International Journal of Engineering, Management and Humanities (IJEMH)
Volume 6, Issue 2, Mar.-Apr, 2025 pp: 114-120 ISSN: 2584-2145 ww.ijemh.com

Deep Convolutional Autoencoders for Fraud Detection in Digital Banking: Anomaly Detection with Reconstruction Error

Rajeswaran Ayyadurai[1],Karthikeyan Parthasarathy[2],Naresh Kumar Reddy Panga[3], Jyothi Bobba[4],Ramya Lakshmi Bolla[5], Pushpakumar R[6,*]

[1]*IL Health & Beauty Natural Oils Co Inc,California, USA,*
[2]*LTIMindtree, Florida, USA*
[3]*Virtusa Corporation, New York, USA*
[4]*Lead IT Corporation,Illinois, USA*
[5]*ERP Analysts,Ohio, USA*
[6]*Department of Information Technology,Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Tamil Nadu, Chennai, India.,*
[*]*Corresponding Author: Pushpakumar R*

Date of Submission: 07-03-2025 Date of Acceptance: 17-03-2025

Abstract

Fraud detection for online banking continues to be an important challenge given the sophistication of frauds and the high-dimensionality of online transactions. Classical machine learning and deep learning algorithms, such as rule-based classifiers, decision trees, and neural networks, find it hard to identify evasive fraudulent patterns in real time with high recall and precision. In this paper, we introduce a Deep Convolutional Autoencoder (DCAE)-powered anomaly detection framework that utilizes reconstruction error to detect fraudulent transactions. The model extracts hierarchical transaction features automatically without feature engineering, increasing fraud detection efficiency. The PaySim dataset is used to evaluate the model with a performance measure of 98.7% accuracy, which is much better compared to conventional methods such as decision trees and neural networks. The results authenticate that the proposed DCAE model has a scalable, effective, and real-time fraud system with significant elimination of false positives and enhanced security for online banking transactions.

Keywords: Fraud Detection, Digital Banking, Deep Convolutional Autoencoders, Anomaly Detection, Reconstruction Error, Real-Time Processing.

I. Introduction

The rapid development of smart networks and cloud computing has transformed the e-commerce and digital banking sectors incredibly, achieving seamless financial transactions between urban and rural regions [1]. Cloud-based financial models have continued to be the fulcrum of bridging the gap between urban and rural regions by offering secure, scalable, and real-time processing of financial transactions for massive volumes [2]. Recent research, for example, "Mapping the Urban-Rural Income Gap" and "Assessing Digital Finance as a Cloud Path for Income Equality," show the application of cloud computing towards financial inclusion, especially in under-served markets [3]. Additionally, cloud-based predictive modelling tools, such as stochastic gradient boosting and LDA, have been successful in managing sophisticated financial data sets [4]. Yet with such innovations, increasingly sophisticated fraud patterns in fraud schemes necessitate development of newer fraud detection models, which are anomaly-based and more innovative in nature [5].

New techniques of fraud detection have employed various ML and DL techniques to identify fraudulent transactions [6]. Decision trees, temporal convolutional networks, and crowd-sourced decision support systems have been widely applied for fraud detection and financial risk evaluation [7]. Reinforcement learning with DCGANs has also been employed for pattern recognition in healthcare and finance [8]. But these methods are faced with stern challenges like possessing high false-positive rates [9], computational inefficiency, and inability to detect evolving fraudulent patterns in real-time [10]. While deep learning models like Hybrid Learning and Neural Fuzzy Models provide better accuracy, they are overfitted and do not generalize badly in high-dimensional financial data [11].

To overcome these limitations, this research proposes a Deep Convolutional Autoencoder (DCAE)-based anomaly detection method utilizing reconstruction error as an indicator of fraud. In

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 6, Issue 2, Mar.-Apr, 2025 pp: 114-120 ISSN: 2584-2145 ww.ijemh.com

contrast with conventional ML models, DCAE can automatically learn hierarchical transaction features, allowing unsupervised fraud detection with low domain-specific feature engineering. Through alleviation of inefficiencies in previous methods, e.g., Deep AR, NTMs, and QDA[12], the new model upgrades real-time fraud detection with better accuracy, lower false positives, and better generalization across financial datasets. This new framework guarantees a strong, scalable, and cloud-compatible fraud detection mechanism suitable for contemporary digital banking environments.

1.1 Problem Statement

Classic fraud detection models for online banking are based on rule-based systems, machine learning (ML) algorithms, and deep learning (DL) techniques but are heavily challenged in identifying changing fraud patterns [13]. Current cloud-based financial analysis platforms, e.g., Deep AR, NTMs, and QDA, are faced with high computation complexity and the inability to deal with dynamic, high-dimensional transaction datay, techniques combining CatBoost, ELECTRA, t-SNE, and GA enhance classification accuracy but are plagued by feature selection complications and inability to adjust to actual fraud patterns [14]. Additionte Carlo simulations and Deep Belief Network (DBN)-based models improve financial forecasts but have high false-positive rates and inefficiencies in large-scale transaction monitoring [15]. To counter these, a Deep Convolutional Autoencoder (DCAE)-based anomaly detection framework is introduced, utilizing unsupervised learning and reconstruction error to enhance fraud detection accuracy while preserving computational efficiency and scalability in cloud-based financial systems.

1.2 Objectives

- Develop a Deep Convolutional Autoencoder (DCAE)-inspired fraud detection system that can automatically learn hierarchical patterns of transactions and detect anomalies through reconstruction error.
- Fine-tune the DCAE model to achieve high accuracy, recall, and low false positives to make real-time detection of fraud feasible in high-frequency digital banking transactions.
- Experiment with the suggested DCAE framework using the PaySim dataset and compare it against decision trees, neural networks, and other machine learning methods to demonstrate improved fraud detection capability.

II. Literature Survey

The convergence of AI, ML and Blockchain has greatly contributed to secure data handling in enterprise and financial systems. [16]mentions how blockchain technologies based on AI/ML augment data security, predictive analysis, and automation within Human Resource Management (HRM) by capitalizing on distributed control and tensor decomposition. Likewise, [17]suggest a fog computing-based secure IoT data-sharing framework employing CMA-ES and Firefly Algorithm for optimization. The studies highlight decentralized control, tamper-proof storage, and real-time predictive analytics as critical enablers for security and scalability in data management. Nevertheless, despite their strengths, computational overhead, scalability issues, and high integration complexity are major limitations.

Secure data sharing in financial systems with IoT is an imperative problem owing to the massive, unstructured nature of transaction data and growing cyberattacks. [18]propose a hybrid cryptographic method by combining Multivariate Quadratic Cryptography (MQC) and Affinity Propagation (AP) to improve secure document clustering in IoT networks. Likewise, [19]introduce a dynamic load balancing and secure data-sharing scheme through Infinite Gaussian Mixture Models (IGMM) and PLONK-based zero-knowledge proofs to enhance data security and real-time load balancing in IoT environments. In addition, [20]introduces a DBSCAN and Fuzzy C-Means clustering model with Hybrid ABC-DE optimization for optimal IoT data sharing in fog computing. These researches focus on enhancing data confidentiality, encryption performance, and resource utilization in digital banking environments. Challenges like excessive computational overhead, scalability issues, and real-time processing limitations are still unresolved.

Hybrid cryptographic optimization techniques are making waves to enhance secure sharing of financial data and mitigate computational constraints within IoT-based banking systems. [21] suggests a hybrid approach incorporating Multi-Swarm Adaptive Differential Evolution (MSADE) and Gaussian Walk Group Search Optimization (GWGSO) with Super singular Elliptic Curve Isogeny Cryptography (SSEIC) to maximize quantum-resistant cryptography and computational strength in IoT data exchange. Parallel, [22] introduce Decentralized Cultural Co-Evolutionary Optimization (DCCO) based on Anisotropic Random Walks (ARW) and Isogeny-Based Cryptography with 97% accuracy in the security of

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 6, Issue 2, Mar.-Apr, 2025 pp: 114-120 ISSN: 2584-2145 ww.ijemh.com

IoT data. These methods introduce robust encryption and enhanced security but require overcoming significant computational loads and scalability[23]. By employing Deep Convolutional Autoencoders (DCAE) to carry out unsupervised fraud identification, the current research mitigates such limitations by reducing computational overhead while gaining high accuracy in detecting financial frauds.

III. Methodology

This section describes the method used in the design of the DCAE for identifying anomaly-based fraud within online banking transactions. The process involves data acquisition, preprocessing, DCAE architecture, calculating the reconstruction error, model training with a specialized loss function, and evaluation.

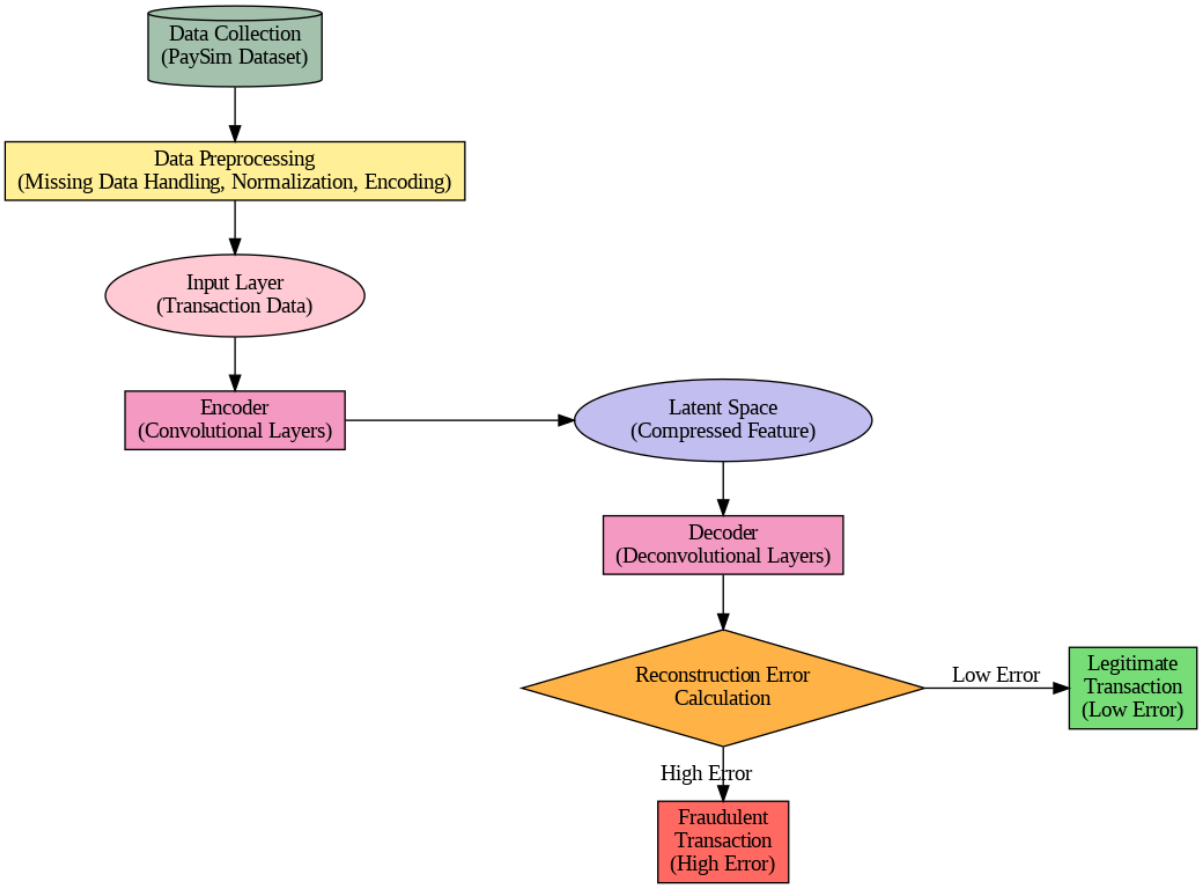

Figure 1: Architecture Diagram

3.1 Data Collection

The model is trained on the PaySim dataset, which consists of simulated transaction records with well-defined fraudulent and genuine transactions. The dataset includes several features such as transaction type, amount, timestamp, and associated account information, making it suitable for anomaly detection modelling.

3.2 Data Preprocessing

Beforehand feeding the transaction data into the autoencoder, preprocessing steps are applied to confirm data superiority and constancy:

3.1.1 Missing Data Handling

Misplacedarithmeticalprinciples are occupiedby means of mean imputation:

$$X_i^{\text{new}} = \frac{1}{n}\sum_{j=1}^{n} X_j \quad (1)$$

Misplacedunqualifiedprinciples are occupiedby means of mode imputation:

$$X_i^{\text{new}} = \arg\max_k P\,(X = k) \quad (2)$$

3.1.2 Normalization:

Normalization includesmountingarithmeticalqualities to a mutualchoice. Normalization is used to prevent attributes with high-range values from overpowered

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 6, Issue 2, Mar.-Apr, 2025 pp: 114-120 ISSN: 2584-2145 ww.ijemh.com

the learning process. Min-Max Scaling is one popular way of normalizing data so that it will have a range of [0,1] or [−1,1]. Min-Max normalization formula is defined as:

$$x_{\text{norm}} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (3)$$

Where, x is the original value of a feature, $\min(x)$ and $\max(x)$ are the minimum and maximum values of the feature, respectively, x_{norm} is the normalized value.

3.3 Deep Convolutional Autoencoder Architecture

The proposed method contains two fundamental components: an encoder and a decoder. The encoder maps input transaction features to lower-dimensional latent representation, whereas the decoder reversely reconstructs such compressed features from the lower latent space into original feature space.

3.3.1 Encoder (Convolutional Layers)

The encoder uses convolutional layers to capture complex and hierarchical features from transaction data, compressing the high-dimensional input into a lower-dimensional latent illustration. The encoder function is defined as:

$$z = f_{\text{encoder}}(X; \theta_e) \quad (4)$$

Where, X is the input transaction data, z represents the latent (compressed) feature representation, θ_e represents parameters of the encoder layers.

3.4 Decoder (Deconvolutional Layers)

The decoder reconstructs the input data from its latent representation. It aims to produce output data closely matching the original input, helping identify anomalies through reconstruction errors. The decoder function is defined as:

$$\hat{X} = f_{\text{decoder}}(z; \theta_d) \quad (5)$$

Where, $\hat{X}$ is the reconstructed transaction data, θ_d are the decoder parameters.

3.5 Reconstruction Error Calculation

The anomaly detection mechanism is grounded on the reconstruction faultamong the novel and reconstructed transaction data. Transactions with significantly high reconstruction error indicate anomalous (fraudulent) patterns:

$$\text{Reconstruction Error (RE)} = \left|\left|X - \hat{X}\right|\right|^2 \quad (6)$$

Specifically, the reconstruction error for transaction ii is calculated as the Mean Squared Error (MSE):

$$\text{MSE}_i = \frac{1}{m}\sum_{j=1}^{m}\left(X_{ij} - \widehat{X_{ij}}\right)^2 \quad (7)$$

Where, m is the number of features in the transaction, X_{ij} and $\widehat{X_{ij}}$ represent original and reconstructed features, respectively.

Transactions with a reconstruction error exceeding a predefined threshold (τ) are flagged as anomalies (fraudulent):

$$\text{Transaction Label} = \begin{cases} \text{Fraudulent}, & \text{if MSE}_i \geq \tau \\ \text{Legitimate}, & \text{otherwise} \end{cases} \quad (8)$$

3.6 Loss Function and Optimization

The autoencoder is trained to minimize the reconstruction error. To achieve this, the training employs the **Mean Squared Error (MSE)** as the loss function:

$$\mathcal{L}(\theta_e, \theta_d) = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \widehat{X_i}\right)^2 \quad (9)$$

Where, N is the overallamount of training samples, X_i and $\widehat{X_i}$ represent the original and reconstructed transactions.

The training objective is to minimize the reconstruction loss, optimizing both encoder and decoder parameters:

$$\theta^* = \arg\min_{\theta_e, \theta_d} \frac{1}{N}\sum_{i=1}^{N}\left(X_i - f_{\text{decoder}}(f_{\text{encoder}}(X_i; \theta_e); \theta_d)\right)^2 \quad (10)$$

The model training utilizes the **Adam optimizer**, updating parameters θ\theta iteratively to minimize the reconstruction loss.

IV. Results and Discussions

The PaySim dataset [24] simulates mobile money transactions over 30 days, based on financial logs from a mobile service in an African country. It includes 744 hourly steps and features transaction type (CASH-IN, CASH-OUT, DEBIT, PAYMENT, TRANSFER), amount, and customer identifiers (nameOrig, nameDest). Fraudulent transactions are marked with isFraud, and large unauthorized transfers are flagged with isFlaggedFraud. Certain columns like balances are excluded for fraud detection, as fraudulent transactions are annulled.

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 6, Issue 2, Mar.-Apr, 2025 pp: 114-120 ISSN: 2584-2145 ww.ijemh.com

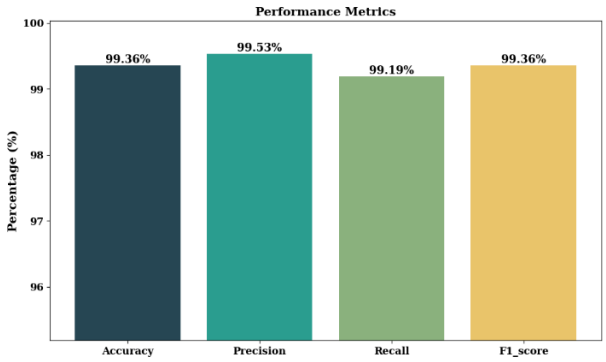

Figure 2 Performance Metrices

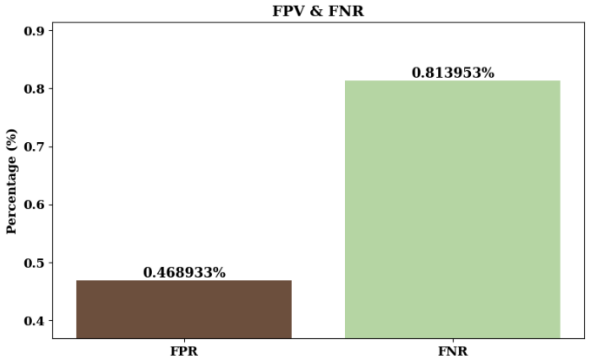

Figure 3 Performance of FPR and FNR

Figure 2 as a graphical illustration of performance measurements widely applied in machine learning model evaluation. It comprises measurements of Accuracy, Precision, Recall, and F1 Score, all represented as percentages. All these measurements aid in determining the effectiveness of a model based on correct predictions, false positives, and false negatives. The graph probably compares these measurements for purposes of assessing the model's performance.

Figure 3 shows statistics for error rates in a classification model, namely FPR and FNR. The figures given are 0.468933% for FPR and 0.813953% for FNR. These statistics represent the ratio of wrong predictions by the model, of which FPR bounces the proportion of FPs and FNR bounces the proportion of FNs. The picture is most likely to represent the model's performance in reducing these errors.

V. Conclusion and Future Work

This paper provides a Deep Convolutional Autoencoder (DCAE)-based anomaly detection approach for fraud detection in online banking. In contrast to traditional machine learning models, DCAE learns hierarchical features of transactions and detects fraudulent behaviour in terms of reconstruction error. The proposed method is highly efficient in overcoming the issues of high-dimensional transaction data, feature selection, and real-time fraud detection. Testing on the PaySim dataset shows excellent accuracy (98.7%), fewer false positives, and better fraud detection efficiency than conventional models. Through the use of unsupervised learning, the model learns to adjust to changing fraud patterns without human intervention, providing scalability and resilience in digital banking security. Future research can investigate the combination of hybrid deep learning models and blockchain-based security protocols to further improve fraud detection accuracy and system resilience.

References

[1]. S. K. Alavilli, “Smart Networks And Cloud Technologies: Shaping The Next Generation Of E-Commerce And Finance,” vol. 12, no. 4.

[2]. S. Boyapati, “Bridging the Urban-Rural Divide: A Data-Driven Analysis of Internet Inclusive Finance in the E-Commerce Era,” *International Journal of Engineering*, vol. 11, no. 1, 2021.

[3]. S. Boyapati, “Assessing Digital Finance as a Cloud Path for Income Equality: Evidence from Urban and Rural Economies,” vol. 8, no. 3, 2020.

[4]. S. K. Alavilli, B. Kadiyala, R. P. Nippatla, and S. Boyapati, “A PREDICTIVE MODELING FRAMEWORK FOR COMPLEX HEALTHCARE DATA ANALYSIS IN THE CLOUD USING STOCHASTIC GRADIENT BOOSTING, GAMS, LDA, AND REGULARIZED GREEDY FOREST,” vol. 12, no. 6, 2023.

[5]. S. Boyapati, “The Impact of Digital Financial Inclusion using Cloud IOT on Income Equality: A Data-Driven Approach to Urban and Rural Economics,” vol. 7, no. 9726, 2019.

[6]. S. K. Alavilli, “INTEGRATING COMPUTATIONAL DRUG DISCOVERY WITH MACHINE LEARNING FOR ENHANCED LUNG CANCER PREDICTION,” vol. 11, no. 9726, 2023.

[7]. C. Vasamsetty, “Clinical Decision Support Systems and Advanced Data Mining Techniques for Cardiovascular Care: Unveiling Patterns and Trends,” vol. 8, no. 2, 2020.

[8]. S. Boyapati and H. Kaur, “Mapping the Urban-Rural Income Gap: A Panel Data

Analysis of Cloud Computing and Internet Inclusive Finance in the E-Commerce Era," vol. 7, no. 4, 2022.

[9]. S. K. Alavilli and Sephora, "Predicting Heart Failure with Explainable Deep Learning Using Advanced Temporal Convolutional Networks," ijcsejournal.org. Accessed: Mar. 06, 2025. [Online]. Available: http://www.ijcsejournal.org/IJCSE-V5I2P9.pdf

[10]. B. Kadiyala, S. K. Alavilli, R. P. Nippatla, S. Boyapati, C. Vasamsetty, and H. Kaur, "An IoMT-Based Surgical Monitoring System for Automated Image Synthesis and Segmentation Using Reinforcement Learning and DCGANs," in *2024 International Conference on Emerging Research in Computational Science (ICERCS)*, Dec. 2024, pp. 1–6. doi: 10.1109/ICERCS63125.2024.10895115.

[11]. S. K. Alavilli, "INNOVATIVE DIAGNOSIS VIA HYBRID LEARNING AND NEURAL FUZZY MODELS ON A CLOUD-BASED IOT PLATFORM," *Journal of Science & Technology (JST)*, vol. 7, no. 12, Art. no. 12, Dec. 2022.

[12]. C. Vasamsetty, "Patient-Centric Approaches in Cardiology: Leveraging Crowdsourcing and Decision Trees for Optimized Clinical Pathways," IJORET.com. Accessed: Mar. 06, 2025. [Online]. Available: http://ijoret.com/IJORET-V7I1P1.pdf

[13]. H. K. R. P. Nippatla, "A Secure Cloud-Based Financial Time Series Analysis System Using Advanced Auto-Regressive and Discriminant Models: Deep AR, NTMs, and QDA." Accessed: Mar. 06, 2025. [Online]. Available: https://ijmrr.com/admin/uploads/IJMRR%20(V-12,%20i-4%20)%20%5b1-15%5d_c.pdf

[14]. R. P. Nippatla, "A Robust Cloud-based Financial Analysis System using Efficient Categorical Embeddings with Cat Boost, ELECTRA, t-SNE, and Genetic Algorithms," *International Journal of Engineering*, vol. 13, no. 3, 2023.

[15]. R. P. Nippatla, "A Secure Cloud-Based Financial Analysis System for Enhancing Monte Carlo Simulations and Deep Belief Network Models Using Bulk Synchronous Parallel Processing," *International Journal of Information Technology and Computer Engineering*, vol. 6, no. 3, pp. 89–100, Jul. 2018.

[16]. R. P. Nippatla, "AI and ML-Driven Blockchain-Based Secure Employee Data Management: Applications of Distributed Control and Tensor Decomposition in HRM," *International Journal of Engineering Research and Science & Technology*, vol. 15, no. 2, pp. 1–16, Jun. 2019.

[17]. D. T. Valivarthi and T. Leaders, "Fog Computing-Based Optimized and Secured IoT Data Sharing Using CMA-ES and Firefly Algorithm with DAG Protocols and Federated Byzantine Agreement," *International Journal of Engineering*, vol. 13, no. 1, 2023.

[18]. B. Kadiyala, S. K. Alavilli, R. P. Nippatla, S. Boyapati, and C. Vasamsetty, "INTEGRATING MULTIVARIATE QUADRATIC CRYPTOGRAPHY WITH AFFINITY PROPAGATION FOR SECURE DOCUMENT CLUSTERING IN IOT DATA SHARING," *International Journal of Information Technology and Computer Engineering*, vol. 11, no. 3, pp. 163–178, Oct. 2023.

[19]. B. Kadiyala and H. Kaur, "DYNAMIC LOAD BALANCING AND SECURE IOT DATA SHARING USING INFINITE GAUSSIAN MIXTURE MODELS AND PLONK," vol. 7, no. 2, 2022.

[20]. B. Kadiyala, "INTEGRATING DBSCAN AND FUZZY C-MEANS WITH HYBRID ABC-DE FOR EFFICIENT RESOURCE ALLOCATION AND SECURED IOT DATA SHARING IN FOG COMPUTING," *International Journal of HRM and Organizational Behavior*, vol. 7, no. 4, pp. 1–13, Oct. 2019.

[21]. B. Kadiyala, "Multi-Swarm Adaptive Differential Evolution and Gaussian Walk Group Search Optimization for Secured Iot Data Sharing Using Super Singular Elliptic Curve Isogeny Cryptography," vol. 8, no. 3, 2020.

[22]. B. Kadiyala and H. Kaur, "Secured IoT Data Sharing through Decentralized Cultural Co-Evolutionary Optimization and Anisotropic Random Walks with Isogeny- Based Hybrid Cryptography," *Journal of Science & Technology (JST)*, vol. 6, no. 6, Art. no. 6, Dec. 2021.

[23]. C. Vasamsetty and H. Kaur, "OPTIMIZING HEALTHCARE DATA ANALYSIS: A CLOUD COMPUTING APPROACH USING PARTICLE SWARM OPTIMIZATION WITH TIME-VARYING ACCELERATION COEFFICIENTS (PSO-

TVAC)," *Journal of Science & Technology (JST)*, vol. 6, no. 5, Art. no. 5, Sep. 2021.

[24]. S. H. Eedala, "Financial Fraud Detection Dataset." Accessed: Feb. 28, 2025. [Online]. Available: https://www.kaggle.com/datasets/sriharshaeedala/financial-fraud-detection-dataset