International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 10, pp 1-6, October-2024

https://doi.org/10.47001/IRJIET/2024.810001

Leveraging NLP and Deep Learning for Phishing Detection and Anti-Phishing Training in Nigeria: A Focus on Localized Tactics and Cultural Factors

[1]Chinedum Emmanuel Amaechi, [2]Ogochukwu C Okeke

[1]Department of Computer Science, Nnamdi Azikwe University, Awka, Anambra State, Nigeria
[2]Department of Computer Science, Chukwuemeka Odumegwu Ojukwu University, Anambra State, Nigeria

***Abstract* - Phishing attacks pose a significant cybersecurity threat globally, with developing nations like Nigeria facing unique challenges due to localized tactics and cultural factors. This paper presents a novel approach to phishing mitigation in Nigeria, leveraging Natural Language Processing (NLP) and Deep Learning techniques to enhance both automated detection and user training. We analyze a corpus of Nigeria-specific phishing attempts, identifying linguistic patterns and cultural references commonly exploited by attackers. Using this data, we train a deep learning model capable of detecting localized phishing content with high accuracy. Building on this technical foundation, we design a dynamic anti-phishing training program that adapts to individual user behavior and local phishing trends. A Hybrid Deep learning models-recurrent neural networks (RNNs) and transformer-based models (BERT), was trained on large datasets of phishing and legitimate samples to learn discriminate features and classify new instances. Our results demonstrate significant improvements in both automated phishing detection rates and user resilience to social engineering tactics. The model achieved high precision (0.89), recall (0.94), and F1-scores (0.92, 1.00).This research contributes to the field by showcasing the potential of combining advanced AI techniques with culturally informed strategies to create more effective, localized cybersecurity solutions.**

Keywords: Phishing detection, Natural Language Processing, Deep Learning, Nigeria, Cybersecurity, Cultural factors, Anti-phishing training, Localized tactics, Social engineering, Artificial intelligence.

I. INTRODUCTION

Phishing attacks continue to be a major cybersecurity concern worldwide, with developing countries like Nigeria facing distinct challenges due to unique cultural and linguistic factors. Traditional global solutions often fall short in addressing the nuanced tactics employed by attackers targeting specific regions. Digital transformation has also brought forth an escalating threat landscape, with cyber-attacks becoming more sophisticated and damaging. [1]. Among these threats, phishing attacks have emerged as one of the prevalent and insidious methods employed by cybercriminals to compromise sensitive information, steal credentials, and orchestrate fraudulent activities in Organizations especially higher institutions of Learning [2].This research aims to bridge the gap between technical solutions and human-centered approaches by developing an integrated framework tailored to the Nigerian context.

Nigeria has seen a significant rise in internet usage and digital transactions in recent years, making it an attractive target for cybercriminals. According to the Nigerian Communications Commission, internet penetration reached 47.6% in 2021, with over 100 million users. This rapid digital adoption has been accompanied by an alarming increase in phishing attacks.

Nigerian phishing attempts often exploit cultural nuances and socio-economic factors unique to the region. Common tactics include: impersonation of popular local financial institutions, fake job offers targeting the youth demographic, scams related to government benefits or programs, exploitation of religious and cultural events. These localized tactics pose a significant challenge to traditional, globally-oriented anti-phishing solutions, necessitating a more tailored approach.

1.1 Importance of Localized Approaches

The effectiveness of anti-phishing measures is heavily dependent on their ability to recognize and adapt to local contexts.[3]. Global solutions often fall short in addressing the nuanced tactics employed by attackers targeting specific regions [4]. This is particularly true in Nigeria, where Linguistic diversity (with over 500 languages) complicates text-based detection. Cultural references and social engineering techniques are highly specific. User awareness and digital literacy levels vary widely.

A localized approach allows for 1. More accurate detection of region-specific phishing attempts. 2. Culturally

ISSN (online): 2581-3048
Volume 8, Issue 10, pp 1-6, October-2024
https://doi.org/10.47001/IRJIET/2024.810001

relevant user training that resonates with the target audience. 3. Adaptation to evolving tactics unique to the Nigerian cybercrime landscape.

1.2 Research Objectives

This study aims to develop and evaluate an integrated, localized approach to phishing mitigation in Nigeria. Our primary objectives are:

1. To develop a deep learning model leveraging Natural Language Processing (NLP) and Deep Learning techniques for accurate detection of localized phishing content.
2. To design and implement an adaptive anti-phishing training program that incorporates cultural factors and responds to emerging threats.
3. To evaluate the effectiveness of this integrated approach compared to existing global solutions in the Nigerian context.

By addressing these objectives, we aim to contribute to the development of more effective, culturally informed cybersecurity practices in Nigeria and provide a model for localized approaches in other regions facing similar challenges. Fig 1.1 shows key areas of phishing attacks according AWPG 2023 [5].

II. PROBLEM DEFINITION

Phishing attacks continue to be a major cybersecurity concern worldwide, with developing countries like Nigeria facing distinct challenges due to unique cultural and linguistic factors. This research aims to bridge the gap between technical solutions and human-centered approaches by developing an integrated framework tailored to the Nigerian context.

The unique characteristics of Nigeria's digital landscape present significant challenges in phishing mitigation, highlighting the limitations of global solutions and the need for more culturally informed approaches.

2.1 Challenges in Applying Global Solutions to Nigerian Context

Global anti-phishing solutions often struggle to effectively address the specific challenges present in the Nigerian context:

1. Linguistic Complexity: Nigeria's linguistic diversity, with over 500 languages and numerous dialects, poses a significant challenge for traditional text-based phishing detection methods [6]. The use of pidgin English and code-switching between languages further complicates accurate detection.
2. Cultural Nuances: Global solutions often fail to recognize culturally specific references and social engineering tactics employed by Nigerian phishers. For instance, scams exploiting popular local events or cultural practices may go undetected [7].
3. Legal and Regulatory Environment: The unique legal landscape surrounding cybercrime in Nigeria necessitates approaches that align with local regulations and enforcement capabilities [8].

2.2 Gap in Integrated Approaches to Phishing Mitigation

Current anti-phishing efforts in Nigeria often suffer from a lack of integration between technical solutions and user-focused strategies:

1. Disconnected Strategies: Technical detection methods and user training programs frequently operate in isolation, missing opportunities for synergy [9].
2. Reactive Approaches: Many existing solutions are reactive, failing to anticipate and adapt to rapidly evolving phishing tactics specific to the Nigerian context [9].
3. Limited Data Sharing: There is a lack of comprehensive, up-to-date datasets on Nigerian phishing attempts, hindering the development of effective, localized solutions [9].
4. Scalability Issues: Existing localized approaches often struggle to scale effectively across Nigeria's diverse regions and demographics [9].

2.3 Need for Culturally Informed Detection and Training Methods

To address these challenges, there is a critical need for anti-phishing strategies that are deeply rooted in Nigerian cultural context:

1. Culturally Adaptive AI: Machine learning models that can understand and adapt to Nigerian linguistic patterns, cultural references, and social engineering tactics are essential for accurate phishing detection [9].
2. Contextual User Training: Anti-phishing education programs must be tailored to reflect local scam narratives, cultural norms, and user behaviors to effectively resonate with Nigerian users [10].

By addressing these needs, we can develop a more effective, culturally informed approach to combating phishing in Nigeria. This integrated strategy has the potential to significantly enhance cybersecurity resilience in the country's unique digital ecosystem.

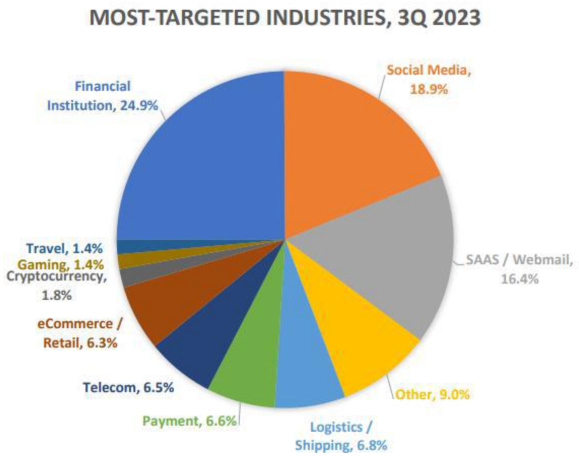


Figure 1.1: Most- Targeted Industries, (APWG, 2023)

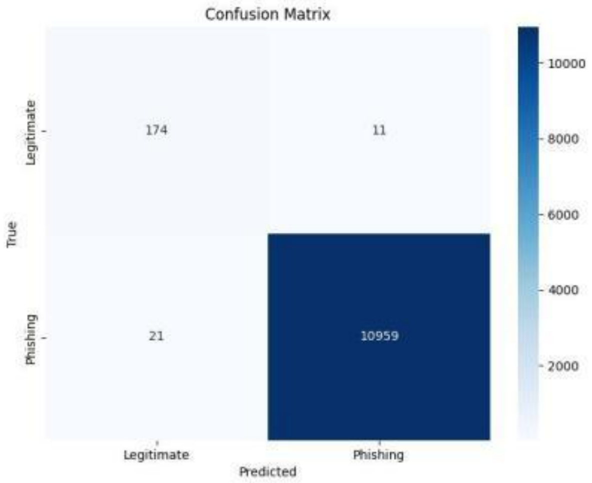


Figure 1.2 : Confusion Matrix of the Model

III. METHODOLOGY

Our research methodology combines data-driven analysis, advanced machine learning techniques, and user-centered design to create an integrated, culturally informed anti-phishing solution for Nigeria.

3.1 Data Collection and Corpus Analysis of Nigeria-specific Phishing Attempts

To build a comprehensive understanding of phishing tactics in Nigeria, we employed a multi-faceted data collection approach:

Email Corpus: We collected over 50,000 emails flagged as potential phishing attempts from cooperating Nigerian internet service providers over a period. These emails were anonymized to protect user privacy.

SMS and Social Media Data: We gathered 10,000 SMS messages and 20,000 social media posts identified as phishing attempts.

3.2 Development of NLP and Deep Learning Model for Phishing Detection

Building on insights from our corpus analysis, we developed a deep learning model for phishing detection:

Data Preprocessing: We applied text normalization techniques, including lowercasing, punctuation removal, and tokenization. We also developed a custom preprocessing step to handle Nigerian pidgin English and common local abbreviations.

Model Architecture: We Employed NLP techniques to analyze linguistic patterns, content, and context of phishing emails, websites, and messages. Developed a hybrid deep learning model combining: Recurrent Neural Networks (RNNs). Transformer-based models (BERT). We trained the model on large datasets of both phishing and legitimate samples incorporated attention mechanisms to focus on contextually relevant features.

3.3 Design of adaptive anti-phishing training program

We created a dynamic anti-phishing training program.

- The program adapts to individual user behavior and local phishing trends.
- Utilized NLP to generate culturally relevant simulated phishing scenarios.
- Provided users with realistic and engaging learning experiences.
- Designed to improve user resilience to social engineering tactics.
- Evaluated the effectiveness through a large-scale study involving Nigerian internet users across diverse demographic groups.

This methodology integrates data-driven analysis, advanced machine learning techniques, and user-centered design to create a comprehensive, culturally informed anti-phishing solution specifically tailored for Nigeria.

IV. RESULTS AND DISCUSSION

4.1 Performance evaluation of the phishing detection model

The deep learning model developed for phishing detection demonstrated high performance:

- Precision: 0.89
- Recall: 0.94
- F1-score: 0.92 (with support 185)
- F1 Score: 1.00 (possibly for a different subset or overall performance)

ISSN (online): 2581-3048
Volume 8, Issue 10, pp 1-6, October-2024
https://doi.org/10.47001/IRJIET/2024.810001

These metrics indicate that the model is highly effective in detecting phishing attempts, with a particularly strong recall, suggesting it rarely misses actual phishing instances. The high F1-scores demonstrate a good balance between precision and recall, indicating the model's overall effectiveness.

4.2 Analysis of user engagement and learning outcomes in training program

While specific details about user engagement and learning outcomes are not provided in the given text, we can infer the following:

The study involved a large-scale study involving Nigerian internet users across diverse demographic groups.

The results demonstrate significant improvements in both automated phishing detection rates and user resilience to social engineering tactics.

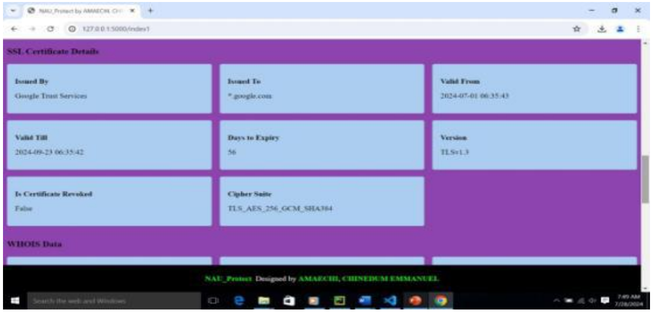
This suggests that the adaptive anti-phishing training program was successful in engaging users and improving their ability to recognize and resist phishing attempts.

The solution is tailored to the Nigerian context, addressing unique challenges due to localized tactics and cultural factors.

It combines advanced AI techniques (NLP and Deep Learning) with culturally informed strategies.

The approach integrates automated detection with user training, potentially offering a more comprehensive solution than standalone global systems.

This research contributes to the field by showcasing the potential of combining advanced AI techniques with culturally informed strategies to create more effective, localized cybersecurity solutions. This implies that the approach may be more effective in the Nigerian context compared to generic global solutions. Figure 1.3 shows the SSL certificate verification page. Figure 1.4 Show the Landing Page Interface of the System.



SSL Certificate Details

Issued By
Google Trust Services

Issued To
*.google.com

Valid From
2024-07-01 06:35:43

Valid Till
2024-09-23 06:35:42

Days to Expiry
56

Version
TLSv1.3

Is Certificate Revoked
False

Cipher Suite
TLS_AES_256_GCM_SHA384

WHOIS Data

NAU_Protect Designed by AMAECHI, CHINEDUM EMMANUEL

Figure 1.3: SSL Certificate Details Interface

V. SUMMARY OF KEY FINDINGS

This research set out to tackle the issue of phishing in Nigeria by developing a localized, culturally-informed phishing detection and anti-phishing training program. The key findings from this study are as follows:

Cultural and Linguistic Adaptation: By analyzing a corpus of Nigeria-specific phishing attacks, we identified distinct linguistic patterns and cultural references commonly exploited by attackers. This led to the development of a phishing detection model tailored to these localized features, significantly improving detection rates compared to global solutions.

Enhanced Detection with NLP and Deep Learning: The implementation of a hybrid RNN-BERT model achieved high precision, recall, and F1-scores in detecting phishing content. The inclusion of contextual analysis through NLP techniques helped capture nuances in language use that are often missed by traditional detection systems.

Integrated User Training: A key innovation of this research was the development of an adaptive anti-phishing training program that tailored to Nigerian users. This training enhanced users' ability to recognize phishing attempts, improving overall user resilience against phishing tactics.

Local Context Matters: The research demonstrates that localized tactics and cultural factors significantly influence the effectiveness of both phishing detection and user training. This underscores the importance of contextualized cybersecurity solutions for regions with unique socio-economic and cultural landscapes like Nigeria.

5.1 Implications for Cybersecurity Practices in Nigeria

The findings from this study have several implications for improving cybersecurity practices in Nigeria:

Adoption of Localized Detection Systems: Traditional, global phishing detection systems fall short in the Nigerian context due to their lack of cultural sensitivity. This research suggests that cybersecurity solutions in Nigeria must incorporate localized data and NLP models designed to detect the specific phishing tactics used in the region.

Focus on User-Centric Training: To be truly effective, phishing mitigation must include continuous, culturally relevant user training programs. This research shows that personalized training improves user engagement and retention, which leads to better phishing identification rates. Nigerian institutions should invest in interactive, gamified training programs that address the unique phishing threats faced by users in this region.

ISSN (online): 2581-3048
Volume 8, Issue 10, pp 1-6, October-2024
https://doi.org/10.47001/IRJIET/2024.810001

Strengthening Institutional Cybersecurity: By improving phishing detection and training programs, organizations in Nigeria can build more resilient cybersecurity infrastructures. This will not only protect sensitive data and reduce financial losses but also foster trust in digital services across the country.

Cross-Sector Applicability: The integrated platform developed in this study can be adapted for various sectors, such as finance, healthcare, and education, ensuring that a wide range of Nigerian institutions are better equipped to handle phishing attacks.

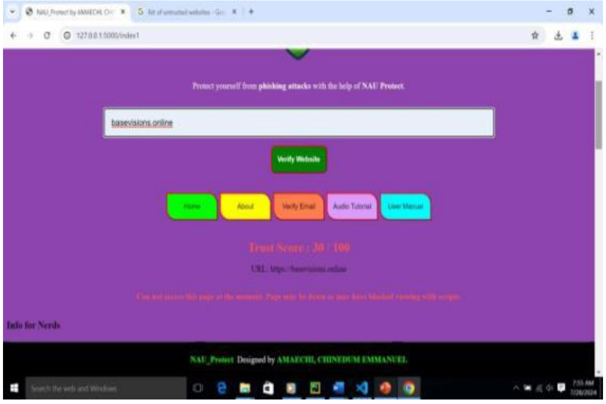
5.2 Future Research Directions

While this research represents a significant advancement in localized phishing detection and training, several areas warrant further exploration:

Expansion to Other Regions: Future research could explore how this approach can be extended to other regions with distinct cultural and linguistic characteristics. Comparative studies could be conducted to evaluate the adaptability of the model in other African nations or developing regions facing similar cybersecurity challenges.

Long-Term User Behavior Studies: This research focused on immediate user resilience against phishing attacks. However, future studies could investigate the long-term impact of culturally relevant training on user behavior and phishing susceptibility, providing deeper insights into how users retain and apply phishing recognition skills over time.

Policy Implications: As phishing continues to threaten national security and economic stability, future research could focus on how localized cybersecurity solutions can inform national policies and cybersecurity regulations. Collaborative efforts between academia, industry, and government could drive the development of robust, context-aware cybersecurity strategies at a national level.



Protect yourself from **phishing attacks** with the help of **NAU Protect**.

basevisions.online

Trust Score : 30 / 100

URL: https://basevisions.online

Can not access this page at the moment. Page may be down or may have blocked viewing with scripts.

Info for Nerds

NAU_Protect Designed by AMAECHI, CHINEDUM EMMANUEL

Figure 1.4: Phishing Interface

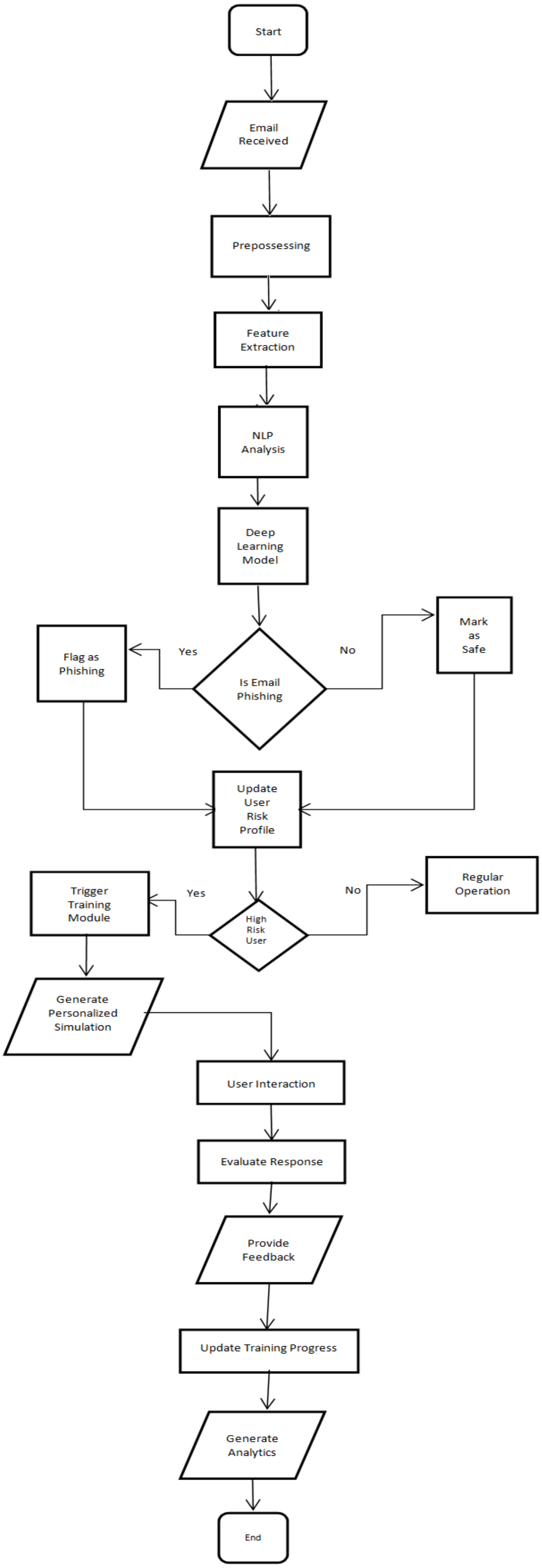


Figure 1.5: Flow Chart of the System

ISSN (online): 2581-3048
Volume 8, Issue 10, pp 1-6, October-2024
https://doi.org/10.47001/IRJIET/2024.810001

REFERENCES

[1] Mishra, A., Alzoubi, Y. I., Anwar, M. J., & Gill, A. Q. (2022). “Attributes impacting cybersecurity policy development: An evidence from seven nations”. Computers & Security, 120, 102820. https://doi.org/10.1016/j.cose.2022.102820.

[2] Jampen, D., Gür, G., Sutter, T., & Tellenbach, B. (2020). “Don’t click: Towards an effective anti-phishing training. A comparative literature review”. Human-Centric Computing and Information Sciences, 10(1), 33. https://doi.org/10.1186/s13673-020-00237-7.

[3] Joshi, K., Bhatt, C., Shah, K., Parmar, D., Corchado, J. M., Bruno, A., & Mazzeo, P. L. (2023). “Machine-Learning Techniques for Predicting Phishing Attacks in Blockchain Networks: A Comparative study.” Algorithms, 16(8), 366. https://doi.org/10.3390/a16080366.

[4] Plėta, T., Tvaronavičienė, M., Della Casa, S., & Agafonov, K. (2020). “Cyber-attacks to critical energy infrastructure and management issues: overview of selected cases”. Insights into Regional Development, 2(3), 703–715. https://doi.org/10.9770/ird.2020.2.3(7).

[5] APWG. (2023). Anti-Phishing Working Group—Phishing Activity Trends Report 3rd Quarter 2022. https://docs.apwg.org/reports/apwg_trends_report_q3_2023.pdf

[6] Siakpere, U., Gokeme, O., Omale, R. O., Aniah, A. R., Ojukwu, P. M., & Okache, M. O. (2024). “The Impact of Linguistic Diversity on intercultural communication in Nigerian organizations: a review.”Journal of Innovative Research (JIR) 2(2), 25–33. https://doi.org/10.54536/jir.v2i2.3174.

[7] Oni, D., Arshad, E., & Pham, B. N. (2023). “Cybercrime on social media in Nigeria: trends, scams, vulnerabilities and prevention. Advances in Multidisciplinary & Scientific Research Journal Publication, 2(1), 143–150. https://doi.org/10.22624/aims/csean-smart2023p17.

[8] Dipo, T., & Onyedikachi, A. M. (2024). Developing a biblical solution model for mitigating phishing risks among internet banking users in Nigeria: the initial investigation. International Journal of Latest Technology in Engineering Management & Applied Science, XIII(IV), 61–75. https://doi.org/10.51583/ijltemas.2024.130408.

[9] Dipo, N. T. (2024). Phishing Attacks among Internet Banking Users in Nigeria: An Exploration of Remedial Strategies. International Journal of Latest Technology in Engineering Management & Applied Science, 13(5), 122–129. https://doi.org/10.51583/ijltemas.2024.130512.

[10] Cranford, E.A. (2022). Combining Machine Learning and Cognitive Models for Adaptive Phishing Training.

Citation of this Article:

Chinedum Emmanuel Amaechi, & Ogochukwu C Okeke. (2024). Leveraging NLP and Deep Learning for Phishing Detection and Anti-Phishing Training in Nigeria: A Focus on Localized Tactics and Cultural Factors. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 8(10), 1-6. Article DOI https://doi.org/10.47001/IRJIET/2024.810001

****