

A Model for Detecting Illegitimate Access in Financial Databases Using Deep Neural Networks

Akinkitan Ajose Allen¹ and Solomon Olalekan Akinola²

^{1,2} Department of Computer Science, Lead City University, Ibadan
akin_allen@yahoo.com¹

Abstract

*This research explores Deep Neural Networks (DNN) as advanced classifiers for detecting illegitimate access to financial databases. **Aim:** The aim of this research is to build a deep learning model capable of identifying illegitimate access to financial databases. The dataset was obtained from the Kaggle Machine Learning Repository (Phishing Dataset) and verified against PhishTank data for accuracy. It contains 10,000 labeled URLs (5,000 phishing and 5,000 legitimate). The implementation was conducted using Anaconda3, Jupyter Notebook with Tensorflow, Keras, Pandas, Numpy and Flask framework. The proposed model achieved accuracy of 99.76%, complemented by a precision of 98.69%, a recall of 99.64%, and an F1-score of 99.50%. Additionally, the ROC-AUC score stands at 0.99.*

Keywords: DNN, System, Datasets, Databases, Model, Illegitimate

1. Introduction

As banking becomes more digital, cybersecurity is currently recognized as a significant security issue in the financial industry (Priya et al., 2026). Security in the context of this research refers to the process of protecting digital information throughout its entire lifecycle from corruption, theft, or unauthorized access. One of the main threats to information security is illegal access, which are maliciously intended to cause damage to digital assets and organizational reputations. In network environments, the threat posed by illegitimate financial database access is particularly severe, as they can spread rapidly across systems, causing widespread damage (Hoque et al., 2025). However, Intrusion Detection Systems (IDS) remain critical for safeguarding information assets in IoT ecosystem (Munshar et al., 2026). Neural networks are computational models inspired by the human brain, consisting of interconnected artificial neurons (Smith, 2021). The deep neural network harnesses the capabilities of neural networks in conjunction with deep learning techniques and accompanying data mining algorithms to work with large datasets. In Deep Learning (DL), networks possess multiple hidden layers that learn hierarchical feature representations, enhancing performance across various tasks. Deep Learning can process high-dimensional data across multiple domains, including signals, texts, images, videos, and audio. It represents an advanced Machine Learning (ML) subset that utilizes Artificial Neural Networks (ANN) to model complex data and automatically detect correlations within large datasets (Johnson and Wu, 2023). Thus, deep learning extends the capabilities of traditional ML, enabling the handling of intricate tasks that are beyond the reach of conventional methods. Given the increasing frequency and complexity of illegitimate databases access in financial systems, there is an urgent need for enhanced detection systems. These systems should be better equipped to adapt to new threats, improve precision in traffic analysis, and increase energy efficiency. This research aims to address these challenges by

exploring advanced machine learning techniques, including classification paradigms and clustering approaches, within Intrusion Detection Systems (IDS) to develop a more resilient defense against evolving cyber threats (Priya et al., 2026). This gap motivates the present research, which integrates Deep Neural Network (Deep-NN) techniques into phishing detection systems to enhance adaptability, precision, and efficiency beyond existing models.

2. Related Works

Priya et al., 2026 developed a hybrid intrusion detection approach combining Support Vector Machine (SVM) along with Isolation Forest. The SVM model effectively classifies known attack patterns, while the Isolation Forest detects anomalous and previously unseen network behaviours, improving zero-day attack detection capability. Training and assessment were carried out using benchmark network cybersecurity datasets for anomaly detection, such as NSL-KDD (or CICIDS dataset, if used), ensuring realistic evaluation including both legitimate and adversarial network activity. The evaluation results confirm that the hybrid model achieved an overall detection correctness of approximately 97-99%, with improved anomaly detection rates and reduced false positives compared to standalone classifiers. Sahu, 2026 constructed and evaluated intrusion detection models using the CICIDS2017 dataset, which contains both benign and harmful traffic. Extensive data preprocessing steps were used to improve the dataset's quality and address the problem of class imbalance. Data purification, normalisation, feature selection, and SMOTE data balancing were all part of these processes. Some common measures used for training and testing LSTM models include accuracy (ACC), precision (PRE), recall (REC), and F1-score (F1). In the best trial result, the model achieved an ACC of 98.51%, PRE of 99.0%, recall of 98.0%, and F1 of 99.0%, proving that it is trustworthy in differentiating between benign and malicious traffic. Reddy and Reddy (2024) developed a model to address phishing attacks. The research aimed to classify websites as legitimate or phishing using machine learning techniques such as Logistic Regression, Random Forest, and XGBoost. The dataset, sourced from Kaggle and PhishTank, included 10,000 transactions from the Ethereum network and 5,000 phishing URLs alongside 5,000 legitimate URLs. It contained attributes like transaction times, Ethereum details, and contract data. Preprocessing involved removing 829 blank data points and balancing the dataset, which initially had 1,350 phishing attempts and 7,662 legitimate ones. The data was divided into 8,000 samples for training and 2,000 for testing. Among the models evaluated, XGBoost achieved the highest accuracy at 94.2%. Bibi et al. (2024) developed an intelligent phishing prediction system using machine learning and deep learning to enhance internet security against phishing attacks. The study employed datasets from the UCI Machine Learning Repository containing 11,055 websites, with 4,898 classified as phishing and 6,157 as legitimate. This method used four machine learning models namely; Decision Tree, Naive Bayes, Support Vector Machine, and Random Forest alongside a Convolutional Neural Network (CNN) as the deep learning model. Among these, the multilayered CNN outperformed all other models, achieving a remarkable accuracy of 99.1%, a precision of 97%, a recall of 96%, and an F1-score of 96%. Syed and Janamolla 2024 investigated the pivotal role of Artificial Intelligence (AI) in the early detection and prevention of financial frauds within the global banking sector. The study delved into the background of financial crimes, reviews relevant literature, explores AI technologies used in intelligent banking, provides recommendations for enhanced prevention strategies, and concludes with the potential impact of AI on global banking. Eze and Shamir (2024) proposed methods to detect AI-generated phishing emails which is a paradigm shift of cyber-

attacks due to the use of generative AI in crafting personalized scams that evade detection. The authors used a publicly available dataset of AI-generated phishing emails, consisting of 600 samples per class for training and 100 for testing, analyzed with 10-fold cross-validation. This research applied three main techniques: UDAT, which identifies style differences between AI-generated and human-written emails; MALLETT, using algorithms like Naive Bayes, Maximum Entropy, and Winnow; and a deep learning model based on Long Short-Term Memory (LSTM) architecture with 100 embedding dimensions and dropout layers. An ensemble classifier combining these methods was also implemented. The Naive Bayes classifier in MALLETT achieved the highest accuracy of 99.3%, with a recall of 0.99, precision of 0.98, and an F1-score of 0.985. Results demonstrated that style-based analysis effectively complements word-based analysis for detecting AI-generated phishing emails.

3. Methodology

The Anaconda3, Jupyter Notebook with Tensorflow, Keras, Pandas, Numpy and Flask framework was used to implement the necessary functions, computational features, and capabilities for experimentations. This work leverages the dimensionality reduction technique to minimize the required features as input variables; thereby extracting only 20% of the features that come with the acquired data points as shown in Table 3.1. The dataset was obtained from the Kaggle Machine Learning Repository (Phishing Dataset) and verified against PhishTank data for accuracy. It contains 10,000 labeled URLs (5,000 phishing and 5,000 legitimate)

Table 3.1: Dataset Source/description

Dataset Name	Phishing Dataset
Dataset File Size	64.4MB
Source	Kaggle Machine Learning Repository
Link to dataset	https://www.kaggle.com/datasets/shashwatwork/phishing-dataset-for-machine-learning
Feature Type	Real, Categorical, Integer
Number of URLs	10,000
Number of Features	65
Missing Values	None
Legitimate cases	5,000
Phishing cases	5,000

3.1 Data Features

The phishing dataset features include various characteristics of URLs and websites that can aid in identifying phishing attacks. Some of the features concentrated on the structure of the URL itself, such as the length of the URL, the number of special characters, and the presence of specific patterns like "?" or "&". Other features examine the domain, such as whether the domain is an IP address, the length of the domain, and whether it uses HTTPS for secure connections. Also, the dataset includes information about the presence of various HTML elements on the page, such as title tags, iframes, and

form submission buttons, which can provide insights into whether the site is suspicious or legitimate. Essential features like NumDots, SubDomainlevel, PathLevel, UriLength, Asynmbol, NumDashInHostname, and others serve as input variables, while phishing emails and/or legitimate emails serve as the output variable being the expected class label or target.

3.2 Features Selection

The Ridge Regression method was used for feature selection. The Ridge Regression operate by adding a penalty term to the loss function, proportional to the square of the coefficients. This discourages large coefficient values and minimizes the impact of less important features. This method shrinks the coefficients of irrelevant features and determine the most significant predictors while maintaining the features in the model to improve accuracy. In this research, Ridge regression was used to reduce the original 25 features in the dataset to the most relevant 15. The selected 15 features captured the critical aspects of phishing detection while eliminating redundancy and noise. By reducing the number of features, computational complexity was significantly reduced during training as, fewer relevant features were required processing. The experimental procedure is shown in Figure 3.1.

3.3 Data Train-Test Split Strategy

The dataset consists of a total of 10,000 instances (with 5,000 legitimate URLs and 5,000 phishing URLs). The training and testing datasets were divided as follows:

- a. 80% Training Set where the total Instances are 10,000, This implies, that the Training Set (80%) = $10,000 \times 0.80 = 8,000$ instances. Thus, the training set will have:
 - Legitimate URLs: $5,000 \times 0.80 = 4,000$
 - Phishing URLs: $5,000 \times 0.80 = 4,000$
- b. 20% Testing Set where the Testing Set are $10,000 \times 0.20 = 2,000$ instances. Thus, the testing set will have:
 - Legitimate URLs: $5,000 \times 0.20 = 1,000$
 - Phishing URLs: $5,000 \times 0.20 = 1,000$

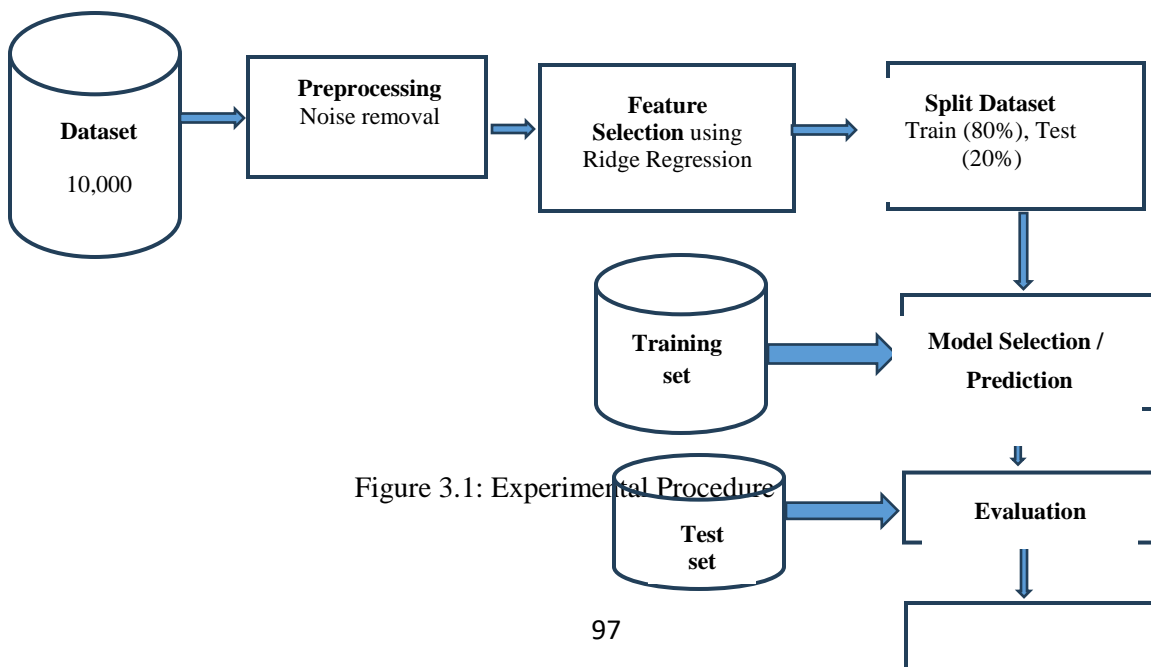


Figure 3.1: Experimental Procedure

The activation function $f(x) = ReLU(x) = \max(0, x)$ was applied in the hidden layers to introduce non-linearity, while a sigmoid function $\sigma(x) = 1/(1 + e^{-x})$ was used in the output layer to classify phishing (1) and legitimate (0) URLs. The loss function used was Binary Cross-Entropy (BCE), minimizing the difference between predicted and actual outcomes.

3.4 Model Evaluation

The proposed model was evaluated based on five (5) metrics namely; Accuracy, Precision, Recall, F1-score, and ROC_AUC. These metrics were computed based on the model's performance on the test set, with their respective formulas provided in Equations (3.1 to 3.5).

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \quad (3.1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.3)$$

$$\text{F1-score} = 2 \times \frac{(\text{Precision} + \text{Recall})}{(\text{Precision} * \text{Recall})} \quad (3.4)$$

$$\text{AUC} = \sum_{i=1}^{n-1} (FPR_i + 1 - FPR_i) \cdot \frac{FPR_i + FPR_{i+1}}{2} \quad (3.5)$$

4. Results and discussion

This section presents results of the model as depicted in Table 4.1. The tabulated data illustrate the performance metrics and provide insights into the effectiveness of the model. It shows exciting results across all metrics by achieving accuracy of 99.76%, complemented by a precision of 98.69%, a recall of 99.64%, and an F1-score of 99.50%. Additionally, the ROC-AUC score stands at 0.99, indicating strong classification ability. These metrics reflect the model's effectiveness; however, further evaluation is needed to ensure it performs consistently on scalable data.

Table 4.1. Classifiers' performance on the Airline dataset (Source: Bajeh et al., 2018)

Accuracy	Precision	Recall	F-1Score	ROC_AUC
99.76%	98.69%	99.64%	99.50%	0.99

The results reflect a solid classification capability suggesting that it is not struggling with better distinction between classes which makes the classifier reliable.

4.1 Comparison with Existing Methods

The proposed model stands out with accuracy of 99.76%, surpassing the accuracy of existing methods. For example, Eze & Shamir (2024) reported accuracy of 99.3%; Reddy & Reddy (2024) showed lower accuracy of 94.2%; Priya et al., 2026 demonstrated accuracy of 99.0% while Bibi et al., 2024 reported accuracy of 99.1%. and Eze and Shamir (2024) showed accuracy of 99.4%.

5. Conclusion

Phishing attacks have become the major avenue by which web intrusion is being perpetrated nowadays by cyber criminals. Various forms of security attacks in which fake websites mimic legitimate ones to steal sensitive information like credit card numbers are easily imposed in the public domain. The data dimension was reduced to ten thousand instances or data points by ten attributes or characteristics, implying 10,000 (elements) X 10 (samples). The attributes of the datasets are the given data samples being predefined in two major classes, which are features/independent variables and target/dependent variables. Essential features like NumDash, NumDashInHostname, AtSymbol, PathLevel, UrlLength, DomainInPaths, HostnameLength, DomainInSubdomains, httpsInHostname, and IP address were carefully extracted using Ridge Regression method as key indicators for Class_Label, which is the target to determine the presence of phishing intruders. The model demonstrated performance with high accuracy, precision, and recall, highlighting its potential as a reliable method for detecting phishing URLs. This research has contributed to the field by offering an improved solution.

Reference

- Bibi, H., Shah, S. R., Baig, M. M., Sharif, M., Mehmood, M., Akhtar, Z. and Siddique, K. (2024). Phishing website detection using improved multilayered convolutional neural networks. *Journal of Computer Science*, 20(9), 1069- 1079.
- Eze, C. S. and Shamir, L. (2024). Analysis and Prevention of AI-Based Phishing Email Attacks. *Electronics*, 13(10), pp. 1839.
- Hoque, F., Shailee, N. M., Hossain, M. S., Patwary, T., Tusher, F. R. and Bari, S. S. (2025). BPSO-IDS: Scalarized Multi-Objective Feature Selection for Intrusion Detection. In 2025 28th International Conference on Computer and Information Technology (ICIT), Bangladesh, pp. 4621-4626. IEEE.
- Johnson, E and Wu, Y. (2023) Advanced Feature Extraction Methods for Phishing Email Detection. *Journal of Information Systems*, 38(3), 2023, pp. 267-284.
- Munshar, H. H. A., Jemili, F., Korbaa, O and Alauthmaan, M. (2026). Comprehensive analysis of intrusion detection systems for enhancing security in internet of things environments. *Discover Applied Sciences*, 1(1), pp 1-38.
- Priya, P. S. and Laasya, M. (2026). Intelligent Hybrid Intrusion Detection System for Financial Network Environments using Cryptography and Network Security. *International Journal of Creative and Open Research in Engineering and Management*, 2(4), pp. 1-11.
- Reddy, V. S. S, and Reddy N. (2024). Machine Learning Techniques for Identifying and Mitigating Phishing Attacks, *International Journal of Artificial Intelligence & Machine Learning (IJAIML)*, 3(2), pp. 116-129.
- Sahu, M. (2026). Data Analytics Approaches for Effective Threat Identification in Cloud Databases. *International Journal of Emerging Research in Engineering and Technology*, 7(2), pp. 1-9.
- Smith R. A. (2021) Evaluating the Impact of Dataset Quality on Phishing Detection Algorithms. *International Journal of Information Security*, 20(4), pp. 321-336.
- Syed, W. K and Janamolla, K. R. (2024). Fight against financial crimes-early detection and prevention of financial frauds in the financial sector with application of enhanced AI. *International Journal of Advanced Research in Computer and Communication Engineering*, 13(1), pp. 59-64.