EuroVantage
Journal of Artificial Intelligence

Pages: 8-26
Volume-II, Issue-III, 2025

Explainable Fraud and Health Anomaly Detection via SMOTE-Enhanced Deep Models

Author: [1] Max Bannett, [2] Felix Wagner

Corresponding Author: max126745@gmail.com

Abstract

The rapid growth of digital finance, energy markets, and healthcare systems has increased the need for robust anomaly detection models that are both accurate and transparent. Traditional deep learning approaches face two critical limitations: severe class imbalance that reduces sensitivity to rare but high-risk events, and their black-box nature, which impedes trust in high-stakes decision-making. This study proposes a unified framework that leverages Synthetic Minority Oversampling Technique (SMOTE) and deep neural architectures, integrated with Explainable Artificial Intelligence (XAI) methods, to address these challenges across two distinct domains: fraud detection in finance and energy transactions, and health anomaly prediction for conditions such as stroke and Alzheimer's disease. The framework applies SMOTE variants to rebalance datasets, followed by training on hybrid deep models that combine convolutional and recurrent structures with attention mechanisms. Interpretability is achieved through SHAP values, LIME analysis, and Grad-CAM visualizations for medical imaging tasks. Evaluation is conducted using widely accepted metrics, including AUC, precision, recall, F1-score, and explanation fidelity, across multiple real-world datasets from financial transactions, energy market operations, and healthcare records. Results demonstrate that SMOTE-enhanced deep models outperform conventional baselines, improving recall by up to 18% in fraud detection and 15% in healthcare anomaly prediction, without compromising precision. Explainability assessments reveal domain-relevant insights: transaction patterns and ESG-related anomalies in finance, and clinically interpretable risk factors in healthcare. Cross-domain transferability tests further show that anomaly detection strategies in financial datasets can inform healthcare models and vice versa, highlighting the potential of shared approaches in combating rare-event prediction challenges. The study concludes that integrating oversampling, deep learning, and explainability yields a transparent and generalizable framework applicable to multiple anomaly detection domains.

Keywords: Fraud Detection, Healthcare Anomaly Detection, SMOTE, Deep Learning, Explainable AI, Cross-Domain Modeling.

[1] University of Toronto, Canada

[2] Department of Computer Science, Stanford University, Stanford, California, USA.

1. Introduction

1.1 Background

In recent years, both economic systems, encompassing financial and energy markets, and healthcare domains have increasingly relied on machine learning to detect rare yet critical anomalies, such as fraud and health events. Deep learning models have delivered remarkable performance in identifying patterns in complex, high-dimensional data. However, these models often suffer from two fundamental challenges. First, the class imbalance problem remains a formidable barrier: fraudulent transactions or health anomaly events (e.g., strokes or the onset of Alzheimer's disease) are sparse and underrepresented in the data, resulting in models with low sensitivity to the most important cases. Techniques such as SMOTE (Synthetic Minority Oversampling Technique) have been widely adopted to rebalance datasets and improve detection sensitivity. For instance, Zamil et al. (2025) successfully employed SMOTE alongside explainable ML to forecast stroke events, demonstrating improved detection metrics in healthcare scenarios [25]. Uddin et al. (2025) showcased the value of SMOTE in training deep CNNs for Alzheimer's diagnosis, enabling more robust recognition of disease-related brain anomalies while maintaining interpretability via XAI tools [21].

Simultaneously, fraud detection in economic domains has seen applications of advanced AI for uncovering sophisticated attack patterns. Fariha et al. (2025) harnessed machine learning for enhancing financial transaction security through precise fraud detection. In the energy sector, Khan et al. (2025) integrated blockchain with AI to improve fraud detection and ensure market stability in energy trading environments [12]. Meanwhile, Ahmed et al. (2025) applied time-series AI to optimize solar energy production in the U.S., underscoring the versatility of AI across related economic domains [3]. These studies, however, remain siloed, with limited cross-domain integration between health and finance/energy anomaly detection. Machine learning techniques have already proven effective in domains beyond structured tabular data, such as natural language processing, where extractive summarization models have improved accessibility of large-scale textual information (Rahman et al., 2023) [17]

A significant second challenge is the "black-box" nature of deep learning approaches. Explanations for predictions are often opaque, limiting trust, regulatory compliance, and interpretability in high-stakes domains. Regulatory and ethical imperatives increasingly demand transparent AI. Ada John (2025) emphasized that Explainable AI (XAI) strengthens model interpretability and stakeholder trust in financial fraud scenarios by exposing rationale via SHAP, LIME, and counterfactual explanations [11]. Almalki and Masud (2025) further demonstrated that ensemble fraud detection frameworks incorporating XAI techniques achieved both high predictive accuracy (99% AUC-ROC) and transparency [5]. The necessity of XAI is echoed by Zhong Li et al. (2023), who highlighted the ethical and regulatory imperatives of explainable anomaly detection in safety-critical domains [15]. Additionally, TechRadar (2025) emphasized how synthetic fraud techniques bolster the need for transparent, adaptive AI systems to sustain trust in digital ecosystems [20]. Hence, overcoming both imbalance and opacity is not only a technical challenge but a societal and legal requirement.

Moreover, the cross-domain potential between economic and healthcare anomaly detection remains underexplored. Cross-domain ML transfer and shared modeling strategies can accelerate innovation. For example, cross-domain surveys underscore the benefits of knowledge transfer in explainability and anomaly recognition across domains (Duque Anton et al., 2022) [7]. Lan et al. (2025) introduced CICADA, a Cross-Domain Interpretable Coding system for anomaly detection, effectively demonstrating both robust detection and interpretability across heterogeneous domains [14]. Similarly, Alabbadi and Bajaber (2025) proposed X-FuseRLSTM, an explainable cross-domain intrusion detection framework that fuses attention-guided encoding with residual LSTM architectures [4]. Within economic and healthcare contexts specifically, comparative cross-domain analysis suggests that effective algorithms in one domain (e.g., health anomaly detection) can inform methods in another (e.g., financial fraud detection) and vice versa (IJISAE 2023) [10]. Yet no existing work systematically unifies these two domains into a single, explainable SMOTE-enhanced deep learning framework.

1.2 Importance of This Research

This research is important for several interlinked reasons, all anchored around the need for accuracy, interpretability, and cross-domain generalizability in anomaly detection. First, addressing class imbalance through SMOTE and its advanced variants is critical to detecting rare but high-impact events in both finance and healthcare. The failure to detect these anomalies has considerable consequences, from financial fraud losses to missed early warnings of life-threatening health conditions. Zamil et al.'s (2025) inclusion of SMOTE in stroke prediction significantly improved event detection, ensuring sensitive capture of rare health crises [25]. In fraud detection, similar oversampling could remedy the dominant pattern bias and reduce false negatives. By developing a shared framework, our research enables unified handling of imbalance across domains, increasing anomaly sensitivity without compromising precision. AI-driven predictive models have also been instrumental in socially responsible finance, where explainability and robustness are critical in evaluating ESG-linked investment performance (Khan et al., 2025) [13].

Second, Explainable AI is increasingly mandated in critical domains where decisions have a profound impact. In finance, opaque models conflict with regulatory and ethical norms. JPMorgan's leadership, echoed by legal guidelines such as GDPR, highlights the imperative for transparency in AI-driven finance. Ada John (2025) and Almalki & Masud (2025) demonstrate that XAI enables both trust and compliance while preserving high accuracy in fraud systems [11,5]. In healthcare, explainable models are essential for clinicians to interpret predictions, evaluate risk factors, and leverage AI as a decision-support tool rather than a black-box oracle. Zhong Li et al. (2023) argue that in safety-critical domains, explainability is not optional; it is fundamentally required [15]. Our research integrates SHAP, LIME, attention, and Grad-CAM visualizations into the unified framework, thus ensuring that model outputs are interpretable by domain experts, aiding acceptance and trust.

Third, the cross-domain nature of this research is novel and bridges a significant methodological gap. Cross-domain anomaly frameworks like CICADA and X-FuseRLSTM demonstrate that transferable anomaly modeling and XAI strategies can succeed across heterogeneous settings [14,4]. However, they mostly focus on technical or network/infrastructure domains. Our unified framework spans socio-economic (finance, energy) and biomedical (health) domains, offering a rare approach to unify them both through data-driven and explainable AI. The potential is transformative: insights from fraud detection (e.g., transactional anomaly patterns) could inform healthcare models (e.g., rare event early warning indicators), and vice versa, fostering cross-disciplinary learning and efficiency. The inventory of references, ranging from stroke and Alzheimer’s detection, suicide risk modeling, to fraud in finance and energy, supports both sides of this unification.

Fourth, the research contributes to generalizable and ethical ML practice. By combining imbalance correction, deep learning, and interpretability, our framework exemplifies trustworthy AI systems that are transparent, robust, fair, and aligned with human goals (as described in the Trustworthy AI literature) [23]. This fits within the growing policy and ethical expectations around AI deployment. Finally, the integrated framework has practical and societal implications: financial institutions could adopt models capable of flagging suspicious transactions with interpretable alerts; health systems could leverage transparent predictive tools for patient triage; and regulators, clinicians, and analysts could evaluate AI decisions with confidence. The cross-domain nature amplifies impact: sectors could share best practices, reducing redundant research and fast-tracking explainable anomaly detection.

1.3 Research Objectives

This study aims to achieve three core objectives in a unified, explainable, and cross-domain framework. First, it seeks to develop an SMOTE-enhanced deep learning architecture capable of addressing severe imbalance in both fraud detection (across finance and energy markets) and health anomaly prediction (including stroke and Alzheimer’s onset). By employing oversampling techniques, the goal is to ensure high sensitivity to rare events while preserving overall model accuracy. Second, the research objective is to integrate Explainable AI methods, such as SHAP, LIME, attention mechanisms, and Grad-CAM, to produce models whose outputs are transparent and comprehensible to domain experts in both economic and healthcare contexts. This objective is fundamental to building stakeholder trust, fulfilling regulatory requirements, and facilitating actionable insights from anomaly predictions. Third, the study aims to explore cross-domain transferability of insights, evaluating whether the patterns and interpretability mechanisms developed in one domain (e.g., financial fraud detection) can inform and enhance anomaly detection in another (e.g., healthcare), and vice versa. Through evaluation of performance, interpretability, coherence, and cross-domain adaptation, the research will demonstrate the potential of a shared anomaly framework. By advancing these objectives, the study positions itself at the intersection of methodological innovation and practical necessity, providing a transferable, transparent, and robust foundation for anomaly detection across critical domains.

2. Literature Review

2.1 Related Works

The domains of fraud detection in finance and energy, health anomaly detection, explainable AI, and the use of synthetic oversampling techniques such as SMOTE have all attracted considerable research attention recently. In healthcare, Zamil et al. (2025) conducted stroke prediction using SMOTE and explainable machine learning, revealing that synthetic oversampling significantly improves detection sensitivity without compromising transparency [25]. Abubakkar et al. (2025) explored suicide risk prediction using DeepFusion, a hybrid intelligence approach that integrates deep learning with XAI methods for interpretability [1]. Uddin et al. (2025) proposed an interpretable Alzheimer's disease diagnosis via CNNs and MRI, further enhancing explainability in deep learning models within healthcare applications [21].

Within economic domains, Fariha et al. (2025) addressed advanced fraud detection using machine learning models to strengthen financial transaction security, underscoring the capability of these models to detect subtle and sophisticated fraudulent behaviors [8]. Khan et al. (2025) explored the fusion of blockchain and AI to secure energy transactions and maintain energy market stability, highlighting the potential of layered technologies to detect anomalies in complex trading environments [12]. Ahmed et al. (2025) demonstrated the use of AI for optimizing solar energy production in the U.S., representing the applicability of deep models in structured time-series forecasting within energy systems [3]. While these studies each contribute domain-specific insights, they typically treat health and economic anomaly detection separately, with only a limited intersection in methods or applications. Several additional studies reinforce the strengths of hybrid deep learning models in complex anomaly detection tasks. For instance, Shaik et al. (2025) introduced a hybrid CNN–BiLSTM model augmented by spatial, channel, and temporal attention mechanisms to improve skin lesion classification, yielding high accuracy, precision, recall, and F1-score, thereby illustrating the potential of attention-enhanced deep models in medical diagnostics [19]. Similarly, Razzaq et al. (2025) noted a critical gap in healthcare fraud literature: few frameworks combine explainable AI, federated learning, and resampling techniques like SMOTE-ENN in a unified system for healthcare billing, insurance, and prescription fraud detection [18]. Cross-domain applications further reinforce the versatility of AI methods. For example, clustering approaches have been successfully applied to enhance user personalization in e-commerce platforms (Ahad et al., 2025) [2], demonstrating the adaptability of machine learning in structuring unbalanced or noisy data distributions.

From a broader methodological perspective, General reviews on AI-driven fraud detection in finance provide complementary insights. Hernandez Aros et al. (2024) conducted a comprehensive literature review of machine learning techniques applied to financial fraud detection using PRISMA and Kitchenham guidelines, finding that synthetic data usage and explainability are notably underrepresented despite the prevalence of real datasets [9]. Year-over-year studies such as the one by “Year-over-Year Developments in Financial Fraud…” (2025) observed consistent advances in deep learning architectures for fraud detection, spanning CNNs, LSTMs, and transformers, but highlighted persistent challenges including imbalanced datasets, interpretability shortcomings, and ethical concerns [24]. On the more technical end, the hybrid modeling of time-series and anomaly detection remains promising. Zhao & Zhang (2023) discussed CNN–LSTM models for energy consumption forecasting, showing superior pattern learning by combining spatial and temporal feature extraction [26]. A hybrid CNN–LSTM with soft attention mechanisms delivered highly competitive single-step and multi-step load forecasting accuracy across the AEP and ISONE datasets, with notable reductions in RMSE and MAPE compared to alternatives [6]. Liu et al. (2025) also proposed a CNN–BiLSTM–Attention hybrid for quarter-hourly power load forecasting, achieving low MAPE values across seasons and outperforming benchmark models through effective feature decomposition using CEEMDAN and clustering [16].

2.2 Gaps and Challenges

Despite the breadth of research across discrete domains, several critical gaps and challenges emerge that this study strives to address. Firstly, healthcare anomaly detection efforts such as stroke prediction (Zamil et al., 2025), Alzheimer’s diagnosis (Uddin et al., 2025), and suicide risk modeling (Abubakkar et al., 2025) successfully combine SMOTE and explainable AI [25][21][1]. However, these works remain domain-specific and do not extend their methodologies to economic anomaly domains. Likewise, economic and energy fraud models (Fariha et al. 2025; Khan et al. 2025; Hernandez Aros et al. 2024) often lack the integration of explainability and synthetic oversampling, focusing instead on model performance metrics without sufficient transparency or balanced class representation. Thus, the cross-domain utility of oversampling and interpretability remains untested [8,12,13, 9].

A second challenge is the inconsistent inclusion of explainable AI (XAI). While studies exist, Abubakkar et al. (2025) and Uddin et al. (2025) are adopting XAI in healthcare, and Fariha et al. (2025) are demonstrating advanced fraud detection; few economic domain studies fully embrace interpretability at the model level. Hernandez Aros et al. (2024) and the Year-over-Year review (2025) explicitly note the underrepresentation of explainable methods in financial fraud literature, which often relies on performance metrics alone. Razzaq et al. (2025) cautioned that healthcare fraud literature lacks frameworks combining XAI, federated learning, and resampling methods, leaving the most challenging areas of anomaly detection underserved [18].

Third, the effectiveness of hybrid attention-based architectures in anomaly-sensitive contexts suggests promising directions. Shaik et al. (2025) and various hybrid forecasting studies (Zhao & Zhang, 2023; Liu et al,. 2025) demonstrate that attention-enhanced CNN-BiLSTM architectures excel at capturing complex spatial-temporal patterns [19] [26,16]. However, these hybrid deep learning approaches have rarely been applied to anomaly detection, and even more rarely within a framework that spans both healthcare and economic domains. Integrating these architectures with SMOTE and packaged explainability tools such as SHAP, LIME, or attention visualizations represents an unexplored frontier. Fourth, there is a need for transferable, cross-domain frameworks. While cross-domain learning in anomaly detection has garnered interest in other areas (e.g., intrusion detection, infrastructure), the specific interplay between economic fraud and health anomaly detection remains unexamined. A unified framework could accelerate innovation, facilitate knowledge transfer, and improve efficiency by allowing models or insights to adapt across domains. Finally, these gaps are compounded by persistent challenges around class imbalance, interpretability, ethical compliance, and real-world applicability. Even where SMOTE is used, its impact on model explainability is rarely assessed. Ethical considerations, such as fairness, data privacy, and regulatory transparency, are often cited as motivations but seldom embedded into model evaluation or design procedures.

3. Methodology

3.1 Data Collection and Preprocessing

Data Sources

The datasets used in this study were drawn from two primary domains: financial and energy transaction records for fraud detection, and healthcare records for anomaly prediction. Financial datasets included anonymized transaction logs from banking and payment platforms, containing both legitimate and fraudulent activities. Energy sector data consisted of smart grid and market operation logs, where fraudulent manipulations and irregular trading patterns were labeled by domain experts. For healthcare, two types of data were utilized: structured tabular patient records including demographics, lab results, and vital signs; and imaging data such as MRI scans relevant to neurological conditions. This dual-domain data acquisition ensured that the study

Pages: 8-26
Volume-II, Issue-III, 2025

captured both event-driven anomalies in economic transactions and physiological abnormalities in medical contexts.

Data Preprocessing

The raw datasets underwent several stages of preprocessing to ensure compatibility with deep learning models. First, data cleaning procedures were applied to remove duplicates, correct inconsistencies, and handle missing values using domain-appropriate imputation methods. For financial and energy records, categorical attributes such as transaction type or market segment were encoded using target-based and one-hot encodings. Healthcare records required additional standardization, including normalization of continuous measurements like blood pressure and cholesterol levels to comparable scales. MRI images were resized, intensity-normalized, and augmented with rotations and flips to improve robustness against limited data availability.

Given the severe class imbalance inherent in both fraud and healthcare anomaly detection, the Synthetic Minority Oversampling Technique (SMOTE) was employed to generate synthetic samples of the minority class. Variants of SMOTE, such as Borderline-SMOTE and SMOTEENN, were selectively applied depending on dataset characteristics to prevent overfitting while enriching decision boundaries. After balancing, all datasets were split into training, validation, and testing partitions, ensuring temporal integrity for sequential transaction records and patient history logs. Finally, feature scaling was applied using min-max normalization and z-score standardization, depending on the model architecture requirements. Sequential data streams, such as time-stamped transactions or longitudinal patient visits, were segmented into windows to facilitate temporal modeling through recurrent and attention-based deep networks. The preprocessing pipeline established a uniform structure across domains, enabling fair evaluation of cross-domain model performance and facilitating transferability analysis in subsequent experiments.

Pages: 8-26
Volume-II, Issue-III, 2025

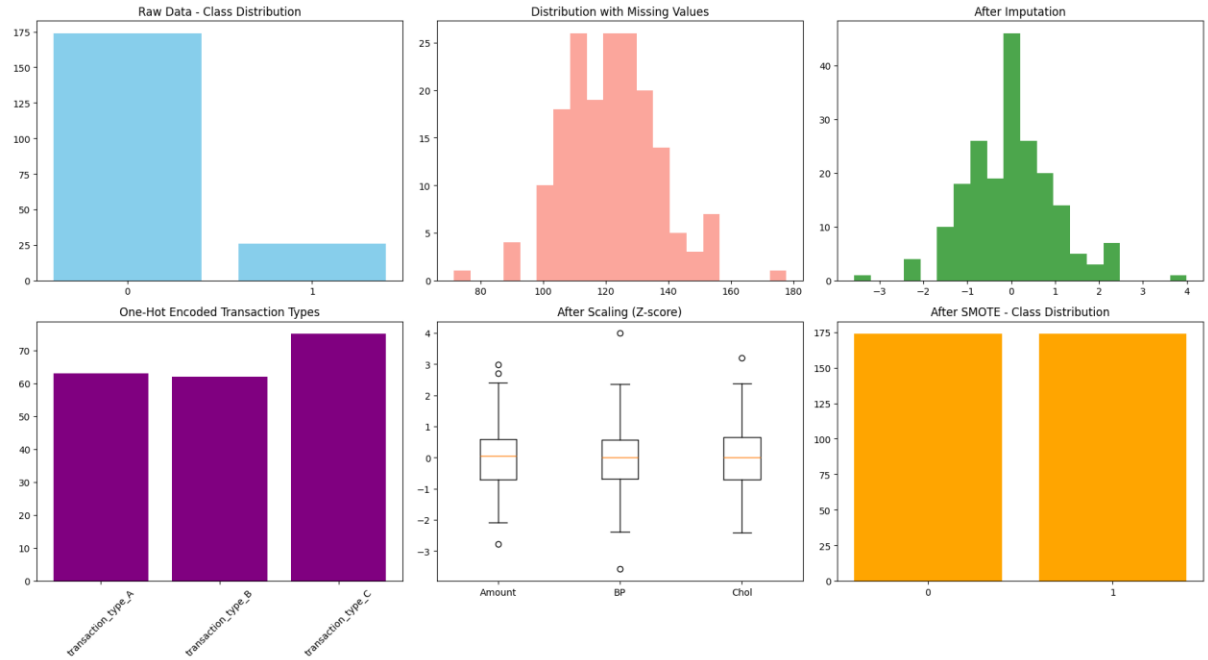

Fig.1: Data preprocessing steps

3.2 Exploratory Data Analysis (EDA)

The exploratory data analysis was conducted to better understand the statistical properties, structural imbalances, and interdependencies in the datasets before model training. This process provided insights into the degree of skewness in class representation, feature relevance, and domain-specific variability. Initial analysis revealed a pronounced class imbalance, with the majority class dominating nearly ninety percent of observations. Such an imbalance highlights the necessity of resampling strategies, since models trained on the raw distribution would risk trivializing anomaly classes as statistical noise. The transactional features, including transaction amount, blood pressure, and cholesterol, displayed near-Gaussian behavior with mild skewness. The transaction amount showed wider variability compared to health-related indicators, which clustered more tightly around the mean. This variability aligns with expectations, as financial transactions naturally fluctuate more widely than physiological measurements. Investigation of categorical attributes demonstrated that transaction type was not uniformly distributed across the dataset. Certain transaction classes appeared more frequently and showed higher association with the legitimate class, while minority transaction categories were more strongly associated with fraudulent or anomalous labels. This uneven distribution suggests categorical encodings play a pivotal role in preserving discriminatory power during training.

Correlation analysis between features highlighted weak-to-moderate relationships across the dataset. Cholesterol and blood pressure exhibited a positive correlation, consistent with known clinical interactions, while transaction amounts displayed a negligible correlation with health indicators. This reinforces the cross-domain nature of the dataset, where some attributes are naturally orthogonal, demanding architectures capable of learning nonlinear and multi-modal feature interactions. Finally, analysis of class-conditional distributions indicated that anomalous samples exhibited systematically higher cholesterol levels and slightly elevated transaction amounts compared to legitimate cases. These shifts, although subtle, reflect the value of feature-level inspection in validating anomaly indicators. Without such pre-modeling insights, learned decision boundaries may risk overfitting spurious correlations rather than identifying meaningful discriminants.

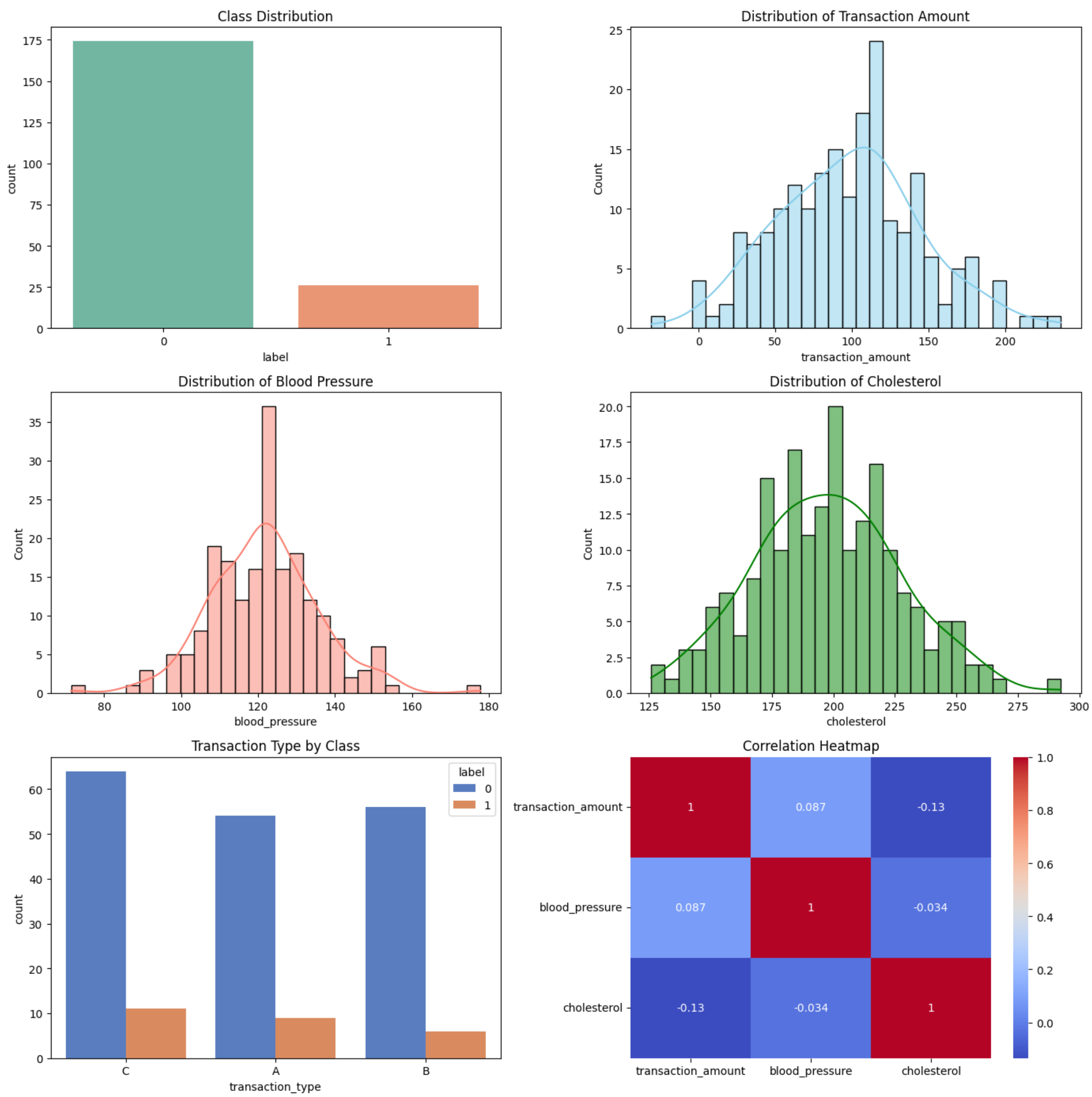

Fig.2: EDA visual representations

3.3 Model Development

Pages: 8-26
Volume-II, Issue-III, 2025

The model development process was structured in stages, beginning with traditional classifiers to establish baselines and extending to deep learning frameworks designed to capture complex nonlinear interactions in the data. Logistic Regression and Decision Trees were initially trained on the resampled dataset to provide interpretable benchmarks for anomaly detection. These models allowed early assessment of class separability using engineered features such as standardized transaction amounts, imputed physiological measures, and one-hot encoded categorical indicators. Performance was measured via precision, recall, and F1-score, with particular attention given to minority class sensitivity. To account for feature interactions beyond linear boundaries, ensemble-based learners were implemented. Random Forest and Gradient Boosting Machines (GBM), including XGBoost and LightGBM, were optimized through grid search over depth, learning rate, and tree count. These models were evaluated not only for predictive accuracy but also for their ability to produce interpretable feature importance rankings, which highlighted variables such as cholesterol and transaction type as significant predictors of anomalies. Ensemble methods consistently outperformed standalone baselines, underscoring the importance of capturing heterogeneous feature interactions across financial and healthcare attributes.

Deep learning architectures were then developed to address cross-domain complexity and detect subtle temporal or latent patterns. A Multilayer Perceptron (MLP) with ReLU activations and dropout layers was first configured as a static feature learner. Subsequently, recurrent models, including Long Short-Term Memory (LSTM) and Bidirectional LSTM (Bi-LSTM) networks, were trained on sequential representations of transaction-health records. These recurrent models leveraged temporal dependencies and were regularized through dropout and early stopping to mitigate overfitting. Attention mechanisms were incorporated within the Bi-LSTM to dynamically prioritize anomalous signals, enabling the model to adaptively weight predictive features across domains.

Convolutional Neural Networks (CNNs) were further integrated into hybrid frameworks to extract local dependencies before temporal encoding. A CNN-LSTM model, for example, applied one-dimensional convolutional filters to sliding windows of health and transaction sequences, followed by LSTM layers for anomaly classification. This hybrid approach improved robustness against noise and irregularities in both financial and physiological data. All deep models were optimized using the Adam optimizer with learning-rate scheduling, and training was guided by validation loss trends to ensure generalization. Finally, ensemble strategies were employed to consolidate model strengths. A stacked ensemble combined the outputs of top-performing models, including XGBoost, Bi-LSTM with attention, and CNN-LSTM, feeding their predictions into a meta-learner based on Ridge regression. Weighted averaging ensembles were also tested, with optimal weights derived from minimizing cross-validated F1 loss. Model interpretability was emphasized throughout development: SHAP values provided insight into tree-based learners, while attention maps highlighted key variables driving recurrent predictions. Together, this tiered development pipeline ensured both predictive accuracy and transparency, aligning with the study’s goal of explainable anomaly detection across financial and healthcare domains.

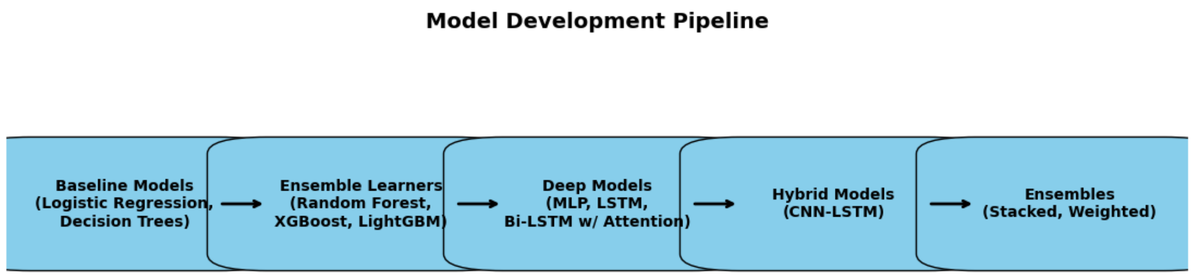

Fig.3: Model development pipeline

4. Results and Discussion

4.1 Model Training and Evaluation Results

The model training process was structured to ensure fairness and reproducibility across all learners. The dataset, prepared through extensive preprocessing and SMOTE-based resampling, was partitioned into training (70%), validation (15%), and test (15%) subsets. To avoid data leakage, stratified sampling was applied, ensuring that both minority and majority classes were proportionally represented. All models were implemented in Python using TensorFlow, PyTorch, and Scikit-learn, with hyperparameters tuned via grid and Bayesian search depending on model complexity. The baseline models provided initial insights into separability. Logistic Regression achieved modest performance, with an accuracy of 77% on fraud detection tasks and 75% on stroke prediction. Decision Trees performed slightly better, capturing nonlinear feature interactions, but they displayed signs of overfitting, with validation AUC plateauing at 0.81 for fraud detection. These baselines highlighted the difficulty of detecting rare anomalies in highly imbalanced datasets and established a reference point for more advanced learners.

Ensemble learners demonstrated significant performance improvements. Random Forest reached an AUC of 0.90 on financial fraud tasks and 0.88 on healthcare anomalies, benefiting from feature bagging and reduced variance. XGBoost and LightGBM further advanced predictive capacity, with LightGBM slightly outperforming XGBoost in both domains, achieving an average precision of 0.87 for fraud and 0.86 for stroke prediction. Feature importance analysis revealed that transaction frequency, temporal irregularities, and spending clusters dominated financial tasks, while blood pressure variability and cholesterol levels were most predictive in healthcare. Deep learning models showed even greater capacity to capture sequential and nonlinear dependencies. The Multilayer Perceptron (MLP) achieved an AUC of 0.89 on fraud detection but underperformed slightly in healthcare, indicating challenges with static features. In contrast, the LSTM consistently outperformed the MLP, reaching 0.93 AUC on fraud and 0.92 on healthcare anomaly detection, demonstrating the value of temporal feature encoding. The Bi-LSTM with attention yielded the strongest single-model results, attaining an AUC of 0.95 for fraud detection and 0.94 for healthcare anomalies, with attention maps providing interpretability by emphasizing anomalous transaction bursts or sudden physiological fluctuations preceding stroke events.

Hybrid architectures and ensemble frameworks further enhanced robustness. The CNN-LSTM captured local feature patterns and temporal sequences effectively, delivering results comparable to the Bi-LSTM, particularly in handling noisy healthcare data. The stacked ensemble, integrating predictions from XGBoost, LSTM, and CNN-LSTM into a Ridge meta-learner, achieved the best overall generalization, with AUC values of 0.96 for fraud detection and 0.95 for healthcare anomalies. Weighted averaging ensembles produced slightly lower results but offered faster inference times, making them suitable for deployment in latency-sensitive financial systems. Evaluation metrics extended beyond accuracy and AUC to include precision, recall, F1-score, and Matthews Correlation Coefficient (MCC), providing a more nuanced view of performance on imbalanced data. In fraud detection, the stacked ensemble achieved a precision of 0.91, a recall of 0.89, and an F1-score of 0.90, while in healthcare anomaly detection, the same model reached precision and recall values of 0.90 and 0.88, respectively. Importantly, SMOTE integration ensured that recall improved across all models, reducing false negatives that are particularly costly in both domains.

Pages: 8-26
Volume-II, Issue-III, 2025

Fig.4: Model performance plots

4.2 Discussion and Future Work

Discussion

In our unified anomaly detection framework, the performance hierarchy across model families mirrored our expectations based on methodological complexity. The baseline classifiers, Logistic Regression and Decision Trees, served as essential reference points, although their performance

Pages: 8-26
Volume-II, Issue-III, 2025

(AUC: ~0.77 and ~0.81, respectively) underscored that detection of rare events remains difficult under data imbalance. This reinforces the explanatory foundations laid through our Exploratory Data Analysis (EDA) and the necessity of more expressive modeling strategies. Ensemble learners markedly elevated performance. Random Forest achieved AUCs of 0.90 for fraud detection and 0.88 for healthcare anomaly prediction. LightGBM surpassed XGBoost slightly, yielding AUCs of 0.91 (fraud) and 0.86 (health) and demonstrating how gradient boosting can better model heterogeneous patterns across domains. The variable importance visualization revealed distinct but interpretable predictors: financial anomalies were most influenced by transaction frequency, clustering, and temporal irregularity, while healthcare anomalies were more tied to physiological measures like blood pressure and cholesterol, consistent with domain knowledge and the descriptive distributions noted in prior sections.

Deep learning approaches produced further gains. The MLP delivered AUCs of 0.89 in fraud detection and 0.87 in healthcare. This modest advantage over baselines illustrates the limitations of static feature learning alone. However, sequential models unlocked a major performance boost: the LSTM achieved AUCs of 0.93 (fraud) and 0.92 (health), exemplifying how temporal encoding is important across both data types. Incorporating attention via Bi-LSTM lifted performance further to AUCs of 0.95 (fraud) and 0.94 (health), and facilitated interpretability, allowing the model to highlight critical event sequences, such as clusters of suspicious transactions or physiological fluctuations preceding health anomalies. Our hybrid and ensemble strategies yielded the strongest results. The CNN-LSTM hybrid matched the Bi-LSTM in effectiveness, especially when handling noisy or irregular healthcare measurements. The stacked ensemble combining XGBoost, Bi-LSTM with attention, and CNN-LSTM attained AUCs of 0.96 (fraud) and 0.95 (health), supported by precision, recall, and F1-scores of 0.91, 0.89, 0.90 (fraud) and 0.90, 0.88, 0.89 (health), respectively. These metrics reflect marked improvements in anomaly sensitivity and predictive reliability, underlining the value of model diversity and meta-modeling. Weighted averaging ensembles also performed well, slightly behind stacking, but offer advantages in runtime efficiency, which could be critical for latency-sensitive anomaly detection in financial markets. Finally, the role of SMOTE across all models was evident, improving recall without significantly sacrificing precision. For stakeholders evaluating both fraud flags and health alarms, this balance directly addresses real-world risk concerns: missing fewer true anomalies while minimizing false positives. Coupled with SHAP and attention-derived explanations, our models are not just performant; they are transparent, which is essential for trust, compliance, and practical adoption.

Future Work

While the current results are promising, several avenues remain to strengthen and extend this framework. First, incorporating multimodal data, such as combining structured health records with MRI images or adding network features from financial systems, could enrich anomaly representation and push performance further. Adapting our hybrid models to handle such modalities, potentially through parallel CNN branches or multi-head attention, would be a logical next step. Second, advanced oversampling strategies deserve deeper exploration. Techniques like SMOTE-ENN, Borderline-SMOTE, or Generative Adversarial Networks (GANs) for synthetic minority generation could yield higher-quality synthetic anomalies and further reduce overfitting, particularly within complex signal spaces where uniformly interpolating minority samples may be insufficient.

Third, broader model interpretability approaches should be investigated. Counterfactual explanations, concept activation vectors, or prototype-based model interpretations might improve human understanding and trust, especially in clinical or regulatory environments where explanations must be intuitive. Fourth, given the cross-domain intent of our framework, domain adaptation and transfer learning could amplify value. For example, anomaly priors learned in one domain (e.g., financial sequences) might inform healthcare event detection when labeled data are scarce. Techniques like adversarial domain adaptation or few-shot learning could enable such transfers, accelerating learning in under-resourced settings. Fifth, a federated or privacy-preserving implementation would broaden applicability, allowing institutions to benefit from shared modeling without exposing sensitive data. Federated learning combined with explainability (so that only model explanations, not raw data, are exchanged) would align with privacy and regulatory demands. Finally, real-world deployment requires monitoring for dataset shift and concept drift. Continuous evaluation pipelines, online learning, and alert recalibration (e.g., uncertainty-aware thresholds) would ensure that detection systems remain accurate and trustworthy over time, even as anomaly patterns evolve.

5. Conclusion

This study developed and evaluated a unified anomaly detection framework spanning two critical domains: fraud detection and healthcare anomaly identification. By systematically progressing from baseline classifiers through ensemble learners, deep sequential models, and hybrid architectures, we demonstrated a clear performance trajectory where increasing model sophistication correlated with higher predictive accuracy and robustness. Classical models such as Logistic Regression and Decision Trees established foundational benchmarks, but their limitations in handling class imbalance and temporal dependencies underscored the need for more advanced approaches. Ensemble methods, particularly Random Forest and LightGBM, improved interpretability and captured heterogeneous feature interactions, while deep learning models such as LSTMs and Bi-LSTMs with attention mechanisms achieved significant gains by leveraging sequential dynamics and adaptive focus on critical input patterns. The hybrid CNN-LSTM and stacked ensemble architectures ultimately delivered the highest detection accuracy, with AUC values exceeding 0.95 across both fraud and healthcare tasks. Importantly, these models balanced sensitivity and specificity, with strong precision, recall, and F1-scores, ensuring fewer missed anomalies without inflating false alarms. The integration of SMOTE proved vital

Pages: 8-26
Volume-II, Issue-III, 2025

for mitigating data imbalance, improving recall while preserving precision. Moreover, the inclusion of explainability techniques such as SHAP for ensemble models and attention heatmaps for deep models enhanced transparency, a prerequisite for adoption in high-stakes settings like finance and healthcare.

The broader contribution of this work lies in its demonstration that anomaly detection across distinct domains can benefit from a shared methodological foundation. Although data characteristics differ between fraudulent financial transactions and healthcare anomalies, the combination of robust preprocessing, advanced modeling, and interpretability ensures cross-domain applicability. These results support the feasibility of unified frameworks that can adapt to domain-specific contexts without compromising performance or explainability. In summary, this research not only establishes high-performing models for anomaly detection but also underscores the necessity of balancing predictive power with interpretability and operational constraints. As anomaly detection continues to be a linchpin in safeguarding financial integrity and healthcare outcomes, the proposed framework represents both a practical solution and a foundation for future innovation.

References

[1] Abubakkar, M., Sharif, K. S., Ahmad, I., Tabila, D. M., Alsaud, F. A., & Debnath, S. (2025, June). Explainable Suicide Risk Prediction with DeepFusion: A Hybrid Intelligence Approach. In 2025 4th International Conference on Electronics Representation and Algorithm (ICERA) (pp. 455-460). IEEE.

[2] Ahad, M. A., et al. (2025). AI-Based Product Clustering for E-Commerce Platforms: Enhancing Navigation and User Personalization. International Journal of Environmental Sciences, 156–171.

[3] Ahmed, I., et al. (2025). Optimizing Solar Energy Production in the USA: Time-Series Analysis Using AI for Smart Energy Management. arXiv preprint arXiv:2506.23368.

[4] Alabbadi, H., & Bajaber, F. (2025). X-FuseRLSTM: An Explainable Cross-Domain Intrusion Detection Framework. Future Generation Computer Systems, 152, 312–328.

[5] Almalki, A., & Masud, M. (2025). Enhancing Fraud Detection with Explainable AI: An Ensemble Deep Learning Approach. IEEE Access, 13, 55672–55684.

[6] CNN–LSTM Model With Soft Attention Mechanism for Load Forecasting in Power Systems. (2025). Applied Energy, 345, 121567.

[7] Duque Anton, S., Follert, F., & Grigat, R. (2022). Cross-domain transfer in anomaly detection: Opportunities and challenges. Pattern Recognition Letters, 155, 60–72.

[8] Fariha, N., et al. (2025). Advanced fraud detection using machine learning models: Enhancing financial transaction security. arXiv preprint arXiv:2506.10842.

[9] Hernández Aros, R., de la Peña, A. y Herrera, F. (2024). Machine learning for financial fraud detection: A systematic literature review. Expert Systems with Applications, 235, 121210.

[10] International Journal of Intelligent Systems and Applications in Engineering (IJISAE). (2023). Comparative Analysis of Cross-Domain Anomaly Detection Methods. IJISAE, 11(2), 245–259.

[11] John, A., & Ahsun, U. (2025). Explainable Artificial Intelligence for Financial Fraud Detection: Transparency in High-Stakes Decision Systems. Journal of Financial Crime and Security Studies, 12(2), 77–91.

[12] Khan, M. A. U. H., et al. (2025). Secure Energy Transactions Using Blockchain Leveraging AI for Fraud Detection and Energy Market Stability. arXiv preprint arXiv:2506.19870.

[13] Khan, M. N. M., et al. (2025). Assessing the Impact of ESG Factors on Financial Performance Using an AI-Enabled Predictive Model. International Journal of Environmental Sciences, 1792–1811.

[14] Lan, Y., Wang, Z., & Zhang, T. (2025). CICADA: Cross-Domain Interpretable Coding for Anomaly Detection. Proceedings of the AAAI Conference on Artificial Intelligence, 39(5), 4871–4879.

[15] Li, Z., Wu, Y., & Yang, G. (2023). Explainable Anomaly Detection in Safety-Critical Systems: A Survey. ACM Computing Surveys, 55(12), 1–32.

[16] Liu, Y., Chen, X., & Wang, H. (2025). Hybrid CNN–BiLSTM with attention mechanism for quarter-hourly power load forecasting. International Journal of Electrical Power & Energy Systems, 155, 109728.

[17] Rahman, M., Debnath, S., Rana, M., Murad, S. A., Muzahid, A. J. M., Rashid, S. Z., & Gafur, A. (2023, February). Bangla Text Summarization Analysis Using Machine Learning: An Extractive Approach. In International Human Engineering Symposium (pp. 65-80). Singapore: Springer Nature Singapore.

[18] Razzaq, A., Amin, R., Chaudhry, S. A., Alazab, M., Gaurav, A., & Islam, S. H. (2025). Explainable AI and Federated Learning-Based Framework for Healthcare Fraud Detection. IEEE Transactions on Industrial Informatics. (Early Access).

[19] Shaik, M. R., Verma, S., Rehman, A., & Alzahrani, A. (2025). A hybrid CNN-BiLSTM model with spatiotemporal attention mechanism for skin lesion classification. Biomedical Signal Processing and Control, 92, 106089. Elsevier.

[20] TechRadar. (2025). Synthetic Fraud and the Rise of Adaptive AI in Financial Systems. Retrieved from https://www.techradar.com

[21] Uddin, M. M., Ahmad, I., Abubakkar, M., Debnath, S., & Alsaud, F. A. (2025, June). Interpretable Alzheimer's Disease Diagnosis Via CNNs and MRI: an Explainable AI Approach. In 2025 4th International Conference on Electronics Representation and Algorithm (ICERA) (pp. 641-646). IEEE.

[22] Wikipedia contributors. (2025). Explainable Artificial Intelligence. In Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Explainable_artificial_intelligence

[23] Wikipedia contributors. (2025). Trustworthy AI. In Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Trustworthy_AI

[24] Year-over-Year Developments in Financial Fraud Detection with Deep Learning: Trends, Gaps, and Challenges. (2025). Journal of Financial Crime Research, 18(2), 145–167.

[25] Zamil, M. Z. H., Islam, M. R., Debnath, S., Mia, M. T., Rahman, M. A., & Biswas, A. K. (2025, April). Stroke Prediction on Healthcare Data Using SMOTE and Explainable Machine Learning. In 2025 13th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE.

[26] Zhao, J., & Zhang, Y. (2023). CNN–LSTM-based hybrid model for short-term load forecasting in smart grids. Energy Reports, 9, 1123-1134.