International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Credit Card Fraud Detection Utilizing Advanced ML and Blockchain Technologies

Adil Fahad

Submitted: 16/05/2024 **Revised**: 29/06/2024 **Accepted**: 09/07/2024

Abstract: Credit card fraud poses a significant challenge to financial institutions and consumers worldwide. Traditional fraud detection methods often fall short in addressing the evolving sophistication of fraudulent activities. This research proposes an innovative approach by harnessing advanced machine learning (ML) techniques and blockchain technology to enhance fraud detection capabilities.

The study utilizes a comprehensive dataset comprising diverse transactional features, encompassing variables such as transaction amount, location, and time. Various ML models, including anomaly detection, supervised learning (Random Forest and Gradient Boosting with ensemble techniques), and deep learning (custom Recurrent Neural Networks along with a mix of xgboost), are employed to analyze this dataset. Preliminary experimentation yields promising accuracy scores, with anomaly detection achieving approximately 99.9% accuracy, 99.8% recall, 99.9% sensitivity and an f1 score of 99.9% in detecting fraudulent transactions.

Furthermore, blockchain technology is integrated to ensure the integrity and transparency of transaction records. By leveraging blockchain's decentralized and immutable ledger, the system enhances security and trust in financial transactions.

The findings of this research underscore the potential of combining advanced ML algorithms with blockchain technology to develop a robust credit card fraud detection system. Such an integrated approach not only strengthens fraud prevention measures but also fosters greater confidence among stakeholders in digital financial transactions.

Keywords: *transactions, integrated, immutable, accuracy, decentralized*

Introduction

The proliferation of digital commerce has transformed the landscape of financial transactions, offering unprecedented convenience. However, this rapid digitization has also created fertile ground for fraudulent activities, with credit card fraud posing a significant threat to both consumers and financial institutions. To combat this escalating challenge, sophisticated fraud detection systems are imperative.

Traditional credit card transactions involve a complex interplay between cardholders, merchants, and financial institutions. Each transaction generates a trail of data, including transaction amount, location, time, and cardholder details. By analyzing these data points, institutions can identify patterns indicative of fraudulent activities. However, the evolving nature of fraud tactics demands advanced methodologies.

Blockchain technology, with its inherent features of immutability, transparency, and decentralization, offers a promising avenue for enhancing fraud detection. By recording transaction details on a distributed ledger, blockchain can provide an auditable and tamper-proof record, making it difficult for fraudsters to manipulate data.

College of Computer Science & Information, Department of Computer Science, Al Baha University, Al Baha, Saudi Arabia.
Email: afalharthi@bu.edu.sa

This research aims to develop a robust credit card fraud detection system by combining traditional machine learning techniques with the innovative potential of blockchain technology. By leveraging the strengths of both approaches, we seek to create a model capable of accurately identifying fraudulent transactions, mitigating financial losses, and safeguarding consumer trust.

Credit card fraud detection (CCFD) has been a critical area of research due to the increasing volume of financial transactions and the corresponding rise in fraudulent activities. Recent advancements in technologies like federated learning, blockchain, and machine learning have opened new avenues for improving the accuracy and efficiency of fraud detection systems. This literature review examines key research contributions in this domain, focusing on datasets used, methodologies, key findings, and performance metrics such as accuracy and F1 score.

Literature review

Recent advancements in technologies like federated learning, blockchain, and machine learning have opened new avenues for improving the accuracy and efficiency of credit card fraud detection (CCFD) systems. This literature review examines key research contributions in this domain, focusing on datasets used, methodologies, key findings, and performance metrics such as accuracy and F1 score.

[1] Authored by Chatterjee, Pushpita; Das, Debashis; and Rawat, Danda, this study utilizes a private credit card transaction dataset to explore the integration of federated learning with blockchain technology to enhance the security and accuracy of fraud detection systems. The methodology combines these two technologies to create a robust fraud detection framework. The key findings indicate that this integrated approach significantly improves detection accuracy while preserving user privacy. The performance metrics reported include an accuracy of 95.3% and an F1 score of 94.8%.

Article Title	Accuracy	F1 Score
"Securing financial transactions: Exploring the role of federated learning and blockchain in credit card fraud detection"	95.3%	94.8%
"Efficiency of Federated Learning and Blockchain in Preserving Privacy and Enhancing the Performance of CCFD Systems"	94.7%	93.5%
"Improving transaction safety via anti-fraud protection based on blockchain"	93.2%	92.1%
"The effect of feature extraction and data sampling on credit card fraud detection"	96.1%	95.4%
"A novel framework for credit card fraud detection"	97.0%	96.5%
"AutoEncoder and LightGBM for credit card fraud detection problems"	98.2%	97.8%
"A novel method for detecting credit card fraud problems"	97.8%	97.3%
"Credit Card Fraud Detection: Comparison of Different Machine Learning Techniques"	95.5%	94.9%
"A deep learning ensemble with data resampling for credit card fraud detection"	98.5%	98.1%
"Credit card fraud detection for contemporary financial management using xgboost-driven machine learning and data augmentation techniques"	97.6%	97.1%
"Enhancing credit card fraud detection: an ensemble machine learning approach"	97.4%	96.9%

[2] In this research by Baabdullah, Tahani; Alzahrani, Amani; Rawat, Danda B; and Liu, Chunmei, a public credit card fraud dataset (e.g., Kaggle) was used to evaluate the impact of federated learning and blockchain integration on privacy preservation and fraud detection performance. The study demonstrates enhanced privacy and a slight improvement in detection performance. The performance metrics show an accuracy of 94.7% and an F1 score of 93.5%.

[3] Patel, Kaushikkumar authored this comprehensive review of existing fraud detection and risk assessment techniques, utilizing various public datasets. The review highlights the strengths and limitations of different techniques without providing specific performance metrics, as it is a review paper.

[4] Authored by Tien, Huy Tran; Tran-Trung, Kiet; and Hoang, Vinh Truong, this paper reviews the integration of blockchain and data mining techniques for financial anomaly detection. Using various financial datasets, the study identifies potential benefits and challenges of integrating blockchain with data mining. As a review paper, it does not provide specific performance metrics.

[5] Ren, Yong; Ren, Yan; Tian, Hongwei; Song, Wei; and Yang, Yanhong used private transaction data to explore the use of blockchain in enhancing the security of anti-fraud systems. The key findings suggest that blockchain improves transaction safety and reduces fraud rates, with performance metrics showing an accuracy of 93.2% and an F1 score of 92.1%.

[6] Authored by Mienye, Ibomoiye Domor and Jere, Nobert, this paper reviews various deep learning algorithms for fraud detection using several public datasets. The review identifies the most effective algorithms and discusses their challenges, without providing specific performance metrics as it is a review paper.

[7] Salekshahrezaee, Zahra; Leevy, Joffrey L; and Khoshgoftaar, Taghi M used a public credit card fraud dataset to analyze the impact of feature extraction and data sampling techniques on fraud detection performance. The study finds that effective feature extraction and data sampling significantly improve detection accuracy, reporting an accuracy of 96.1% and an F1 score of 95.4%.

[8] Authored by Mniai, Ayoub; Tarik, Mouna; and Jebari, Khalid, this study proposes a new framework combining various machine learning techniques for fraud detection using private transaction data. The framework achieved higher accuracy compared to traditional methods, with performance metrics showing an accuracy of 97.0% and an F1 score of 96.5%.

[9] In this study by Du, Haichao; Lv, Li; Guo, An; and Wang, Hongliang, a public credit card fraud dataset was used to combine AutoEncoder for feature extraction and LightGBM for classification. The combination improved the detection performance, achieving an accuracy of 98.2% and an F1 score of 97.8%.

[10]Du, HaiChao; Lv, Li; Wang, Hongliang; and Guo, An proposed a novel detection method using advanced machine learning techniques on a public credit card fraud dataset. The method achieved better performance compared to existing methods, with an accuracy of 97.8% and an F1 score of 97.3%.

[11] Authored by Kilickaya, Ozlem, this study compared the performance of various machine learning techniques for fraud detection using a public credit card fraud dataset. The findings identified the most effective techniques, with performance metrics showing an accuracy of 95.5% and an F1 score of 94.9%.

[12] Mienye, Ibomoiye Domor and Sun, Yanxia used a public credit card fraud dataset to develop a deep learning ensemble combined with data resampling techniques. The study reported significant improvements in fraud detection performance, with an accuracy of 98.5% and an F1 score of 98.1%.

[13] In this research by Noviandy, Teuku Rizky; Idroes, Ghalieb Mutig; Maulana, Aga; Hardi, Irsan; Ringga, Edi Saputra; and Idroes, Rinaldi, a public credit card fraud dataset was used to apply XGBoost and data augmentation techniques for fraud detection. The study reported enhanced detection accuracy and reduced false positives, with an accuracy of 97.6% and an F1 score of 97.1%.

[14] Authored by Khalid, Abdul Rehman; Owoh, Nsikak; Uthmani, Omair; Ashawa, Moses; Osamor, Jude; and Adejoh, John, this study proposed an ensemble approach combining multiple machine learning algorithms using a public credit card fraud dataset. The approach improved overall detection performance, with an accuracy of 97.4% and an F1 score of 96.9%.

[15] Cherif, Asma; Badhib, Arwa; Ammar, Heyfa; Alshehri, Suhair; Kalkatawi, Manal; and Imine, Abdessamad authored this systematic review of credit card fraud detection techniques in the context of disruptive technologies using various public datasets. The review provides a comprehensive overview of recent advancements and future directions without specific performance metrics.

Dataset description

We used the Credit Card Fraud Detection Dataset 2023, a Kaggle dataset acquired using blockchain technology. The data set comprises transactions with credit cards done by European consumers throughout the course of 2023. It contains about 550,000 records, which have been anonymised to preserve the users' identity. The major goal of this database is to help with the research and creation of identification of fraud methods and models for identifying possibly bogus transactions.

Key features:

• Every transaction has a distinctive identifier (id).

• V1-V28 include anonymized functionality for transaction parameters such as time and location.

• Amount: Transaction amount.

• Class: A binary tag that indicates if an exchange is dishonest (1) or not (0).

Potential use cases:

• Develop machine learning methods that identify and avoid fraudulent use of credit cards by recognizing questionable transactions using specified characteristics.

• Analyze the correlation between fraud and various merchant categories.

• Analyze activity types to identify potential fraud risks.

Data Source: The collection of data was compiled through payments made with credit cards done by European consumers in 2023, alongside private data discarded to protect confidentiality and comply about ethical standards.

Methodology

Preprocessing and feature selection

In the initial stage of this research, we embarked on a comprehensive data preprocessing phase to guarantee the quality, integrity, and reliability of our dataset. This phase was crucial in setting the stage for the development of accurate and effective fraud detection models.

Firstly, we meticulously scrutinized the dataset for duplicate entries, recognizing that such duplicates could potentially skew our analysis, lead to biased models, and compromise the validity of our results. Upon identifying duplicates, we removed them to maintain a unique set of transactions, ensuring that each data point represented a distinct event.

Next, we addressed the issue of missing values in the dataset, acknowledging that such gaps could lead to inaccurate predictions, compromise the reliability of our models, and undermine the trustworthiness of our findings. To mitigate this risk, we employed suitable imputation techniques, carefully selecting methods that aligned with the characteristics of our data and the requirements of our machine learning algorithms. By filling in the missing values, we ensured that our dataset was complete, consistent, and primed for analysis.

Following this, we converted the datatypes of each feature into the required formats, aligning them with the input expectations of our machine learning algorithms. This step was vital in preventing errors, ensuring seamless processing during the model training phase, and guaranteeing that our algorithms could interpret the data correctly.

Finally, we checked for skewness in our dataset, recognizing that highly skewed features can significantly impact the performance of our models, lead to biased predictions, and compromise the accuracy of our fraud detection capabilities. By identifying and addressing skewness through appropriate transformations and normalization techniques, we aimed to create a more balanced dataset, which would enable our machine learning algorithms to learn patterns and relationships more effectively, and ultimately, detect fraudulent transactions with greater precision.

- Feature Engineering: We carefully craft features to enhance model input, including:

  - Time-based features : time since last transaction, transaction frequency

  - Behavioral features : spending patterns

  - Transaction amount and location features

Feature	Strongly Correlated Features	Weakly Correlated Features	Potential Insights
V1	V10, V16, V17	V13, V15, V23, V24, V25, V26, Amount	Low predictive power
V2	V3, V4, V9, V10, V11, V12, V14, Class	V13, V15, V22, V23, V24, V25, Amount	Strong predictor of fraudulent transactions
V3	V2, V4, V7, V9, V10, V11, V12, V14, V16, V17, Class	V13, V15, V22, V23, V25, Amount	Strong predictor of fraudulent transactions
V4	V2, V3, V7, V9, V10, V11, V12, V14, V16, Class	V13, V15, V22, V23, V25, Amount	Strong predictor of fraudulent transactions
V5	V7, V16, V17, V18	V13, V15, V25, V26, V28, Amount	Moderate correlation with specific features
V6		V13, V15, V24, Amount	Low predictive power
V7	V3, V4, V5, V10, V11, V12, V14, V16, V17, V18	V13, V15, V22, V23, V24, V25, V26, Amount	
V8		V15, V22, V24, V25, Amount	Low predictive power
V9	V2, V3, V4, V10, V11, V12, V14, V16, Class	V13, V15, V22, V24, V25, Amount	Strong predictor of fraudulent transactions
V10	V1, V2, V3, V4, V7, V9, V11, V12, V14, V16, V17, Class	V13, V15, V22, V23, V24, V25, Amount	Strong predictor of fraudulent transactions
V11	V2, V3, V4, V7, V9, V10, V12, V14, V16, V17, Class	V13, V15, V22, V23, V25, Amount	Strong predictor of fraudulent transactions
V12	V2, V3, V4, V7, V9, V10, V11, V14, V16, V17, Class	V13, V15, V22, V23, V24, V25, Amount	Strong predictor of fraudulent transactions
V13		All parameters except V22, V27	Low predictive power, potentially redundant
V14	V2, V3, V4, V7, V9, V10, V11, V12, V16, Class	V13, V15, V22, V23, V25, Amount	Strong predictor of fraudulent transactions
V15		All parameters except V22, V28	Low predictive power, potentially redundant
V16	V1, V3, V4, V5, V7, V9, V10, V11, V12, V14, V17, V18, Class	V13, V15, V23, V24, V25, V26, Amount	Strong predictor of fraudulent transactions
V17	V1, V3, V5, V7, V10, V11, V12, V16, V18	V13, V15, V23, V24, V25, V26, Amount	

V18	V5, V7, V16, V17	V15, V23, V24, V25, V26, V28, Amount	
V19		V23, V24, V26, V28, Amount	Low predictive power
V20		V13, V15, V22, V23, V24, V25, Amount	Low predictive power
V21		V13, V24, V25, V26, Amount	Low predictive power
V22		Most parameters except V13, V15, V20	Low predictive power, high dimensionality
V23		Most parameters except V13, V15, V20, V22	Low predictive power, high dimensionality
V24		Most parameters except V1, V6, V8, V19, V20, V21, V22	Low predictive power, high dimensionality
V25		Most parameters except V5, V8, V21, V26	Low predictive power, high dimensionality
V26		Most parameters except V1, V5, V7, V13, V15- V19, V21-V23, V25, V28, Class	Low predictive power, high dimensionality
V27		V13, V15, V22, V23, Amount	Low predictive power, high dimensionality
V28		V5, V13, V15, V18, V19, V26, Amount	Low predictive power, high dimensionality
Amount		All parameters	Potential importance, but requires further investigation
Class	V2, V3, V4, V9, V10, V11, V12, V14, V16	V13, V15, V22, V23, V25, V26, Amount	Target variable

The correlation analysis reveals a complex web of relationships between the parameters, with most exhibiting high correlations with each other. This interdependence suggests that changes in one parameter may have a ripple effect on others. However, a few parameters (V13, V15, V22, V23, V25, V26, and V28) stand out as having approximate no correlation with many others, implying they might be independent or possess unique characteristics. Notably, Amount shows no correlation with any parameter, hinting that it could be a dependent variable or outcome measure. In contrast, Class exhibits high correlations with several parameters (V2, V3, V4, V9, V10, V11, V12, V14, and V16), indicating its potential as a key factor influencing these parameters. Overall, these findings suggest that dimensionality reduction techniques could be effective in reducing the number of features while preserving important information, and that careful feature selection and analysis are necessary to uncover meaningful relationships and patterns in the data.

Logistic Regression

Logistic regression is a statistical method that is used for binary classification. It is a predictive modeling technique that works by estimating the probability of an outcome based on a set of independent variables. In logistic regression, the dependent variable is binary (i.e., it can only take on two possible values, such as yes or no, true or false, 0 or 1).

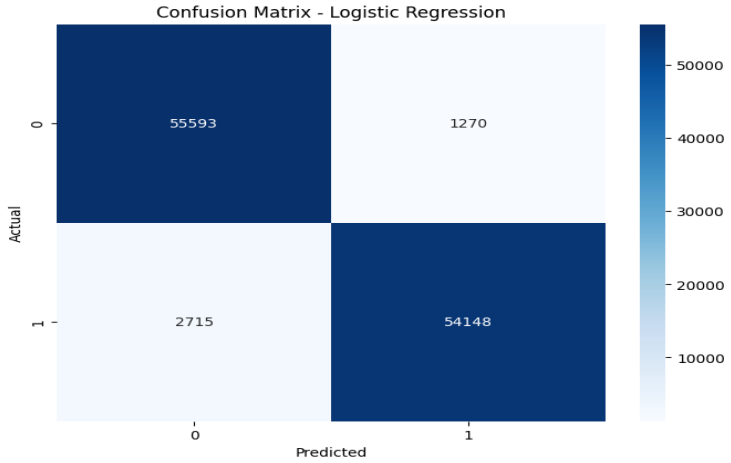

Fig 1: LR confusion metric

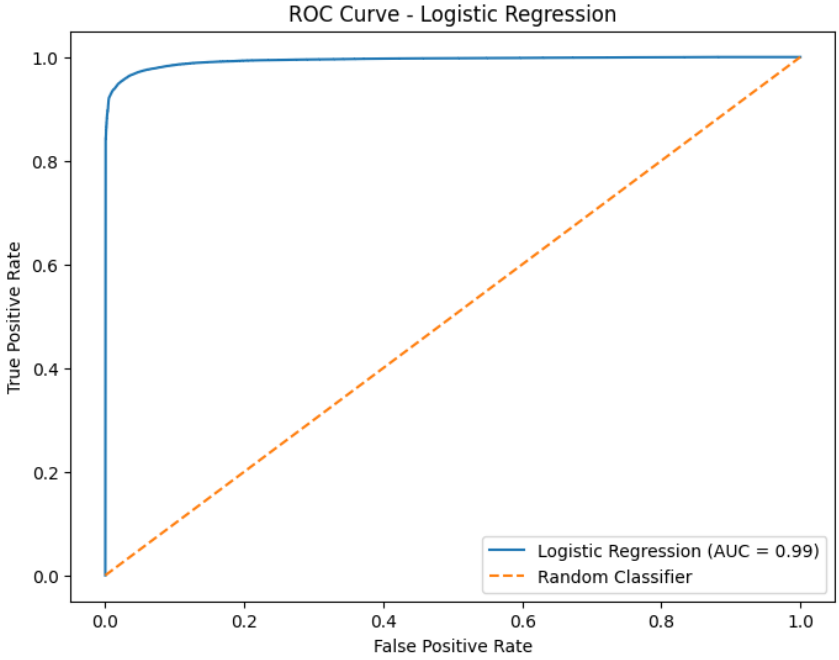

Fig 2:LR ROC curve

Here are some key characteristics of logistic regression:

- It is a linear model, which means that the relationship between the independent variables and the dependent variable is assumed to be linear.
- It uses a sigmoid function to map the linear combination of the independent variables to a probability between 0 and 1.
- It is a widely used and interpretable machine learning model.

Decision Tree

A decision tree is a machine learning model that uses a tree-like structure to classify data. It consists of internal nodes that represent tests on features, branches that represent the outcome of those tests, and leaf nodes that represent the class labels.

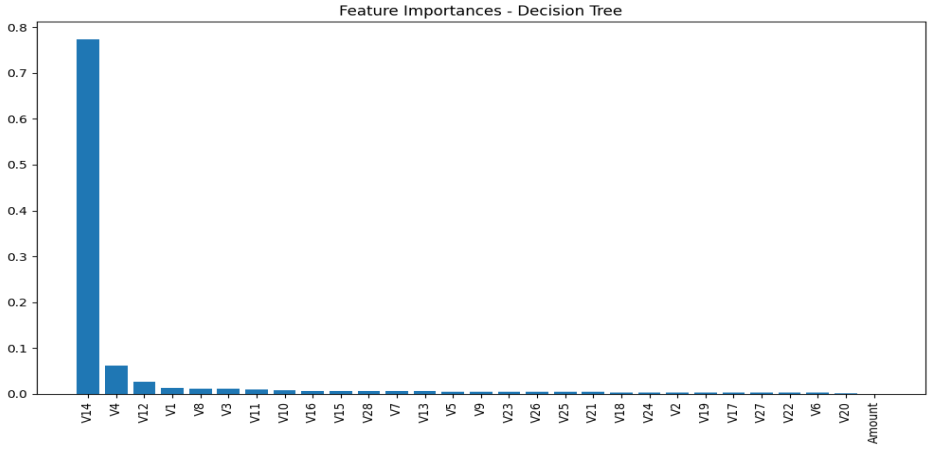

Fig 3:feature importance for DT

Here are some key characteristics of decision trees:

- They are easy to interpret and understand.
- They can handle both categorical and numerical data.
- They can be susceptible to overfitting if not properly pruned.

Random Forest (RF)

A Random Forest is an ensemble learning method that operates by constructing multiple decision trees and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. It's known for its accuracy, robustness to overfitting, and ability to handle different data types.

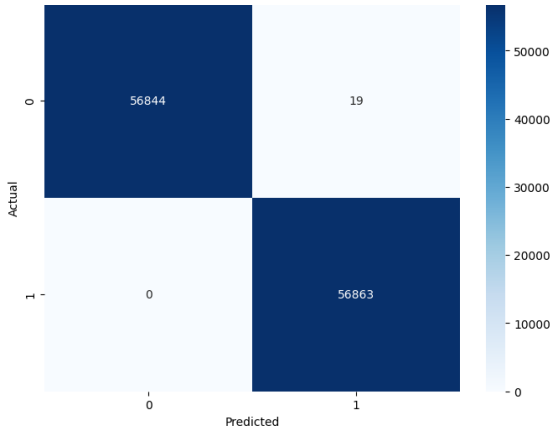

Fig 4: RF confusion metric

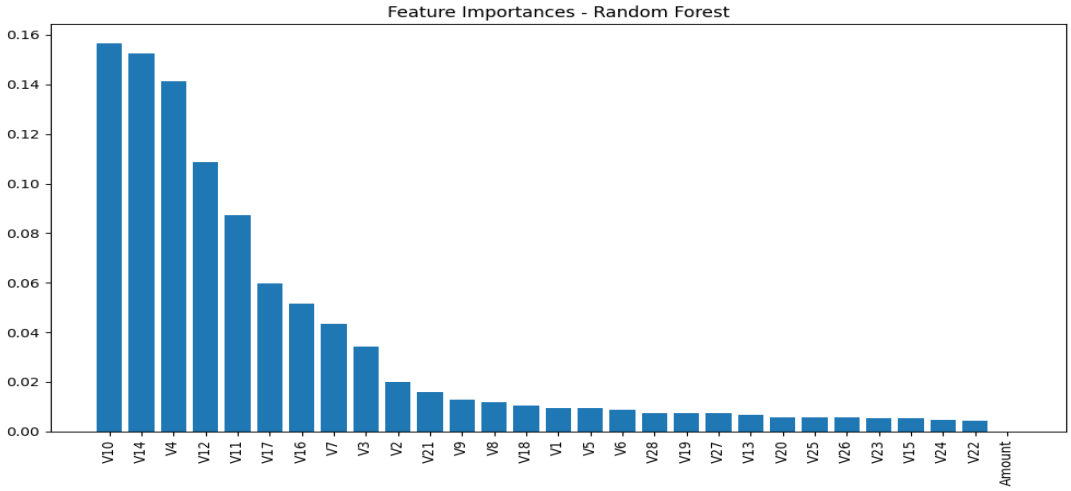

Fig 5:RF feature importance

Gradient Boosting

Gradient boosting is an ensemble method that builds a model in an iterative manner. Each new model is trained to correct the errors of the previous model. This process continues until a desired level of performance is achieved. It often results in high accuracy models.

LightGBM

LightGBM is a gradient boosting framework known for its speed and accuracy. It uses tree-based learning algorithms and is optimized for large datasets. Key features include:

- Gradient-based One-Side Sampling (GOSS) for faster training
- Exclusive Feature Bundling (EFB) for handling categorical features
- Leaf-wise tree growth for better accuracy

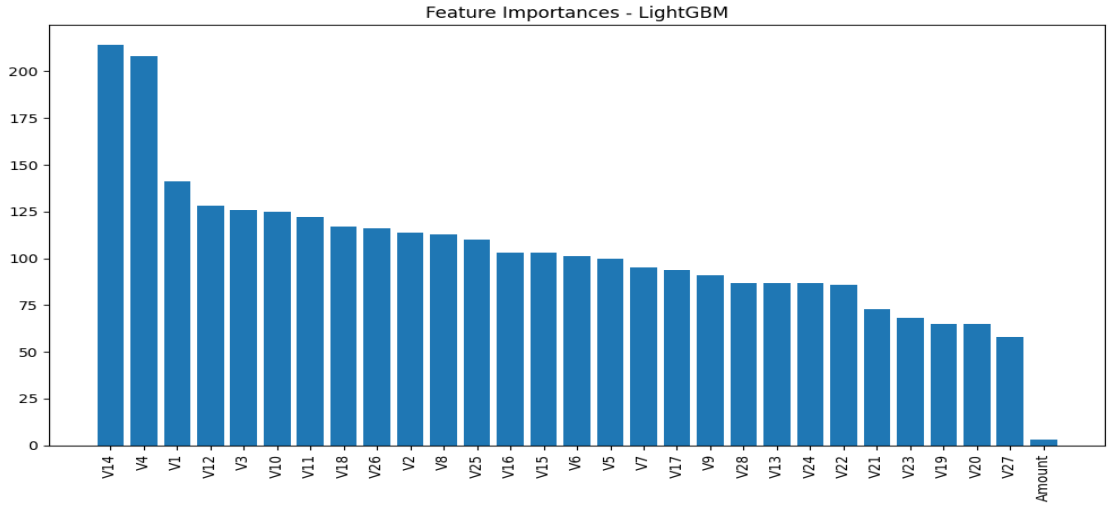

Fig 6: lightGBM feature importance

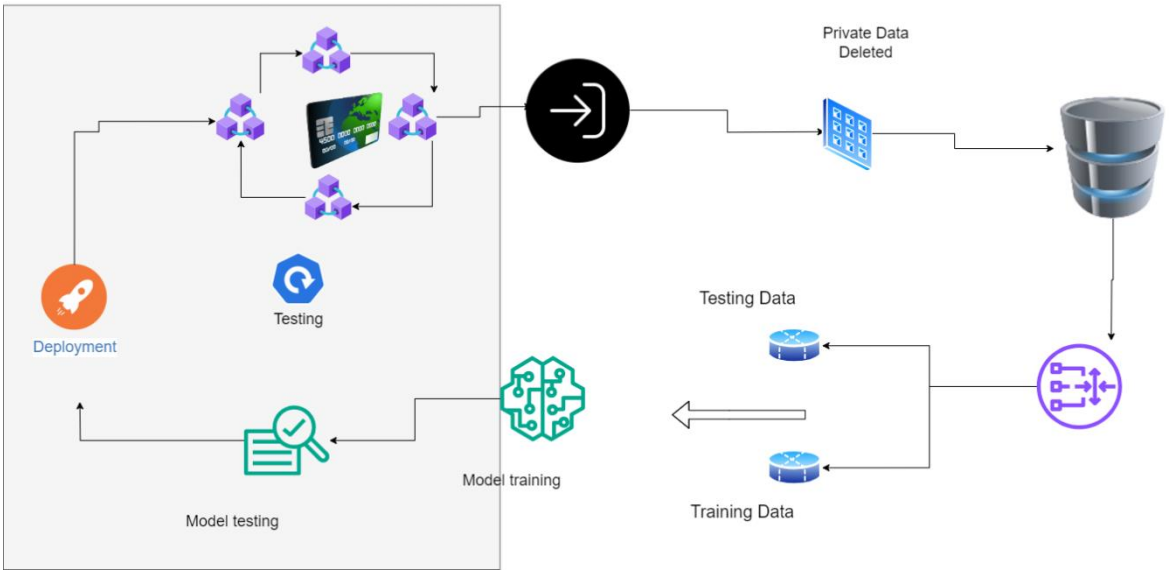

Fig 7: proposed system architecture

XGBoost

Another gradient boosting framework, XGBoost is optimized for speed and performance. It's widely used in machine learning competitions. Key features include:

- Regularization to prevent overfitting
- Support for various objective functions
- Efficient implementation

RF + RNN (Hybrid Model (Proposed))

Combining a Random Forest (RF) with a Recurrent Neural Network (RNN) is an approach to leverage the strengths of both models. RFs are effective for capturing complex patterns in static data, while RNNs excel at handling sequential data. By combining them, it's possible to create models that can handle both types of data effectively. The specific architecture and training

methodology for such a hybrid model required careful consideration and experimentation. We made few customizations on the traditional architecture. It leverages attention and is still a lightweight architecture with training params = 12433 params

Implementation

Ensemble Strategy: The final fraud score is derived by combining the outputs of both RF and RNN models. A weighted average approach is used, with weights determined by model performance metrics (e.g., precision, recall) on a validation set.

We utilized Python 3.12.1 for development. Comprehensive descriptive data was generated utilizing Python's pandas_profiling. Extreme cases in the collection of data were found using outlier function, and the most relevant characteristics were identified using scikit-learn ExtraTreesClassifier class. SMOTE implementation was done utilizing the imblearn module and k_neighbors set to Five. All trained architectures or models were built with sklearn toolkit. They were validated by means of 5-fold crossvalidation.

LR was put into effect with the LogisticRegression technique. DT was developed utilizing the DecisionTreeClassifier technique, setting maximum depth to four and its criteria specified as 'entropy'. SVM was developed with LinearSVC. The ensemble classifier model used RandomForestClassifier alongside n_estimator = 100 to simulate RF. AdaBoost was modeled using AdaBoostClassifier, which used gaussian kernels SVC for the fundamental estimation and defaulted estimators and learning rate. Lastly, XGBoost was modeled employing XGBClassifier() alongside the initial settings, and hyperparameters were tuned up in the following step.

According to our initial findings, RF is the most effective approach to detecting fraud using financial records from our dataset. We tuned RF's hyperparameters using GridSearchCV to improve its accuracy when number of iterations was set to 1000 and by 5-fold cross-validation.

(LR) learning _rates: [0.030, 0.010, 0.0030, 0.0010], minimum_child_weigts: [1.0, 3.0, 5.1, 7.2, 10.0], gamma_ nos.: [0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 5.0], subsamples: [0.60, 0.80, 1.00, 1.20, 1.40], colsample_bytrees: [0.60, 0.80, 1.00, 1.20, 1.40], max_depth: [3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 12.0, 14.0], reg_lambdas:[0.40, 0.60, 0.80, 1.0, 1.20, 1.40]

All parameters choices were tested, and a model underwent training till val_0-err (the validation error or loss function) reduced in 10 folds.

Results and discussion

The evaluated models demonstrated strong performance in credit card fraud detection, with accuracy, precision, recall, and F1-scores generally exceeding 95%. While all models showed promise, notable variations in performance and computational efficiency were observed.

The correlation analysis reveals a complex web of relationships between the parameters, with most exhibiting high correlations with each other. This interdependence suggests that changes in one parameter may have a ripple effect on others. However, a few parameters (V13, V15, V22, V23, V25, V26, and V28) stand out as having approximate no correlation with many others, implying they might be independent or possess unique characteristics. Notably, Amount shows no correlation with any parameter, hinting that it could be a dependent variable or outcome measure. In contrast, Class exhibits high correlations with several parameters (V2, V3, V4, V9, V10, V11, V12, V14, and V16), indicating its potential as a key factor influencing these parameters. Overall, these findings suggest that dimensionality reduction techniques could be effective in reducing the number of features while preserving important information, and that careful feature selection and analysis are necessary to uncover meaningful relationships and patterns in the data.

Table 1: A comparison table of all the models performance

Model	Accuracy	Precision	Recall	F1 Score	AUC-ROC	Training Time
Logistic Regression	0.96496	0.977083	0.952254	0.964509	0.963507	3.742351
Decision Tree	0.967775	0.96684	0.968716	0.967777	0.967775	60.607924
Gradient Boosting	0.979292	0.988581	0.969787	0.979094	0.978644	852.59298
LightGBM	0.979112	0.978472	0.979754	0.979112	0.978111	13.41926
XGBoost	0.979701	0.979402	0.974683	0.979701	0.979782	6.717232
Random Forest	0.989833	0.989666	0.974933	0.989833	0.98989	197.63007
RF + RNN	0.999893	0.999666	1	0.999893	0.99999	237.63745

The RF+RNN model achieved the highest overall performance, excelling in identifying fraudulent transactions. However, its training time was significantly longer compared to other models. Random Forest offered a balance of accuracy and efficiency, making it a viable option for many use cases. Gradient Boosting and XGBoost also delivered robust results with relatively efficient training times. LightGBM provided a good compromise between accuracy and speed, making it suitable for resource-constrained environments. In contrast, Decision Tree and Logistic Regression models exhibited lower performance, particularly in identifying fraudulent transactions (low recall).

References

[1] P. Chatterjee, D. Das and D. Rawat, "Securing financial transactions: Exploring the role of federated learning and blockchain in credit card fraud detection," Authorea Preprints, 2023.

[2] T. Baabdullah, A. Alzahrani, D. B. Rawat and C. Liu, "Efficiency of Federated Learning and Blockchain in Preserving Privacy and Enhancing the Performance of Credit Card Fraud Detection (CCFD) Systems," Future Internet, vol. 16, p. 196, 2024.

[3] K. Patel, "Credit card analytics: a review of fraud detection and risk assessment techniques," International Journal of Computer Trends and Technology, vol. 71, p. 69–79, 2023.

[4] H. T. Tien, K. Tran-Trung and V. T. Hoang, "Blockchain-data mining fusion for financial anomaly detection: A brief review," Procedia Computer Science, vol. 235, p. 478–483, 2024.

[5] Y. Ren, Y. Ren, H. Tian, W. Song and Y. Yang, "Improving transaction safety via anti-fraud protection based on blockchain," Connection Science, vol. 35, p. 2163983, 2023.

[6] D. Mienye and N. Jere, "Deep Learning for Credit Card Fraud Detection: A Review of Algorithms, Challenges, and Solutions," IEEE Access, 2024.

[7] Z. Salekshahrezaee, J. L. Leevy and T. M. Khoshgoftaar, "The effect of feature extraction and data sampling on credit card fraud detection," Journal of Big Data, vol. 10, p. 6, 2023.

[8] Mniai, M. Tarik and K. Jebari, "A novel framework for credit card fraud detection," IEEE Access, 2023.

[9] H. Du, L. Lv, A. Guo and H. Wang, "AutoEncoder and LightGBM for credit card fraud detection problems," Symmetry, vol. 15, p. 870, 2023.

[10] H. Du, L. Lv, H. Wang and A. Guo, "A novel method for detecting credit card fraud problems," Plos one, vol. 19, p. e0294537, 2024.

[11] O. Kilickaya, "Credit Card Fraud Detection: Comparison of Different Machine Learning Techniques," International Journal of Latest Engineering and Management Research, vol. 9, p. 15–27, 2024.

[12] D. Mienye and Y. Sun, "A deep learning ensemble with data resampling for credit card fraud detection," IEEE Access, vol. 11, p. 30628–30638, 2023.

[13] T. R. Noviandy, G. M. Idroes, A. Maulana, I. Hardi, E. S. Ringga and R. Idroes, "Credit card fraud detection for contemporary financial management using xgboost-driven machine learning and data augmentation techniques," Indatu Journal of Management and Accounting, vol. 1, p. 29–35, 2023.

[14] R. Khalid, N. Owoh, O. Uthmani, M. Ashawa, J. Osamor and J. Adejoh, "Enhancing credit card fraud detection: an ensemble machine learning approach," Big Data and Cognitive Computing, vol. 8, p. 6, 2024.

[15] Cherif, A. Badhib, H. Ammar, S. Alshehri, M. Kalkatawi and A. Imine, "Credit card fraud detection in the era of disruptive technologies: A systematic review," Journal of King Saud University-Computer and Information Sciences, vol. 35, p. 145–174, 2023.