Detecting Fraud in Automated Teller Machine Transactions in the Nigerian Bank System Using Unsupervised Deep Learning

Ojulari Hakeem Olusegun, Oke Alice Oluwafunke & Arulogun Oladiran Tayo
Department of Computer Science and Engineering
Ladoke Akintola University of Technology
Ogbomoso, Oyo State, Nigeria
E-mail: ojulari.hakeem@microconcept.com.ng & update_ak47@yahoo.com, **Ph:** +234-8120808094
Email: aooke@lautech.edu.ng, **Ph:** +234-830729018

ABSTRACT

Automated Teller Machines (ATMs) are globally adopted for financial transactions. However, the exponential rise in ATM fraud poses a significant risk to Nigeria's digital payment and banking systems, with many banks failing to provide intelligent services to ATM customers. Moreover, existing research predominantly employed supervised learning on labelled data for ATM fraud detection, overlooking unlabelled customer-clustered transaction data. In this study, customer-clustered transaction attributes were employed with unsupervised autoencoder deep learning models to develop a fraud detection system. Three years of historical transaction data were obtained from a selected Nigerian bank's ATM transaction repository, consisting of 1.2 million transaction records from 10 bank customers. The transaction features considered in this work included CardNo, Amount, TransType, BranchCode, and TransDateTime. H2O autoencoder deep learning models, and anomaly detection algorithms were adopted to uncover underlying patterns in ATM transactions. The results indicated a remarkable overall system performance, with an average Mean Squared Error (MSE) of 0.0057, accuracy of 97.6%, precision of 94.8%, recall of 93.5%, and an F1-score of 94.0%. Individual model assessments revealed slightly lower performance for specific customer clusters. The study's findings emphasized the efficacy of unsupervised deep learning in combatting ATM fraud within the Nigerian banking system.

Keywords: ATM, H2O.ai, Unsupervised Learning, Deep Learning, Autoencoder, Fraud detection

CISDI Journal Reference Format
Ojulari, H.O., Oke, A.O. & Arulogun, O.T. (2024): Detecting Fraud in Automated Teller Machine Transactions in the Nigerian Bank System Using Unsupervised Deep Learning. Computing, Information Systems, Development Informatics & Allied Research Journal. Vol 15 No 2, Pp 1-14. dx.doi.org/10.22624/AIMS/CISDI/V15N2P1.
Available online at www.isteams.net/cisdijournal

1. INTRODUCTION

Financial institutions globally play a pivotal role in shaping economies, expanding their influence alongside technological advancements (Fethi and Pasiouras, 2010). In Nigeria, the banking sector has witnessed significant growth, propelled by the introduction of Automated Teller Machines (ATMs) and related services. Strategic investments in self-service and omnichannel ATMs have transformed payment activities, enabling customers to initiate diverse transactions and cross-channel activities previously reliant on manual processes. These services include cheque and cash deposits, bill payments, account inquiries, inter-bank transfers, and loan applications. ATMs have evolved into miniature banks, especially benefiting rural areas with services like account opening, cash deposits and withdrawals. These advantages notwithstanding, the card information printed on the ATM cards aids in financial identity theft and fraudulent transactions (Gupta et al, 2023). Hence, ATM technology poses challenges making ATM fraud a significant concern in the Nigerian banking industry.

ATM fraud is a global phenomenon causing great concern in the banking industry (Prakash et al., 2018) and substantial loss in millions to banks (Sakharova, 2012). Generally, many organizations are aware of the importance of utilizing advanced technologies to provide efficient fraud detection systems, thereby preventing fraudulent transactions from occurring (Rangineni and Marupaka, 2023). The prevalence of financial fraud in Nigeria specifically underscores the need for advanced fraud detection mechanisms. There was an increase in the frequency of ATM fraud cases from 7,181 in 2014 to 8,039 in 2015, an increase of 11.95%, which summed up to a ₦1.5 billion loss (Ifeanyi, 2016). The number of bank fraud and forgery cases in 2020 increased to 146,183 and stood at a ₦120.79 billion loss (James, 2022), and ATM card-related fraud had the highest frequency, accounting for 39.81%. ATM fraudsters steal money from customer accounts by card cloning, using skimming devices, shoulder surfing, or stealing card details to perform unauthorized transactions (Braimah and Okonkwo, 2016). Though banks have installed anti-skimming measures on ATMs to prevent card detail theft, threats like lost cards, card theft, and shoulder surfing persist.

Despite several efforts to control ATM fraud, previous studies primarily relied on supervised learning approaches with labelled data. However, these approaches have limitations in the Nigerian banking system, such as the scarce availability of labelled data, which is both costly and time-consuming to obtain. These constraints hinder the effective capture of evolving patterns of fraudulent activities in individual customer transactions, particularly when faced with large volumes of unlabelled transaction data. Additionally, the approaches may suffer from class imbalance problems, affecting their generalization.

The core objective of this study was to develop an intelligent fraud detection system using individual customer historical transaction data. The empirical research methodology centered on the utilization of unsupervised learning of deep neural network models, with a specific emphasis on autoencoders using the H2O.ai library. These autoencoder models were employed to intricately learn individual customer transaction patterns, enabling them to accurately detect anomalies within real-time transactions that may indicate malicious intent. This approach addresses the research gap and enhances the capabilities of ATM fraud detection in the Nigerian banking system.

2. LITERATURE REVIEW:

Previous studies predominantly concentrated on employing supervised learning techniques for fraud detection in financial transactions, particularly those involving credit cards. The majority of experimental research on fraud detection utilized credit card datasets. Ojulari et al, (2017) underscored the significance of directly extracting transaction information from ATM electronic journals (Ejs) within the Nigerian banking system. However, a noticeable gap exists in the literature regarding the application of extracted information from ATM transaction EJ, particularly in the context of lacking fraud detection techniques. Researchers have addressed the complex challenges in ATM services by primarily focusing on enhancing operational efficiency and mitigating fraud risks.

Venkatesh et al, (2014) pioneered a model that combined clustering and neural network techniques to forecast cash demand in ATMs, offering valuable insights for optimizing cash replenishment. It's important to note, however, that fraud detection was not within the scope of the study. Rajwani et al, (2017) concentrated on predicting ATM cash flow using regression techniques, showcasing the effectiveness of the Long Short-Term Memory (LSTM) model. The study, while offering advancements in transaction forecasting, primarily centred on successful withdrawals and did not explore unsupervised learning for fraud detection. Tsegaye (2017) employed K-means clustering and classifiers for real-time ATM card fraud detection, yet the study's classifiers were not suitable for unlabelled data, revealing a limitation in the broader applicability of the model.

Prakash et al, (2018) investigated machine learning techniques for ATM card fraud detection, achieving high accuracy with labelled data. However, the study did not address the practical challenge of detecting fraud in unlabelled data, limiting its real-world utility. Misra et al. (2020) proposed an autoencoder-based model for credit card fraud detection, demonstrating superior performance with labelled data. However, both Prakash and Misra did not address the challenge of unlabelled data, narrowing its application scope.

Lokanan and Sharma (2022) addressed investment fraud using machine learning techniques, showcasing high accuracy. The work implemented four machine learning techniques which were logistic regression, Decision Tree Classifier (DTC), Random Forest Classifier (RFC), and CatBoast and GridSearchCV. It was observed in the work that RFC and GSCV obtained the highest precision and accuracy of 99%, meaning that 99% of the observations classified as fraud were actually fraudulent. Similarly, Megdad et al. (2022) compared various machine learning algorithms such as Multi-layer Perceptron (MLP) Regressor, Complement Naïve Bayes (NB), MLP Classifier, Gaussian NB, Bernoulli NB, Linear Gradient Boosting Model (LGBM) Classifier, Ada Boost Classifier, Logistic Regression, Bagging Classifier, DTC, and RFC for predicting fraudulent transactions.

The study found that RFC emerged as the best choice for unbalanced datasets with 99.97% of accuracy and precision. However, it's important to note that the classification techniques utilized in both studies are supervised learning methods, which are limited to labelled data requiring a large amount of labelled data, which is costly and time-consuming to obtain. Moreover, the supervised learning techniques failed to effectively capture the dynamic and complex patterns of fraudulent activities in individual customer transactions, especially when dealing with large volumes of unlabelled transaction data.

In a recent study, Gupta et al, (2023) employed Regression Model, Decision Tree, XGBoost, and Artificial Neural Network (ANN) classification techniques. The study applied the Random Over-sampling technique to balance the dataset, aiming to enhance the overall performance of the selected classification algorithms. The results indicated that XGBoost outperformed the others, achieving 99% accuracy and prediction scores. However, despite this success, the reliance on labelled data poses constraints on real-world applicability in the Nigerian banking system, where the available data is predominantly unlabelled. Considering this challenge, supervised learning techniques are not suitable for handling unsupervised learning problems. While supervised learning methods have demonstrated efficacy in fraud classification, the reviewed studies notably lacked the modelling of unlabelled customer transactions to effectively detect ATM fraud at the individual account level.

3. METHODOLOGY

Design Approach:

It was crucial to select machine learning tools for data analytics that best suited the requirements of the intended research, allowing for the rapid development of various prototypes. The choice of the R language was motivated by its status as a robust statistical programming language with a wide range of libraries and extensive adoption within the data analysis and machine learning community. R is particularly well-suited for big data analytics due to its utilization of multicore-threading for resource management and algorithm speed RStudio was chosen as the Integrated Development Environment (IDE). H2O.ai (or simply H2O or H2O deep learning) supports the R language and provides an easy-to-use package for developing neural network models. Its intrinsic one-hot encoding for categorical variables makes it a valuable choice.

In the prototyping process, H2O allows for the design of neural network architectures and training models with customizable parameters. It enhances model creation and facilitates data manipulation, including splitting and shuffling data into groups for training, testing, and validation, as well as data reshaping to better suit neural network training. The use of H2O streamlines the prototyping process, reducing the amount of code required, enhancing portability, and minimizing the time spent on writing non-essential code that wouldn't significantly impact the research. The research approach began with data acquisition, followed by data preprocessing, segmentation, modelling, testing, and result evaluation. In the modelling, testing, and evaluation stages, H2O was employed to train, test, and evaluate the MSE in autoencoders. The model's performance was further evaluated using accuracy, precision, recall and f-score.

Dataset and Preprocessing:

The fraud detection process involved the extraction of datasets from an ATM transaction data repository encompassing multiple Nigerian bank customers, originating from Electronic Journals (EJs). This dataset comprised of extracted and selected features of multinomial values of categorical and continuous variables, including TerminalId, BranchCode, CardNo, TransType, Amount and DateTime (Ojulari et al., 2017) using ATM domain knowledge. Generally, domain knowledge is a contributory factor in data preprocessing within any research space. Therefore, it contributed to understanding the dataset as a mixed-effects distribution (Hedeker, 2003) of customer transactions, as shown in Figure 1. The figure shows how customers (C_1 – C_n) are distributed across several ATMs (A_1 – A_n) in multiple branches ($Branch_1$ – $Branch_n$) of a bank.

The dataset is a 3-year-old collection of ten customers (CardNo 1 to 10) with 1.2 million transactions, that was extracted from EJs and stored in a database. The transaction data contained various transaction types, such as bill payments, fund transfers, third-party payments, mobile top-ups, and cash withdrawals; while transaction metadata lack direct geographical information of individual ATMs, TerminalId and BranchCode identify physical ATMs. This section also outlines the essential steps taken for data cleaning and preprocessing. Missing and invalid values in the dataset were treated as outliers to ensure accuracy and completeness. The InterQuartile Range (IQR) for each numerical variable was calculated, and a threshold for outliers was defined—typically 1.5 times the IQR above the third quartile or below the first quartile. Therefore, outliers were removed from the dataset based on this threshold. After handling outliers using the IQR method, z-score normalization was applied to standardize the data.

The DateTime feature was transformed into Year, Month, Day, and Hour, excluding the year to focus on habitual transaction patterns. These steps collectively prepared the dataset for optimal utilization in the subsequent autoencoder deep learning model. A sample of ATM transaction for a typical customer (CardNo 1) data is shown in Table 1.

Table 1. A Sample of Transaction Data

BranchCode	TransType	Amount	Year	Month	Day	Hour
384	3	20000	19	1	25	7
384	3	20000	19	1	25	7
217	3	5000	19	1	25	7
141	3	4000	19	1	15	2
375	3	4000	19	1	3	9
458	3	4000	19	1	25	4
384	3	3000	19	1	25	7
156	3	20000	19	1	28	9
259	3	1500	19	2	18	5
18	3	1000	19	2	5	1
440	3	8000	19	2	12	11
518	3	10000	19	1	10	1
518	3	10000	19	1	10	1
518	3	10000	19	1	10	1
518	3	10000	19	1	10	1
518	3	10000	19	1	10	1
327	3	10000	19	1	2	9
327	3	20000	19	1	2	9
327	3	20000	19	1	2	9
384	3	3000	19	1	23	5
141	3	1000	19	1	19	10

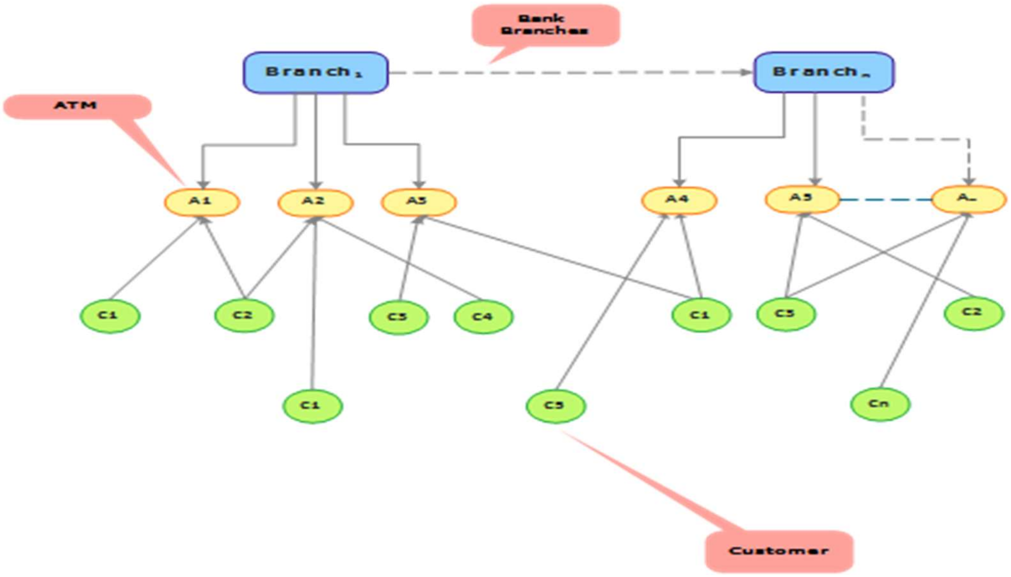

Figure 1. Distribution of Customer Transactions across ATMs

Customer Segmentation:

The purpose of segmentation was to mitigate the complexity of the mixed-effects data distribution of customer transactions. This stage aimed to group the data based on unique CardNo identifiers, thereby effectively dissecting the dataset into individual customer transaction clusters. This created a cognitive customer transaction profile within a cluster. Therefore, given the relatively small number of customers (10) in this study, a simple segmentation process that involves grouping the data by 'CardNo' was found to be the most suitable and appropriate approach for these customer size. This method avoids the complexities associated with categorical clustering algorithms and offers an intuitive way to segment the data based on distinct customer identifiers. Each group reveals the behavioural transaction patterns of individual customers. This process identified related customers with similar transaction patterns and grouped the similar data points into clusters using hierarchical clustering. However, if the number of customers was large, building a model for each customer would be resource-intensive and impractical, and model maintainability would be difficult. Since the number of active customers, as the case is for a typical Nigerian commercial bank ranges from five hundred thousand to more than two million customers.

Model Design

An *H2O* deep learning framework was used for a deep learning autoencoder design as shown in Figure 2 for the individual customer-based transaction clusters with input features, namely: BranchCode, TransType, Amount, Month, Day, and Hour.

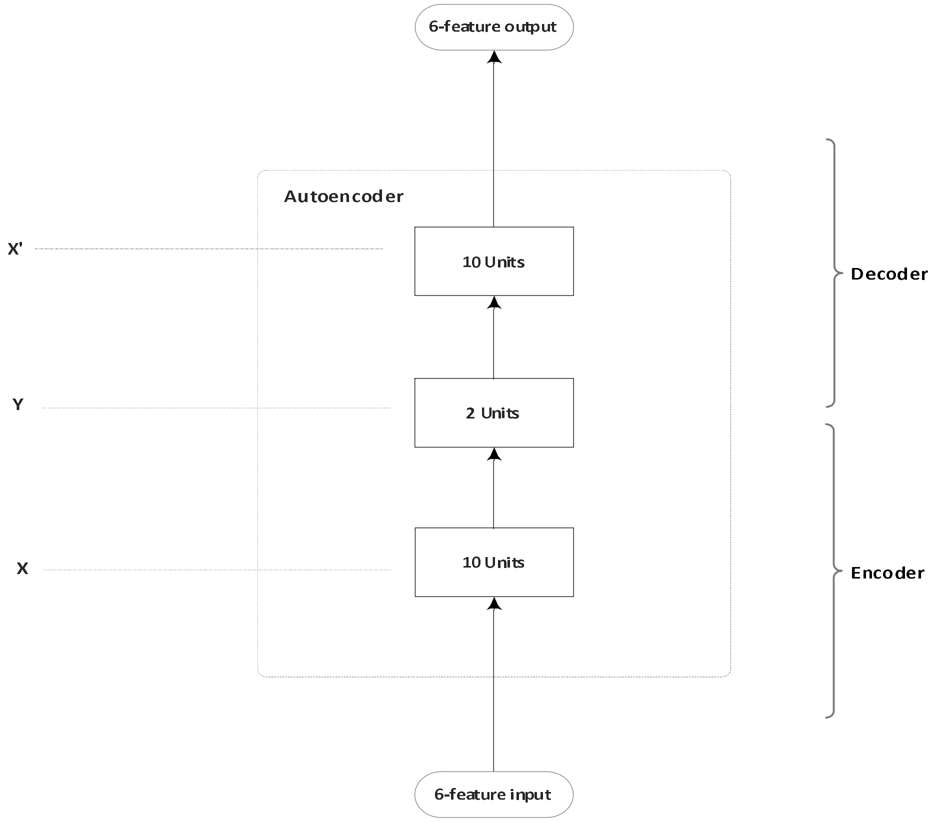

Figure 2. A Deep Learning Autoencoder Design

This autoencoder deep learning architecture involved transforming the input features into lower-dimensional representations at the bottleneck, Y, for effective dimensionality reduction. Subsequently, the transformed features were decoded back to their original transaction input features at the output, X'. The primary objective of this design was to train an autoencoder model for the purpose of anomaly detection. The following shows the design mathematically at each stage.

1. Data input description
   BranchCode – Categorical variable (e.g., 203, 115, 290, etc. not in any order)
   TransType – Categorical variable (e.g., 1, 2, 3 as defined as 1 – Inquiry, 2 – Unknown, 3 – Withdrawal, 4 – Pin Change, 5 – Interbank Transfer, 6 – Third Party Payment, Fund Transfer)
   Amount – Continuous variable (e.g., 100.00, 2000.00, 1000.00, 5000.00, 10000.00, etc.)
   Month – Categorical variable (1 – 12, ordinal values, i.e. Jan - Dec)
   Day – Categorical variable (1 – 31)
   Hour – Categorical variable (0 – 23)
2. Network Architecture:
   a. Input Layer:
      Let $\mathrm{X} = (\mathrm{n}, \{\mathrm{branchCode}, \mathrm{transType}, \mathrm{amount}, \mathrm{month}, \mathrm{day}, \mathrm{hour}\})$
      where n is the number of observations.
   b. Encoder:
      Calculate the weighted sum at the encoder layer by applying Equation 1.
      $$\mathrm{y\_encoder} = \mathrm{X} * \mathrm{w\_enc} + \mathrm{b\_enc} \quad (1)$$
      where $\mathrm{w\_enc}$ represents the weights and $\mathrm{b\_enc}$ represents the biases of the encoder layer. The encoder is expressed by applying an activation function on equation 1 as shown in equation 2.
      $$\mathrm{y} = \mathrm{Tanh}\ (\mathrm{y\_encoder}) \quad (2)$$
      where y is the encoded input at the bottleneck layer
   c. Decoder:
      Calculate the weighted sum at the decoder layer by adapting Equation 3
      $$\mathrm{y\_decoder} = \mathrm{y} * \mathrm{w\_dec} + \mathrm{b\_dec} \quad (3)$$
      where $\mathrm{w\_dec}$ represents the weights and $\mathrm{b\_dec}$ represents the biases of the decoder layer. The decoder is expressed by applying an activation function as shown in Equation 4.
      $$\mathrm{X}' = \mathrm{Tanh}\ (\mathrm{y\_decoder}) \quad (4)$$
      where X' is the reconstructed output of the input X
   d. Output Layer:
      The output layer's activation function is linear, meaning that X' is directly used as the output. The autoencoder model was named $\mathrm{CustModel}$.
   e. Loss Function:
      Calculate the loss between the input data X and the reconstructed data X'. A common loss function for autoencoders is the Mean Squared Error (MSE) as shown in Equation 5:
      $$\mathrm{Loss}(\mathrm{X}, \mathrm{X}') = (1/\mathrm{n}) * \sum(\mathrm{X} - \mathrm{X}')^2 \quad (5)$$
      This calculates the average squared difference between the input and the reconstructed output for all samples n. Hence, the reconstruction MSE is used to select a threshold for anomaly detection for each model.
   f. Threshold:
      Threshold (τ) is selected using the maximum value of MSE and multiplied by 0.85.

Anomaly Detection:

An anomaly detection algorithm from H2O.ai deep learning was applied to the test set to detect potential fraudulent transactions. This algorithm detected deviations from established transactional patterns in the trained model, thereby identifying potential instances of fraud. The anomaly detection process is critical for distinguishing abnormal transactions within the test set. To identify anomalies (fraud) based on specific categories within the transaction features of potential customers, X was considered as the input features, as defined in the autoencoder model design. $\mathrm{CustModel_m}$ represents the customer-based transaction model, and $\mathrm{AtmTrans_j}$ represents new ATM customer transactions as shown in Equation 6.

$$\mathrm{AtmTrans_j} = \{\, \mathrm{x_i}:\ \mathrm{i} = 1, ..\, \mathrm{n}\} \qquad (6)$$

For each unique category i within X which is $\mathrm{x_i}$, the probability $\mathrm{P(X = x_i)}$ was calculated that a transaction belongs to category i. The threshold value, τ, was set to a reasonable value and it represented an acceptable probability range for normal transactions and was based on the reconstruction MSE derived during the training of the autoencoder. Transactions with probabilities outside this range was considered anomalies. In mathematical terms, Anomaly detection can be represented in Equation 7 as:

$$\mathrm{Anomaly}\left(\mathrm{AtmTrans_j},\ \mathrm{CustModel_m}\right) = \{\, \mathrm{P}\big(\mathrm{X} = \mathrm{x_j}\big) < \tau \text{ or } \mathrm{P}\big(\mathrm{X} = \mathrm{x_j}\big) > 1 - \tau\,\} \qquad (7)$$

$\mathrm{AtmTrans_j}$ represents the j-th transaction to be detected, $\mathrm{P}\big(\mathrm{X} = \mathrm{x_j}\big)$ represents the probability that transaction j belongs to category $\mathrm{x_j}$ within the categorical variables in X. This probability is compared with the threshold τ to determine whether the transaction is an anomaly or normal. If $\mathrm{P(X = x_i)} < \tau$ or $\mathrm{P(X = x_i)} > 1 - \tau$, then $\mathrm{AtmTrans_j}$ is classified as an anomaly. It is worth noting that the H2O library offers a robust intrinsic function to detect categorical data anomalies to represent this mathematical approach.

Model Training

Each clustered dataset was split into a training set (70%) and a test set (30%). Autoencoder deep learning model from H2O.ai was utilized to build a model for each cluster using each training set iteratively. The number of models depends on the number of identified clusters. This helped the models to learn more rules and patterns of the dataset. It generated the customer transaction patterns per customer within the transaction time space. The model was trained by optimizing the loss function with respect to the weights and biases in both the encoder and decoder layers. The objective was to minimize the loss function, typically the MSE between the input and the output. The Training involved using backpropagation and gradient descent. The following steps were performed for the model training.

1. Input Presentation:
   The training dataset having preprocessed ATM transaction features serves as the input to the autoencoder deep learning network.
2. Forward Pass:
   The input data was fed forward through the autoencoder neural network inputs. There are two main parts: an encoder and a decoder. The encoder compresses the input data into a lower-dimensional representation, called the latent space or encoding. The decoder then attempts to reconstruct the input data from this compressed representation.

3. Loss Calculation:
   The difference between the original input and the reconstructed output is calculated. This is known as the reconstruction error or loss which is key for calculating reconstruction MSE. The goal during training was to minimize this loss, encouraging the autoencoder to learn a meaningful representation of the input data in the latent space.
4. Backward Pass (Backpropagation):
   The error is propagated backward through the network using a process called backpropagation. The neural network's weights were adjusted to minimize the error, using optimization algorithms like gradient descent as shown in the hyperparameter tunning.
5. Iterative Training:
   Steps 2-4 were repeated for multiple epochs or iterations over the entire training dataset. The iterative nature of training allowed the autoencoder to gradually improve its ability to encode and decode the input data.
6. Model Tuning:
   The performance of the autoencoder is monitored on a validation dataset, and hyperparameters may be tuned to enhance its generalization capabilities. The detection model was tuned using parameters described in Table 2 to minimize the probability of having deniable outcomes of legitimate transactions and suggest the possible result of customer transactions based on transaction features. The training parameters in Table 2 were according to the H2O documentation for hyperparameter tuning.
7. Evaluation:
   Individual trained models were built and saved upon completion of training, with evaluation based on the default performance metric (MSE) in the hyperparameter setting for H2O autoencoder training. Additional performance metrics, including accuracy, precision, recall, and F-score, were also observed for comprehensive evaluation.

Table 2. H2O Autoencoder Design Parameters

Parameter	Value
Activation function	Tanh
Hidden layers	[10, 2, 10]
Number of epochs	100
Standardization	Enabled
Stopping metric	Mean Squared Error (MSE)
Loss function	Automatic
Training samples per iteration	32
Training data shuffle	Enabled
Autoencoder mode	Enabled
L1 regularization	10e-5
Input features	All input features
Training frame	Training set
Model ID	model_1

Testing Models

30% of each customer-clustered dataset was allocated to test the performance of the trained models. The anomaly detection algorithm was utilized to detect anomalies in the test datasets that deviated from regular transaction patterns found in the training sets. Thus, the anomaly detector observed deviations or outliers in customer behaviours related to the test transaction datasets and compared the patterns to the established patterns in the trained models. When suspicious transactions were detected, the observed transactions were marked as fraudulent. These suspicious transactions are either rejected or accepted with verification codes to continue the transactions in the actual ATM environment, depending on the bank's decision.

The testing process for the trained models was as follows:

1. Used the test datasets as the new data points different from the training sets.
2. Applied the trained autoencoder models to reconstruct new data points based on the insights within the models.
3. Utilized the H2O deep learning anomaly detection algorithm to identify potential fraud in the new data points (test sets). It confirmed whether the requesting transaction features deviated from the regular transaction patterns of the customers.
4. Used the training data/sets to test the trained models and compared the patterns with those of the test sets.
5. In the final step, decision-making was reserved for the bank to determine the appropriate course of action. If by chance the suspicious transaction is deemed legitimate, the bank may proceed to generate an authorization code facilitating the continuation of the transaction.

Performance Evaluation:

Performance evaluation was carried out to ascertain the effectiveness of the developed fraud detection system. Utilizing validation datasets, carefully curated to include simulated/intended fraud cases, Mean Squared Error (MSE), accuracy, precision, recall, and F1-score were the metrics used. The inclusion of simulated fraud cases enabled a comprehensive assessment of the model's ability to differentiate between genuine and fraudulent transactions.

Ethical Considerations

This research adhered to ethical standards, ensuring the confidentiality and privacy of individual transactional data. The study complied with relevant data protection regulations and institutional ethical guideline such as General Data Protection Regulations (GDPR) and Payment Card Industry Data Security Standard (PCI-DSS) to adhere to consumer protection policies.

Implementation of the Fraud Detection System

The implementation stage entailed seamlessly integrating the trained autoencoder models into an existing ATM environment, ensuring a smooth deployment of the developed fraud detection system. In Figure 3, the fraud detection system integration involved key components: a Fraud Detection Server (FDS), a Core Banking System (CBS), an ATM Host, and an ATM labelled 1,2,3, and 4 respectively. The integration began with the FDS profiling the bank's customers from CBS and obtaining historical customer's transactions from the ATM Host to train autoencoder models. When a customer initiates a transaction at any ATM, the request is sent to the FDS for validation to determine its legitimacy. If the transaction is deemed legitimate, the FDS instructs the ATM host to proceed with the transaction to the CBS stage for debiting the customer's account and instructing the ATM to dispense cash. In the event of a detected fraud during validation, the ATM host is instructed to decline the transaction.

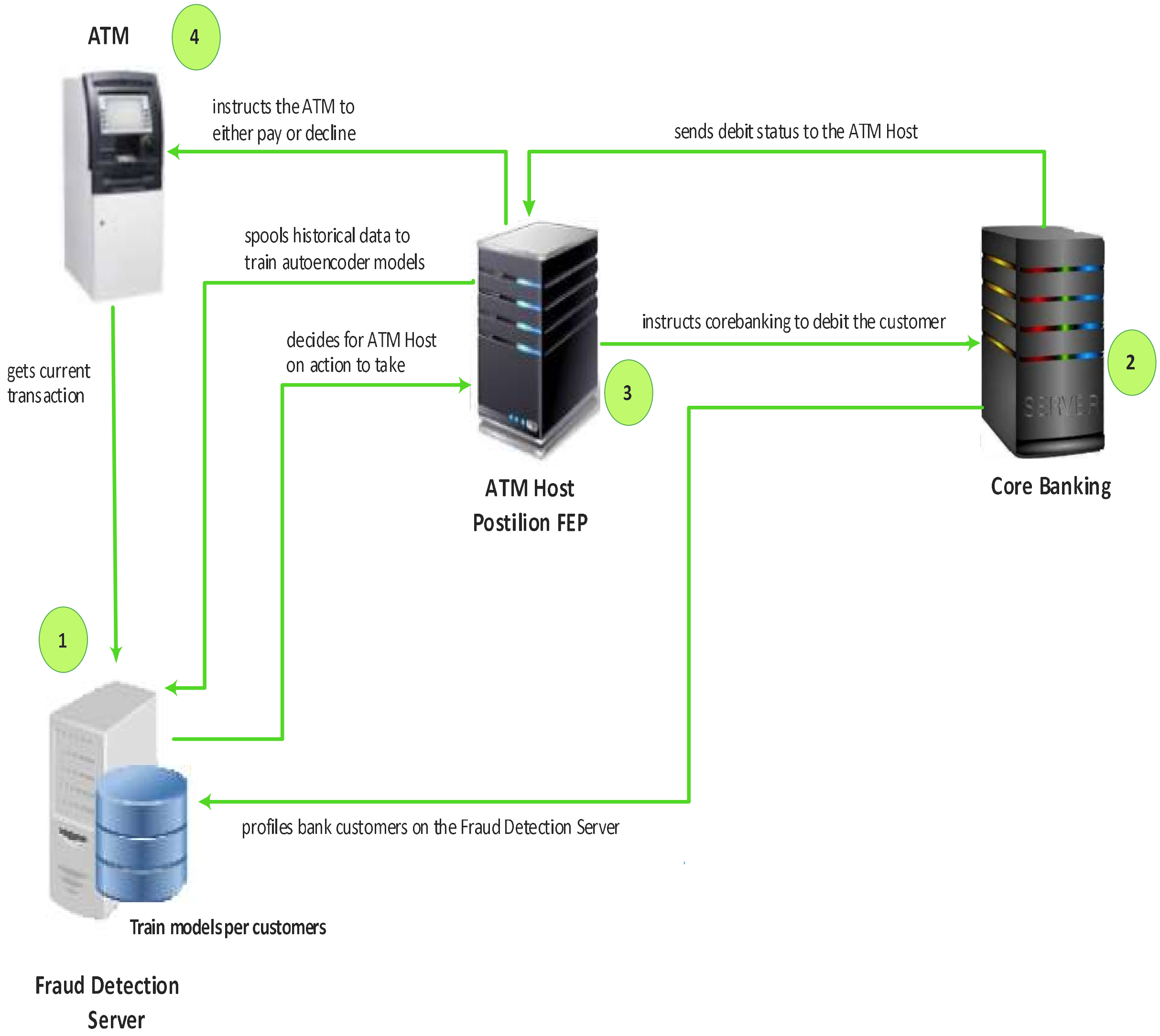

Figure 3: Fraud Detection System Implementation

4. RESULTS AND DISCUSSION:

The H2O.ai library has built-in algorithms to evaluate the Mean Squared Error (MSE) of autoencoder reconstructions. The reconstruction MSE values of the autoencoder models were graphically compared between the training and test sets as shown in Figure 4 and 5. Figure 4 shows the graphical behaviour of the trained model on the training sets, while Figure 5 depicts its behaviour on the test sets. Figure 5 closely resembles 4 in terms of patterns. This proves that the models are not overfitted. The overall performance metrics of the autoencoder deep learning models demonstrated impressive results as shown in Table 3, with an average MSE of 0.0057, accuracy of 97.6%, precision of 94.8%, recall of 93.5%, and F1-score of 94.0%.

While the overall performance is exceptional, individual models, such as CardNos 3 and 6, exhibit slightly lower performance. CardNo 3 shows results of 95.0% accuracy, 89.5% precision, 85.0% recall, and an F1-score of 87.2%, with an MSE of 0.0085. CardNo 6 demonstrates 95.0% accuracy, 85.7% precision, 90.0% recall, and an F1-score of 87.8%, with an MSE of 0.0079.

The results indicate the effectiveness of the autoencoder deep learning model in identifying fraudulent ATM transactions, providing a robust defence against unauthorized activities. It is pertinent to contextualize this research within the broader landscape of fraud detection studies. Comparing the findings with existing literature underscores the innovative contribution of unsupervised learning in ATM fraud detection, emphasizing the unique advantages of autoencoder models.

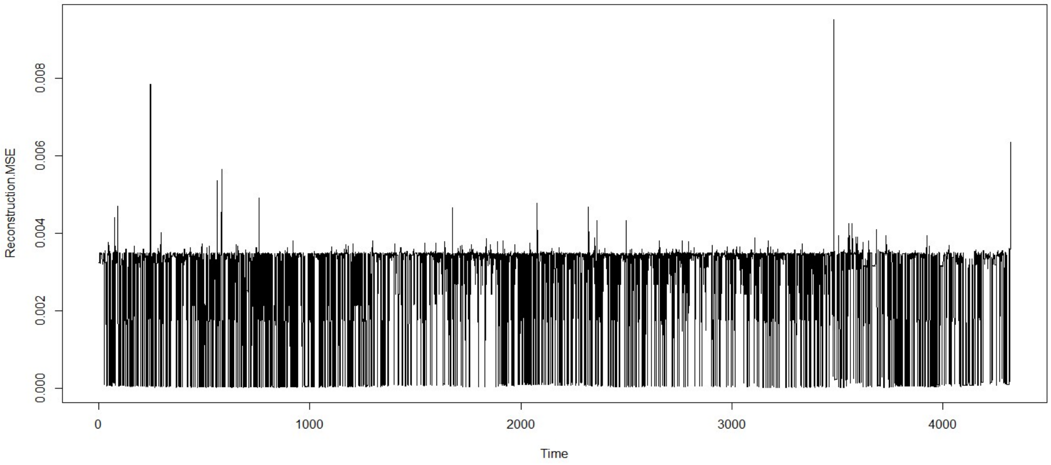

Figure 5: Graphical behaviour of the Trained model to training data

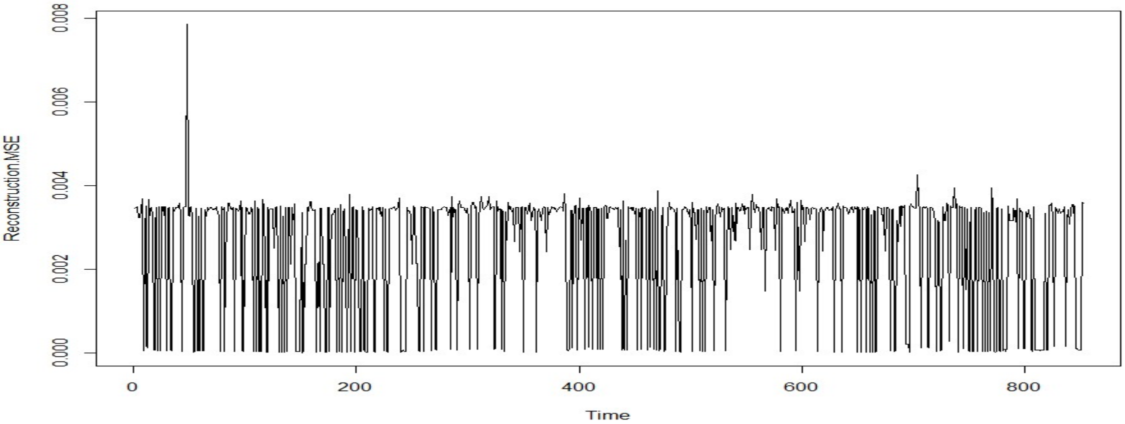

Figure 6: Graphical behaviour of the trained model to the test data

Table 3. Performance Metrics for 10 Autoencoder models

CardNo	MSE	Accuracy	Precision	Recall	F1-Score
1	0.0063	97.00	100.00	85.00	91.90
2	0.0026	100.00	100.00	100.00	100.00
3	0.0085	95.00	89.47	85.00	87.18
4	0.0072	97.00	94.74	90.00	92.31
5	0.0056	99.00	100.00	95.00	97.44
6	0.0079	95.00	85.71	90.00	87.80
7	0.0031	100.00	100.00	100.00	100.00
8	0.0078	95.00	82.60	95.00	88.37
9	0.0070	98.00	95.00	95.00	95.00
10	0.0013	100.00	100.00	100.00	100.00

Remarkably, a comparison was made between the performance of the model and research conducted outside the Nigerian banking system. On average, the H2O autoencoder models exhibited performances of 97.6% accuracy, 94.8% precision, 93.5% recall, and 94.0% F1-score. The model's performance is favorable when compared to industry standards, which primarily utilize supervised learning on labeled data for fraud detection. Notably, the model achieved significant accuracy and precision using unsupervised learning techniques on unlabeled data. However, there were variations in the balance between precision and recall, as well as the F1-score. Several factors could contribute to these variations, such as differences in data sources, data preprocessing, modelling techniques, or even the characteristics of the underlying fraud patterns specific to the regions under study. It is crucial to acknowledge the specific strengths and limitations of each approach and consider their relevance to the practical requirements of fraud detection in this context. The accomplishments of research conducted outside Nigeria contextualize this work within the ongoing developments in the field of fraud detection, highlighting the dynamic nature of research and the pursuit of optimal model performance.

5. CONCLUSIONS

In summary, the research successfully designed a fraud detection system using unsupervised learning with autoencoder deep neural network models. The design considered individual customer transaction clusters to protect legitimate account holders from unauthorized ATM transactions. With unavailability of labelled transaction data in Nigerian banking system, the methodology still performed averagely above 90% of accuracy and precision. The study significantly contributes to the field by showcasing the exceptional performance of the autoencoder deep learning model in combating ATM fraud within the Nigerian banking system. Hence, it addresses financial security concerns and combines relevant areas such as fraud detection, ATM transactions, and the application of advanced machine learning techniques. Given the importance of financial security and the increasing use of Artificial Intelligence (AI) in fraud detection, this study could generate interest and contribute valuable insights to both the academic community and the banking industry. Future research endeavours may explore additional techniques or enhancements to further strengthen the fraud detection system on unlabelled data.

REFERENCES:

Braimah, O. J., and Okonkwo, I. A. (2016). Statistical Monitoring (SM) of Electronic Fraud Occurring in Nigerian Banks. Advances in Multidisciplinary Research Journal, 2(3), 93-104.

Fethi, M. D., and Pasiouras, F. (2010). Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey. European journal of operational research, 204(2), 189-198.

Gupta, P., Varshney, A., Khan, M. R., Ahmed, R., Shuaib, M., and Alam, S. (2023). Unbalanced credit card fraud detection data: a machine learning-oriented comparative study of balancing techniques. Procedia Computer Science, 218, 2575-2584.

Hedeker, D. (2003). A mixed-effects multinomial logistic regression model. Statistics in medicine, 22(9), 1433-1446.

James, E. (2022). NDIC: Bank Fraud, Forgeries Amounted to N120.79bn in 2020. This Day. https://www.thisdaylive.com/index.php/2022/10/03/ndic-bank-fraud-forgeries-amounted-to-n120-79bn-in-2020/

Lokanan, M. E., and Sharma, K. (2022). Fraud prediction using machine learning: The case of investment advisors in Canada. Machine Learning with Applications, 8, 100269.

Megdad, M. M., Abu-Naser, S. S., and Abu-Nasser, B. S. (2022). Fraudulent Financial Transactions Detection Using Machine Learning. International Journal of Academic Information Systems Research (IJAISR), 6(3).

Misra, S., Thakur, S., Ghosh, M., and Saha, S. K. (2020). An Autoencoder Based Model for Detecting Fraudulent Credit Card Transaction. Procedia Computer Science, 167, 254-262. https://doi.org/10.1016/j.procs.2020.03.219

Ojulari, H. O., Arulogun, O. T., and Oke, A. O. (2017). Transaction Information Extraction from Automated Teller Machine Electronic Journal Using Regular Expression. International Journal of Research in Engineering and Technology (IJRET), 5(8), 29-40

Prakash, B., Murthy, G. V. M., Ashok, P., Prithvi, B. P., and Kira, S. S. H. (2018). ATM Card Fraud Detection System Using Machine Learning Techniques. International Journal of Research, 5(12), 4010-4016

Rajwani, A., Syed, T., Khan, B., and Behlim, S. (2017). Regression analysis for ATM cash flow prediction. In 2017 International Conference on Frontiers of Information Technology (FIT) (pp. 212-217). IEEE.

Rangineni, S., and Marupaka, D. (2023). Analysis Of Data Engineering For Fraud Detection Using Machine Learning And Artificial Intelligence Technologies. International Research Journal of Modernization in Engineering Technology and Science, 5(7), 2137-2146.

Sakharova, I. (2012). Payment card fraud: Challenges and solutions. In 2012 IEEE international conference on intelligence and security informatics, 227-234

Tsegaye, N. (2017). Constructing a predictive model for Real-Time ATM CARD Fraud Detection (Dissertation). Retrieved from http://etd.aau.edu.et/handle/123456789/14041.

Venkatesh, K., Ravi, V., Prinzie, A., and Van den Poel, D. (2014). Cash demand forecasting in ATMs by clustering and neural networks. European Journal of Operational Research, 232(2), 383-392.