Preventing Cyber-Fraud in Nigeria’s Banking System Using Fraudaeck-AI

Solomon M. Ekolama, Sunny Orike, and Promise Elechi

Abstract — Preventing Cyber-Fraud in Nigeria's Banking System Using Fraudaeck-AI (Fraud analysis environment for cyber-fraud check) is an in-depth analysis of artificial intelligence (AI) assisted identity verification and authentication systems to prevent cyber fraud in Nigeria's banking system. The Fraudaeck (fraud analysis environment for cyber-fraud check) is a machine learning model developed to learn the interconnected subsystems of the communication network and the banking applications and how they function seamlessly to provide customers with ease and comfort of banking even on the go. An investigation revealed how this created a vulnerability in the system, allowing malicious software attacks such as the Xenomorph Trojan to gain access to the system, paving the way for cyber fraudsters—“Yahoo-boys”—to gain access. This paper proposes a solution to cyber-fraud during electronic banking transactions by using an artificial neural network model called Fraudaeck, which can be interfaced with both the telecommunication protocols and the banking application to detect network intrusion, customer identity theft, and prevent cyber-fraud to obtain hitch-free banking activities in Nigeria.

Keywords — Artificial Neural Network, Communication, Cyber-Attack, Cybercrime, ISP, Network; Network Intrusion.

I. INTRODUCTION

An Internet service provider (ISP) refers to a company that grants its customers access to the internet [1]. ISPs and the respective banks rely on data for their operations. As such, the security of data is of primary concern as data forms an integral resource for these organizations as well as a target for fraudsters, which in essence is cybercrime [2].

Cybercrime can be described as activities that involve the use of a computer or mobile device as a tool and the internet as a connection to achieve illegal objectives such as fraud, phishing or identity theft, attacks, digital signature forgery, etc [3]. In Nigeria, one person can have access to more than one mobile line, which gives him or her more than one means of connecting to the internet. This creates a situation that makes it difficult to track cybercrimes linked to a particular person, especially as all Nigerian ISPs keep biometric information about subscribers in their own separate databases [4]. Another challenge is that, presently, biometric data can be fooled with the help of AI-generated synthetic biometric data [2], so, to solve this problem, AI and biometrics must be combined and used together to make a secure authentication system that can help confirm a customer’s real identity during a bank transaction and stop fraud.

II. NIGERIA’S CYBER SPACE ATTACKS

Cybercrime has always been a concern to individual Internet users and organizations, as imposters fake their identities and convince their victims to hand over their money. The advent of the Internet further placed greater emphasis on associating people with bits of data, such as codes, passwords, and personal identification numbers (PIN). This information has become the target of theft by cyber-fraudsters who employ several means, such as “phishing” and “hacking”, to gain access to their victims’ login credentials. This has led to the global net loss due to cybercrime being estimated at more than $400 billion, and Nigeria stands at the risk of losing US$6 trillion by 2030 to cybercrime [5, 6].

The Economic and Financial Crimes Commission (EFCC) reported that in Nigeria, cybercrimes are perpetrated by individuals, hackers or connected networks of criminals [7]. The Digital Economy for Africa (DE4A) Transformation Initiative offered a helping hand with a programme to initiate inclusiveness and a collaborative structure for Nigeria to leverage the opportunities. Thus, the dream of achieving a reliable digital economy in Nigeria can only be realized through adequate technology and policy implementation by the Nigerian government. As Nigeria moves towards digitization and a digital economy, it is imperative that digital identity becomes the focus of the growing trend and a valid concern. The shift required in this digital age is to improve efficiency and transparency across all sectors of government and private organizations that are in charge of identity management [8].

As a result of the incoherence in the management of databases by the respective biometric data capture agencies, this creates room for cybercrime to thrive in Nigeria, as criminals present multiple identities across-the-board [8]. This proliferation engendered a weak digital identity management system for the Federal Government. For instance, access to the digital market is through internet connection, which research revealed contributes 8% of the global cybercrime index through the use of mobile devices. In addition, Nigeria has the highest African mobile teledensity and is served by four major operators: MTN, Globacom, Airtel, and 9mobile, with 71% of Nigeria’s mobile lines being used as a primary platform to access the Internet [9, 10].
Consequently, as businesses rely more on the Internet to deliver their services these days, cybercrime has become a huge source of illicit wealth creation [10]. This became possible because the GSM network, no matter how secure it appears to be, has a number of loopholes that criminals exploit and attack [11]. For instance, the COMP128 algorithms used in the implementations of the A3 and A8 functions that define the GSM standard have a number of flaws that make it possible to obtain another SIM card. The process of duplicating a SIM card's identity is known as cloning, in which the identity of the original SIM card is transferred into a secondary SIM card and used in another phone. Although cracking the IMSI is relatively easy, finding the Ki can prove to be very difficult. However, with the application of an AI-tool, extracting the Ki is just a matter of minutes [12], [13].

In electronic banking and identity verification, passwords play a key role in restricting unauthorized access to "various computing systems such as ATMs, internet services, Windows login, automated access doors, etc." As such, passwords have become a target of attack in order to compromise the privacy and confidentiality of the system [14]. Also, eavesdropping attacks are used to steal identities. This is when a hacker intercepts, deletes, or changes information that is being sent between two devices.

To achieve the gains of the digital economy, and especially to ride on the comfort of electronic banking, there is a need to prevent cyber-attacks so as to protect all categories of data from theft and damage [16]. This is why the Federal Government of Nigeria wanted to give its citizens a national identity. In 1978, the Department of National Civic Registration was created, but it was only able to register 52.6 million people and give national identity cards to 37.3 million Nigerians [17].

The inability to get the national identity to all Nigerians created a huge challenge for the system. Another challenge also lies in recognizing what biometric traits to measure and the confidence level ascribed to the adopted metrics [18]. As such, Nigeria is currently facing data harmonization challenges which have affected the digital identity management system. Also, biometrics have proved to be far more efficient in authenticating individuals compared to the usual means of passwords and PIN numbers, which are susceptible to hacking. But biometric systems employ sensors that convert the biometric traits of a person to an electrical signal, which is used to distinguish individuals. So, since artificial intelligence technologies can use these electrical signals to create fake patterns that fool biometric systems, we must also use biometric authentication that is combined with AI to solve this problem [2].

AI-aided biometric authentication in this era is a must as over 4.6 billion people globally have mobile phones, of which 52.7% browse the Internet on their phones. And mobile technology makes it easy for cybercrimes like identity theft, harassment, stalking, and revenge porn to happen. Because of this, some countries have made surveillance programs that track how their citizens use their mobile phones and use AI to verify their identities [19].

III. Scope

The scope of the research was limited to collection and analysis of data from three major bodies: National Identity Management Commission (NIMC), six selected Banks, and the Economic and Financial Crime Commission (EFCC).

A. Data Collection

The data collected from National Identity Management Commission (NIMC), were used to form the National Identity Number (NIN) dataset, the data collected from the six selected Banks were used to form the financial transaction and fraud pattern dataset and the data collected the Economic and Financial Crime Commission (EFCC) was used to form fraudulent transaction dataset. The three datasets were combined to produce two distinct sets of datasets: Fraudeack-img and Fraudeack-csvs as shown in Fig. 1 and Table I below.

As seen in Fig. 1, additional images such as animals, birds, peoples in cars and ships were introduced to the images from the original three datasets for form the Fraudeack-img dataset.

The Fraudeack-csv dataset as seen in Table i, constitutes biometric information from three major data capture agencies and the financial transaction patterns from the selected banks.

![Fig. 1 Fraudeack-img dataset.](image-url)
B. Data Engineering

The raw data was converted into numerical data through a data engineering process to suit the machine learning and deep learning models. The categorical data represented variable data with labeled values rather than numerical values. Two processes were employed in converting the categorical data into numerical data, which are integer encoding and one-hot encoding. One-hot encoding allowed the categorical feature with three unique classes to represent three columns. Each column represents a binary flag for a particular categorical value.

The image data to be processed for the model were encoded in two batches and integrated into one as inputs. The raw images from the dataset to be fed in as inputs for the training required a fixed resolution of 256×256. Since the model expects input images to have the form 224×224, this was accomplished by employing training augmentation.

IV. FRAUDAECK ARCHITECTURE

The architecture of Fraudaeck employs a component-based engineering approach, which is aimed at separating concerns with respect to the wide range of intended purposes the system is designed to provide. The aim here is a reuse-based approach: to define, implement, and compose loosely coupled independent components of the application into a single system. This method is meant to give both the system itself and the organizations that will use it a wide range of short-term and long-term benefits to varying degrees. As this system is aimed at providing software-as-a-service (SaaS) delivery, the engineering approach in the design target components is part of the starting platform for service-orientation, which plays the role of an API, as shown in Fig. 2.

The component-based architecture of Fig. 2 allows Fraudaeck to seamlessly allow integration of third-party applications, by using three major components or sections: Convolutional section, Siamese section, and Fraud detection section.

As seen in Fig. 3, the convolutional section is made up of the convolutional layers, pooling layers, and fully connected (FC) layers. In addition, there is the dropout layer that helps modify this section of the network in conjunction with the activation function that guides the network’s learning process.

The dropout layer was introduced to prevent the model from overfitting by avoiding complex co-adaptations on the training data. Each hidden neuron is randomly omitted from the network with a probability of 0.5 so that a hidden neuron cannot rely on other hidden neurons being present. On the other hand, the Siamese section employs a twin network to simultaneously compute comparable output vectors from two distinct input vectors, which makes it possible for similarity measures to be applied in situations where dual inputs of identical features are supplied to be distinguished which may include signatures, fingerprints, passport photos, etc.

In this design, the model in Fig. 3 works in two stages: first to classify the images, and second to identify and distinguish between two images and measure their similarities. The first stage is the convolutional stage which classifies the input images. This enables high-dimensional information of the image to be compressed into the low-dimensional space for the model to learn. The convolutional section is designed to learn through gradient descent to classify images clearly. During the learning stage of the model, the image's nonlinearity is improved by using a LeRu activation function on the hidden layers and a Softmax activation function on the output layer to map the image feature.

In the second and final stage, the fully connected layer passes the classified and predicted image to the Siamese layer, which uses this information to clearly distinguish and identify different images. Thus, if we look at categories such as fingerprints, the model can distinguish between the fingerprint of one person and another. The same is also applicable to passports and signature classes. Because as we go into deeper layers, more information-rich features are being extracted through various matrix operations, which give rise to the high intelligence of Fraudaeck as shown in Fig. 4, Fig. 5, and Fig. 6, respectively.

Consequently, to identify an image, pixel matching requires that each pixel in the original image must be formed in the matching image with the same intensity and geometric relationship. While semantic matching requires an establishment of a correspondence between the two images based on semantic consistency. These requirements render pixel matching and semantic matching less accurate during image verification and authentication.

B. Identifying a Potential Fraudster

To identify a potential fraudster, we first design a solution pattern to enable the model to learn. We employed the “Situation Complication Question” (SCQ) approach for the model to learn from the designed solution.

To design the SCQ, we put the end user in mind. The end-users in this case are the banks which are currently facing cyber-attacks during the process of providing services to their customers. To overcome this challenge and provide effective means of service, the banks and other financial institutions are now exploring data-driven approach to verify and authenticate a customer before the transaction is granted. Within this context, the SCQ was design.
Fig. 3. Siamese Convolutional Neural Network of Fraudaeck.

Fig. 4. Convolutional section of Fraudaeck Model: Adapted from [20].

Fig. 5. Visualized version of Fraudaeck Convolutional Stage (Adapted from [21]).

Fig. 6. Visualization version of Fraudaeck Siamese Stage: Adapted from [22].
The SCQ divides the problem statement into four components and expands each component with the right question and connects them to the desired future state as shown in Fig. 7.

i. Desired Future State: The target results once the problem is solved.

ii. Situation: A brief narrative of the overall problem statement that details the current challenges faced by the end-user, which is wrapped in the question statement. Complication: Defines the major roadblock that hinders the successful transition from the current situation to the desired future situation.

iii. Question: The key question that needs to be addressed in order to mitigate fraud.

C. Fraud Detection Network

The fraud detection section of the Fraudaeck model was trained with the bank transaction dataset to have knowledge of bank transaction patterns. The model studied how money sent to a customer or a fraudster constitutes a genuine or fraudulent transfer, and how money withdrawn by a customer, or an agent of a fraudster constitutes genuine or fraudulent cash out. The model used this knowledge to predict genuine or fraudulent transactions, as shown in Fig. 8.

D. Implementation and Deployment of the Model

The complete design was implemented in three distinct yet integrated algorithms: an algorithm for the Convolutional Neural Network, an algorithm for the Siamese network, and an algorithm for the Fraud Detection Network as shown in Fig. 9.

E. Third Simulation

Through the model’s API, all external agents such as the Banks, EFCC, NIMC, etc., can be integrated and have access to the model and use it for real-time authentication, verification, detection and stop any fraudulent financial transaction that is likely to occur as shown in Fig. 10.
Fig. 9. Authentication Web Service (AWS).

Fig. 10. Third party integration and model usage.

Fig. 11. Fraudaeck learning Epochs.
V. Results and Discussion

The results of the classification showed that the Fraudaeck model can classify images into authentic passports, signatures, and fingerprints. It also classifies images into animals, birds, and other-object classes respectively.

Results also showed that model can distinguish between different fingerprints, passports, and signatures. Also, the model can distinguish between these classes and other classes of images such as bird, animals, etc. The Fraudaeck model can also distinguish between real human passport photograph and when the passport photograph is combined with other objects such as humans in airplane, car, and Ship. The model can also distinguish between forged signatures and real signatures, real fingerprints, and fake fingerprints.

A. Validation Accuracy Test

The result shows that as the validation-loss is decreasing from initial 1.15 to 0.65, the validation-accuracy is increasing from initial 0.62 to 0.95 which gave a validation accuracy of 92.97%, as shown in Fig. 11, Fig. 12a and 12b.

B. Image Prediction and Customer Verification

The result of the third-party simulation showed that the model can match an input image with that of the datasets and predict the correct image with 96.67% accuracy, as shown in Fig. 13.

On predicting the image, the model takes another step to authenticate and verify the true identity of the individual, as shown in Fig. 14.

On authenticating the true identity of the customer, the model sends a request with the image handle to the Fraudaeck-csv dataset to release the Customer’s BVN and NIN. With the BVN and NIN information of the customer, the model matches the passport’s features then retrieves and verify the signature and fingerprint of the customer and present them as shown in Fig. 15 as a report.

C. Imposter Detection and Fraud Alert

The model was simulated with an imposter from a third-party application. The result showed that the passport image of the imposter has dissimilar features from those of Joana_Passport, and immediately marked presence of an imposter as shown in Fig. 16.

On receiving a failed hand from the Fraudaeck image match, as shown in Fig. 16, the model tags the customer as an imposter, and call the Fraudaeck Get_Sysinfo algorithm to fetch the imposter’s system information. The result of the Get_Sysinfo is shown in Fig. 17.
imposter and retrieve his or her true identity. With the true identity information, the model can track the fraudster's operations and prevent any form of fraud-related transactions with 93% accuracy.

REFERENCES


VI. CONCLUSION

Yahoo is an electronic-based cybercrime that has taken on a new dimension as a result of the introduction of sophisticated software aided by artificial intelligence. In this research, we developed a machine learning model called Fraudaeck (Fraud Analysis Environment for Cybercrime Check). The model was trained to understand the intrinsic patterns employed by fraudsters to commit cyber fraud. Using this knowledge, the model could detect a fraudster or an imposter’s NIN by extracting it from the internet connectivity information of the ISP as shown in Fig. 18.

With NIN, the model sends a request to the Fraudaeck-csv dataset to release the imposter’s BVN from the Bank datasets using NIN handle. With the BVN and NIN information of the imposter, the model retrieves the true identity of the imposter and verifies the true passport, signature and fingerprint of the imposter, and presents as report as shown in Fig. 18.

D. Genuine and Fraudulent Transactions

The result of the third-party simulation revealed two types of transactions that fall under the categories of genuine or fraudulent transaction: Transfers and Cash outs.

The result revealed that fraud occurs when both "transfer" and "cash out" are sent to a customer/fraudster but ultimately end up in the account of a merchant who pays the customer/fraudster in cash.

The result of the model showed that accounts linked to fraudsters have specific patterns of unusual old and new balances, as the transfers and cash outs involve huge sums of money. This finding corroborates with the respective reports of fraudulent transaction in Nigeria's Banking sector by security agencies, in which huge money is transferred to another account before being withdrawn in cash and paid into the fraudster’s account, or direct transfer from an agent to the fraudster’s account.

VI. CONCLUSION

Yahoo is an electronic-based cybercrime that has taken on a new dimension as a result of the introduction of sophisticated software aided by artificial intelligence. In this research, we developed a machine learning model called Fraudaeck (Fraud Analysis Environment for Cybercrime Check). The model was trained to understand the intrinsic patterns employed by fraudsters to commit cyber fraud. Using this knowledge, the model could detect a fraudster or an imposter and retrieve his or her true identity. With the true identity information, the model can track the fraudster's operations and prevent any form of fraud-related transactions with 93% accuracy.

REFERENCES


DOI: http://dx.doi.org/10.24018/ejece.2022.6.6.480
Vol 6 | Issue 6 | December 2022
Engr. S. M. Ekolama holds a Bachelor of Technology in Engineering (B.Tech.) Degree in Agricultural Engineering from Rivers State University of Science and Technology, (Now Rivers State University, Port Harcourt), Post Graduate Diploma (PGD) in Electrical Engineering, from Rivers State University of Science and Technology, (Now Rivers State University, Port Harcourt). He is currently pursuing Postgraduate studies (M. Tech) in Electrical Engineering (Communications Option) at River State University. He is a COREN Registered Engineer, a Corporate Member of Nigeria Society of Engineers (NSE), a member of the Institute of Electrical and Electronics Engineers (IEEE), a member of the Nigerian Institution of Electrical and Electronics Engineers (NIEEE), a member of the International Association of Engineers (IAENG). He is presently at the Department of Agricultural and Environmental Engineering, Niger Delta University, Wilberforce Island, Bayelsa State, Nigeria.

Engr. Dr. Sunny Orike received his Bachelor of Technology (B.Tech) in Computer Engineering from Rivers State University of Science and Technology (now Rivers State University); Master of Science (M.Sc) in Computing from The Robert Gordon University, Aberdeen, United Kingdom; and Doctor of Philosophy (PhD) in Artificial Intelligence from Heriot-Watt University, Edinburgh, United Kingdom. Dr. Orike is currently a Reader (Associate Professor) in the Department of Computer Engineering; and Deputy Director of Information and Communication Technology Centre, Rivers State University, Port Harcourt, Nigeria. He is a Registered Engineer with the Council for the Regulation of Engineering in Nigeria (COREN); and an active member of several professional bodies, including: Nigerian Society of Engineers (NSE), Institute of Electrical and Electronics Engineers (IEEE), Nigerian Institution of Electrical and Electronics Engineers (NIEEE) and International Association of Engineers (IAENG). Dr. Orike has published in several reputable journals, presented at both local and international conferences, and acts as a reviewer to several local and international journals.

Engr. Dr Promise Elechi received his B.Eng. and M.Eng. degrees in Electrical/Electronic Engineering from the University of Port Harcourt, Choba, Rivers State, Nigeria in 2006 and 2011 respectively. He obtained PhD in Electronic and Telecommunication Engineering from the University of Benin, Benin City, Nigeria in 2019. He is a corporate member of the Nigerian Society of Engineers (NSE), the International Association of Engineers (IAEng), the Institute of Engineering and Technology (IET), and the Nigerian Institute of Electrical/Electronic Engineers (NIEEE). He is also a registered practising Engineer with the Council for the Regulation of Engineering in Nigeria (COREN). Between August 2018 and October 2019, He was the consultant to Nigeria’s House of Representatives committee to investigate the health implications of mounting GSM masts close to buildings. Dr Elechi has 68 research publications in reputable journals across the world. He has more than 175 citations on google scholar. His current research interest is in Radio Propagation for Mobile Communications, GSM Technology, Microwave Propagation, Signal Analysis, NanoTechnology and Embedded systems. He is at present, a Senior Lecturer in the Department of Electrical and Electronic Engineering, Rivers State University, Port Harcourt. Elechi can be reached at: elechi.promise@ust.edu.ng, and +2348005476478.