A Peer Reviewed Journal

Available online at http://www.mbbcollege.in

VIDYAPATTAN: A JOURNAL OF SOCIAL AND SCIENTIFIC RESEARCH

Vidyapattan: A Journal of Social and Scientific Research
Volume 1, Issue, 1, pp. 1-8, June 2026

RESEARCH ARTICLE

Unmasking Hidden Anomalies in Financial Data: A Deep Autoencoder-Based Approach for Detecting Fraudulent Transactions

Durga Sowmya K V N[1], Anurag De[1], Gautam Pal[2*], and Ananya Mukherjee[3]

[1]School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh- 522241, India
[2] Computer Science and Engineering Department, Tripura Institute of Technology, Narsingarh, Tripura-799015, India
[3]Electronics and Communication Engineering Department, Tripura Institute of Technology, Narsingarh, Tripura-799015, India

ARTICLE INFO

Article History:
Received: 13.05.2026
Revised: 22.05.2026
Accepted: 15.06.2026
Published: 30.06.2026

Corresponding author:
Gautam Pal, Associate Professor,
Tripura Institute of Technology
Agartala, Trpiura-799015, India
Email: gautamtit2008@gmail.com

KEYWORDS:
Deep Autoencoders,
Anomaly Detection,
Financial Transactions,
Fraud Detection,
Real-Time Detection

ABSTRACT

The risk of financial fraud keeps rising, making it increasingly important to build efficient anomaly detection systems. This research examines how effective an algorithm is at identifying fraudulent financial transactions. For example, this can include unusual payment patterns or transactions that sharply differ from a customer's normal behavior. The proposed method uses a deep autoencoder to reduce the data's dimensionality and reconstruct it through several hidden layers. To address class imbalance, the intended method applies oversampling methods as well as increasing generalization of the model by using dropout and batch normalization during the model training process. Performance scores show that when the model is only trained using legitimate transactions and tested on an independent dataset, it achieved an AUC-ROC score of .99 and a precision of 95%. Given these results, the model has demonstrated very strong performance in the area of anomaly detection. Although this model performs exceptionally well in detecting non-linear relationships within the data, many of the challenges faced in this regard will remain, for example interpretability issues and reliance on high quality data.

1. INTRODUCTION

Fraudulent credit card transactions have developed at a staggering rate with the increased number of digital financial transactions leading to tremendous problems for customers and banks. Current statistics show that fraud represents billions of dollars in losses annually and there is a pressing need for effective fraud detection systems. In addition, traditional methods are unable to quickly respond to the rapidly changing techniques of the fraudsters; therefore, it is imperative that banks and other institutions implement more advanced forms of machine learning technology. To address this issue, this study develops a new methodology for detecting fraudulent transactions based on the use of deep autoencoder neural networks (Su et al., 2019). This research uses over-sampling methods to find solutions for class imbalanced data using deep autoencoder networks to detect anomalies in the data. The model was trained on valid transactions and deviations from the reconstructed transactions indicated fraudulent activities. Thus, creating an easyto-use and scalable approach for fraud detection is highly essential. The significant contributions of this research include:

- The use of a deep autoencoder model is applied to identify fraudulent transactions by identifying any anomalies in the reconstruction errors through patterns identified within those transaction types.
- By employing an oversampling approach to handle the imbalanced data within the datasets, the model will have enhanced capability to detect those very rare instances of fraud.
- This methodology produces a greater than average AUC-ROC score which signifies the effectiveness of this methodology to identify fraudulent transactions within a real-world application.

The remainder of the paper is structured in the following manner: Section 2 provides insights into the dataset and the preprocessing methods applied in this research. Section 3 reviews pertinent literature on fraud detection. Section 4 details the proposed approach, encompassing framework and the training methodology. Section 5 analyzes the experimental outcomes and assesses the effectiveness of the model. Lastly, Section 6 wraps up the paper by highlighting significant findings and suggesting possible future enhancements.

2. DATA DESCRIPTION

The dataset that is used in this research study is taken from Kaggle (https://www.kaggle.com, 2023) and contains anonymized credit card transaction records. It comprises 21,878 transactions, with only 86 classified as fraudulent, highlighting a significant class imbalance. The dataset consists of 33 features. Key attributes include 'Time' (seconds since the first recorded transaction), 'V1–V28' (PCA-transformed features), 'Amount' (transaction value), and 'Class' (0 for legitimate and 1 for fraudulent transactions).

Table 1. Statistical Summary of Transaction Amounts for Legitimate and Fraudulent Transactions

Transaction Type	Mean Amount	Standard Deviation
Legitimate Transactions	12,026.31	6,929.50
Fraudulent Transactions	12,057.60	6,909.75

Table 1 provides the mean transaction amount and standard deviation for both legitimate and fraudulent transactions, offering insights into their statistical distribution. Analysis reveals that fraudulent transactions, on average, have slightly higher **amounts than legitimate ones, though** their standard deviation remains similar. This indicates that fraud cases do not significantly deviate in monetary distribution. However, other behavioural patterns in transaction data aid in fraud detection.

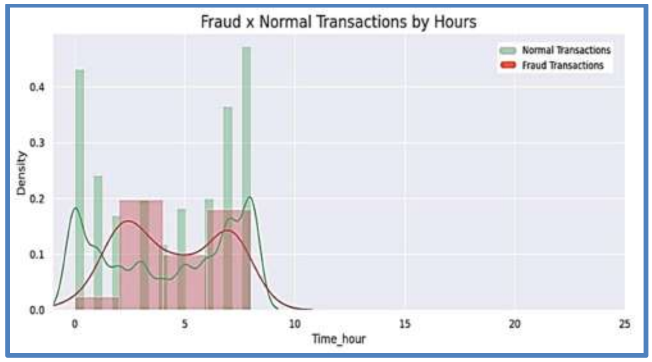

Figure 1. Distribution of fraudulent and normal transactions over different hours of the day

Figure 2. Distribution of fraudulent and normal transactions over different minutes of the day.

Figures 1 and 2 illustrate the distribution of fraudulent and legitimate transactions based on time, specifically across different hours and minutes of the day. The density plots indicate patterns in transaction behaviour, highlighting potential time-based trends in fraudulent activities. These insights help in identifying high-risk time frames for fraud detection.

3. REVIEW OF LITERATURE

In the field of credit card fraud analysis, an investigation was carried out into anomaly detection, applying an autoencoder model to the detection of fraudulent transactions on an imbalanced dataset.

The designed autoencoder was, contrary to the design objective, unsuccessful in capturing slight patterns of fraud as a consequence of the large imbalance ratio and demonstrated a large false-negative rate. The necessity for further enhancement of fraud detection using advanced resampling algorithms or hybridization methods were confirmed (Thimonier et al., 2024).

A comparative evaluation of anomaly detection methods for fraud detection assessed various techniques on real-world credit card transaction data. The study has shown that many methods were unable to detect fraudulent activity because of the differences in distribution between the actual data and the distributed data, resulting in lower detection capabilities. As concluded, it is essential that fraud detection models dynamically train and monitor in order to have accurate fraud detection results due to the ever-changing characteristics of fraud (Islam et al., 2023).

A study looking at resampling techniques for detecting credit card fraud identified and analyzed an array of techniques to resolve the class imbalance problem. The findings of this study indicate that many of the leading resampling techniques did not provide considerable gains when it came to performance due to the limits placed on their use by computational aspects, or simply did not provide the necessary improvement. Consequently, resampling techniques have the potential to provide balance to the dataset; however, they will be insufficient on their own when it comes to improving fraud detection rates, thereby requiring the development of additional feature engineering and anomaly detection methods (Chugh et al., 2025).

Many problems have been found when using machine learning algorithms for identifying fraudulent activity due to a big difference in the number of fraudulent and legitimate financial transactions (class imbalance); lack of consistent benchmark data; and no access to private financial records. Some issues related to using machine learning techniques to identify fraud were studied, including extreme class imbalance, lack of standard benchmark datasets, and restricted access to confidential transaction records. These issues make it difficult to develop effective models used for detecting fraud and require new ways to improve the performance of algorithms (Kulatilleke et al., 2022). A systematic review of existing credit card fraud detection methods revealed limitations of machine learning technologies caused by class imbalance and dataset limitations. The review also indicated that there is difficulty in obtaining a high-performance level from a model because of insufficient amounts of quality datasets, subsequently restricting the performance level of several models by limiting their size (Priscilla & Prabha, 2020).

Looking at recent improvements for deep learning on financial fraud detection shown to have issues with propriety research unincluded, lack of standardization in evaluating metrics, and regulation for implementation can slow further progress. The issues of unbalanced data and no interpretation are yet to be addressed and further work is needed in order to improve accuracy in credit card fraud detection (Chen et al., 2025). In this work, we offer an in depth look at the leading AI techniques for detecting credit card fraud with improved accuracy and reduced false alarms. The investigation found a number of shortcomings in the paper to include the lack of real-world application, inconsistency in the evaluation metrics, and little mention of adversarial attack. Inability to scale to large financial systems should also be explored further (Hafez et al., 2025).

4. METHODOLOGY

In this paper, deep autoencoder networks are used in conjunction with unsupervised learning for anomaly detection in financial transactions. A high reconstruction error indicates anomaly because fraudulent transactions are uncommon, and thus deviate significantly from usual transactions. This method first performs data cleaning such as filling missing values in the data, normalizing the numeric features and removing outliers (Dash et al., 2023). Important features are selected based on amount, frequency, and the occurrence of prior fraud. In the detection of anomaly, there are three stages; coding the transaction, evaluating the loss in reconstruction and determining anomaly with the threshold computed with the statistics of the loss. We optimize using gradient based methods and apply dropout and batch normalization for a better generalization. The dataset is divided into training, validation, and testing sets in order to improve reliability. Model accuracy is evaluated by calculating precision, recall and AUC. Deep autoencoder successfully detects fraud in financial transactions.

4.1. Data Preprocessing

The most significant step in getting the financial transactions ready for anomaly detection is the preparation of the data. Raw financial transactions can suffer from incomplete data, varying scales and outliers, which can severely reduce the effectiveness of a deep autoencoder network. In order to maximize accuracy of the model, a number of methods of pre-processing is applied prior to training.

Missing Value Imputation

There could be a variety of reasons why transaction values might be missing from the financial transaction records. These include errors in the system's log, partially logged transactions, or conditions specific to a user, none of which could lead the learning process misleading. For any of the missing numerical attributes, mean or median imputation is applied. Mean imputation is used for normal distributed attribute whereas for a non-normal distributed attribute median imputation is employed. For an attribute which has strong dependency on other attributes k-Nearest Neighbors imputation algorithm (K-NN imputation) is employed and missing attribute values are estimated. Missing category in categorical attribute is given the modal class or unknown class.

Data Cleaning and Inconsistency Removal

The data quality is extremely important for the success of any deep learning model. Financial transaction data obtained from different sources and covers a considerably large amount of time may include various types of inconsistent data (Priya & Chahar, 2023). Inconsistent data can be in terms of invalid characters, date inconsistent data, or illogical transactions amounts. Incorrect data may cause the error in training the deep auto encoder. Therefore, we are going to check the inconsistencies in the financial transaction data and fix it; for example, by cleaning the invalid characters and removing illogical amounts or data by ensuring the data validation rules. It may also include identifying and correcting the date inconsistencies.

4.2. Deep Autoencoder Architecture

In this work, we implemented deep auto-encoder which can be represented as a sequence of encoder and decoder to learn a latent representation of the normal transaction and detects the anomaly based on reconstruction error. It was built with TensorFlow (Géron, 2022) and Keras, which are composed by two parts, encoder and decoder. They are all constructed with fully connected layers (Silva et al., 2025).

The encoder is used to reduce the input feature to a lower-dimension latent space. The first layer is a fully connected layer of 200 neurons with ReLU activation (Zhang et al., 2024), followed by a dropout layer with rate=0.1 to prevent over-fitting (Bejani and Ghatee 2021). Then, a layer of 100 neurons with ReLU activation, and finally a bottleneck layer with 5 neurons which has captured the characteristic of the transactions.

The decoder is designed to reconstruct the transaction feature from a low-dimension latent space to the original feature space. The first decoder layer is trained to recover the 5 neuron latent vector to a 100 neuron vector, followed by a second layer with 100 neurons as the previous layer with ReLU activation. Another dropout layer is included to improve generalization. The last output layer utilizes a sigmoid activation function to reconstruct the original feature set. The autoencoder is compiled using the Mean Squared Error (MSE) loss function in the equation (1) given below (Hodson et al., 2021).

$$L(x,x^{\wedge}) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x_i})(1)$$

where x is the input transaction and $x^{\wedge}$ is the reconstructed transaction. Transactions with high reconstruction errors are flagged as anomalies.

To optimize training, Early Stopping (Ji et al., 2021) is applied to halt training if the validation loss does not improve after five epochs. The model is trained on normal transactions and evaluated using reconstruction error thresholds to detect fraudulent activity.

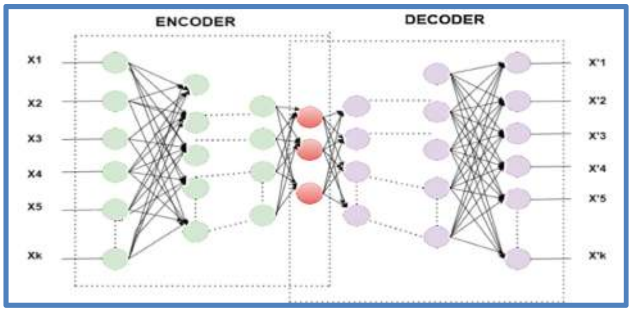

Figure 3. Model Architecture

This architecture enables robust and scalable fraud detection, effectively distinguishing anomalies from normal financial transactions. In Figure 3 shows the architecture of a deep autoencoder depicted with an encoder and a decoder. The encoder compresses input features (X1, X2, ..., Xk) into a lower-dimensional latent representation, while the decoder reconstructs them as (X′1, X′2, ..., X′ k).

4.3.Training Process

We have designed the architecture of the deep autoencoder and the training dataset of financial transactions has been processed. In this section, we focus on the training procedure of the autoencoder to learn the normal transaction patterns and then differentiate them from anomalies which represent fraud activities. The procedure is as follows:

Data Splitting: In order to ensure that the trained autoencoder is generalized enough, so as to not overfit on the training data, we shall use three disjoint sets: Training Set, Validation Set, and Test Set, to split the preprocessed financial transaction data more thoroughly:

Training Set (usually 70-80): The data set which will actually be used to train the autoencoder. It is on this data set that the autoencoder will discover the underlying structure of "normal" financial transactions by working through its examples and minimizing the reconstruction error.

Validation Set (10-15): The purpose of this set is to avoid overfitting the model on the training data. This set will be used during the training process to measure the performance of the autoencoder and thus determine any signs of overfitting. By evaluating the training model on the validation data set we will discover the most optimal point to cease the training of the autoencoder, making sure it does not just memorize the examples from the training data.

Testing Set (10-15): This data set is finally used as an objective measure to test how well our trained autoencoder perform on unseen data. By testing it on this data set, we gain insight on how efficient is our trained autoencoder in discovering the anomalies corresponding to financial fraud in real world conditions.

Model Training: The central component in extracting the underlying characteristics of typical financial transactions is the deep autoencoder, consisting of the encoder and the decoder. We would employ an effective optimizer such as Adam or RMSprop (Ahda et al., 2024) to supervise learning, where optimizers adjust the learning rate in the process of training, leading to much faster learning, and avoidance of reaching non-optimum solutions. The autoencoder would then be trained on the training set; the autoencoder will receive financial transaction data points iteratively. On each iteration, the process can be described as:

Forward Pass: A single financial transaction example is given as the input to the encoder. The encoder would gradually compress the data while extracting key characteristics that constitute the proper transactions and send this to bottleneck layer, this layer now contains only the essence of normal financial transactions.

Backward Pass and Reconstruction: The latent variable at the bottleneck layer will now be fed into the decoder where it will be decompressed so that the original transaction is regenerated at the output layer. *Parameter Update:* The weights and biases of the autoencoder's network will be adjusted using the loss calculated, based on the optimizer. This step would cause the weights and biases of the network to minimize the loss function.

4.4. Anomaly Detection

Once trained effectively, the deep autoencoder can then be applied for the anomaly detection purposes. That's because the deep autoencoder, trained on normal financial transactions, can reconstruct normal financial transactions accurately. A normal transaction has small reconstruction error while a fraudulent transaction is expected to have high reconstruction error. By looking at the reconstruction error for each transaction, the autoencoder works as a protector and identifies anomalous transaction that should be investigated further. Because the anomalous transactions by nature don't conform to normal transaction pattern they won't be learned properly and thus lead to high reconstruction error. So for the transactions with reconstruction error beyond some threshold can be marked as potential anomalies, a sensitive fraud detection approach of the deep autoencoder. There are advantages. A supervised technique (Shetty et al., 2022) can be applied only if the labeled data indicating fraud transactions can be available while the autoencoder can learn from the abundant of normal financial transactions. The autoencoder "is capable of learning complicated, non-linear relations" (Schuster et al., 2022) within data. Compared with rules designed for the particular known fraud patterns, the autoencoder can easily find new fraud schemes, because the definition of normalcy becomes updated in each training step and the rules should be adapted as new kind of fraud comes up.

5. RESULTS AND DISCUSSION

The autoencoder model's performance was assessed using a financial transaction dataset, which was preprocessed to address missing values, standardize numerical features, and eliminate unimportant attributes. The major goal of this assessment was to find out how well deep autoencoder networks (DAEs) can detect fraudulent transactions by examining reconstruction errors.

The deep autoencoder was trained exclusively on normal transaction data and evaluated on a test dataset containing both normal and fraudulent transactions. The results demonstrate that the model effectively differentiates between fraudulent and legitimate transactions, achieving an AUC-ROC score of 0.99, which indicates excellent discriminatory ability. The precision score of 0.95 signifies that most of the flagged anomalies were indeed fraudulent transactions. From the results, DAEs performed highly efficiently for detecting fraudulent financial transactions with an optimal balance between false positive rate (FPR) and false negative rate (FNR) and a high level of accuracy and precision. One of the strengths of the model is that it can identify usual transactional features with-out any labeled fraudulent data. This feature is more beneficial in real financial scenarios, in whichactual occurrences of fraud are rare and hard to categorize. The train and test loss under different epochs are illustrated in **Figure 4**.

The train loss (blue line) decreases constantly, suggesting that the model learns the data very quickly. The test loss (orange line) decreases as well but fluctuates more, implying that the model generalizes well. Since both the curves are going downwards and getting closer to the lower value, we conclude that overfitting does not occur. **Figure 5** shows the ROC curve, which is the way of measuring how successful the classifier is by looking at true and false positive rates. It suggests that it performs better as it gets close to the top left corner (high specificity and high sensitivity). The AUC under the curve is 0.99 and it has outstanding performance in anomaly detection.

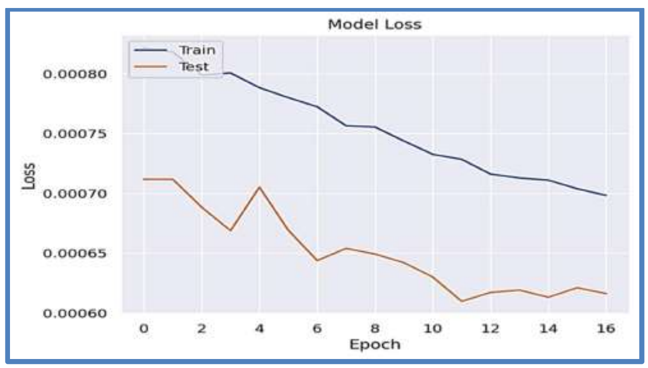

Figure 4. Model train and test loss

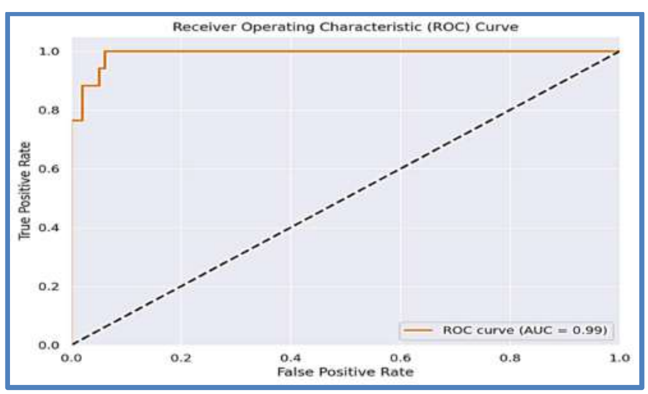

Figure 5. ROC curve for the model's performance.

6. CONCLUSION

The deep autoencoder has played a significant role in financial fraud detection. It exploits massive, unlabeled datasets to find anomalous patterns. Deep autoencoders can identify anomalous deviations due to the non-linear and complex relationships it can learn between different attributes in a transaction. They serve as a passive fraud detection system. The performance of deep autoencoders heavily relies on data quality; bias in the dataset will affect detection rate. Ongoing monitoring of models is required, as is the investigation into methods to increase the explainability of deep auto-encoders.

Combining the deep autoencoder with other statistical anomaly detection methods and rule-based systems would enhance the detection rate of a fraud detection system. Use of domain knowledge through feature engineering can lead to improved anomaly classification. Future work on the deep autoencoder that addresses privacy concerns by using federated learning could allow financial institutions to train their model on decentralized data. Additionally, real-time anomaly detection by improving efficiency of deep autoencoder architectures would contribute to immediate intervention of fraudulent activities.

ACKNOWLEDGEMENTS

We convey our sincere thanks to the School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh, India and Computer Science and Engineering Department, Tripura Institute of Technology, Narsingarh, Tripura, India for providing us the infrastructural support to carry out this research work. .

CONFLICTS OF INTEREST

The author declared that there is no conflict of interest.

FUNDING STATEMENT

No funding has been received to conduct this research work.

REFERENCES

Ahda, F.A.I., Wibawa, A.P., Prasetya, D.D., Sulistyo, D.A. (2024). Comparison of Adam Optimization and RMS prop in Minangkabau-Indonesian Bidirectional Translation with Neural Machine Translation. JOIV: International Journal on Informatics Visualization 8 (1), 231-238. https://dx.doi.org/10.62527/joiv.8.1.1818.

Bejani, M.M., Ghatee, M. (2021). A systematic review on overfitting control in shallow and deep neural networks. Artificial Intelligence Review 54 (8), 6391-6438. https://doi.org/10.1007/s10462-021-09975-1

Chen, Y., Zhao, C., Xu, Y., Nie, C. (2025). Year-over-Year Developments in Financial Fraud Detection via Deep Learning: A Systematic Literature Review. arXiv preprint arXiv:2502.00201. https://doi.org/10.48550/arXiv.2502.00201

Chugh, B., Malik, N., Gupta, D., Alkahtani, B.S. (2025). A probabilistic approach driven credit card anomaly detection with CBLOF and isolation forest models. Alexandria Engineering Journal 114, 231-242. https://doi.org/10.1016/j.aej.2024.11.054

Dash, C.S.K., Behera, A.K., Dehuri, S., Ghosh, A. (2023). An outliers detection and elimination framework in classification task of data mining. Decision Analytics Journal 6, (2023): 100164. https://doi.org/10.1016/j.dajour.2023.100164

Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, Inc.

Hafez, I.Y., Hafez, A.Y., Saleh, A., Abd El-Mageed, A.A., Abohany, A.A. (2025). A systematic review of AI-enhanced techniques in credit card fraud detection. Journal of Big Data 12 (6). https://doi.org/10.1186/s40537-024-01048-8

Hodson, T.O., Over, T.M., Foks, S.S. (2021). Mean squared error, deconstructed. Journal of Advances in Modeling Earth Systems 13 (12), e2021MS002681. https://doi.org/10.1029/2021MS002681

https://www.kaggle.com/datasets/nelgiriyewithana/credit-card-fraud-detection-dataset-2023. (Credit card fraud detection dataset: last accessed 2025/04/10).

Islam, M.A., Uddin, M.A., Aryal, S., Stea, G. (2023). An ensemble learning approach for anomaly detection in credit card data with imbalanced and overlapped classes. Journal of Information Security and Applications 78. https://doi.org/10.1016/j.jisa.2023.103618

Ji, Z., Li, J., Telgarsky, M. (2021). Early-stopped neural networks are consistent. Advances in Neural Information Processing Systems 34, 1805-1817.

Kulatilleke, G.K. (2022). Challenges and complexities in machine learning based credit card fraud detection. arXiv preprint arXiv:2208.10943. DOI: https://doi.org/10.48550/arXiv.2208.10943

Priscilla, C.V., Prabha, D.P. (2020). Credit card fraud detection: A systematic review. In Intelligent Computing Paradigm and Cutting-edge Technologies: Proceedings of the First International Conference on Innovative Computing and Cutting-edge Technologies. ICICCT 2019, Istanbul, Turkey, pp. 290-303. Springer International Publishing, (2020). https://doi.org/10.1007/978-3-030-38501-9_29

Priya, V.L., Chahar, A. (2023). Artificial intelligence as a tool for enhanced data integrity and data security. In 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), pp. 781-785. IEEE. https://doi.org/10.1109/AISC56616.2023.10085250

Schuster, N.A., Rijnhart, J.J., Twisk, J.W., Heymans, M.W. (2022). Modeling non-linear relationships in epidemiological data: the application and interpretation of spline models. Frontiers in Epidemiology 2, 975380. https://doi.org/10.3389/fepid.2022.975380.

Shetty, S.H., Shetty, S., Singh, C., Rao, A. (2022). Supervised machine learning: algorithms and applications. Fundamentals and methods of machine and deep learning: algorithms, tools and applications, 1-16. https://doi.org/10.1002/9781119821908.ch1

Silva, J.F., Faraggi, V., Ramirez, C., Egaña, A., Pavez, E. (2025). Understanding encoder-decoder structures in machine learning using information measures. Signal Processing 234, 109983. https://doi.org/10.1016/j.sigpro.2025.109983

Su, Y., Li, J., Plaza, A., Marinoni, A., Gamba, P., Chakravortty, S. (2019). DAEN: Deep autoencoder networks for hyperspectral unmixing. IEEE Transactions on Geoscience and Remote Sensing 57 (7), 4309-4321. https://doi.org/10.1109/TGRS.2018.2890633

Thimonier, H., Popineau, F., Rimmel, A., Doan, B.L., Daniel, F. (2024). Comparative evaluation of anomaly detection methods for fraud detection in online credit card payments. In International Congress on Information and Communication Technology, pp. 37-50. Singapore: Springer Nature Singapore, (2024).

Zhang, Z., Song, Y., Yu, G., Han, X., Lin, Y., Xiao, C., Song, C., Liu, Z., Mi, Z., Sun, M. (2024). ReLU^2 Wins: Discovering Efficient Activation Functions for Sparse LLMs. arXiv preprint arXiv: 2402.03804, (2024). https://doi.org/10.48550/arXiv.2402.03804

Cite this article: Durga Sowmya, K.V.N, De, A., Pal, G., & Mukherjee, A. (**2026**). Unmasking Hidden Anomalies in Financial Data: A Deep Autoencoder-Based Approach for Detecting Fraudulent Transactions. *Vidyapattan: A Journal of Social and Scientific Research*, Vol.1, Issue 1, pp. 1-8.