International Journal of Economics and Financial Issues

ISSN: 2146-4138

available at http: www.econjournals.com

International Journal of Economics and Financial Issues, 2025, 15(4), 8-18.

Towards Mitigating Cyberfraud in the South African Financial Institutions: A Deep Learning Approach

Oluwatoyin Esther Akinbowale*, Mulatu Fekadu Zerihun, Polly Mashigo

Faculty of Economics and Finance, Tshwane University of Technology, Pretoria, South Africa. *Email: oluwatee01@gmail.com

Received: 22 November 2024 **Accepted:** 05 May 2025 **DOI:** https://doi.org/10.32479/ijefi.18685

ABSTRACT

This study demonstrates the application of deep learning approach specifically the deep learning for cyberfraud incidence classification and time series prediction in the South African financial institutions. Secondary data from the South African Banking Risk Information Centre (SABRIC) was employed and the data was trained under the deep learning paradigm using the Long Short-Term Memory (LSTM) model and adaptive moment estimation (ADAM) algorithm for fraud incidence classification and time series prediction of fraud incidences. Overall, there were 94.1% correct classifications as opposed to 5.9% incorrect classifications. Moreover, the accuracy, precision, recall and F1-score of the LSTM classification model were 71.668%, 87.5%, 99.1% and 78.78% respectively. This indicates that the developed LSTM model is suitable for classification purposes. In addition, the model's performance improves as new datasets are fed in. This is evident as the root mean square error (RMSE) reduced from 253.5116 obtained initially to 150.9 after new data was fed in. This study contributes conceptually, theoretically and empirical to knowledge on cyberfraud mitigation. The results show that the LSTM model can be deployed for fraud classification and time series analysis of fraud incidences. The outcome of this study may promote cyber resilience and sustain the fight against the perpetration of cyber-related fraud in South Africa's financial institutions. The use of the LSTM model for cyberfraud classification and time series prediction of cyberfraud incidences in the South African financial institutions demonstrated in this study is unique.

Keywords: Cyberfraud, Financial Institutions, LSTM, Machine Learning, Time Series Prediction
JEL Classifications: G2, G3

1. INTRODUCTION

The trend of cyberfraud is growing with the increasing number of internet users and reliance on the cyberspace for the operations of financial institutions. The digitalization of the operations of the financial institutions opened up new opportunities, creates improved customer services and relations and promotes easy access to the services of the financial institutions (Jara et al., 2014). It has also enabled the financial institutions to expand their customer base and to operate remotely with improved efficiency, convenience, time and cost effectiveness, convenience (Nel and Boshoff, 2014; Maduku, 2016; Tran and Corner, 2016; Singh and Srivastava, 2018). Nevertheless, the digital and remote operations of the financial institutions also promote the rate of cyberfraud perpetration resulting in significant losses including financial loss, loss of goodwill and reputation amongst others (Koto et al., 2021). This increases the burden on the financial institutions and regulatory bodies as the need to combat cyberfraud via an improved prevention and detection systems becomes necessary. The problem of cyberfraud is compounded with the shortage of cybercrime experts globally that can develop and implement sustainable solutions geared towards cyberfraud detection and prevention. This is coupled with the fact that some financial institutions do not have effective technological solutions and expert systems mitigate cyberfraud thus the implications of cyberfraud impact directly on their bottom line and reputation.

Existing studies indicate that the perpetration of cyber related fraud in South Africa financial institutions impacts negatively on their level of profitability, customers' satisfaction, reputation and good will amongst others (Akinbowale et al., 2021; 2022).

Although the South Africa Reserve Bank (SARB) devotes effort to financial regulation and coordination of the banking activities in South Africa. This is to ensure uniformity, robustness and safety of operation according to the Banks Act (No. 94 of 1990), or the Mutual Banks Act (No. 124 of 1993) (SARB, 2020), however, the rate of cyberfraud perpetration in South Africa is still high.

Kaspersky's report (2023) reported that the African is one of the most targeted continents in 2023 by the cyber criminals. Perpetrators leverage on the increasing number of Internet users in Africa coupled with the lack of capacity to effectively combat cybercrime. South Africa is the leading country in the African continent and placed in the world in terms of cybercrime density (Surfshark's Report, 2022). Therefore, South Africa is considered in this study as a reference. However, the deep application of deep learning model for cyberfraud mitigation is not limited to the South Africa context, as other countries implement the model based on their peculiarities for cybercrime mitigation.

According to SABRIC report (2021 and 2022), the number of reported digital banking fraud reduced by 18%, nevertheless, there was 4 increase in gross losses, from the sum of R310,484,349 in 2020 to R438,238,743 in 2021 accounting for 41.14% increase. In 2022 an increase of 24% in digital banking fraud was reported compared 2021. This increased the gross losses incurred from the sum of R440,123,125 in 2021 to R740,847,488 in 2022, accounting for 68% increase. Social engineering activities contributed significantly to the perpetration of digital banking fraud in 2022.

Table 1 summarizes some statistics on cyberfraud incidences in as reported by SABRIC.

Table 1 shows the classification of the nature of cyberfraud perpetrated in the South African financial institutions categorised as banking applications, online banking and mobile banking. The statistics in Table 1 shows that there is increase in cyberfraud incidences yearly except for mobile banking coupled with the fact that the gross loss due to cyberfraud incidences also increases steadily except for a few cases (online banking in 2022 and mobile banking in 2021). This indicates that cyberfraud perpetration is still a major threat to the South African financial institutions. The number of reported cases of online fraud incidences are lesser than that of banking applications and mobile banking but the impact in terms of gross loss is substantial. For mobile banking, the number of reported cases was high but with low impact in terms of gross loss SABRIC (2021 and 2022) indicated that enhanced detection measures employed by the banks were effective in mitigating fraud losses through mobile banking.

Mbelli and Dwolatzky (2016) indicated that lack of supervision, control and analysis of the security reaches as well as inadequate response to cyberattacks are some of the reasons why the South African banking remains vulnerable to cyberattack.

The study aims to demonstrate the application of the deep learning to cyberfraud mitigation in the South African financial institutions. This study is significant in that it may assist financial institutions to detect intrusions and forecast the trends of cyberfraud incidences and their associated losses over time. Furthermore, it may also assist in the investigation of the effect of changes cyberfraud perpetration on the losses incurred by financial institutions. This may promote the fight against cyberfraud in financial institutions and also assist in forecasting the financial loss incurred over time. The LSTM employed has the capacity to update the trends and analysis with the arrival of new data set, thus, making it a viable real-time decision-making tool.

This paper is structured as follows: section 2 details on the literature review while section 3 presents the methodology employed (the LSTM deep learning approach). Section 4 discusses the results obtained followed by some policy implications. This study ends with the conclusion, recommendations as well as direction for future work.

2. LITERATURE REVIEW

Cyberfraud remains a threat to the financial institutions and to the global economy Almuhammadi and Alsaleh, 2017; PwC Report, 2018; PwC Report, 2020). The reliance on technology and digitalisation of financial and banking services coupled with the increasing number of Internet users as well as the adoption of remote operations aggravated the vulnerability of individuals and financial institutions to cyberthreats (Yu and Guo, 2008; Raza and Hanif, 2013; Raza et al., 2015; Nkoyi et al., 2019; Raza et al., 2020). Abbas et al. (2024) indicated that the threat actors primarily

Table 1: Cyberfraud incidences in South Africa

Nature of cyberfraud	Year	Reported cases	Gross loss (Rands)	% difference in gross loss
Banking applications	2019	10,668	108,389,041	-
	2020	10,281	123,990,231	14
	2021	12,254	219,248,397	55.5
	2022	16,638	363,322,114	50
Online banking	2019	3,304	171,705,112	-
	2020	3,943	139,786,621	-20
	2021	5,866	198,055,406.25	35
	2022	9,455	348,198,319.36	55
Mobile banking	2019	12,575	28,245,948	-
	2020	21,083	45,786,257	47.38
	2021	11,040	17,529,549.72	-89.25
	2022	10,077	29,633,899.52	51.32

Source: SABRIC report (2020, 2021 and 2022) target the Internet of Things (IoT) devices of organisations being the devices that enables the connection and facilitate exchange of data with other systems or devices over the Internet or other communications networks. Maduku (2013) opined that the digitalization of the banking operations makes it efficient, easy and more robust but with the challenge of cyberthreats. The e-banking system was into the South African banking system around 1996, and since then it has radically transformed the sector in terms of operational efficiency (Redlinghuis and Rensleigh, 2010). However, the South African financial institutions are still faced with the challenge of cyberthreats due to their reliance on technology for business operations which sometimes result in cyberattack with the consequences of financial losses (Coovadia, 2011; Cassim, 2016; Kundu et al., 2018; Hubbard, 2019; Koto et al., 2021; Madiba, 2021). Moatshe (2023) reported that cybercrime cost South Africa R2.2 billion annually.

According to SABRIC (2021 and 2022), the common form of cyberthreat faced the South African financial institutions is social engineering, prevalent in the forms of phishing, smishing, vishing, email hacking, spamming, sim swap and business email compromise etc. Van Niekerk (2017) from the synthesis literature found that cyberattack resulting from data exposure has the most negative impact on financial institutions and individuals in South Africa.

Cassim (2016) acknowledged the efforts of the stakeholders in mitigating cyberfraud in South Africa, yet the crime is still increasing with detrimental impacts to the organization's reputation and goodwill, as well as customers' satisfaction (Bhasin, 2011; Skalak et al., 2011; Saini et al., 2012; Kraemer-Mbula et al., 2013; Lagazio et al., 2014; Akinbowale et al., 2020; 2021; 2022; 2024).

Some efforts aimed at mitigating cyberfraud in the financial institutions in South Africa include legislation. For instance, the Section 6(5) of the Bank Act 94 of 1990 specifies alignment of the South African banks to cyber risk management measures to prevent cyber attack, and promote effective response cyber disruptions (SARB, 2017). The South African banking sector has also adopted the "BASEL" regulatory framework to ensure cyber resilience and a secure operation.

Existing studies agree that the implementation of sound cyber risk management measures is necessary for cyber risk mitigation (Kopp et al., 2017; Evdokimova et al., 2019). Some authors suggested collaboration among the stakeholders and increase in cybersecurity awareness initiatives as well as effective implementation of cyber laws and the use of experts for fraud investigation (Dlamini and Modise, 2012; Dzomira, 2017; Mtuze and Musoni, 2023). Mbelli and Dwolatzky (2016) proposed a cybersecurity framework which comprises of four key elements, namely: threat optimization and validation, control of operational conditions threat entry points control, and network security.

Abbas et al. (2024) employed the deep learning models for intrusion detection in IoT devices. The study trained and validated three different variants of the DNN, CNN and RNN. The results obtained indicated that the first variant of the RNN outperform others with 96.61% accuracy, 98.55% precision and F1-score of 98.57% thus validating the suitability of the RNN deep learning architecture for intrusion detection IoT devices.

Shende and Throat (2020) employed the LSTM model for training of KDD99 dataset to detect attack. The binary classification gave 99.2% accuracy while 96.9% accuracy was obtained for the multiclass classification. In order to reinforce the IoT security, Jony and Arnob (2024) employed the LSTM model for intrusion detection strategy using the CIC-IoT2023 dataset, which represents a mixed array of IoT network traffic situations. The proposed LSTM model was able to detect trends of cyber-attack with an accuracy of 98.75% and F1 score of 98.59%. HaddadPajouh et al. (2018) employed the Recurrent Neural Network (RNN) deep learning approach for the detection of malware in IoT devices using a data set comprising of 281 malware and 270 benign ware. The proposed model was validated using new malware samples with three LSTM configurations. The results obtained indicated the LSTM configuration having 2-layer neurons outperformed others with 98.18% accuracy in the detection malware.

Other similar approaches employed for features extraction and intrusion detection have been reported. For instance, the deep feed-forward neural network was used for features extraction in Windows application binary files and the outcome gave a detection accuracy of 95% with a 0.1% false positive rate (Saxe and Berlin, 2015). Kolosnjaji et al. (2016) reported on the combination of the CNN and RNN for hierarchical feature extraction, as well as the N-gram technique for malware detection which gave 89% detection accuracy. Rhode et al. (2018) employed the combination of the RNN and LSTM for malware detection and the result gave 98% detection accuracy with a 1.41% false alarm rate. The deep belief network (DBF) has also been used for malware detection and classification with F1 score of maximum precision of 97.42%, and recall of 97.33% (Chandran et al., 2022). Ibitoye et al. (2019) conducted a comparative analysis of outcomes of the Feedforward Neural Network (FFN) and the Self-normalizing Neural Network (SNN) using BoT-IoT dataset. The results indicated that the FNN outperformed the SNN when employed for intrusion detection in the IoT considering some metrics such as accuracy, precision, recall and Copen Cappa Score. However, when the FNN was validated with three adversarial samples, the accuracy dropped from 95.1% to 24%, 18%, and 31% respectively while the SNN model showed better stability and resilience against the introduction of the adversarial samples.

To improve the accuracy of intrusion detection, Awad et al. (2023) proposed the use of improved Long Short-Term Memory (ILSTM) model which integrates the chaotic butterfly optimization algorithm (CBOA) and particle swarm optimization (PSO). The intrusion detection accuracy of the proposed ILSTM was evaluated using the NSL-KDD dataset and LITNET2020 dataset for binary and multi-class classifications. The results obtained showed that the proposed ILSTM model outperformed the ordinary LSTM model with an accuracy of 93.09% and precision of 96.86% compared to the LSTM which gave an accuracy of 82.74% and a precision of 76.49%.

Limited studies have applied the deep learning approach for the fraud incidences classifications as well as time series prediction of fraud cases and the corresponding losses in the South African financial institutions. Furthermore, the application of the LSTM deep learning paradigm that has the capacity to update in real time with the entry of new data set and visualise changes in the rate of cyberfraud perpetration and the consequences in terms of financial losses has not been widely reported in the literature. These are the research gaps explored in this study to aid the uunderstanding of the progress made regarding cyberfraud mitigation and to also help financial institutions plan and make informed decisions and responses to cyber threats in real time. Thus, this study proposes a LSTM deep learning architecture that can help financial institutions to classify fraud incidences and update their information relating to cyberfraud incidences and corresponding losses in real time.

3. METHODOLOGY

A LSTM is form of deep neural network that can be used for the analysis of historical information of time series data and for long-term nonlinear series prediction. The LSTM was employed in this study for intrusion detection and to predict and update the data relating to cyberfraud incidences. The choice of the LSTM stems from its ability to capture complex trends, and process long sequences of data with long-term dependencies or relationships. It is also suitable for capturing information from previous time steps and store it over a long period (Al-Garadi et al., 2020; Jony and Arnob, 2024). In the context of cyberattack, the LSTM model can detect the details of communication, network traffic or intrusions to aid the risk mitigation plans. Existing studies indicated that the LSTM neural network delivers better detection accuracy when compared with other machine learning models (HaddadPajouh et al., 2018; Abbas et al., 2024).

3.1. Overview of Dataset

Dataset from the SABRIC reports (2020, 2021 and 2022) were employed in this study. It consists of three major classes of digital crime prevalent in the South African financial institutions categorized as banking applications, online banking and mobile banking (Table 1). The dataset provided details about the nature of cyberfraud perpetrated, reported cases, gross loss and differences in gross loss from 2019 to 2022. The various datsaset from 2019 to 2022 contained in the 2020, 2021 and 2022 are consolidated into a single file. The data is classified into inputs (prior time-series window) and outputs (predicted next value) with the inputs fed into the developed LSTM model to generate predicted outputs. Furthermore, the data is scaled for enable effective training process and divided into training dataset which makes up 80% of the total dataset while the validation dataset comprises of 20%. This will also enable the model to fit properly with negligible error.

Figures 1 and 2 present the gross loss due to cyberfraud and the corresponding gross loss from 2019 to 2022 as reported by SABRIC (2020, 2021 and 2022). This dataset was employed for classification, time series analysis and prediction using the LSTM model.

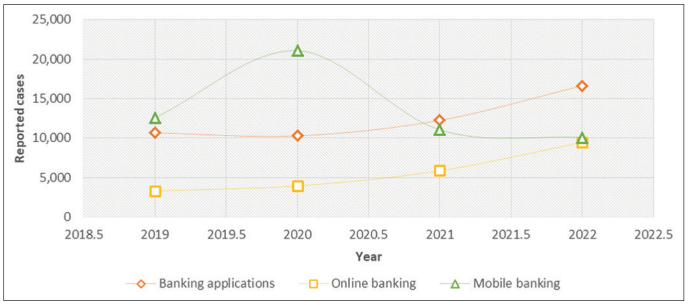
As reported by SABRIC (2020) bank app fraud involves the use of stolen or compromised credentials to access the bank app or multiple bank apps by fraudsters through various social engineering techniques such as phishing etc. In some cases, the credentials were compromised through the weaknesses exploited such as ineffective management of sensitive information such as the one that occurs during SIM swap. An online banking fraud takes place when a fraudster accesses and transfer funds from an online bank account (whether individual or corporate) via account takeover or automatic transfer system. Some social engineering techniques used include malware, phishing, smishing, spamming, SIM swap etc.

Mobile banking fraud involves methods such as phishing, vishing or SIM swap, that allow fraudsters intrude into personal mobile banking account for fraud perpetration (SABRIC, 2020).

3.2. The Architecture of LSTM Model

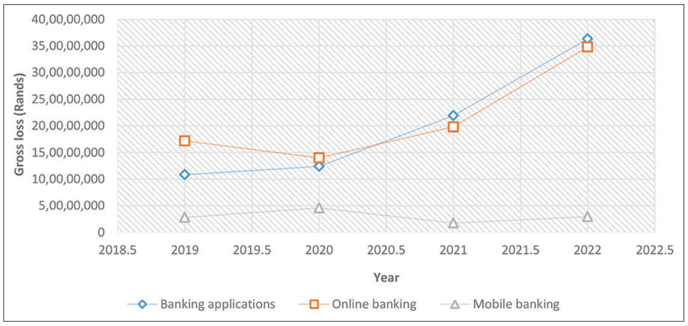
Figure 3 shows the architecture of LSTM model at time step *t*. It comprises of four gates and illustrate the operation of the gates in

Figure 1: The reported cases of cyberfraud in South Africa (2019-2022)



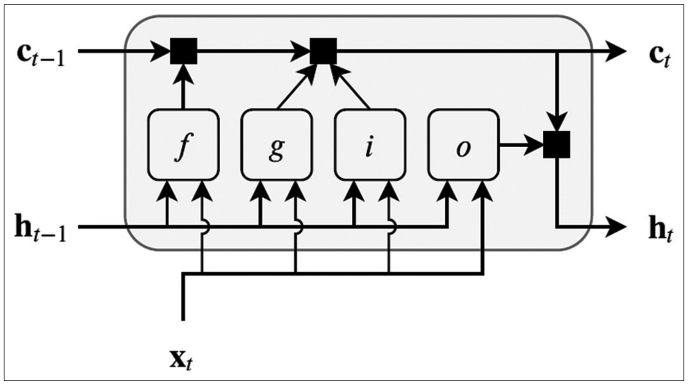
Source: Authors (Raw data extracted from SABRIC reports, 2020, 2021 and 2022)

Figure 2: The gross loss due to cyberfraud (2019-2022)



Source: Authors (Raw data extracted from SABRIC reports, 2020, 2021 and 2022)

Figure 3: The architecture of LSTM model



terms of how they forget, update, and output the cell and the hidden states. The input gate (*i*) decides the information to be stored in the cell state from the current state and updates it accordingly while the "forget gate" (*f*) decides the information to be deleted from the cell state. The function of the cell candidate (*g*) is to add information to cell state while the output gate (*o*) regulates the level of cell state added to hidden state. The implementation of the LSTM model was done in the MATLAB 2022 environment. For the classification problem, the network was trained to detect and classify banking application fraud, online banking fraud as well as mobile banking fraud using features such as malicious files, fake links, location, e-mail compromise, destination IP address, login credentials, SIM swap occurrence, multiple access to bank application, frequency of transaction and amount, sequence of transactions amongst others.

The LSTM layer weight comprise of the following: input weights denoted as *W*, the recurrent weights represented by *R*, and the bias denoted as *b*. Equation 1 shows the chains of the weights *W*, *R*, and *b* respectively for each component.

$$W = \begin{bmatrix} W_i \\ W_f \\ W_g \\ W_o \end{bmatrix}, R = \begin{bmatrix} R_i \\ R_f \\ R_g \\ R_o \end{bmatrix}, b = \begin{bmatrix} b_i \\ b_f \\ b_g \\ b_o \end{bmatrix} \tag{1}$$

Equations 2 and 3 present the cell state and the hidden state at time t respectively.

$$c_t = f_t .\ c_t\text{-}1 + i_t . g_t \tag{2}$$

$$h_t = o_t .\ \sigma_c\ (c_t) \tag{3}$$

Where σ_c is the activation function and the LSTM layer uses the hyperbolic function (tanh) to calculate the state activation function.

Equations 4-7 present the mathematical expressions for the input gate, forget gate, cell candidate and output gate respectively.

$$i_t = \sigma_g\ (W_i x_t + R_i h_{t\text{-}1} + b_i) \tag{4}$$

$$f_t = \sigma_g\ (W_f x_t + R_f h_{t\text{-}1} + b_f) \tag{5}$$

$$g_t = \sigma_c\ (W_g x_t + R_g h_{t\text{-}1} + b_g) \tag{6}$$

$$o_t = \sigma_g\ (W_o x_t + R_o h_{t\text{-}1} + b_o) \tag{7}$$

Where *σ* is the sigma represents the activation functions.

Table 2 presents the parameters for the LSTM model.

During the training, the training dataset is fed into the LSTM model where the parameters are updated recurrently. A progress plot indicates the process of iterative training and displays the root-mean-square error (RMSE) computed during the data. This is followed the evaluation of the performance of the model.

Table 2: The LSTM model parameters

Parameter	Multi-class
Optimizer/algorithm	Adaptive moment estimation (Adam)
Learning rate	0.01
LSTM 1 hidden nodes	200
LSTM 2 hidden nodes	100
Epoch	250
L2 regularization	0.0001
Loss function	Cross entropy
State activation function	tanh
Gate activation function	Sigmoid
Gradient threshold	1

The trained LSTM model was evaluated to determine its accuracy in classifying data using evaluation criteria such as accuracy (A), precision (P), recall (R) and F1 score calculated using equation 8-11 respectively.

$$A = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

$$P = \frac{TP}{TP + FP} \tag{9}$$

$$R = \frac{TP}{TP + FN} \tag{10}$$

$$\text{F1 score} = 2.\frac{\text{Precison} + \text{Recall}}{2} \tag{11}$$

3.3. Forecasting Time Series Data using LSTM

The future time steps forecasting was conducted by training a sequence-to-sequence regression LSTM network, having the training sequences as the responses whose values are adjusted by one time step. The LSTM network is trained to forecast the value of the next time step for each time step of the input sequence, and the state of the network is updated according after each prediction.

The LSTM network predicts the number of cyberattacks incidences by using the information relating to the number of cyberfraud incidences in the past years. The dataset comprises of a single time series, having time steps that corresponds to the years and values which tally to the number of cyberfraud incidences. The data is prepared as a row vector with the output comprising of a cell array, in which each of the elements is a single time step. The data is scaled for enable effective training process and divided into training dataset which makes up 80% of the total dataset while the validation dataset comprises of 20%. The performance evaluation of the model was conducted after the training by validation and performance metrics evaluation.

The training data was standardized to zero mean and unit variance to ensure a good fit and avert divergence during training.

To predict the future time steps values of a sequence, the output of the training sequences are indicated with the values adjusted by one time step. In other words, for each time step of the input sequence, the LSTM models learns and predict the next time step value while the training sequence serves as the predictors.

The initialization of the network state was achieved by using the training data and predictions are made using the previous time step of the training response. This is looped over the outstanding predictions and the predicted values are compared with the test data.

One of the significance of this LSTM model is that it can updated in real time as new data arrives or as new values are obtained. The model is reset to make new predictions so as to prevent past predictions from influencing the outcome of the new predictions.

4. RESULTS AND DISCUSSION

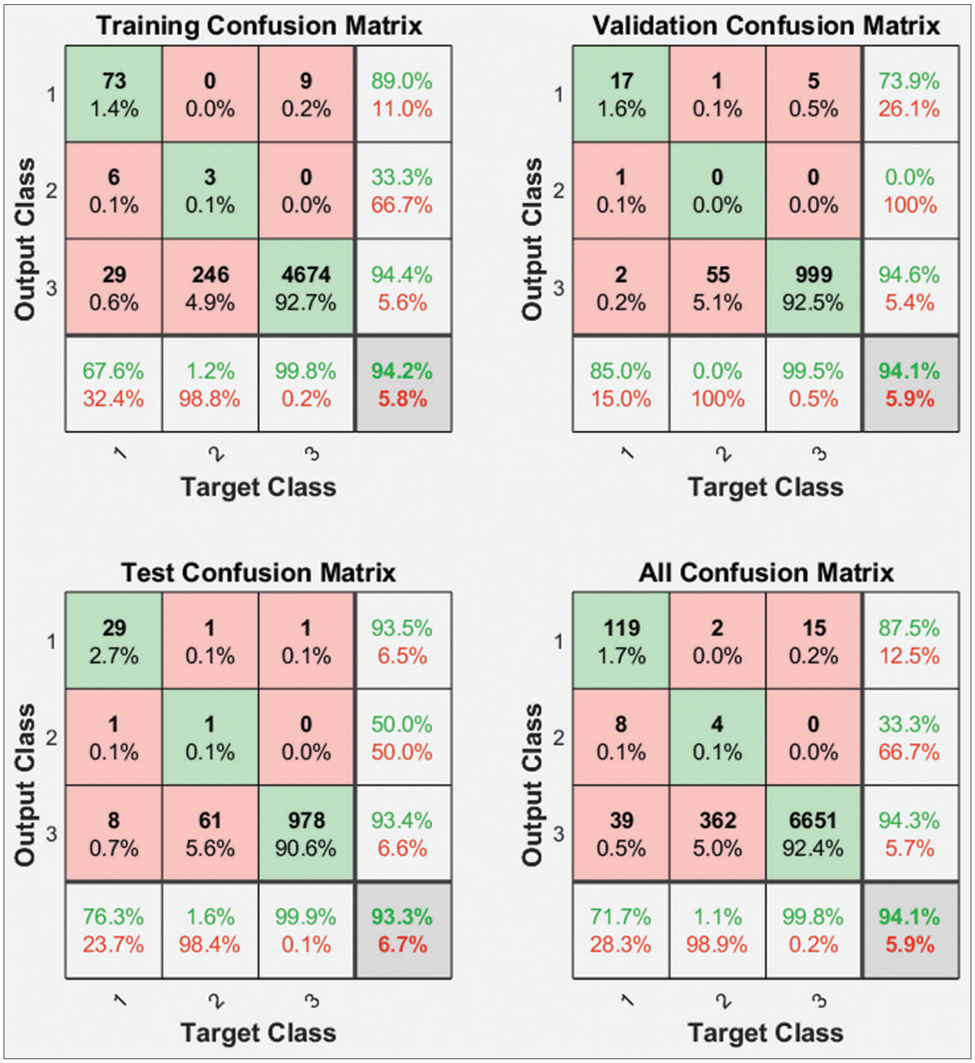
4.1. Multi Classification

Figure 4 presents the confusion matrix of the LSTM model. The output ad target class “1”, “2” and “3” represent “banking applications,” “online banking,” and “mobile banking” respectively. The first row and first column of the overall confusion matrix represent the classification of fraud under “banking applications” while second row and second column represent “online banking” and the third row and third column for “mobile banking.” The LSTM model correctly predicted 50 fraud cases as “Banking Applications” without any incorrect predictions. Looking at the “banking applications” column (first column), 119 predictions for “banking applications” are correct (true positive) while 8 fraud cases that actual belong to the “banking applications” class were wrongly classified as “online banking and 39 fraud cases that also belong to the “banking applications” were grouped as mobile banking.” For the banking applications,” the percentage of correct classification 71.7% with 28.3% incorrect classifications.

Looking at the second column, 2 fraud cases were incorrectly classified as “banking applications” instead of “online banking” while 4 fraud cases were correctly classified as “online banking.” 36 fraud cases were incorrectly classified as “mobile banking”. Considering the second column, there are 3 fraud cases that belong to the “mobile banking” instead of “online banking. Therefore, the percentage of correct classification was 1.1% with 98.9% incorrect classifications for “online banking.”

For the third column, 15 fraud cases were incorrectly classified as “banking applications” instead of “mobile banking” while 6651 fraud cases were correctly classified as “mobile banking.” Thus, the percentage of correct classification was 99.8% with 0.2% incorrect classifications for “mobile banking.” On the overall,

Figure 4: The confusion matrix of the LSTM model



Training Confusion Matrix

Output Class	1	2	3	
1	73 1.4%	0 0.0%	9 0.2%	89.0% 11.0%
2	6 0.1%	3 0.1%	0 0.0%	33.3% 66.7%
3	29 0.6%	246 4.9%	4674 92.7%	94.4% 5.6%
	67.6% 32.4%	1.2% 98.8%	99.8% 0.2%	94.2% 5.8%

Target Class

Validation Confusion Matrix

Output Class	1	2	3	
1	17 1.6%	1 0.1%	5 0.5%	73.9% 26.1%
2	1 0.1%	0 0.0%	0 0.0%	0.0% 100%
3	2 0.2%	55 5.1%	999 92.5%	94.6% 5.4%
	85.0% 15.0%	0.0% 100%	99.5% 0.5%	94.1% 5.9%

Target Class

Test Confusion Matrix

Output Class	1	2	3	
1	29 2.7%	1 0.1%	1 0.1%	93.5% 6.5%
2	1 0.1%	1 0.1%	0 0.0%	50.0% 50.0%
3	8 0.7%	61 5.6%	978 90.6%	93.4% 6.6%
	76.3% 23.7%	1.6% 98.4%	99.9% 0.1%	93.3% 6.7%

Target Class

All Confusion Matrix

Output Class	1	2	3	
1	119 1.7%	2 0.0%	15 0.2%	87.5% 12.5%
2	8 0.1%	4 0.1%	0 0.0%	33.3% 66.7%
3	39 0.5%	362 5.0%	6651 92.4%	94.3% 5.7%
	71.7% 28.3%	1.1% 98.9%	99.8% 0.2%	94.1% 5.9%

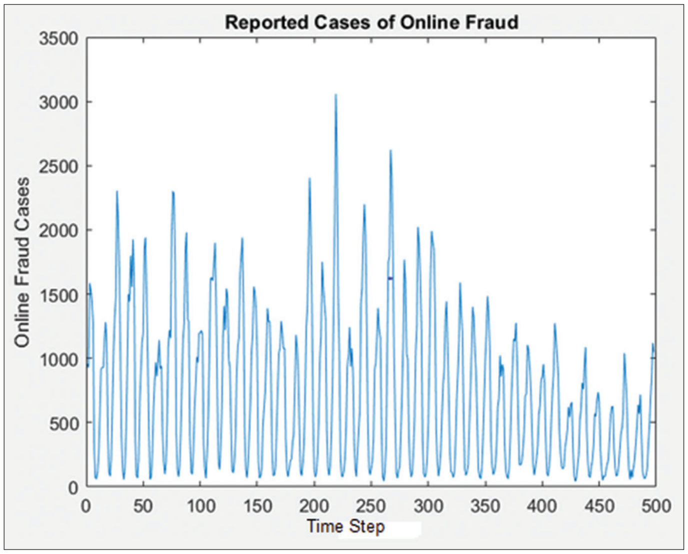
Target Class

there was 94.1% correct classification as opposed to 5.9% incorrect classifications.

From Figure 4, True Positive (TP) was 119, while True Negative (TN) equals to 4+6651 = 6655. False Positive (FP) equals to 8+39 = 47, while False Negative (FN) equals to 2+15 = 17.

Therefore, using equations 8-11, the accuracy, precision, recall and F1-score of the LSTM classification model are calculated as 71.668%, 87.5%, 99.1% and 78.78% respectively. The results obtained for cyberfraud classification in this study fell within the range of the results obtained in the literature for intrusion detection using similar performance metrics reported by Kolosnjaji et al., 2016, Rhode et al., 2018, Shende and Throat, 2020; Awad et al., 2023; Jony and Arnob, 2024. These studies considered deep learning models with high values of accuracy, precision, recall and F1-score (>70%) as a high performing model.

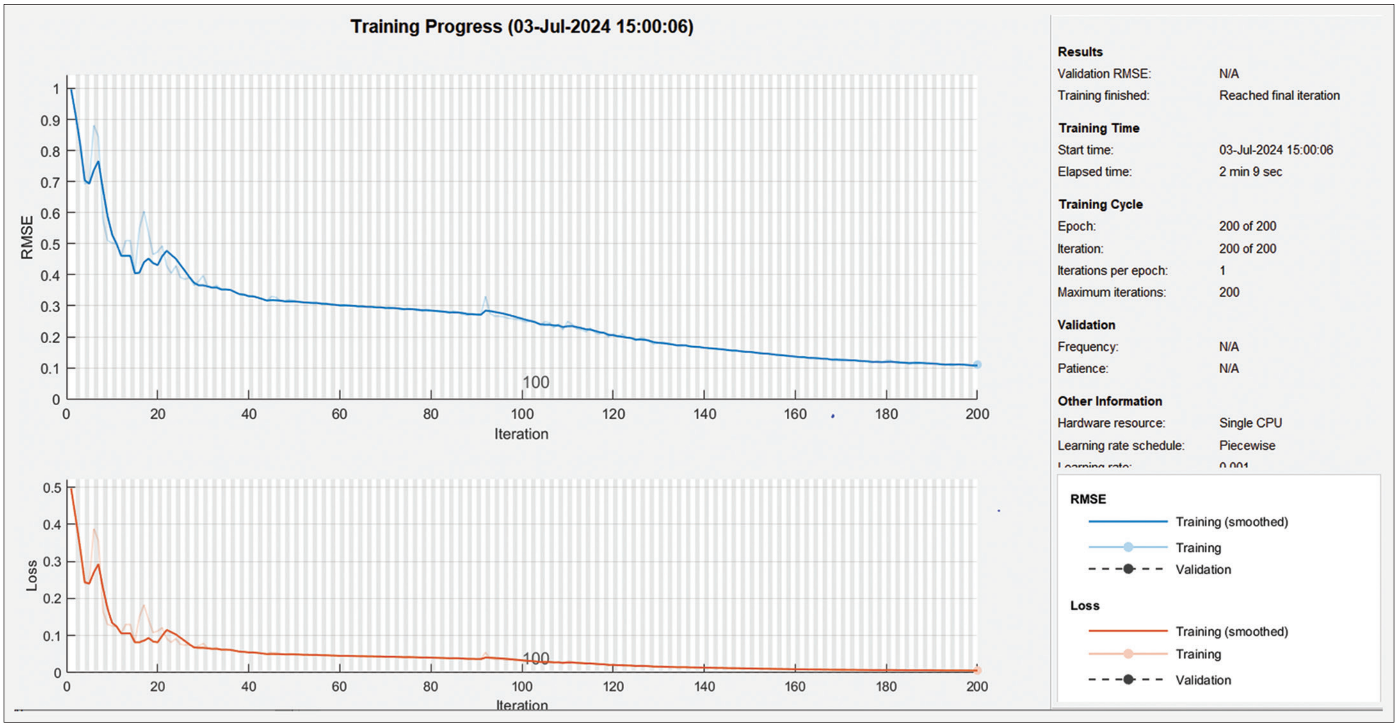
Figure 5: The time series analysis of the online fraud cases from 2019 to 2022 using LSTM model



4.2. Time Series Prediction

Figure 5 displays the time series analysis of the online fraud cases from 2019 to 2022 using the LSTM model. The essence is to visualise the trend of online fraud cases from 2019 to 2022. Figure 5 shows that the occurrence of online fraud activities varies with time depending on the proactiveness of individual, financial institutions and the sophistication of the threat actors in exploiting the vulnerabilities of individuals or financial institutions. Thus, over time an increasing and decreasing trends in cyberfraud perpetration was observed ad in addition, cyclic patterns that repeats after a certain interval of time were also observed. This analysis agrees significantly with the SABRIC report (2000, 2001 and 2022) that the rate of cyberfraud perpetration depends on the proactiveness and control measures implemented by the financial institutions. Threat actors usually employ social engineering techniques such as malware, phishing, smishing, vishing, SIM swap etc. to obtain sensitive information from individuals or organisations for fraud perpetration across the digital channels. Threat actors also exploit susceptibilities in the management of sensitive information, and source for login credentials that are saved on systems or devices or multiple applications to commit fraud. A common form of vishing employed by the threat actors, is to disguise as a bank official or service provider and manipulate individuals into divulging personal information, which will later be used for cyberfraud perpetration (SABRIC report, 2000; 2001 and 2022). Figure 6 shows the progress of the training of the LSTM model for 200 iterations (epoch). The figure shows that the model reached the maximum iteration without any sign of

Figure 6: The progress of the training of the LSTM model


Training Progress (03-Jul-2024 15:00:06)

Results
Validation RMSE: N/A
Training finished: Reached final iteration

Training Time
Start time: 03-Jul-2024 15:00:06
Elapsed time: 2 min 9 sec

Training Cycle
Epoch: 200 of 200
Iteration: 200 of 200
Iterations per epoch: 1
Maximum iterations: 200

Validation
Frequency: N/A
Patience: N/A

Other Information
Hardware resource: Single CPU
Learning rate schedule: Piecewise
Learning rate: 0.001

Figure 7: The time series prediction of the online fraud cases from 2019 to 2022

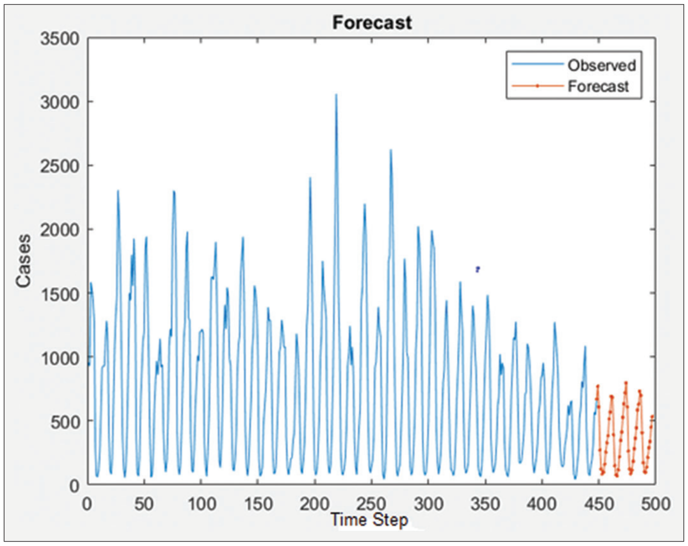
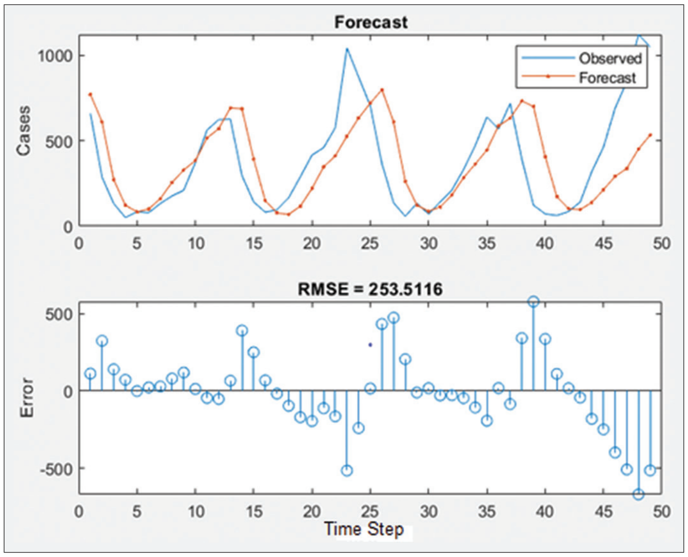


Figure 8: The time series prediction of the online fraud cases and the error generated



over fitting. Each iteration estimates the gradient and updates of the network parameters.

Figure 7 shows the time series prediction of the fraud cases and the using the LSTM model. The forecast shows the tendency for reduction in the cases of online fraud with time. Figure 8 shows the time series prediction of the fraud cases, the RMSE generated. The RMSE is the measure of the deviation of the predicted values from the actual value. The lower the RMSE, the better the model for predictive purpose and vice versa. The developed LSTM model showed tendency for reduction in the value of RMSE with the introduction of new inputs as updates as shown in Figure 9.

Figure 9: The time series prediction with updates

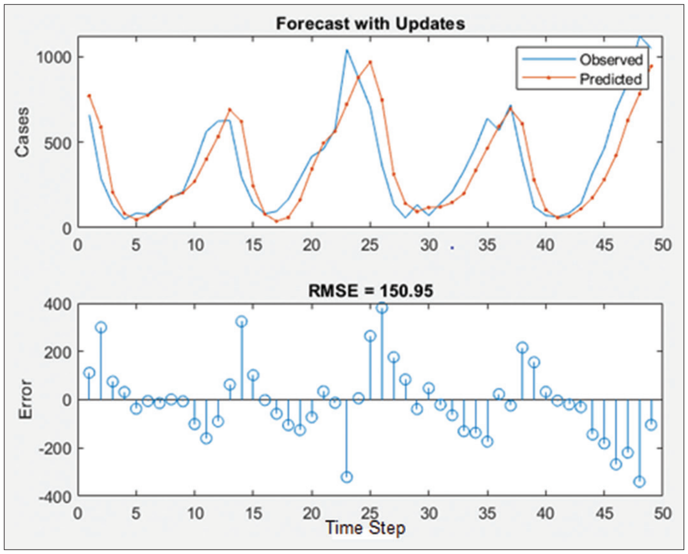


Figure 9 shows the time series prediction with updates. The arrival of new dataset fed into the developed LSTM model shows that the model has the ability to update automatically. Comparing Figures 8 and 9, it can be deduced that the model performance improves as new dataset is fed in. This is evident as the RMSE reduced from 253.5116 obtained initially to 150.9 after new data was fed in. This result shows improvement over the one obtained by Ibitoye et al. (2019) for intrusion detection using Feedforward Neural Network (FFN) which shows that the model's performance reduced with the introduction of three adversarial samples.

5. CONCLUSION AND RECOMMENDATIONS

The purpose of this study was to demonstrate the application of machine learning approach for the mitigation of cyberfraud perpetration in the South African financial institutions. This was achieved with the use of secondary data obtained from the South African Banking Risk Information Centre (SABRIC). The data was trained under the deep learning paradigm using the LSTM model and adaptive moment estimation (ADAM) algorithm for fraud incidence classification and time series prediction of fraud incidences.

On the overall, there was 94.1% correct classifications as opposed to 5.9% incorrect classifications. Furthermore, the accuracy, precision, recall and F1-score of the LSTM classification model were 71.668%, 87.5%, 99.1% and 78.78% respectively. This indicates that the developed LSTM model is suitable for classification purpose. In addition, the model performance improves as new dataset is fed in. This is evident as the RMSE reduced from 253.5116 obtained initially to 150.9 after new data was fed in.

The outcome of this study provides insights that can be helpful to the South African financial institutions in combating cyberfraud. The classification and time series analysis demonstrated in this study offer a roust solution geared towards cyberfraud mitigation. The potentials of these deep learning approaches can be fully harnessed with the availability of robust dataset, consistent model updates, integration with other compatible data analytic techniques amongst others. This study is limited to the use of the LSTM algorithm for fraud incidences classification, future works can consider a comparative analysis of different hybrid deep learning and recurrent neural models. The fraud detection and predictive capability of the LSTM model can be further explored to promote the proactiveness of financial institutions in combating cyberfraud.

The classification of fraud using the LSTM model can help financial institutions understand the scope of fraud thereby leading to a more strategic management of fraud. It provides insights the type of fraud perpetrated ad how it was perpetrated. Thus, financial institutions can employ the developed LSTM fraud classification model to achieve reduction in manual classification or classification error while ensuring greater consistency in fraud classification as this can lead to better fraud tracking. From the understanding gained through fraud classification, financial institutions should sensitize their customers effectively on the current cyberthreats and activities of the threat actors as well as some proactive steps that customers can take against cyberfraud. This will promote customer service and relationship as well as public trust. Financial institutions can improve the classification model to minimize the incorrect classifications to promote classification accuracy.

For the time series analysis, financial institutions should embrace this technique for transaction monitoring. Through real time series data analysis, unusual or malicious patterns that may lead to fraud can be identified early and mitigated. Financial institutions can also adopt the time series analysis as a complementary technique to monitor the risk profiles of transactions. The analysis of transactions history, and the magnitude of loss incurred may assist financial institutions in the identification and classification of high-risk transactions and to take proactive steps before they escalate into major fraud cases. Furthermore, financial institutions can use the information acquired from the time series analysis to generate comprehensive and accurate reports that details on suspicious and fraudulent activities required by the regulatory bodies. In addition, financial institutions can employ the time series analytic technique as a predictive tool to project into the future likely scenarios so that practical steps can be taken to address the major risks before they occur. The example demonstrated in this study integrates adaptive algorithm with the deep learning technique to capture new data trends ad update accordingly in real time as new data arrives. Therefore, financial institutions must ensure regular update of the LSTM model with new data set to keep up to date with the dynamics of the cyberfraud and the threat actors. Financial institutions can embrace this approach as part of the continuous monitoring and improvement that allows for real time alert for suspicious transactions or reach of information to engineer swift ad effective response. Thus, the outcome of this study contributes to the automation of the fraud mitigation and risk management processes. With the time series analysis of complex and voluminous datasets, financial institutions can achieve improved operational excellence and sustain the fight against cyberfraud. As different types of cyberfraud continue to emerge with increasing risks and financial losses, the need for financial institutions to leverage on machine learning approach for fraud classification and prediction becomes imperative. The machine learning approach can promote the integrity and security of operations of the financial institutions.

As financial institutions grow the use of effective data processing technique and optimization algorithms can promote the competitiveness and ensure financial institutions remains efficient and responsive to cyberthreats. However, financial institutions must sure the qualitative and quantitative attributes of the data fed into the machine learning model. The effectiveness of classification and the time series model is a function of the volume ad type of data employed for analysis. Financial institutions often deal with vast amount of data which may vary in quality. Poor data quality usually leads to inaccurate analysis and poor decision making. Thus, the implementation of effective data management practices is central to the use of the machine learning approach for fraud classification and mitigation. Effective data management practices encompass data acquisition processes, preprocessing and storage. Periodic quality checks will promote the integrity of the data employed for prediction purpose. It is noteworthy to mention that cyberfraud is an evolving crime as the threat actors usually devise novel method for fraud perpetration. Hence, the use of time series analysis alone may not sufficiently guarantee effective cyberfraud mitigation. Nevertheless, the time series analysis technique can be incorporated into the existing forecasting tool where compatible without the violation of regulatory guidelines.

REFERENCES

Abbas, S., Alsubai, S., Ojo, S., Sampedro, G.A., Almadhor, A., Al Hejaili, A., Bouazzi, I. (2024), An efficient deep recurrent neural network for detection of cyberattacks in realistic IoT environment. The Journal of Supercomputing, 80(10), 13557-13575

Akinbowale, O.E., Klingelhöfer, H.E., Zerihun, M.F. (2020), Analysis of cyber-crime effects on the banking sector using balance score card: A survey of literature. Journal of Financial Crime, 27(3), 945-958.

Akinbowale, O.E., Klingelhöfer, H.E., Zerihun, M.F. (2021), The integration of forensic accounting and the management control system as tools for combating cyberfraud. Academy of Accounting and Financial Studies Journal, 25(2), 1-14.

Akinbowale, O.E., Klingelhöfer, H.E., Zerihun, M.F. (2022), Analytical hierarchy process decision model and Pareto analysis for mitigating cybercrime in the financial sector. Journal of Financial Crime, 29(3), 884-1008.

Akinbowale, O.E., Klingelhöfer, H.E., Zerihun, M.F. (2024), The assessment of the impact of cyberfraud in the South African banking industry. Journal of Financial Crime, 31(2), 287-301.

Al-Garadi, M.A., Mohamed, A., Al-Ali, A.K., Du, X., Ali, I., Guizani, M. (2020), A survey of machine and deep learning methods for Internet of Things (IoT) security. IEEE Communications Surveys and Tutorials, 22(3), 1646-1685.

Almuhammadi, S., Alsaleh, M. (2017), Information security maturity model for NIST cyber security framework. Computer Science and Information Technology, 2017, 52-62.

Awad, A.A., Ali, A.F., Gaber, T. (2023), An improved long short-term memory network for intrusion detection. PLoS One, 18(8), e0284795.

Bhasin, M.L. (2011), Corporate governance disclosure practices in India: An empirical study. International Journal of Contemporary Business Studies, 2(4), 34-57.

Cassim, F. (2016), Addressing the Growing Spectre of Cybercrime in Africa: Evaluating Measures Adopted by South Africa and other Regional Role Players. In: School of Law, University of South Africa. Based on a paper presented at the First International Conference of the South Asian Society of Criminology and Victimology (SASCV) at Jaipur, India from 15-17 January 2011. p126-138.

Chandran, P.P., Rajini, N.H., Jeyakarthic, M. (2022), Optimal deep belief network enabled malware detection and classification model. Intelligent Automation and Soft Computing, 35(3), 3349-3364.

Coovadia, C. (2011), Banking Sector Overview. Available from: https://www.epiccommunications.co.za/sites/epic/files/cascoovadia_1_0.pdf [Last accessed on 2019 Aug 01].

Dlamini, Z., Modise, M. (2012), Cyber Security Awareness initiatives in South Africa: A Synergy Approach. 7th International Conference on Information Warfare and Security, Seattle, USA. p1-10.

Dzomira, S. (2017), Internet banking fraud alertness in the banking sector: South Africa. Banks and Bank Systems, 12(1), 143-151.

Evdokimova, Y., Shinkareva, O., Egorova, E. (2019), Banking Information Technology as an Element of the Information Society. 54th International Scientific Conference on Economic and Social Development-XIX International Social Congress, Moscow. p550-555.

HaddadPajouh, H., Dehghantanha, A., Khayami, R., Choo, K.K.R. (2018), A deep Recurrent Neural Network based approach for Internet of Things malware threat hunting. Future Generation Computer Systems, 85, 88-96.

Hubbard, J. (2019), SA Business Underplaying the Danger of Cybercrime? Finweek. Vol. 4. Available from: https://hdl.handle.net/10520/EJC-1444bed59d

Ibitoye, O., Shafiq, O., Matrawy, A. (2019), Analyzing Adversarial Attacks against Deep Learning for Intrusion Detection in IoT Networks. In: 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 2019. p1-6.

Jara, A.J., Parra, M.C., Skarmeta, A.F. (2014), Participative marketing: Extending social media marketing through the identification and interaction capabilities from the internet of things. Personal and Ubiquitous Computing 18(4), 997-1011.

Jony, A.I., Arnob, A.K.B. (2024), A long short-term memory based approach for detecting cyber attacks in IoT using CIC-IoT2023 dataset. Journal of Edge Computing, 3(1), 28-42.

Kaspersky. (2023), Africa remains one of the regions most targeted by cybercrime in 2023. Available from: https://kaspersky.africa-newsroom.com/press/africa-remains-one-of-the-regions-most-targeted-by-cybercrime-in-2023?lang=en#:~:text=In%20QQ3%20 2023%2C%20according%20to,ICS%20machines%20in%20Q3%20 2023 [Last accessed on 2024 Jun 04].

Kolosnjaji, B., Zarras, A., Webster, G., Eckert, C. (2016), Deep Learning for classification of malware system call sequences. In: Kang, B.H., Bai, Q., editors. AI 2016: Advances in Artificial Intelligence. Lecture Notes in Computer Science. Vol 9992. Cham: Springer. pp.137-149.

Kopp, E., Kaffenberger, L., Wilson, C. (2017), Cyber Risk, Market Failures and Financial Stability. IMF Working Paper. Available from: https://www.imf.org/en/publications/wp/issues/2017/08/07/cyber-risk-market-failures-and-financial-stability-45104 [Last accessed on 2023 Apr 27].

Koto, C., Smith, R.J., Schutte, B. (2021), Cyber Risk Management Frameworks for the South African Banking Industry. In: 6th Annual International Conference on Public Administration and Development Alternatives, South Africa. p80-89.

Kraemer-Mbula, E., Tang, P., Rush, H. (2013), The cybercrime ecosystem: Online innovation in the shadows? Technological Forecasting and Social Change, 80(3):541-555.

Kundu, S., Islam, K.A., Jui, T.T., Rafi, S., Hossain, M.A., Chowdhury, I.H. (2018), Cyber Crime Trend in Bangladesh, An Analysis and Ways Out to Combat the Threat. In: 2018 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon, Korea (South), 2018. p474-480.

Lagazio, M., Sherif, N., Cushman, M. (2014), A multi-level approach to understanding the impact of cybercrime on the financial sector. Computers and Security, 45, 58-74.

Madiba, A. (2021), The Scary Nature of Cybercrime and the Strain of Bringing Perpetrators to Book. Sunday Independent News. Available from: https://twitter.com/thesundayindep1/status/1399032513977257986 [Last accessed on 2023 Apr 27].

Maduku, D.K. (2013), Predicting retail banking customers' attitude towards internet banking services in South Africa. Southern African Business Review, 17(3), 76-100.

Maduku, D.K. (2016), The effect of institutional trust on Internet banking acceptance: Perspectives of South African banking retail customers. South African Journal of Economic and Management Sciences, 19(4), 533-548.

Mbelli, T. M., Dwolatzky, B. (2016), Cyber security, a threat to cyber banking in South Africa: An approach to network and application security. In: 2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud), China. pp. 1-6.

Moatshe, R. (2023), Bleak Picture Painted as Cybercrime Cost SA R2.2 Billion Annually. Available from: https://www/iol.co.za/pretoria-news/news/bleak-picture-painted-as-cybercrime-costs-sa-r22-billion-annually-6abb1eaf-5dba-4852-ad7a-be774de68595 [Last accessed on 2023 Jul 19].

Nel, J., Boshoff, C. (2014), Enhancing the use of internet banking in an emerging market. South African Journal of Economic and Management Sciences, 17(5), 624-638.

Nkoyi, A., Tait, M., Van der Walt, F. (2019), Predicting the attitude towards electronic banking continued usage intentions among rural banking customers in South Africa. South African Journal of Information Management, 21(1), a1016.

PwC Report. (2018), Global Economic Crime Survey: Pulling Fraud out of the Shadows. p1-30. Available from: https://www.pwc.org [Last accessed on 2019 Jan 05].

PwC Report. (2020), Global Economic Crime and Fraud Survey. 7th ed. p1-32. Available from: https://www.corruptionwatch.org.za/wp-content/uploads/2020/06/global-economic-crime-survey-20201.pdf [Last accessed on 2021 Jan 17].

Raza, A., Hanif, N. (2013), Factors affecting internet banking adoption among internal and external customers: A case of Pakistan. International Journal of Electronic Finance, 7(1), 82-96.

Raza, S.A., Jawaid, S.T., Hassan, A. (2015), Internet banking and customer satisfaction in Pakistan. Qualitative Research in Financial Markets, 7(1), 24-36.

Raza, S.A., Umer, A., Qureshi, M.A., Dahri, A.S. (2020), Internet banking service quality, e-customer satisfaction and loyalty: The modified e-SERVQUAL model. The TQM Journal, 32(6), 1443-1466.

Redlinghuis, A., Rensleigh, C. (2010), Customer perceptions on internet banking information protection. South African Journal of Information Management, 12(1), 1-6.

Rhode, M., Burnap, P., Jones, K. (2018), Early-stage malware prediction using recurrent neural networks. Computers and Security, 77, 578-594.

Saini, H., Rao, Y.S., Panda, T.C. (2012), Cyber-crimes and their impacts: A review. International Journal of Engineering Research and Applications, 2(2), 202-209.

Saxe, J., Berlin, K. (2015), Deep Neural Network Based Malware Detection Using Two Dimensional Binary Program Features. 2015 10th International Conference on Malicious and Unwanted Software (MALWARE), Fajardo, PR, USA, 2015. p11-20.

Shende, S., Thorat, S. (2020), Long short-term memory (LSTM) deep learning method for intrusion detection in network security. International Journal of Engineering Research and Technology, 9(6), 1615-1620.

Singh, S., Srivastava, R.K. (2018), Predicting the intention to use mobile banking in India. International Journal of Bank Marketing, 36(2), 357-378.

Skalak, S.L., Alas, M.A., Sellito, G. (2011), Fraud: An Introduction. In: Golden, T.W., Skalak, S.L., Clayton, M.C., editors. A Guide to Forensic Accounting Investigation. Hoboken, NJ: John Wiley and Sons. Inc., US. p1-23.

South African Banking Risk Information Centre (SABRIC). (2020), Annual Crime Statistics. Available from: https://www.sabric.co.za/media/20oouwbg/sabric-annual-crime-stats-2020.pdf [Last accessed on 2022 Jun 20].

South African Banking Risk Information Centre (SABRIC). (2021), Annual Crime Statistics; 2021. Available from: https://www.banking.org.za/news/sabric-annual-crime-stats-2021 [Last accessed 2023 Jan 27].

South African Reserve Bank (SARB). (2020), Management of the South African Money and Banking System. Available from: https://www.resbank.co.za/aboutus/functions/pages/management-of-the-south-african-money-and-banking-system.aspx [Last accessed on 2020 Feb 02].

South African Banking Risk Information Centre (SABRIC). (2020). Annual Crime Statistics. Available from: https://www.sabric.co.za/media/gq4hmbjw/sabric-annual-crime-stats-2022.pdf [Last accessed on 2024 Aug 07].

Surfshark. (2022), Cybercrime statistics. Available from: https://surfshark.com/research/data-breach-impact/statistics [Last accessed on 2024 Aug 07].

Tran, H.T.T., Corner, J. (2016), The impact of communication channels on mobile banking adoption. International Journal of Bank Marketing, 34(1), 78-109.

Yu, J., Guo, C. (2008), An exploratory study of ubiquitous technology in retail banking. Academy of Commercial Banking and Finance, 8(1), 7-12.