, find its k nearest neighbors (NNs) .Randomly select a neighbor x_{zn} Generate a synthetic sample xnew along the line segment between x_i and x_{zn}

$$x_{new} = x_i + rand(0,1) * (x_{zn} - x_i)$$

Where rand(0,1) is a random number between 0 and 1. This process is repeated until the class distribution is balanced. Now a scikit-learn Pipeline is used for each model to all the frauds.

The pipeline first Standardizes the features and then applies the chosen Classifier. This ensures that scaling is performed consistently and correctly on both the training and testing data.

The next step is Standard Scaling (Normalization), where Standardization transforms the features to have a mean of 0 and a standard deviation of 1.

$$z = \frac{x - \mu}{\sigma}$$

Where x is the original feature value, μ is the mean of the training data, and σ is the standard deviation of the training data.

Evaluation Metrics: Model performance is evaluated using metrics suitable for imbalanced classification: Classification Report Metrics, Metrics are calculated for the positive class (Fraud, C=1).

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

1. Precision (P): The proportion of positive predictions that were actually correct (True Positives).

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} = \frac{TP}{TP + FP}$$

2. Recall (R): The proportion of actual positive cases that were correctly identified (also known as Sensitivity or True Positive Rate).

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

3. F1-Score (F1): The harmonic mean of Precision and Recall. It provides a single score that balances both metrics. This is the primary metric used to select the best model.

$$\text{F1} - \text{Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

4. Support: The number of occurrences of the class in the true responses (TP+FN).

The Receiver Operating Characteristic (ROC) curve plots the True Positive Rate (Recall) against the False Positive Rate (FPR) at various threshold settings.

$$False\ Positive\ Rate(FPR) = \frac{False\ Positives}{False\ Positives + True\ Negatives} = \frac{FP}{FP + TN}$$

- Area Under the Curve (AUC): The area under the ROC curve (AUC) measures the overall ability of the model to distinguish between positive and negative classes. An AUC of 1.0 is perfect.

Precision-Recall Curve and Average Precision (AP)

Plots Precision against Recall at various threshold settings.

- Crucial for imbalanced data because it doesn't involve the True Negatives (TN), which are often dominant in fraud detection.
- Average Precision (AP): Summarizes the P-R curve as the weighted mean of precisions achieved at each threshold, with the increase in recall from the previous threshold used as the weight.

$AP = \sum(R_n - R_{n-1})P_n$

Where P_n and R_n are the precision and recall at the n-th threshold.

The model with the highest F1-Score is selected as the best one. The entire pipeline, including the scaler and the best classifier, along with the feature names, is saved to the disk using joblib.dump() as a single file (creditcard_fraud, upi_fraud, loan_fraud). This allows the model to be loaded and used later for prediction on new, unseen data without retraining.

The system employs a preventative phase following the detection of possible fraudulent transactions using machine learning (ML) pipelines for loan, credit card, and UPI data. In order to enforce security, transparency, and auditability, this step makes sure that transactions that are identified as fraudulent are banned and permanently documented using a blockchain-based ledger.

The prevention module functions as a comprehensive system that connects tamper-proof enforcement and fraud detection. Blockchain Layer (Ganache) A private Ethereum blockchain was simulated using Ganache. It is appropriate for research and testing since it offers test accounts and Ether for deployment and interaction. This blockchain guarantees the immutability, auditability, and tamper-proof storage of all fraud detection judgments, in contrast to a centralized storage system.

On Ganache, a smart contract built with Solidity was implemented. It has two mappings, One that records if a transaction is authentic or fraudulent. Another one documents transactions that are expressly permitted (safe). To promote openness and ease monitoring, two events are released. By doing this, post-factum manipulation is avoided and every decision made by the ML model is guaranteed to be permanently saved on-chain. Which connects blockchain enforcement and machine learning predictions. First it load the details of the deployed contract. And then Uses python script to connect to Ganache. After that Process

the transaction using the relevant machine learning model (loan, credit card, or UPI). then it Forward the choice to the smart contract for long-term documentation. Provides the transaction hash back as evidence that the blockchain was executed.

The smart contract is queried for recorded results by a verification script. The system outputs either of the following explicitly based on values stored on the blockchain. For fraudulent transactions, it has been blocked.

- 🚨 **Blocked** – for fraudulent transactions.
- ✔✔ **Allowed** – for legitimate transaction

This ensures that fraud protection is implemented at the blockchain level rather than only being forecasted by machine learning models.

6. Result and Analysis

The performance analysis of credit card fraud detection:

Model	Precision(Fraud)	Recall(Fraud)	F1-Score(Fraud)	Support(Fraud)
Logistic Regression	0.97	0.97	0.97	90
Random Forest	0.96	0.97	0.96	90
Decision Tree	0.88	0.91	0.90	90
SVM	0.97	0.93	0.95	90

Logistic Regression is the top performing model in the 'Fraud' category, with an F1-Score of 0.97, which is a balanced measure of precision and recall. Maximum F1 score (0.97): The F1-Score is determined as the harmonic mean of recall and precision. In these scenarios, accurately detecting fraud cases is often critical; thus, a high F1-Score indicates that the model is highly good at decreasing false positives (high accuracy) and false negatives (high recall) for the minority 'Fraud' category. Highest Recall (0.97): It shares the highest recall with Random Forest. The model's high recall for the 'Fraud' class suggests that it is highly good at detecting the majority of genuine fraud instances, hence reducing missed fraud.

Highest Precision (0.97): It shares first place with SVM. High accuracy decreases false alarms by raising the likelihood that the model will be correct when it predicts a case as "Fraud". Based only on these classification data, Logistic Regression is the most dependable model for the minority 'Fraud' class, exceeding all other models in all three crucial criteria (precision, recall, and F1-score).

The performance analysis of Loan Fraud Detection:

Model	Class	Precision	Recall	F1-Score	Support
Logistic Regression	Legit	1.00	0.98	0.99	1270777
	Fraud	0.05	0.99	0.10	1639
	macro avg	0.53	0.99	0.54	1272416
	weighted avg	1.00	0.98	0.99	1272416
XGBoost	Legit	1.00	0.99	0.99	1270777

	Fraud	0.08	0.97	0.15	1639
	Macro avg	0.54	0.98	0.57	1272416
	weighted avg	1.00	0.99	0.99	1272416

XGBoost is regarded as the superior model for this extremely imbalanced dataset (with 'Fraud' as the minority class) because it strikes a better balance between successfully identifying fraud and minimizing false positives, as seen by its metrics for the 'Fraud' class and macro averages: 'Fraud' has the highest F1-Score (0.15): The harmonic mean of recall and precision is the F1-score. The F1-score of XGBoost is greater for the critical minority class ('Fraud') (0.15 vs. 0.10). A higher F1-score indicates that the model performs better and more evenly when it comes to detecting fraud situations.
Maximum Accuracy for 'Fraud' (0.08): The accuracy for 'Fraud' is the percentage of all anticipated fraud cases that are fraud. Because of its increased precision (0.08 vs. 0.05), XGBoost is more likely to identify a transaction as fraudulent when it does so, resulting in fewer false alarms and less time and money wasted looking into situations that aren't fraudulent.
F1-Score with the highest macro average (0.57): For the 'Legit' and 'Fraud' classes, the macro average F1-score is the unweighted average of the F1-scores. A critical evaluation statistic for unbalanced datasets, XGBoost's greater macro F1-score (0.57 vs. 0.54) suggests it has a more balanced performance across all classes because the macro average handles both classes equally.

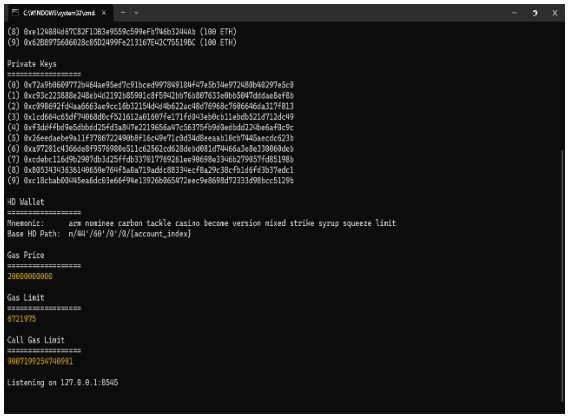
The performance analysis of UPI fraud detection:

Model	Class	Precision	Recall	F1-Score	Support	Accuracy
KNN	Legit	0.86	0.62	0.72	99	0.63
	Fraud	0.36	0.68	0.47	31	
	macro avg	0.61	0.65	0.59	130	
	weighted avg	0.74	0.63	0.66	130	
SVM	Legit	0.85	0.95	0.90	99	0.83
	Fraud	0.74	0.45	0.56	31	
	Macro avg	0.79	0.70	0.73	130	
	weighted avg	0.82	0.83	0.82	130	

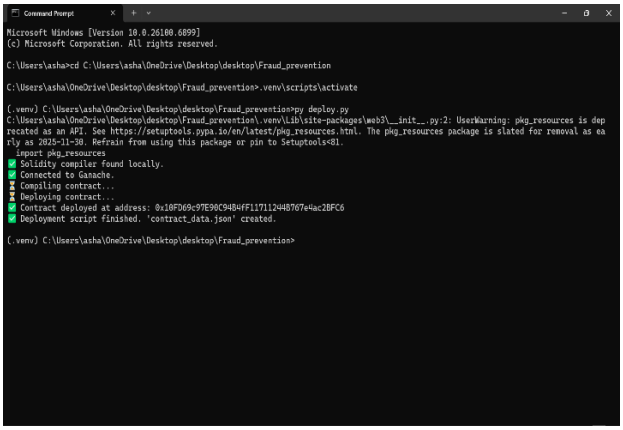
SVM (Support Vector Machine) is regarded as the best model. Because it performs better overall across the main aggregated measures, Maximum Total Accuracy (0.83): KNN's accuracy of 0.63 is much lower than SVM's 0.83.

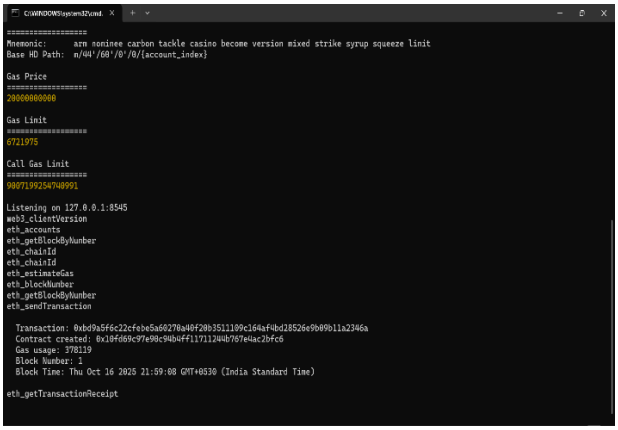
The proportion of correctly identified samples to all samples is known as accuracy. F1-Score with the highest macro average (0.73): Regardless of class imbalance, the macro average F1-score is an essential indicator for measuring performance evenly across classes. KNN's 0.59 macro average F1-score is significantly lower than SVM's 0.73, suggesting a more evenly distributed performance across the "Legit" and "Fraud" classes.
Highest Weighted Average Metrics: SVM's F1-score (0.82 vs. 0.66), recall (0.83 vs. 0.63), and precision (0.82 vs. 0.74) all have higher weighted averages. SVM performs significantly better at overall classification on the dataset when the weighted average takes into consideration the amount of samples in each class. Greater Precision for 'Fraud' (0.74): SVM consistently predicts fraud correctly, reducing false positives for the 'Fraud' class (high precision of 0.74 vs. KNN's 0.36).

Now the prevention part using Blockchain Technology:



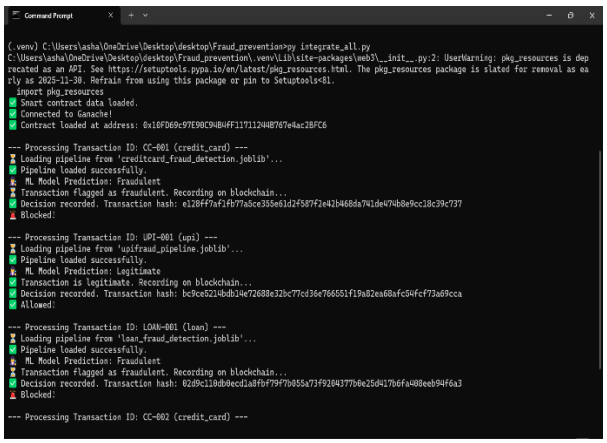
The graphic above depicts an Ethereum wallet's testing setup, that is Ganache, which is a personal, in-memory blockchain that developers use to test and deploy smart contracts without relying on a real network such as the Ethereum mainnet or testnet. This demonstrates that the deployment script has connected to this local development environment successfully.
which includes generated accounts, private keys, the recovery phrase, and network configuration settings. The mnemonic and private keys need to be safeguarded right away if any of the funds this data represents are real.

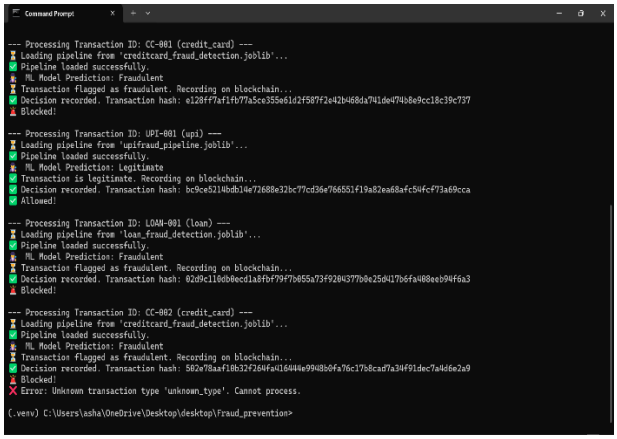




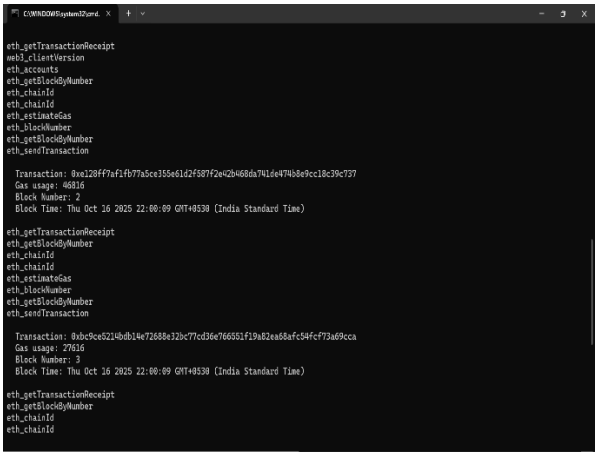
The Solidity compiler was located locally and connected to Ganache, as seen in the graphic on the left above, which describes the Smart Contract Deployment. It illustrates a successful DevOps cycle for a blockchain-based application, pertaining to "Fraud Prevention.. Deployment script finished. contract_data.json created."

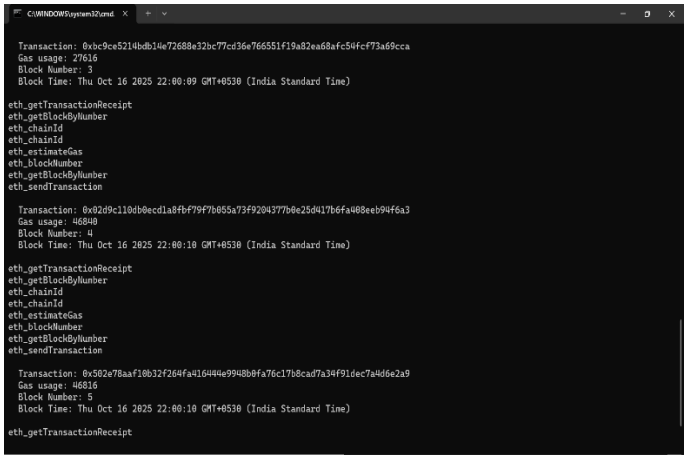
We deployed a Solidity smart contract to a local Ganache test blockchain using a Python script in a virtual environment, enabling testing and integration of the contract. The image to the right shows a record of the successful deployment process on a local Ethereum development chain (Ganache). It logs the resource use (Gas Usage), the transaction hash, and the address of the recently generated smart contract.





A strong decentralized fraud detection system is shown in the graphic on the left above, which: analyzes various financial transactions using specific machine learning models (Credit Card, UPI, Loan). records the ML judgment (Legitimate/Fraudulent) as an unchangeable record on a blockchain smart contract (Ganache). This architecture is a potent use of blockchain technology since it offers an auditable, transparent, and permanent record of all fraud decisions. The decentralized fraud detection system's ability to handle various transaction types and record four choices (three blocks, one allowance) on the blockchain is demonstrated in the screenshot on the right above, which demonstrates the system's operational integrity. The script encountered an unexpected or unhandled data type in its input stream, as indicated by the logical error that ends the execution and stops additional processing.

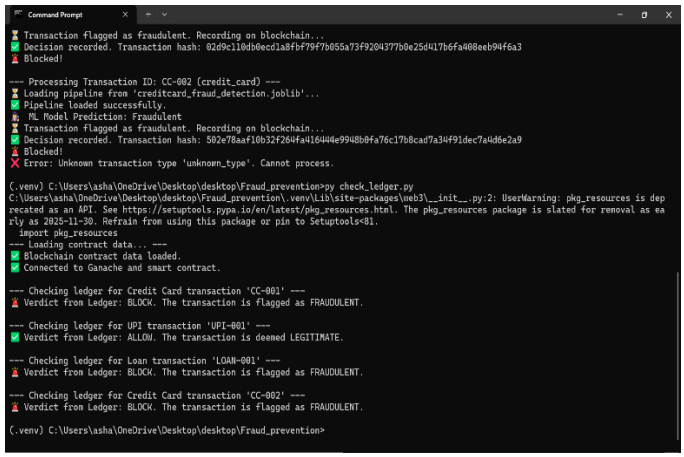


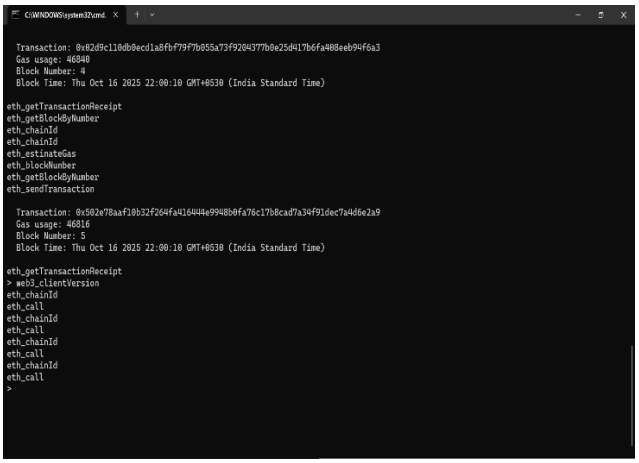


The low-level, on-chain confirmation logs for the fraud detection system's initial two decisions are shown in the screenshot on the left above. The above right one indicates that the log completes the record of the four successful fraud decisions made by the ML system and permanently stored on the smart contract.

It displays the transaction hashes, the precise gas costs, and the block numbers where these decisions were permanently recorded on the local Ganache blockchain. It offers conclusive on-chain evidence for LOAN-001 and CC-002 blocking, including transaction hashes, gas consumption, and block numbers. A DApp's

deployment, integration of an ML model, and unchangeable documentation of operational decisions are all depicted in the sequence of screenshots.





The verification and audit stage of the decentralized fraud detection system is seen in the screenshot on the left. The check_ledger.py script reads the recorded decisions for the four tested transactions after successfully connecting to the smart contract (the immutable ledger). An entire DApp workflow is depicted in the series of screenshots: Deployment: Putting the smart contract into action and configuring the environment. Integration/Execution: Executing the machine learning models and permanently documenting their findings on the blockchain (via transactions).Verification/Audit: Reading the blockchain's unchangeable decisions to demonstrate that the system operates in a transparent and impenetrable manner. The last piece of technical evidence is the right-side log segment, which demonstrates that all four decision-recording transactions were completed successfully. More significantly, it offers the low-level evidence, the repeated eth_call commands, that the follow-up script was successful in reading the immutable fraud decisions back from the smart contract ledger. The decentralized fraud detection system has now been fully demonstrated.

7. Conclusion

This study offers a thorough framework for detecting and preventing financial fraud that combines block-chain technology with machine learning (ML) to improve the security, dependability, and transparency of online financial transactions. Using classifiers including Logistic Regression, Random Forest, XGBoost, KNN, and SVM, several machine learning pathways were created for credit card, UPI, and loan transactions. SMOTE handled imbalanced datasets well, and F1-score, ROC-AUC, and Precision-Recall curves were used to thoroughly assess model performance.

A smart contract installed on a local Ganache Ethereum network continuously records fraud judgments, guaranteeing that legal transactions are permitted while fraudulent ones are banned.

This system ensures a transparent, auditable, and impenetrable ledger, allowing for authoritative confirmation of every decision. The integrity, transparency, and dependability of the system were validated during the verification phase, which was carried out using the script for testing the ledger audit. It was verified that the recorded verdicts could be successfully retrieved from the blockchain. Low-level transaction logs offered useful proof of system operation by confirming successful on-chain finalization and comprehensive gas utilization. The suggested approach shows how combining blockchain-enabled enforcement with machine learning-based detection results in a reliable, transparent, and scalable real-time fraud management solution. It lowers the chance of monetary loss, provides end-to-end safety, and increases confidence in digital transaction systems. The existing solution offers a scalable blueprint for real-world deployment, including possible integration with public blockchain networks, even if it functions in a local test environment.

8. Future work

Incorporating zero-knowledge proofs to preserve data privacy without sacrificing auditability, optimizing gas consumption for mainnet deployment, investigating decentralized governance for ML contract updates, and expanding the framework to accommodate real-time streaming transactions and sophisticated ensemble or deep learning models are some future directions. All things considered, this study confirms a robust, end-to-end architecture that improves financial integrity and security in digital finance by fusing distributed ledger technology with predictive analytics.

References

1. Ozili, Peterson. (2023). Digital finance research and developments around the World: a literature review. International Journal of Business Forecasting and Marketing Intelligence. 1. 10.1504/IJBFMI.2023.127698
2. Omogbeme, Angela & Atoyebi, Iyabode & Soyele, Adesola & Ogunwobi, Emmanuel. (2024). Enhancing fraud detection and prevention in fintech: Big data and machine learning approaches. World Journal of Advanced Research and Reviews. 24. 2301-2319. 10.30574/wjarr.2024.24.2.3617
3. (Mohite), Vaishali & Meher, Kunal & Dass, Ryan & Jonista, Athisaya & D'Souza, Jeston & Victor, Raymun. (2023). Fraud Detection Using Machine Learning and Blockchain. International Journal on Recent and Innovation Trends in Computing and Communication. 11. 584-590. 10.17762/ijritcc.v11i6s.6970. : Fraud Detection Using Machine Learning and Blockchain
4. L. Singh, A. Chirputkar and P. Ashok, "Risk Management in the Digital Age: Fintech Security Strategies," 2024 1st International Conference on Sustainable Computing and Integrated Communication in Changing Landscape of AI (ICSCAI), Greater Noida, India, 2024, pp. 1-7, doi: 10.1109/ICSCAI61790.2024.10866839. Risk Management in the Digital Age: Fintech Security Strategies
5. Jubiter, Francis. "Blockchain and Machine Learning Integration for Real-Time Fraud Detection in Fintech." *International Journal of Emerging Trends in Computer Science and Information Technology* (2025): 539-548.
6. Bello, Halima Oluwabunmi, Courage Idemudia, and Toluwalase Vanessa Iyelolu. "Integrating machine learning and blockchain: Conceptual frameworks for real-time fraud detection and prevention." *World Journal of Advanced Research and Reviews* 23.1 (2024): 056-068.

7. Onteddu, Abhishake Reddy, et al. "Integrating Blockchain Technology in FinTech Database Systems: A Security and Performance Analysis." *Asian Accounting and Auditing Advancement* 11.1 (2020): 129-142.
8. Tatineni, Sumanth. "Enhancing Fraud Detection in Financial Transactions using Machine Learning and Blockchain." *International Journal of Information Technology and Management Information Systems (IJITMIS)* 11.1 (2020): 8-15.
9. Bhattacharyya, Siddhartha, et al. "Data mining for credit card fraud: A comparative study." *Decision support systems* 50.3 (2011): 602-613.
10. Yedukondalu, G., et al. "Antifraud Model For Internet Loan Using Machine Learning." *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2021.
11. Fang, Weiwei, et al. "Deep learning anti-fraud model for internet loan: Where we are going." *Ieee Access* 9 (2021): 9777-9784.
12. Wen, Hanlin, and Fangming Huang. "Personal loan fraud detection based on hybrid supervised and unsupervised learning." *2020 5th IEEE international conference on big data analytics (ICBDA)*. IEEE, 2020.
13. Kamble, Vitthal B., et al. "Enhancing UPI Fraud Detection: A Machine Learning Approach Using Stacked Generalization." *International Journal of Multidisciplinary on Science and Management* 2.1 (2025): 69-83.
14. Rani, Rupa, Adnan Alam, and Abdul Javed. "Secure UPI: Machine learning-driven fraud detection system for UPI transactions." *2024 2nd International Conference on Disruptive Technologies (ICDT)*. IEEE, 2024.
15. Valli, Mudunuri Sri Uma Satya Naga, and A. Durga Devi. "UPI FRAUD DETECTION USING MACHINE LEARNING." *International Journal of Management Research and Reviews* 15.2s (2025): 149-154.
16. Stojanović, Branka, et al. "Follow the trail: Machine learning for fraud detection in Fintech applications." *Sensors* 21.5 (2021): 1594.
17. Trivedi, Sonal, Kiran Mehta, and Renuka Sharma. "Systematic literature review on application of blockchain technology in E-finance and financial services." *Journal of technology management & innovation* 16.3 (2021): 89-102.
18. Javaid, Mohd, et al. "A review of Blockchain Technology applications for financial services." *BenchCouncil transactions on benchmarks, standards and evaluations* 2.3 (2022): 100073.
19. Adejumo, Adetunji Paul, and Chinonso Ogburie. "Blockchain for Fraud Prevention: Transforming Accounting and Finance." *Futurity Proceedings* 2 (2025).
20. Ajish, Deepa. (2024). A COMPREHENSIVE STUDY ON BENEFITS AND CONCERNS OF BLOCKCHAIN IN SECURITY AND COMPLIANCE IN BANKS. International Research Journal of Modernization in Engineering Technology and Science. 06. 2251-2265. 10.56726/IRJMETS48632.
21. Balagolla, E. M. S. W., et al. "Credit card fraud prevention using blockchain." *2021 6th international conference for Convergence in Technology (I2CT)*. IEEE, 2021.
22. Mabel, Emmanuel. (2025). Fortifying Financial Transaction Security: A Hybrid Approach Using Artificial Intelligence and Distributed Ledger Technologies.