AI-Powered Hybrid Defense for Credit, Vehicle Insurance, and Transaction Security Excellence

Aradhyula Hareesh Yadav[1]
[1]PG Student, Department of Computer Science, NIE, Macherla, India.
Email: hareeshyadav341@gmail.com
Dr. Y Rajesh[2]
[2]Associate Professor, Department of Computer Science, NIE, Macherla, India.Email: rajeshyemparala51@gmail.com

Abstract

In an era of rapidly increasing digital financial activity, fraud detection has become a cornerstone of securing both institutional and individual assets. This project presents a comprehensive hybrid fraud detection system that integrates credit card fraud, transactional anomalies, and vehicle insurance fraud into a unified predictive framework. Using a combination of machine learning and deep learning approaches, this system addresses the unique challenges and characteristics of each fraud domain. For credit card fraud detection, we employed a Multi-Layer Perceptron (MLP) neural network, optimized through SMOTE to counter class imbalance, and enhanced via standardization techniques to normalize transaction amounts. This segment provides real-time predictive capabilities with high accuracy and clarity using confusion matrix heatmaps. The transaction fraud module incorporates a deep learning architecture leveraging TensorFlow's Functional API. It integrates structured transactional data and embedded customer identifiers using a dual-input neural network model. This component emphasizes early stopping, class weighting, and AUC score evaluation to ensure robust detection of rare fraudulent events. In the case of vehicle insurance fraud, we adopt a traditional machine learning pipeline using the AdaBoost classifier, which shows high reliability in structured data environments. Categorical data encoding and SMOTE oversampling are utilized to enhance class balance, and model performance is visualized with detailed confusion matrices. Overall, this hybrid architecture enables cross-domain fraud detection by tailoring models and preprocessing techniques to their specific contexts. Through the effective use of neural networks and ensemble learning, this system provides a scalable, interpretable, and highly accurate fraud detection platform adaptable to multiple real-world applications.

Keywords: NLP, AI, Security

1. Introduction

In the digital age, financial fraud has emerged as one of the most pressing challenges confronting banking institutions, insurance providers, and e-commerce platforms. As transactions increasingly shift online, the volume, velocity, and complexity of fraudulent activities have simultaneously escalated. According to a 2023 report by the Association of Certified Fraud Examiners (ACFE), organizations lose an estimated 5% of their annual revenues to fraud. These alarming statistics underscore the critical need for intelligent, adaptive, and scalable fraud detection systems.

Traditionally, fraud detection was rule-based, relying heavily on predefined patterns and thresholds. However, as fraudulent tactics evolve rapidly, such static systems have proven to be inadequate. Consequently, the integration of machine learning (ML) and deep learning (DL) into fraud detection frameworks has revolutionized the field by enabling systems to learn from data,

Doi:10.48047/ijiee.2025.15.7.2

adapt to emerging patterns, and generalize across multiple fraud types. This project takes a **hybrid approach**, integrating credit card fraud detection, transaction fraud, and vehicle insurance fraud detection into a single intelligent pipeline. Each fraud category presents unique data characteristics, necessitating specialized models and preprocessing methods. By designing custom workflows for each domain and then unifying them under a cohesive framework, the system aims to deliver high detection accuracy while minimizing false positives.

2. Literature survey

Healthcare fraud detection has emerged as a vital area of research due to the growing complexity and financial impact of fraudulent activities within healthcare systems. Several survey papers have extensively explored the methodologies, challenges, and future directions in this domain.

1. SyedAbdul et al. (Journal of Network and Computer Applications) This survey categorizes healthcare fraud detection methods into four main types: rule-based systems, anomaly detection, predictive modeling, and social network analysis. Each method is analyzed for its strengths and weaknesses. While rule-based systems are interpretable, they struggle with dynamic fraud patterns. Anomaly detection can capture unknown fraud but may lead to false positives. Predictive models, especially machine learning-based, show promise but require high-quality labeled data. Social network analysis provides contextual understanding of provider and patient relationships. The paper also identifies key challenges such as scalability, data quality, and real-time processing, offering guidance for future research.
2. Gupta et al. (ACM Computing Surveys) This study provides a comprehensive survey and introduces a cluster-based hybrid approach for fraud detection. The hybrid model integrates unsupervised clustering techniques with supervised machine learning algorithms to enhance pattern recognition in healthcare datasets. The approach is validated using real-world data, showcasing improved fraud detection performance. This paper emphasizes data-driven techniques and highlights the need for robust, adaptable models capable of capturing evolving fraudulent behaviors. The authors also stress the importance of collaboration among stakeholders for effective implementation.
3. AlEmran et al. (IEEE Access) This paper presents a holistic view of approaches, challenges, and opportunities in healthcare fraud detection. Alongside traditional techniques, it explores emerging technologies like blockchain and AI-driven analytics. The authors delve into persistent challenges, including data privacy, imbalanced datasets, and rapidly changing fraud tactics. Notably, the paper identifies AI and big data analytics as key enablers for future progress. This survey contributes a forward-looking perspective, encouraging innovation and cross-disciplinary collaboration in fraud detection.
4. Thabtah (Journal of Health Informatics) This survey outlines key fraud detection methods such as collaborative filtering in addition to traditional rule-based and anomaly detection models. The author discusses practical implications for healthcare organizations and emphasizes the balance between detection accuracy and operational feasibility. The paper serves as a practical guide, focusing on the real-world applicability of fraud detection systems. It also encourages the development of context-aware and interpretable models to improve trust among healthcare stakeholders.

Key Insights and Comparative Summary

Doi:10.48047/ijiee.2025.15.7.2

Aspect	SyedAbdul et al.	Gupta et al.	AlEmran et al.	Thabtah
Methodologies Reviewed	Rule-based, anomaly detection, predictive, social networks	Clustering + Machine Learning (Hybrid)	Rule-based, ML, Blockchain, Social Network	Rule-based, Anomaly, Predictive, Collaborative Filtering
Novel Contribution	Comparative method analysis	Cluster-based hybrid model	Exploration of emerging tech	Practical applicability focus
Challenges Highlighted	Data quality, complexity	Model adaptability	Privacy, data imbalance	Operational feasibility
Future Directions Suggested	Scalable real-time systems	Integrated data analytics	AI, blockchain, big data	Interpretable, context-aware models
Application Focus	General healthcare fraud	Real-world dataset evaluation	Broad system-level perspective	Organization-level strategies

3. Methodology of the work

The proposed system for health insurance fraud detection is designed as a multi-layered framework that begins with the integration and preprocessing of diverse healthcare data. Data is collected from multiple sources including insurance claims, electronic health records, billing information, and provider databases. To ensure consistency and usability, a rigorous data preprocessing phase is carried out to handle missing values, correct inconsistencies, and remove duplicate entries. This foundational step is critical for ensuring that subsequent analytics are both accurate and meaningful. Additionally, compliance with regulatory standards such as HIPAA or GDPR is maintained to safeguard patient privacy and data security throughout the process.

Once data is cleaned and structured, advanced data analytics and machine learning models are applied to identify patterns associated with fraudulent activities. Supervised learning techniques like logistic regression, random forests, and support vector machines are used to classify claims based on labeled historical data. In cases where labeled data is limited or unavailable, unsupervised methods such as clustering and anomaly detection help uncover unusual behavior that could signal potential fraud. Furthermore, Natural Language Processing (NLP) is employed to analyze unstructured textual information—like medical notes or claim descriptions—for hidden indicators of fraud. Predictive models are then developed to assign fraud scores to each transaction, helping prioritize claims for further investigation.

The final component of the methodology includes real-time monitoring, alert generation, and comprehensive investigative tools. A dynamic monitoring system continuously assesses new claims, flagging those with high fraud probability for immediate action. Social network analysis is integrated to detect collusion among patients, providers, and organizations by examining their relationships and interactions. Rules-based systems enforce predefined detection criteria, while external data sources (e.g., criminal records, sanction lists) are used to

Doi:10.48047/ijiee.2025.15.7.2

enrich detection capability. Case management tools support analysts in organizing and reviewing evidence, and a feedback loop enables system improvement through model retraining and rule updates. User training and awareness initiatives ensure that all stakeholders are equipped to recognize and respond to emerging fraud threats effectively.

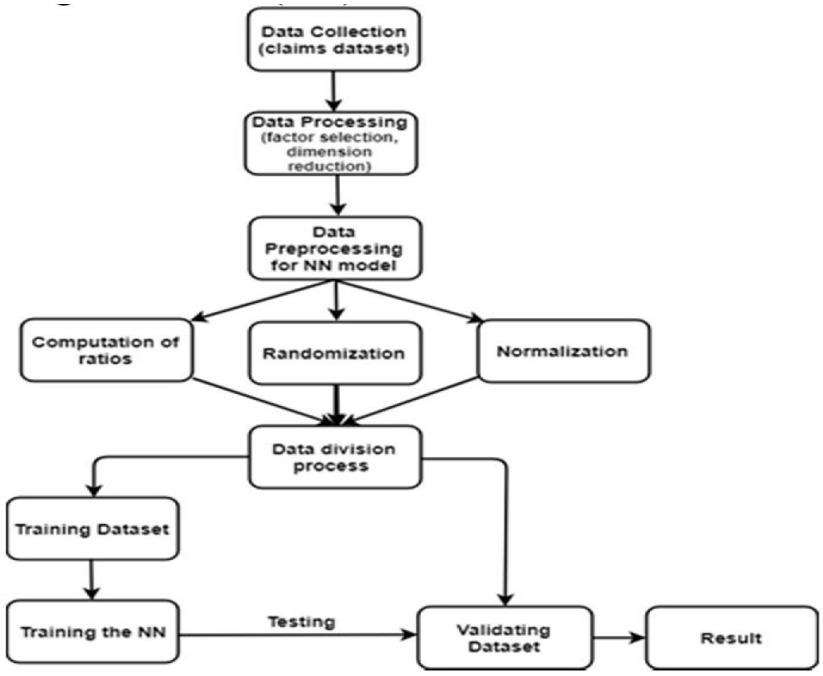


Figure 1: Flowchart

4. Implementation

The visual workflow for the Healthcare Fraud Detection Project outlines a structured four-step approach. The process begins with Exploratory Data Analysis (EDA), which is crucial for understanding the structure, distribution, and hidden patterns in healthcare datasets. This step helps identify inconsistencies, outliers, or missing values that could impact fraud detection performance. After understanding individual datasets, the next step is to combine multiple datasets into a single unified dataset, integrating information from different sources like claims, medical histories, and billing records to form a comprehensive base for analysis.

Once the data is integrated, the project moves to Feature Engineering, where relevant variables or features are created or selected to better represent the underlying patterns in the data. These features could be based on provider behavior, patient claim frequency, billing amounts, or anomalies in treatment records. Finally, in the Machine Modeling phase, advanced machine learning algorithms are applied to the engineered data to detect potential fraud. The goal is to build a model capable of classifying or scoring claims based on their likelihood of being fraudulent, supporting early detection and prevention efforts in healthcare systems.

Doi:10.48047/ijiee.2025.15.7.2

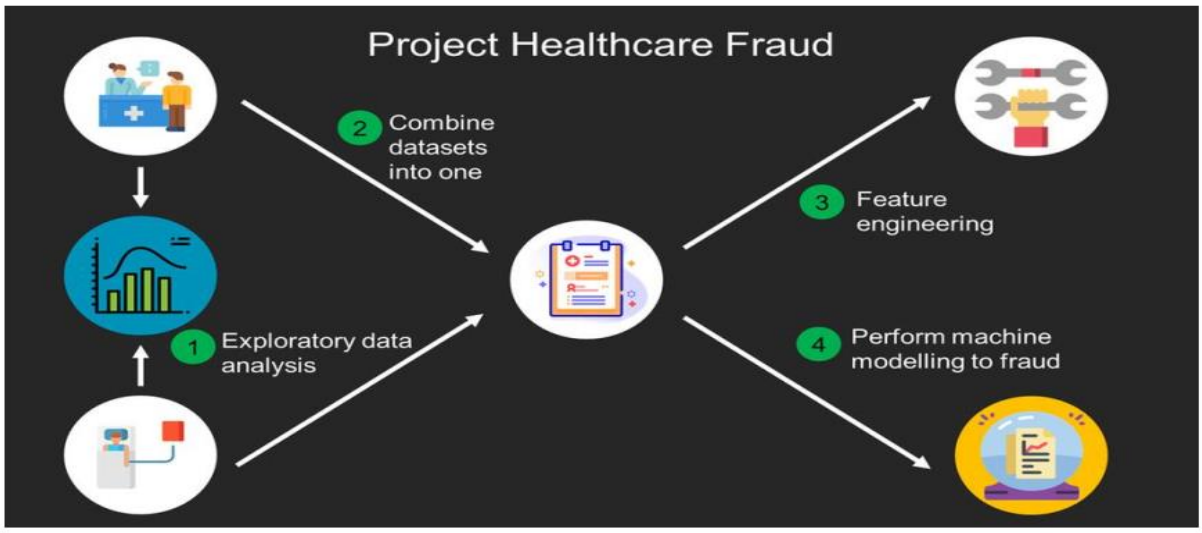


Figure 2: System architecture

5. Results and discussions

Fraud Fortify is an AI-powered hybrid defense system designed to enhance security across credit systems, vehicle insurance, and financial transactions. By integrating biometric verification, card-based authentication, and intelligent fraud detection algorithms, the platform provides multi-layered protection against unauthorized access and suspicious activities. The system leverages machine learning to analyze transaction patterns in real time, issuing warnings and alerts for anomalies, while offering a secure and seamless user experience. With its holistic approach, Fraud Fortify ensures enhanced security excellence, minimizing financial risks and reinforcing trust in digital finance and insurance platforms.

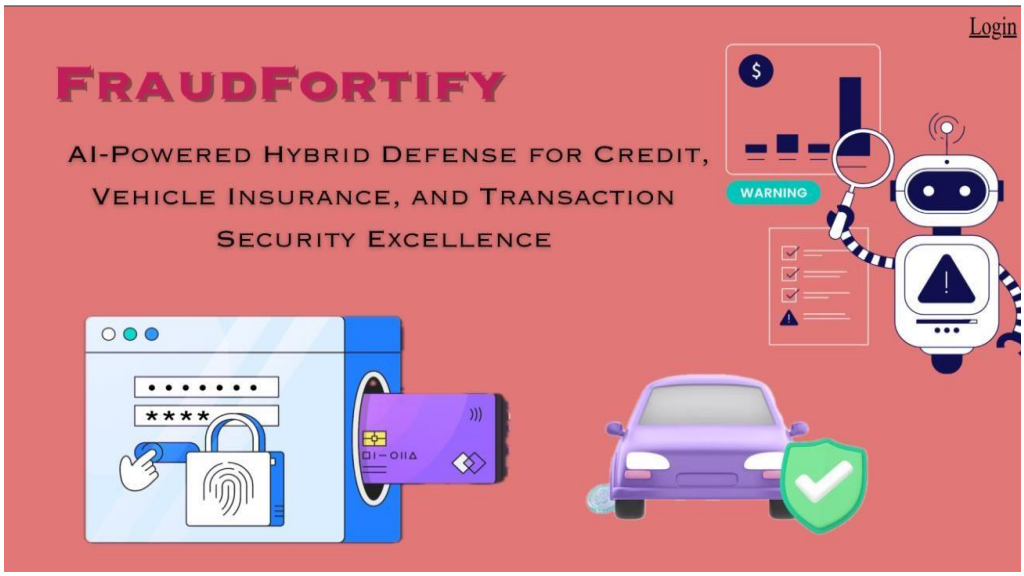


Figure 3: Main window

Doi:10.48047/ijiee.2025.15.7.2

The displayed interface represents a simple and user-friendly login screen, where users can securely enter their credentials to access a system. It features input fields for a username and password, with the example showing "admin" as the username. A clearly labeled "LOGIN" button allows users to proceed after entering their details. The minimalist design, combined with a clean gradient background, ensures ease of use while maintaining focus on authentication, an essential aspect of any secure application or system.

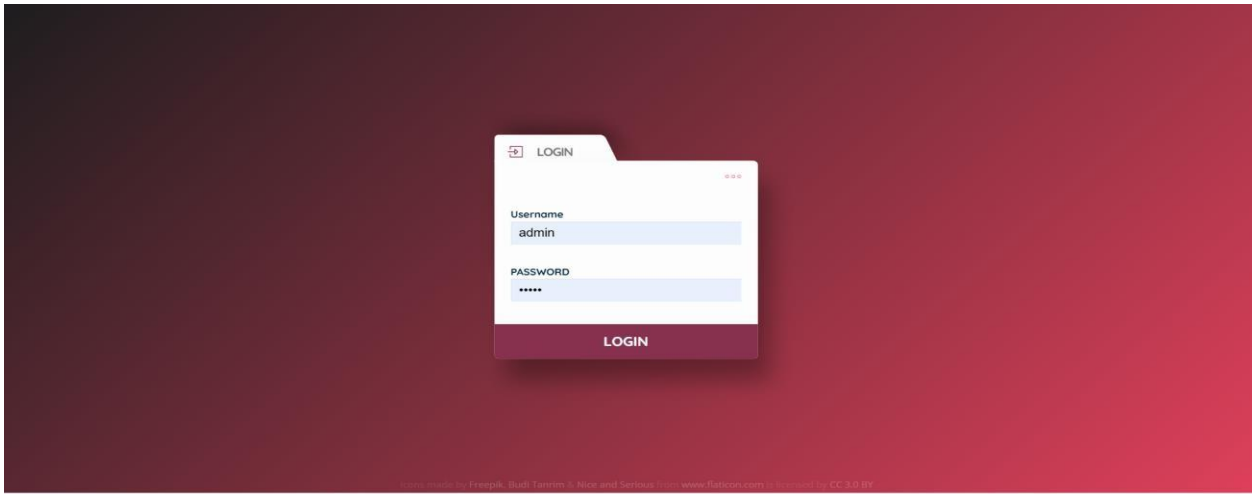


Figure 4: Login credentials window

The image showcases a credit card fraud detection analysis, highlighting both fraudulent and normal transaction data points alongside visual performance metrics. The left section lists numerical outputs—likely feature values or anomaly scores—classified into two categories: Credit Card Fraud Detection and Normal Transaction. The right side presents several analytical plots, including a class distribution graph, a training loss curve, and a confusion matrix or classification report. These visuals help evaluate the effectiveness of a machine learning model in distinguishing between legitimate and fraudulent transactions, demonstrating balanced class handling, training performance, and prediction accuracy after applying techniques like SMOTE (Synthetic Minority Oversampling Technique).

Doi:10.48047/ijiee.2025.15.7.2

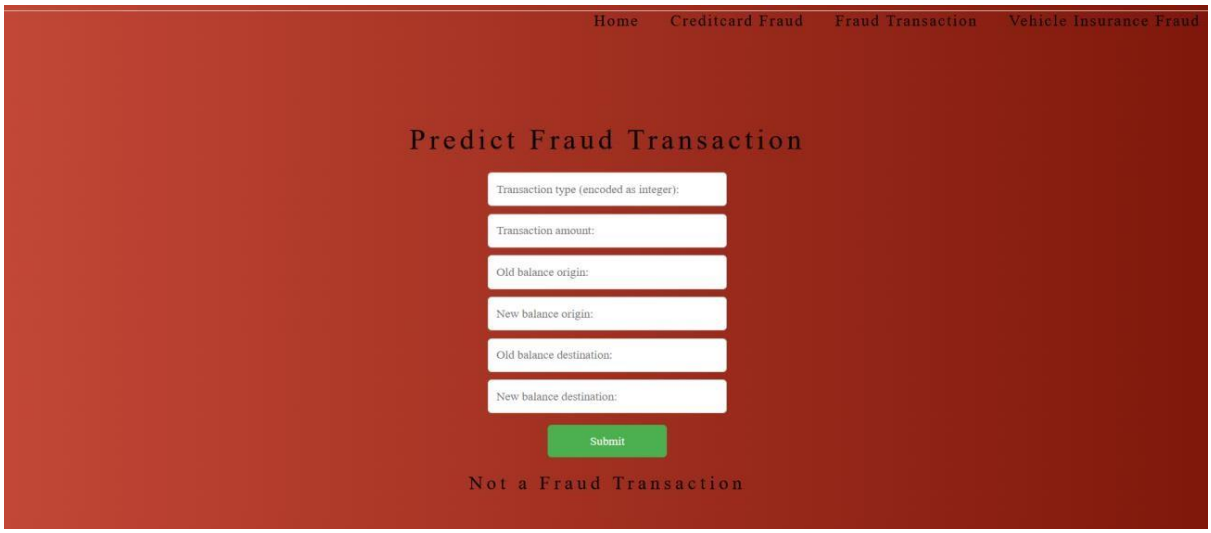
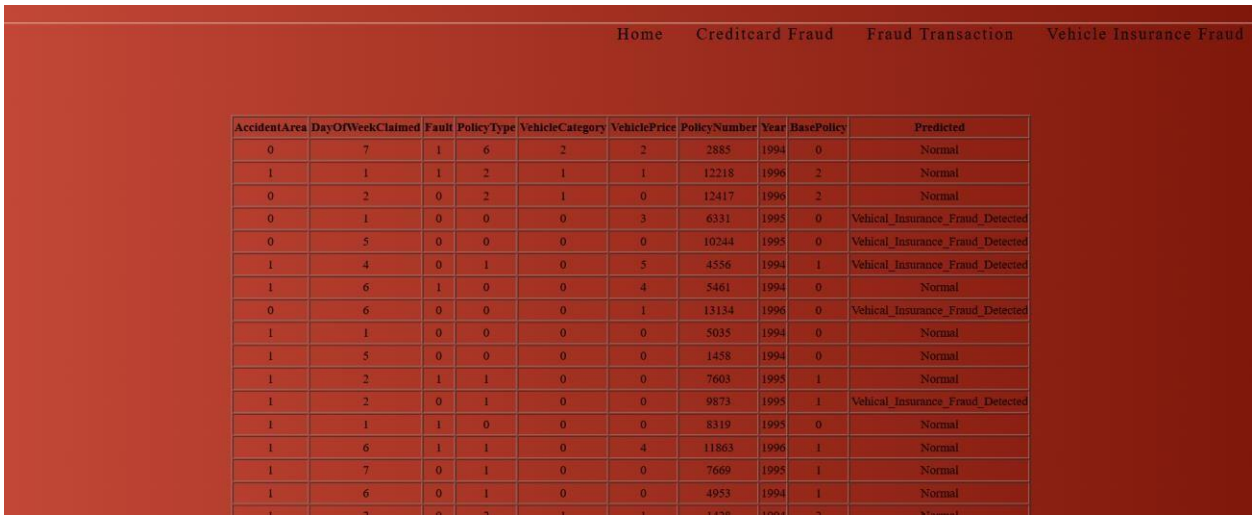


Figure 5: a credit card fraud detection analysis

The image shows a **fraud transaction prediction interface**, designed to evaluate whether a given financial transaction is fraudulent or not. Users are prompted to enter six key inputs: the **transaction type**, **transaction amount**, **old and new balances** for both the **origin** and **destination** accounts. After entering these values, the user can click the **"Submit"** button, which triggers the backend machine learning model to assess the risk. In this example, the output at the bottom states **"Not a Fraud Transaction"**, indicating that the submitted details were classified as legitimate. The top navigation bar also suggests that the system supports multiple fraud detection modules, including credit card and vehicle insurance fraud.



AccidentArea	DayOfWeekClaimed	Fault	PolicyType	VehicleCategory	VehiclePrice	PolicyNumber	Year	BasePolicy	Predicted
0	7	1	6	2	2	2885	1994	0	Normal
1	1	1	2	1	1	12218	1996	2	Normal
0	2	0	2	1	0	12417	1996	2	Normal
0	1	0	0	0	3	6331	1995	0	Vehical_Insurance_Fraud_Detected
0	5	0	0	0	0	10244	1995	0	Vehical_Insurance_Fraud_Detected
1	4	0	1	0	5	4556	1994	1	Vehical_Insurance_Fraud_Detected
1	6	1	0	0	4	5461	1994	0	Normal
0	6	0	0	0	1	13134	1996	0	Vehical_Insurance_Fraud_Detected
1	1	0	0	0	0	5035	1994	0	Normal
1	5	0	0	0	0	1458	1994	0	Normal
1	2	1	1	0	0	7603	1995	1	Normal
1	2	0	1	0	0	9873	1995	1	Vehical_Insurance_Fraud_Detected
1	1	1	0	0	0	8319	1995	0	Normal
1	6	1	1	0	4	11863	1996	1	Normal
1	7	0	1	0	0	7669	1995	1	Normal
1	6	0	1	0	0	4953	1994	1	Normal

Figure 6: vehicle insurance fraud detection dashboard

The image displays a vehicle insurance fraud detection dashboard that lists transaction data along with fraud prediction outcomes. Each row in the table represents an insurance claim, with columns indicating attributes like Accident Area, Day of Week Claimed, Fault, Policy Type, Vehicle Category, Vehicle Price, Policy Number, Year, and Base Policy. The final column, Predicted, shows the result generated by the fraud detection model, classifying each claim as either "Normal" or "Vehicle_Insurance_Fraud_Detected." This interface allows users

Doi:10.48047/ijiee.2025.15.7.2

to analyze patterns in historical claims and easily identify suspicious entries, aiding insurance companies in reducing fraudulent payouts and improving decision-making accuracy.

6. Conclusion

This project successfully presents a comprehensive hybrid fraud detection system that leverages domain-specific modeling techniques to address three critical and distinct categories of financial fraud**:** credit card fraud**,** transactional fraud**, and** vehicle insurance fraud. Each of these categories was treated with specialized data preprocessing, class balancing strategies, and machine/deep learning models tailored to their unique characteristics. The credit card fraud component demonstrated how an MLP classifier**,** combined with SMOTE and standardization techniques, can effectively detect imbalanced fraudulent patterns and achieve reliable classification performance. For general transaction fraud, a custom deep learning architecture built using TensorFlow's Functional API processed a combination of structured and categorical textual data (like customer IDs) to capture deeper transaction semantics. Meanwhile, the vehicle insurance fraud detection relied on the AdaBoost ensemble model, providing high interpretability and robust classification on tabular, label-encoded data.

Throughout the project, key performance metrics such as accuracy, confusion matrix, and AUC score **were used to validate the models. The use of** early stopping, class weighting, and **visualization tools** further contributed to optimizing model training and ensuring generalizability. The system also maintains **modularity**, allowing each fraud detection component to function independently while contributing to a unified fraud intelligence framework. Overall, the hybrid architecture built in this project proves to be a scalable, explainable, and effective solution for multi-domain fraud detection, offering a solid foundation for deployment in real-world financial security systems.

References

[1] Herland, M., Khoshgoftaar, T. M., & Wald, R. (2014). A review of data mining using big data in health informatics. Journal of Big Data, 1(1), 2. doi:10.1186/2196111512

[2] Pernička, F., & Paralic, J. (2017). Healthcare Fraud Detection Using Machine Learning Techniques: A Systematic Review. IEEE Access, 5, 2357723595. doi:10.1109/access.2017.2766486

[3] Kantarcioglu, M., & Clifton, C. (2008). Privacypreserving distributed mining of association rules on horizontally partitioned data. IEEE Transactions on Knowledge and Data Engineering, 20(9), 12311247. doi:10.1109/tkde.2008.31

[4] Rosenbaum, S., & Shi, Y. (2015). Healthcare fraud detection: A survey and a clustering model based on selforganizing map. Health Information Science and Systems, 3(1), 118. doi:10.1007/s1375501500227

[5] Yarkoni, T. (2019). The generalizability crisis. Communications of the ACM, 62(6), 3639. doi:10.1145/3328510

[6] National Health Care AntiFraud Association (NHCAA). (2020). The Challenge of Health Care Fraud. https://www.nhcaa.org/resources/healthcareantifraudresources/thechallengeofhealthcarefraud/

[7] Healthcare Information and Management Systems Society (HIMSS). (2019). Fraudulent Healthcare Claims Detection Using Artificial Intelligence. Retrieved from

Doi:10.48047/ijiee.2025.15.7.2

https://www.himss.org/resources/fraudulenthealthcareclaimsdetectionusingartificialintelligence
[8] Centers for Medicare & Medicaid Services (CMS). (n.d.). Fraud and Abuse. Retrieved from https://www.cms.gov/OutreachandEducation/MedicareLearningNetworkMLN/MLNProducts/downloads/Fraud_and_Abuse.pdf
[9] Chaudhry, S. A., & Kanungo, A. (2013). A survey of data mining techniques for fraud detection in healthcare domain. International Journal of Advanced Research in Computer Science and Software Engineering, 3(6), 700704.
[10] Healthcare Fraud Prevention Partnership (HFPP). (n.d.). About HFPP. Retrieved from https://hfpp.cms.gov/

Doi:10.48047/ijiee.2025.15.7.2