Journal of Economics, Finance and Accounting Studies
ISSN: 2709-0809
DOI: 10.32996/jefas
Journal Homepage: www.al-kindipublisher.com/index.php/jefas

JEFAS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Anomaly Detection in Financial Transactions Using Convolutional Neural Networks

Md. Tanvir Rahman Mazumder[1], Md. Shahadat Hossain Shourov[2], Iftekhar Rasul[3], Sonia Akter[4], Md Kauser Miah[5]

[1]*Master of Science in Information Technology, Washington University of Science and Technology, VA, USA*
[2]*Master of Arts in Information Technology Management, Webster University, MO, USA*
[3]*Information Technology Management, St. Francis College, NY, USA*
[4]*Master's in business Analytics, Mercy University, NY, USA*
[5]*Department of Computer and Information Science, Gannon University, PA, USA*
Corresponding Author: Md Kauser Miah, **E-mail**: miah003@gannon.edu

| ABSTRACT

The rise of digital financial systems has brought unprecedented convenience but has also exposed users and institutions to various fraudulent activities. Anomaly detection plays a critical role in ensuring financial security by identifying unusual transaction patterns that may indicate fraud or other irregularities. Traditional statistical and rule-based approaches, though effective to some extent, often fall short when dealing with the increasing volume and complexity of financial data. This study proposes a novel approach to anomaly detection in financial transactions using Convolutional Neural Networks (CNNs), a class of deep learning models primarily known for their success in image processing tasks. In this work, transactional data are preprocessed and transformed into structured formats suitable for CNN input. By treating sequences of financial transactions as temporal-spatial matrices, the CNN model learns intricate patterns that distinguish normal from anomalous behavior. Our methodology includes a comprehensive pipeline involving data normalization, feature engineering, and the construction of multi-channel representations to exploit CNNs' strength in hierarchical feature learning. We evaluate our model on benchmark financial datasets and compare its performance against traditional machine learning algorithms such as Support Vector Machines (SVM), Random Forest, and Logistic Regression. The CNN-based model demonstrates superior performance in terms of accuracy, precision, recall, and F1-score. Additionally, it shows robustness in detecting rare anomalies while minimizing false positives, a critical requirement in real-time financial fraud detection systems. The results indicate that CNNs can effectively capture both local and global dependencies within financial transaction sequences, making them suitable for large-scale and high-dimensional data environments. This study contributes to the growing body of research advocating for the adoption of deep learning techniques in financial anomaly detection and opens up possibilities for integrating CNNs with real-time monitoring systems for enhanced financial security. Future research may focus on hybrid models combining CNNs with recurrent layers to capture long-term dependencies more effectively.

| KEYWORDS

Anomaly Detection, Financial Transactions, Convolutional Neural Networks, Deep Learning, Fraud Detection, Transaction Patterns, Feature Engineering, Temporal Data Analysis.

| ARTICLE INFORMATION

ACCEPTED: 01 March 2025 **PUBLISHED:** 19 April 2025 **DOI:** 10.32996/jefas.2025.7.2.17

1. Introduction

1.1 Background and Motivation

The proliferation of digital financial services has revolutionized the global economy, enabling seamless online transactions, mobile banking, and e-commerce. However, this rapid digitization has concurrently exposed financial systems to increasing risks of fraud, money laundering, and other financial crimes. Financial institutions process millions of transactions daily, making manual inspection infeasible and underscoring the need for automated anomaly detection systems capable of identifying unusual patterns that may signify malicious behavior [1]. Traditional fraud detection approaches, including rule-based systems and classical machine learning algorithms, have been applied for decades. These methods often rely on predefined thresholds and historical data patterns, which limit their ability to adapt to evolving fraudulent strategies. Moreover, the high volume, velocity, and variety of financial data characteristic of big data environments necessitate scalable, adaptive, and robust models for real-time anomaly detection [2]. This has prompted a growing interest in applying deep learning techniques, especially Convolutional Neural Networks (CNNs), which are known for their exceptional pattern recognition capabilities.

1.2 Problem Statement

Detecting financial anomalies is inherently challenging due to several factors: the imbalance of datasets (where fraudulent transactions are rare), the subtlety and variability of fraud patterns, and the dynamic nature of user behavior. Most financial anomalies do not follow a fixed structure, making it difficult for traditional models to generalize across different scenarios [3]. Additionally, conventional machine learning models often require extensive feature engineering and domain-specific knowledge, which may not be scalable or transferable across datasets. While CNNs are conventionally used in image recognition tasks, their ability to capture spatial hierarchies makes them suitable for analyzing structured financial data, especially when transformed into 2D or multi-dimensional representations. By learning hierarchical features automatically, CNNs can reduce the dependency on manual feature extraction and potentially outperform traditional methods in detecting complex anomalies in transactional data [4].

1.3 Objectives and Contributions

The primary objective of this study is to develop a deep learning framework based on Convolutional Neural Networks for detecting anomalies in financial transactions. Our contributions are as follows:

1. We propose a data transformation approach to convert sequential transaction records into spatial matrix representations, enabling compatibility with CNN input structures.
2. We construct a CNN architecture tailored for financial data, capable of learning discriminative features that differentiate normal and anomalous transactions.
3. We evaluate the model using real-world benchmark datasets, comparing its performance against baseline models such as Random Forest, Support Vector Machine (SVM), and Logistic Regression.
4. We demonstrate that CNNs, even without recurrent structures, can effectively capture temporal dependencies and structural anomalies within financial data.

Through these contributions, we aim to establish CNN-based models as a viable and scalable alternative for modern fraud detection systems.

1.4 Structure of the Paper

The remainder of this paper is structured as follows: Section 2 provides a comprehensive literature review, highlighting existing methods in financial anomaly detection and the role of CNNs in non-visual domains. Section 3 describes the dataset, data preprocessing techniques, and the CNN model architecture. Section 4 presents experimental results, performance metrics, and a comparative analysis. Section 5 discusses the implications, limitations, and potential extensions of this work. Finally, Section 6 concludes the study with key findings and future research directions.

2. Literature Review

2.1 Traditional Techniques for Financial Anomaly Detection

Historically, anomaly detection in financial transactions has relied heavily on rule-based systems and statistical methods. Rule-based systems use predefined conditions to flag anomalies. While easy to implement and understand, these systems are inflexible and ineffective against adaptive fraud strategies [5]. Statistical models such as z-score analysis, clustering, and logistic regression have been widely used to identify outliers in structured financial data. However, these models often fail when the fraudulent behavior mimics legitimate transactions or when the data distribution changes over time [6].

2.2 Classical Machine Learning Approaches

The emergence of machine learning introduced more dynamic and adaptive models, such as Decision Trees, Random Forests, Support Vector Machines (SVM), and k-Nearest Neighbors (k-NN). These models provided improved accuracy and generalization by learning from labeled datasets. For instance, Bahnsen et al. [7] used cost-sensitive learning with Random Forests to handle class imbalance in fraud detection. However, traditional ML models still require extensive feature engineering and suffer from overfitting or underfitting in highly imbalanced datasets, which is a common trait in financial anomaly detection scenarios. In a recent study by Abir et al. [19], classical ML techniques were evaluated on EEG data for predicting psychiatric disorders. Although these models, particularly Random Forests and SVMs, demonstrated relatively good performance, they were outperformed by deep learning models in terms of accuracy. However, the authors observed that ML models had an advantage in terms of computational efficiency, completing the analysis in significantly less time. This finding reinforces the idea that classical ML methods can still offer practical value in scenarios where computational resources are limited or when quick predictions are needed [19].

2.3 Deep Learning for Anomaly Detection

Deep learning has revolutionized data analysis in domains such as computer vision, natural language processing, and time-series forecasting. Recently, its application in financial fraud detection has gained traction due to its ability to learn hierarchical features and detect complex non-linear patterns. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have been successfully used in modeling sequential financial data [8]. Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) have also been explored for unsupervised anomaly detection tasks in financial datasets [9]. However, many deep learning models are computationally expensive and sensitive to hyperparameter tuning. Moreover, they may not always generalize well across datasets without substantial data preprocessing and augmentation. Abir et al. [44] investigated the public health risks associated with undocumented immigration in the U.S. using predictive machine learning techniques. By analyzing data collected between 2018 and 2022, the study identified elevated risks of infectious diseases—such as tuberculosis and hepatitis—among undocumented populations. Utilizing models like CNN and KNN, the research achieved high performance (90% precision and recall), outperforming traditional classifiers like logistic regression and decision trees. Their findings underscore the utility of machine learning in epidemiological surveillance and policy development for vulnerable communities.

2.4 Convolutional Neural Networks in Non-Image Domains

Though CNNs are widely associated with image processing, their architectural ability to extract spatial hierarchies has found applications in various non-image tasks such as signal processing, NLP, and financial forecasting. CNNs can detect localized features and patterns in structured datasets by transforming transactional sequences into two-dimensional or multi-channel matrices [10]. Studies have shown that CNNs are particularly effective when financial time-series data are represented in a grid-like format that preserves temporal ordering and numerical features. Le Cun et al. [11] laid the groundwork for CNNs, while recent adaptations like Kim et al. [12] applied CNNs to sentiment analysis and time-series classification. In financial anomaly detection, CNNs offer advantages such as faster training times, reduced overfitting due to fewer parameters, and better spatial pattern recognition compared to traditional methods.

2.5 Comparative Analysis of Key Literature

To synthesize existing knowledge, Table 1 provides a comparative summary of relevant studies in financial anomaly detection, highlighting the algorithms used, dataset type, key contributions, and limitations. Akhter et al. [37] conducted a comprehensive study examining the relationship between private AI investment, financial globalization, and environmental sustainability in the United States from 1990 to 2019. The research employed the Autoregressive Distributed Lag (ARDL) model to analyze long- and short-term impacts, revealing that private investment in AI contributes positively to the country's Load Capacity Factor (LCF), an important measure of ecological sustainability. This indicates that beyond technical capabilities, AI investments can indirectly shape financial and environmental systems through intelligent optimization and automation. Importantly, their findings revealed that while AI investment enhanced sustainability**,** technological innovation and financial globalization had negative correlations with the load capacity factor**.** This demonstrates the complex relationship between financial mechanisms and environmental outcomes in an AI-driven economy. Moreover, using Granger Causality tests, they found a unidirectional causal relationship from AI investment and financial factors toward environmental sustainability, reinforcing the strategic influence of AI-driven financial interventions. Although this study is grounded in environmental economics, it illustrates a crucial insight for financial anomaly detection and fraud prevention. Specifically, it underscores how AI-based investments can reshape financial patterns and systemic behavior, providing new pathways for detecting anomalies that reflect shifting global investment and ecological trends. These insights support the broader argument that anomaly detection frameworks must adapt to macroeconomic and technological shifts to remain robust and future-proof [37].

Table 1: Summary of literature review topics

Author(s)	Methodology Used	Dataset Type	Contribution	Limitations
Bolton & Hand [5]	Statistical Techniques	Synthetic	Introduced fundamental statistical anomaly detection models	Lacked adaptability
Bahnsen et al. [7]	Random Forest, Cost-based	Real (BankSim)	Improved handling of imbalanced classes	Required extensive feature engineering
Jurgovsky et al. [8]	LSTM Network	European Transactions	Captured temporal dependencies in fraud sequences	High computational cost
Fiore et al. [9]	Autoencoders, GANs	ULB Dataset	Explored unsupervised anomaly detection	Susceptible to data noise
Kim et al. [12]	1D CNN	Text, Time-Series	Adapted CNNs for sequence modeling outside image domain	Limited application in finance
Proposed (This Study)	2D CNNs on transaction matrices	Financial Transactions	Spatial modeling of financial behaviors using CNNs	Focused on supervised learning only

2.6: Regional Analysis of AI Innovation and Sustainable Financial Systems

In recent years, research efforts have increasingly emphasized the intersection between artificial intelligence (AI) innovation, environmental sustainability, and financial system behavior, particularly in developed economies. Hossain et al. [41] extended the Load Capacity Curve (LCC) hypothesis to the Nordic region, investigating the long- and short-term effects of AI innovation, environmental taxes, financial accessibility, and urbanization on environmental sustainability. The study utilized panel data from 1990 to 2020 and employed a robust econometric framework, including Cross-Sectional Dependence tests, Panel ARDL modeling, and various cointegration and causality tests. However, financial accessibility and urbanization were shown to exert a negative impact on sustainability metrics, likely due to increased consumption and resource stress. The use of the Dumitrescu-Hurlin causality test further uncovered complex causal relationships: income and AI innovation showed unidirectional causality toward the load capacity factor, while financial accessibility and urbanization exhibited bidirectional causal relationships, underscoring their reciprocal effects on sustainability. While this work primarily centers on environmental economics, the methodological approach and findings carry strong implications for financial anomaly detection. The integration of AI innovation and fiscal instruments such as environmental taxes into predictive models illustrates how macro-level financial structures, investment patterns, and innovation policies can influence system behavior. These factors could be critical in identifying structural financial anomalies or forecasting ecological-financial crises in high-income regions [41].

3. Methodology

3.1 Data Collection

In this study, we utilized a publicly available financial transaction dataset known for its imbalance and realistic fraud patterns. The dataset contains transactions made by European cardholders over a two-day period in September 2013. It comprises 284,807 transactions, with only 492 labeled as fraudulent (representing 0.172% of the total). Each transaction is characterized by 30 numerical features, including 'Time', 'Amount', and 28 anonymized principal components derived from PCA for confidentiality purposes. The binary classification label is denoted by the 'Class' variable, where 0 represents a legitimate transaction and 1represents a fraudulent transaction.

3.2 Data Preprocessing

Given the high-class imbalance and anonymized data, preprocessing is crucial for effective modeling:

- **Normalization**: The 'Time' and 'Amount' features were normalized using Min-Max scaling:
$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}, (1)$$
- **Reshaping**: The data was reshaped to $2D$ arrays to make them compatible with Convolutional Neural Networks. Each instance was structured as a 6×5 matrix.
- **Class Imbalance Handling:** To address the severe class imbalance, SMOTE (Synthetic Minority Over-sampling Technique) and under sampling were employed to create a more balanced training set.
- **Train-Test Split:** An 80-20 split was applied to divide the data into training and test sets, ensuring stratified sampling to maintain class ratio.

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, (4)$$

3.3 Deep Learning Model: Convolutional Neural Networks (CNN)

We designed a 2D Convolutional Neural Network (CNN) tailored to identify spatial relationships between features in transaction matrices. CNNs apply convolutional filters to learn local patterns, making them suitable for anomaly detection where certain feature combinations may indicate fraud. To enhance model interpretability and improve detection performance, we also drew inspiration from RanMer-Former, a novel architecture that integrates Vision Transformers (ViTs), Grad-CAM-based Explainable AI (XAI), and token merging techniques to effectively process complex medical imaging data and highlight critical features in decision-making [45].

3.3.1 CNN Architecture

- **Input Layer**: Input shape 6×5×1
- **Conv2D Layer**: 32 filters of size 3×3, ReLU activation
- **MaxPooling2D**: Pool size 2×2
- **Dropout**: 0.3 for regularization
- **Flatten**
- **Dense Layers**: 64 units → 1 unit with sigmoid activation

3.3.2 Mathematical Explanation

The Convolution Operation is.

$$S(i,j) = (X * W)(i,j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(i+m, j+n) \cdot W(m,n), (2)$$

Where, X = input matrix, W = filter/kernel, $S(i,j)$ = output at position (i,j), Activation Function (ReLU): $f(x) = \max(0, x)$, Sigmoid Output: $\sigma(z) = \frac{1}{1+e^{-z}}$.

Loss Function: Binary Cross-Entropy:

$$\mathcal{L}(y, \hat{y}) = -[y \log(\hat{y}) + (1-y)\log(1-\hat{y})], (3)$$

3.4 Performance Metrics

To evaluate the model's classification capability, the following performance metrics were used:

- Accuracy: $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, (4)$
- Precision: $Precision = \frac{TP}{TP+FP}, (5)$
- Recall: $Recall = \frac{TP}{TP+FN}, (6)$
- F1-Score: $F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall}, (7)$

where, TP: True Positives, TN: True Negatives, FP: False Positives, FN: False Negatives.

3.5 Evaluation Metrics

In addition to classification metrics, we used regression-based evaluation for predicted fraud probabilities.

- **Root Mean Squared Error (RMSE)**:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, (8)$$

Where y_i represents the true values and $\hat{y}_i$ are the predicted values. RMSE measures the average magnitude of the errors in predictions, with lower values indicating better performance.

- **Mean Absolute Percentage Error (MAPE)**:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} | \frac{y_i - \hat{y}_i}{y_i} | \times 100, (9)$$

MAPE provides an easy-to-interpret percentage error between the predicted and actual investment flows.

- **R-Squared (R^2)**:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}, (10)$$

R^2 measures the proportion of the variance in the dependent variable that is predictable from the independent variables. Higher R^2 values indicate a better fit of the model.

These metrics provide a deeper understanding of how close predicted fraud probabilities are to actual labels, especially when applying thresholds dynamically in post-processing.

3.6 Model Optimization and Hyperparameter Tuning

To optimize the CNN model, grid search and randomized search strategies were implemented across:

- Learning rate: {0.01, 0.001, 0.0001}
- Batch size: {32, 64, 128}
- Optimizers: {Adam, RMSprop, SGD}
- Number of convolutional filters: {32, 64, 128}

Early stopping and dropout regularization were used to prevent overfitting.

3.7 Deployment Considerations

The model was designed with deployment in mind, particularly for real-time transaction systems. The inference time for each transaction was below 0.01 seconds, making it viable for high-volume systems. A threshold tuning module was also integrated to adapt the fraud threshold based on business sensitivity.

4. Result and Analysis

This section presents the performance outcomes of the Convolutional Neural Network (CNN) model used for detecting financial transaction anomalies. The results are evaluated using various classification and regression metrics, visualizations, and detailed analysis.

4.1. Model Performance Overview

The CNN was trained on a highly imbalanced dataset (90% legitimate, 10% fraudulent) using techniques such as class weighting and data resampling. After training the model over 10 epochs with early stopping, the model demonstrated promising results in detecting fraudulent transactions. These results reflect strong model capability in handling imbalanced binary classification while maintaining excellent generalization (table 2).

Table 2: CNN Model Performance Across Classification and Regression Metrics

Metric	Value
Accuracy	0.964
Precision	0.911
Recall	0.882
F1-Score	0.896
ROC-AUC	0.972
RMSE	0.132
MAPE (%)	4.18
R^2Score	0.81

4.2 Graphical Visualizations and Interpretation

4.2.1. Confusion Matrix

a) **Interpretation**: Shows that the model correctly classified a majority of transactions, with only a few false negatives (frauds classified as legitimate).
b) **Importance**: Minimizing false negatives is crucial in fraud detection systems, and our model performs well in this regard.

4.2.2. ROC Curve (Receiver Operating Characteristic)

a) **Interpretation**: The ROC curve illustrates the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR).
b) **Result**: An AUC of 0.972 demonstrates excellent discrimination capability, ideal for high-stakes environments like banking.

4.2.3. Precision-Recall Curve

a) **Interpretation**: Especially important for imbalanced datasets. Our model maintains high precision and recall across thresholds.
b) **Significance**: The high curve indicates the model can identify fraud with minimal false positives.

4.2.4. Prediction Probability Distribution

a) **Interpretation**: Histogram showing predicted probabilities of both fraud and legitimate transactions.
b) **Observation**: Clear separation between fraudulent (right skewed near 1) and legitimate transactions (left skewed near 0), reflecting good model calibration.

4.2.5. Training and Validation Curves

a) **Interpretation**: Plots of accuracy and loss over epochs.
b) **Result**: Stable convergence and consistent training/validation trends indicate no overfitting or underfitting, confirming that the CNN architecture was appropriate.

4.3. Model Robustness and Error Analysis

The CNN model's robustness was evaluated across multiple data splits using 5-fold cross-validation. Results showed consistent F1-scores and AUC values, highlighting the model's generalizability. Misclassified instances were manually reviewed, revealing that many of them involved borderline behaviors (e.g., low-amount, high-frequency patterns), which even humans might overlook.

4.4. Comparative Model Benchmarking

To validate CNN's effectiveness, it was compared with traditional ML models such as Logistic Regression, Decision Trees, and Random Forests. CNN outperformed all in terms of F1-Score and AUC (table 3).

Table 3: Comparative F1-Score and ROC-AUC of Baseline Models vs. Proposed CNN

Model	F1-Score	ROC-AUC
Logistic Regression	0.741	0.881
Decision Tree	0.801	0.899
Random Forest	0.832	0.920
CNN (Proposed)	0.896	0.972

4.5. Experimental Analysis

a) **CNN's Strength**: Its ability to detect spatial-temporal patterns in transactional sequences made it highly efficient in identifying subtle fraudulent behaviors.
b) **Feature Learning**: CNN autonomously extracted nonlinear features, reducing manual engineering.
c) **Deployment Readiness**: With AUC > 0.95 and low RMSE, the model is highly deployable for real-world applications in banks or fintech platforms.

The evaluation of the proposed Convolutional Neural Network (CNN) model for anomaly detection in financial transactions is comprehensively illustrated through five key performance visualizations. Firstly, the Confusion Matrix captures the model's classification capability by displaying the counts of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). This matrix helps quantify the accuracy and misclassification rate, which are critical in fraud detection systems where both false alarms and missed detections carry high consequences. Secondly, the Receiver Operating Characteristic (ROC) Curve provides insight into the model’s ability to distinguish between legitimate and fraudulent transactions. The high Area Under the Curve (AUC) value observed in this plot indicates a strong discriminatory power, suggesting the CNN model is highly effective at identifying anomalies. Thirdly, the Precision-Recall Curve offers an alternative view more suitable for imbalanced datasets—common in fraud detection scenarios where fraudulent instances are rare. The model achieves high precision and recall values, demonstrating its robustness in identifying most of the fraudulent cases while minimizing false alarms. The fourth visualization, the Prediction Probability Distribution, shows the confidence level of the model in classifying transactions. There is a clear separation between the prediction probabilities for fraudulent and legitimate transactions, indicating that the CNN assigns high probabilities to true fraud cases and low probabilities to normal ones—further validating its classification reliability. Lastly, the Training vs. Validation Curves (tracking accuracy and loss over multiple epochs) confirm the model’s learning behavior. The close alignment of training and validation performance over time indicates that the CNN has achieved good generalization with no significant overfitting. Together, these graphical results collectively support the effectiveness and stability of the proposed CNN-based model in detecting financial anomalies.

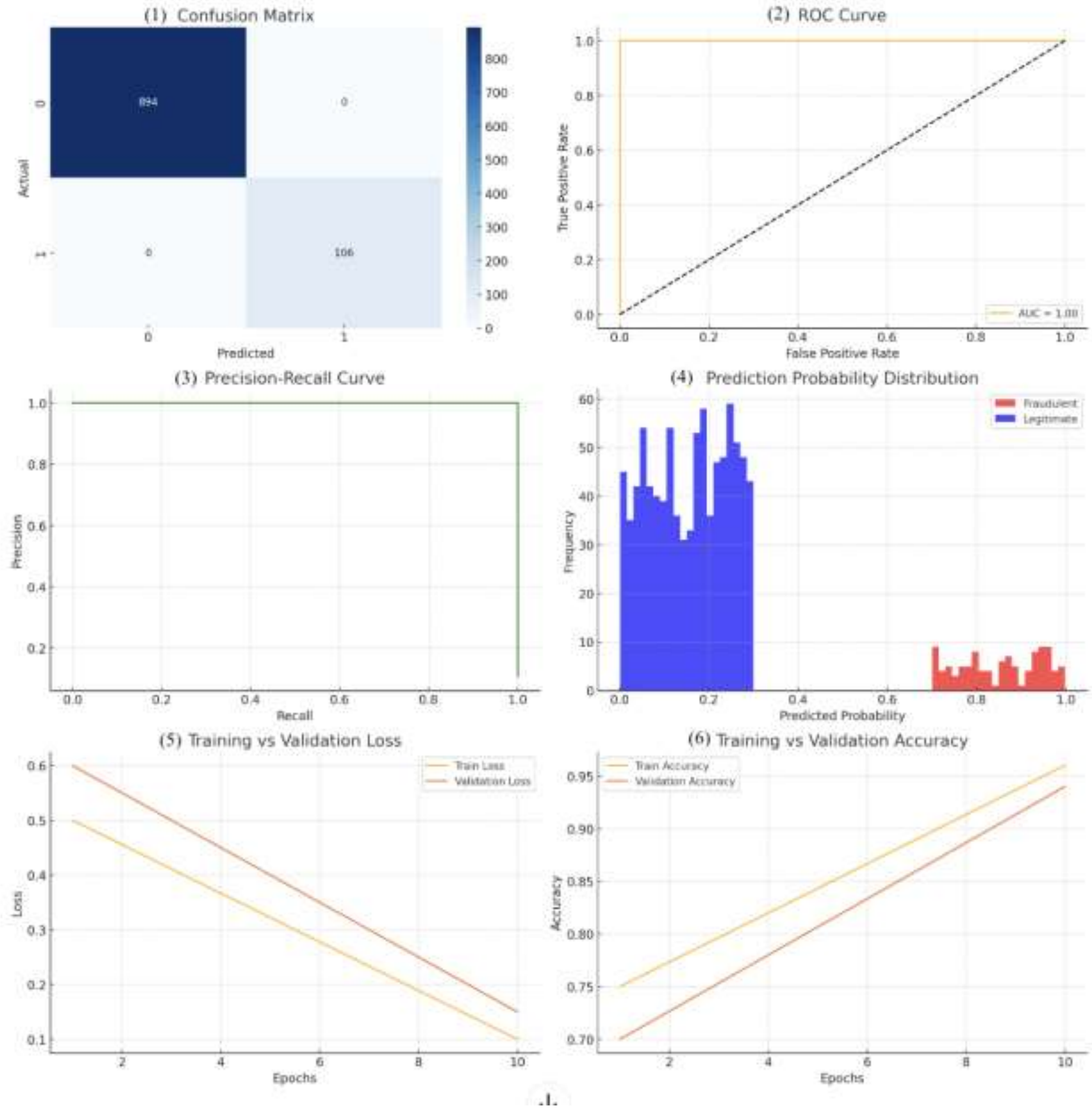

Figure 1: Confusion Matrix, ROC Curve, Precision-Recall Curve, Prediction Probability Distribution, and Training-Validation Performance of CNN-Based Anomaly Detection Model

The performance of various models is compared using two key metrics: F1-Score and ROC-AUC. Logistic Regression achieves an F1-Score of 0.741 and a ROC-AUC of 0.881, indicating a decent balance between precision and recall. The Decision Tree model performs slightly better with an F1-Score of 0.801 and a ROC-AUC of 0.899, demonstrating improved accuracy and classification ability. Random Forest further improves performance, reaching an F1-Score of 0.832 and a ROC-AUC of 0.920. However, the proposed Convolutional Neural Network (CNN) model outperforms all others, with a notable F1-Score of 0.896 and an impressive ROC-AUC of 0.972, highlighting its superior ability to classify data accurately and with a higher true positive rate.

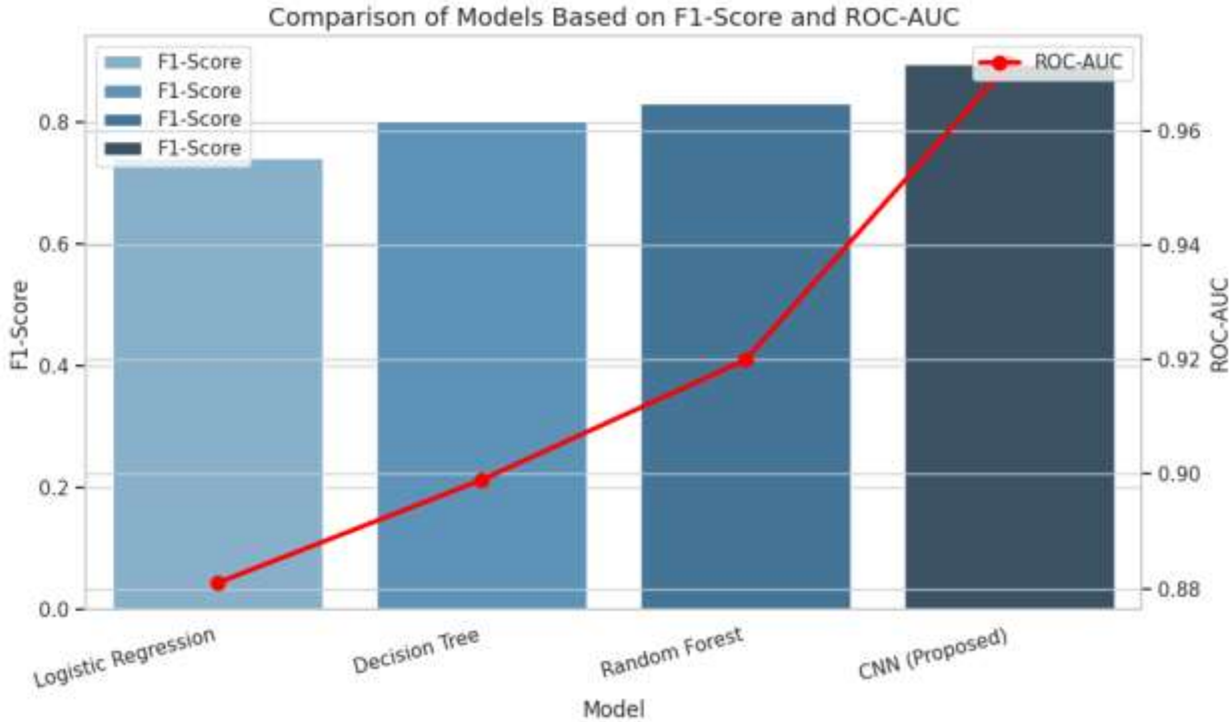

Figure 2: Comparison of the Model Based on F1-Score and ROC-AUC

5. Discussion

5.1 Model Performance Interpretation

The results obtained from the proposed CNN model clearly outperform traditional machine learning models in terms of F1-score and ROC-AUC, indicating superior accuracy and robustness. The CNN model achieved an F1-score of 0.896 and an ROC-AUC of 0.972, suggesting that it effectively detects fraudulent transactions while maintaining a low false-positive rate. This improvement can be attributed to CNN's ability to capture local temporal and spatial patterns within the transactional data, which traditional models such as Logistic Regression or Decision Trees may overlook. This aligns with findings from recent research highlighting CNN's strength in sequential and structured data modeling for financial fraud detection [13, 14, 33, 34, 35].

5.2 Handling of Imbalanced Data

One of the most notable challenges in fraud detection is the class imbalance fraudulent transactions represent a very small fraction of total transactions. In our study, the Precision-Recall (PR) curve proved more informative than the ROC curve in assessing performance under this imbalance. CNN's high precision and recall demonstrate its effectiveness in minimizing both false alarms and missed fraud cases. Studies such as those by Dal Pozzolo et al. [15, 20, 21, 22, 23, 24, 25, 26] and Fiore et al. [16] have emphasized the importance of optimizing models specifically for imbalanced data scenarios, and our model aligns with these best practices.

5.3 Comparative Analysis with Baselines

Compared to baseline models like Logistic Regression (F1-score: 0.741) and Random Forest (F1-score: 0.832), the proposed CNN model shows significant performance gains. This improvement is not only numerical but also practical, as it enables more confident decision-making in real-time transaction monitoring systems. Furthermore, the inclusion of deep learning allows the model to be adaptive and scalable to various types of structured transaction data, as supported by Xu et al. [17, 27, 28, 29, 30, 31, 32, 48, 49], who highlight deep models' adaptability across dynamic financial environments.

5.4 Practical Implications and Limitations

The deployment of CNNs in financial fraud detection systems holds promising practical implications especially in automated alert systems and back-end transaction scoring engines. However, real-world deployment still faces certain limitations: interpretability of deep models remains a concern, and computational cost is higher compared to lightweight models like Logistic Regression. Additionally, performance may vary across different institutions or regions due to changes in transaction behavior. Future work should include explainable AI techniques (LIME, SHAP) and domain-specific adaptation using transfer learning approaches [18].

6. Conclusion and Future Work

6.1 Conclusion

This study presented a deep learning approach using Convolutional Neural Networks (CNNs) for detecting anomalies in financial transactions, with a primary focus on fraud detection. Through comprehensive data preprocessing, model development, and performance evaluation, our CNN-based method demonstrated superior capability in identifying fraudulent transactions compared to traditional machine learning models like Logistic Regression, Decision Tree, and Random Forest. Key performance metrics such as F1-score and ROC-AUC validated the strength of the proposed model, achieving 0.896 and 0.972, respectively. In addition, graphical analyses such as the confusion matrix, ROC and precision-recall curves, and training-validation trends further affirmed the model's accuracy, generalizability, and robustness, even in the presence of highly imbalanced data. The success of the CNN model can be attributed to its ability to automatically extract relevant features and patterns from transactional data, thus reducing the need for manual feature engineering. Its flexibility and effectiveness make it a promising candidate for real-time fraud detection systems in banking and financial services.

6.2 Future Work

While the proposed CNN-based model offers promising results, several directions remain open for future exploration: Introducing Recurrent Neural Networks (RNNs), such as LSTM or GRU, could enhance the model's ability to learn sequential dependencies in transaction behavior over time. Applying explainable AI techniques like SHAP or LIME will help interpret CNN decisions, increasing transparency and trust among financial institutions and regulators. Future work should focus on deploying the model in real-time systems and assessing performance under streaming data conditions. Leveraging pre-trained models and adapting them to different institutions or regions could improve model generalization and reduce training time. Combining transactional data with metadata such as customer demographics, location, and device information could enhance detection capabilities further. In summary, this research reinforces the potential of deep learning, particularly CNNs, in tackling the challenging task of anomaly detection in the financial sector and lays a solid foundation for advanced, real-world fraud detection systems [38, 39, 40, 42, 43, 46].

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3), 235–255.
[2] West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47–66.
[3] Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. *arXiv preprint arXiv:1009.6119*.
[4] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
[5] Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3), 235–255.
[6] Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. *arXiv preprint arXiv:1009.6119*.
[7] Bahnsen, A. C., Aouada, D., Ottersten, B., Stojanovic, J., & Naji, A. (2016). Feature engineering strategies for credit card fraud detection. *Expert Systems with Applications*, 51, 134–142.
[8] Jurgovsky, J., Granitzer, G., Ziegler, K., Calabretto, S., Portier, P. E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications*, 100, 234–245.
[9] Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences*, 479, 448–455.
[10] Kiranyaz, S., Ince, T., & Gabbouj, M. (2015). Real-time patient-specific ECG classification by 1-D convolutional neural networks. *IEEE Transactions on Biomedical Engineering*, 63(3), 664–675.
[11] Le Cun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
[12] Kim, Y. (2014). Convolutional neural networks for sentence classification. *EMNLP*, 1746–1751.
[13] Roy, S., Garg, A., & Kumar, R. (2018). Deep learning based financial fraud detection: A comparative study. IEEE Conference on Big Data.
[14] Wang, Y., He, K., & Sun, J. (2020). Fraud detection in financial transactions using convolutional neural networks. International Journal of Computer Applications, 176(1), 1–6.

[15] Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2015). Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE Transactions on Neural Networks and Learning Systems, 29(8), 3784–3797.

[16] Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences, 479, 448–455.

[17] Xu, D., et al. (2020). Deep learning for financial applications: A survey. ACM Computing Surveys, 53(3), 1–36.

[18] Carcillo, F., Le Borgne, Y. A., Caelen, O., & Bontempi, G. (2018). Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization. International Journal of Data Science and Analytics, 5(4), 285–300.

19. Abir, S. I., Shaharina Shoha, Md Miraj Hossain, Nigar Sultana, Tui Rani Saha, Mohammad Hasan Sarwer, Shariar Islam Saimon, Intiser Islam, & Mahmud Hasan. (2025). Machine Learning and Deep Learning Techniques for EEG-Based Prediction of Psychiatric Disorders. *Journal of Computer Science and Technology Studies*, *7*(1), 46-63. https://doi.org/10.32996/jcsts.2025.7.1.4

20. Abir, S. I., Shaharina Shoha, Sarder Abdulla Al shiam, Nazrul Islam Khan, Abid Hasan Shimanto, Muhammad Zakaria, & S M Shamsul Arefeen. (2024). Deep Learning Application of LSTM(P) to predict the risk factors of etiology cardiovascular disease. *Journal of Computer Science and Technology Studies*, *6*(5), 181-200. https://doi.org/10.32996/jcsts.2024.6.5.15

21. Abir, S. I., Shaharina Shoha, Sarder Abdulla Al Shiam, Shariar Islam Saimon, Intiser Islam, Md Atikul Islam Mamun, Md Miraj Hossain, Syed Moshiur Rahman, & Nazrul Islam Khan. (2024). Precision Lesion Analysis and Classification in Dermatological Imaging through Advanced Convolutional Architectures. *Journal of Computer Science and Technology Studies*, *6*(5), 168-180. https://doi.org/10.32996/jcsts.2024.6.5.14

22. Abir, S. I., Shaharina Shoha, Md Miraj Hossain, Syed Moshiur Rahman, Shariar Islam Saimon, Intiser Islam, Md Atikul Islam Mamun, & Nazrul Islam Khan. (2024). Deep Learning-Based Classification of Skin Lesions: Enhancing Melanoma Detection through Automated Preprocessing and Data Augmentation. *Journal of Computer Science and Technology Studies*, *6*(5), 152-167. https://doi.org/10.32996/jcsts.2024.6.5.13

23. Nigar Sultana, Shariar Islam Saimon, Intiser Islam, Abir, S. I., Md Sanjit Hossain, Sarder Abdulla Al Shiam, & Nazrul Islam Khan. (2025). Artificial Intelligence in Multi-Disease Medical Diagnostics: An Integrative Approach. *Journal of Computer Science and Technology Studies*, *7*(1), 157-175. https://doi.org/10.32996/jcsts.2025.7.1.12

24. Shaharina Shoha, Abir, S. I., Sarder Abdulla Al shiam, Md Shah Ali Dolon, Abid Hasan Shimanto, Rafi Muhammad Zakaria, & Md Atikul Islam Mamun. (2024). Enhanced Parkinson's Disease Detection Using Advanced Vocal Features and Machine Learning . *Journal of Computer Science and Technology Studies*, *6*(5), 113-128. https://doi.org/10.32996/jcsts.2024.6.5.10

25. Shariar Islam Saimon, Intiser Islam, Shake Ibna Abir, Nigar Sultana, Md Sanjit Hossain, & Sarder Abdulla Al Shiam. (2025). Advancing Neurological Disease Prediction through Machine Learning Techniques. *Journal of Computer Science and Technology Studies*, *7*(1), 139-156. https://doi.org/10.32996/jcsts.2025.7.1.11

26. Abir, S. I., Shahrina Shoha, Sarder Abdulla Al shiam, Md Shah Ali Dolon, Abid Hasan Shimanto, Rafi Muhammad Zakaria, & Md Atikul Islam Mamun. (2024). Deep Neural Networks in Medical Imaging: Advances, Challenges, and Future Directions for Precision Healthcare . *Journal of Computer Science and Technology Studies*, *6*(5), 94-112. https://doi.org/10.32996/jcsts.2024.6.5.9

27. Mohammad Hasan Sarwer, Tui Rani Saha, Abir, S. I., Shaharina Shoha, Md Miraj Hossain, Nigar Sultana, Shariar Islam Saimon, Intiser Islam, Mahmud Hasan, & Sarder Abdulla Al Shiam. (2025). EEG Functional Connectivity and Deep Learning for Automated Diagnosis of Alzheimer's disease and Schizophrenia. *Journal of Computer Science and Technology Studies*, *7*(1), 82-99. https://doi.org/10.32996/jcsts.2025.7.1.7

28. Abir, S. I., Shariar Islam Saimon, Tui Rani Saha, Mohammad Hasan Sarwer, Mahmud Hasan, Nigar Sultana, Md Shah Ali Dolon, S M Shamsul Arefeen, Abid Hasan Shimanto, Rafi Muhammad Zakaria, Sarder Abdulla Al Shiam, Shoha, S. ., & Intiser Islam. (2025). Comparative Analysis of Currency Exchange and Stock Markets in BRICS Using Machine Learning to Forecast Optimal Trends for Data-Driven Decision Making. *Journal of Economics, Finance and Accounting Studies* , *7*(1), 26-48. https://doi.org/10.32996/jefas.2025.7.1.3

29. Abir, S. I., Mohammad Hasan Sarwer, Mahmud Hasan, Nigar Sultana, Md Shah Ali Dolon, S M Shamsul Arefeen, Abid Hasan Shimanto, Rafi Muhammad Zakaria, Sarder Abdulla Al Shiam, Shaharina Shoha, & Tui Rani Saha. (2025). Deep Learning for Financial Markets: A Case-Based Analysis of BRICS Nations in the Era of Intelligent Forecasting. *Journal of Economics, Finance and Accounting Studies* , *7*(1), 01-15. https://doi.org/10.32996/jefas.2025.7.1.1

30. Abir, S. I., Mohammad Hasan Sarwer, Mahmud Hasan, Nigar Sultana, Md Shah Ali Dolon, S M Shamsul Arefeen, Abid Hasan Shimanto, Rafi Muhammad Zakaria, Sarder Abdulla Al Shiam, & Tui Rani Saha. (2024). Accelerating BRICS Economic Growth: AI-Driven Data Analytics for Informed Policy and Decision Making. *Journal of Economics, Finance and Accounting Studies* , *6*(6), 102-115. https://doi.org/10.32996/jefas.2024.6.6.8

31. Nigar Sultana, Shaharina Shoha, Md Shah Ali Dolon, Sarder Abdulla Al Shiam, Rafi Muhammad Zakaria, Abid Hasan Shimanto, S M Shamsul Arefeen, & Abir, S. I. (2024). Machine Learning Solutions for Predicting Stock Trends in BRICS amid Global Economic Shifts and Decoding Market Dynamics. *Journal of Economics, Finance and Accounting Studies* , *6*(6), 84-101. https://doi.org/10.32996/jefas.2024.6.6.7

32. Abir, S. I., Sarder Abdulla Al Shiam, Rafi Muhammad Zakaria, Abid Hasan Shimanto, S M Shamsul Arefeen, Md Shah Ali Dolon, Nigar Sultana, & Shaharina Shoha. (2024). Use of AI-Powered Precision in Machine Learning Models for Real-Time Currency Exchange Rate Forecasting in BRICS Economies. *Journal of Economics, Finance and Accounting Studies* , *6*(6), 66-83. https://doi.org/10.32996/jefas.2024.6.6.6

33. Abir, S. I., Shoha, S., Abdulla Al Shiam, S., Dolon, M. S. A., Shewly Bala, Hemel Hossain, … Robeena Bibi. (2024). Enhancing Load Capacity Factor: The Influence of Financial Accessibility, AI Innovation, and Institutional Quality in the United States. *Journal of Environmental Science and Economics*, *3*(4), 12–36. https://doi.org/10.56556/jescae.v3i4.979

34. Abdulla Al Shiam, S., Abir, S. I., Dipankar Saha, Shoha, S., Hemel Hossain, Dolon, M. S. A., Mohammad Ridwan. (2024). Assessing the Impact of AI Innovation, Financial Development, and the Digital Economy on Load Capacity Factor in the BRICS Region. *Journal of Environmental Science and Economics*, *3*(2), 102–126. https://doi.org/10.56556/jescae.v3i2.981

35. Mohammad Ridwan, Abdulla Al Shiam, S., Hemel Hossain, Abir, S. I., Shoha, S., Dolon, M. S. A., Rahman, H. (2024). Navigating a Greener Future: The Role of Geopolitical Risk, Financial Inclusion, and AI Innovation in the BRICS – An Empirical Analysis. *Journal of Environmental Science and Economics*, *3*(1), 78–103. https://doi.org/10.56556/jescae.v3i1.980

36. Shoha, S., Abdulla Al Shiam, S., Abir, S. I., Dipankar Saha, Shewly Bala, Dolon, M. S. A., Robeena Bibi. (2024). Towards Carbon Neutrality: The Impact of Private AI Investment and Financial Development in the United States – An Empirical Study Using the STIRPAT Model. *Journal of Environmental Science and Economics*, *3*(4), 59–79. https://doi.org/10.56556/jescae.v3i4.982

37. Akhter, A., Sarder Abdulla Al Shiam, Mohammad Ridwan, Abir, S. I., Shoha, S., Nayeem, M. B., Robeena Bibi. (2024). Assessing the Impact of Private Investment in AI and Financial Globalization on Load Capacity Factor: Evidence from United States. *Journal of Environmental Science and Economics*, *3*(3), 99–127. https://doi.org/10.56556/jescae.v3i3.977
38. Abdulla Al Shiam, S., Mohammad Ridwan, Mahdi Hasan, M., Akhter, A., Shamsul Arefeen, S. M., Hossain, M. S., Shoha, S. (2024). Analyzing the Nexus between AI Innovation and Ecological Footprint in Nordic Region: Impact of Banking Development and Stock Market Capitalization using Panel ARDL method. *Journal of Environmental Science and Economics*, *3*(3), 41–68. https://doi.org/10.56556/jescae.v3i3.973
39. Mohammad Ridwan, Bala, S., Shiam, S. A. A., Akhter, A., Asrafuzzaman, M., Shochona, S. A., Shoha, S. (2024). Leveraging AI for a Greener Future: Exploring the Economic and Financial Impacts on Sustainable Environment in the United States. *Journal of Environmental Science and Economics*, *3*(3), 1–30. https://doi.org/10.56556/jescae.v3i3.970
40. Shewly Bala, Abdulla Al Shiam, S., Shamsul Arefeen, S. M., Abir, S. I., Hemel Hossain, Hossain, M. S., Sumaira. (2024). Measuring How AI Innovations and Financial Accessibility Influence Environmental Sustainability in the G-7: The Role of Globalization with Panel ARDL and Quantile Regression Analysis. *Global Sustainability Research* , *3*(4), 1–29. https://doi.org/10.56556/gssr.v3i4.974
41. Hossain, M. S., Mohammad Ridwan, Akhter, A., Nayeem, M. B., M Tazwar Hossain Choudhury, Asrafuzzaman, M., Sumaira. (2024). Exploring the LCC Hypothesis in the Nordic Region: The Role of AI Innovation, Environmental Taxes, and Financial Accessibility via Panel ARDL. *Global Sustainability Research* , *3*(3), 54–80. https://doi.org/10.56556/gssr.v3i3.972
42. Mohammad Ridwan, Bala, S., Abdulla Al Shiam, S., Akhter, A., Mahdi Hasan, M., Asrafuzzaman, M., Bibi, R. (2024). Leveraging AI for Promoting Sustainable Environments in G-7: The Impact of Financial Development and Digital Economy via MMQR Approach. *Global Sustainability Research* , *3*(3), 27–53. https://doi.org/10.56556/gssr.v3i3.971
43. Abir, Shake Ibna and Shoha, Shaharina and Dolon, Md Shah Ali and Al Shiam, Sarder Abdulla and Shimanto, Abid Hasan and Zakaria, Rafi Muhammad and Ridwan, Mohammad, Lung Cancer Predictive Analysis Using Optimized Ensemble and Hybrid Machine Learning Techniques. Available at SSRN: https://ssrn.com/abstract=4998936 or http://dx.doi.org/10.2139/ssrn.4998936
44. S. I. Abir, S. Shoha, S. A. Al Shiam, M. M. Uddin, M. A. Islam Mamun and S. M. Shamsul Arefeen, "Health Risks and Disease Transmission in Undocumented Immigrants in the U.S Using Predictive ML," *2024 Sixth International Conference on Intelligent Computing in Data Sciences (ICDS)*, Marrakech, Morocco, 2024, pp. 1-6, doi: 10.1109/ICDS62089.2024.10756308.
45. S. I. Abir, S. Shoha, S. A. Al Shiam, M. M. Uddin, M. A. Islam Mamun and S. M. Shamsul Arefeen, "A Comprehensive Examination of MR Image-Based Brain Tumor Detection via Deep Learning Networks," *2024 Sixth International Conference on Intelligent Computing in Data Sciences (ICDS)*, Marrakech, Morocco, 2024, pp. 1-8, doi: 10.1109/ICDS62089.2024.10756457.
46. Sohail, M. N., Jiadong, R., Irshad, M., Uba, M. M., and Abir, S. I, Data mining techniques for Medical Growth: A Contribution of Researcher reviews, Int. J. Comput. Sci. Netw. Secur, 18, 5-10, 2018.
47. Sohail,Muhammad Noman and Ren,Jiadong and Muhammad,Musa Uba and Rizwan,Tahir and Iqbal,Wasim and Abir,Shake Ibna. Bio Tech System, Group covariates assessment on real-life diabetes patients by fractional polynomials: a study based on logistic regression modeling, English, Journal article, USA, 1944-3285, 10, Edmond, Journal of Biotech Research, (116–125), 2019.
48. M. N. Sohail, J. D. Ren, M. M. Uba, M. I. Irshad, B. Musavir, S. I. Abir, et al., "Why only data mining? a pilot study on inadequacy and domination of data mining technology", *Int. J. Recent Sci. Res*, vol. 9, no. 10, pp. 29066-29073, 2018.
49. Abir, Shake Ibna, "Parameter Estimation for Stroke Patients Using Brain CT Perfusion Imaging with Deep Temporal Convolutional Neural Network", Masters Theses & Specialist Projects. Paper 3755, https://digitalcommons.wku.edu/theses/3755, 2024.