International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Building a Cloud-Based Machine Learning System for Automated Phishing Detection and Email Security in Cloud-Based Email Platforms.

Karakadzai Sithole

NFSU

ABSTRACT

This survey report examines the development of a cloud-powered machine learning system for automated phishing detection and email security in cloud-based email platforms. It examines current research, development processes, and technological advancements to create practical, scalable solutions to combat the increasing risk of phishing attacks and data breaches. Emails are a popular form of communication for both business and personal needs. Sensitive and confidential information, like financial data, credit reports, login credentials, and the like, is also sent via mail. Because of this, cybercriminals might utilize them to their advantage and utilize this information for illicit activities. Phishing is a tactic employed by con artists to obtain personal information from people by seeming to come from reliable sources. The sender of a phished email may pose as a trustworthy source to trick recipients into divulging personal information. Phishing emails have also been widely employed in consumer fraud and financial institutions. This paper presents an overview of the various methods that are currently being employed in email filtering for phishing email detection and protection. A comparison and evaluation of various methods

Keywords: Phishing, cloud-powered, machine learning, email platforms, data breaches, cybercriminals

INTRODUCTION:

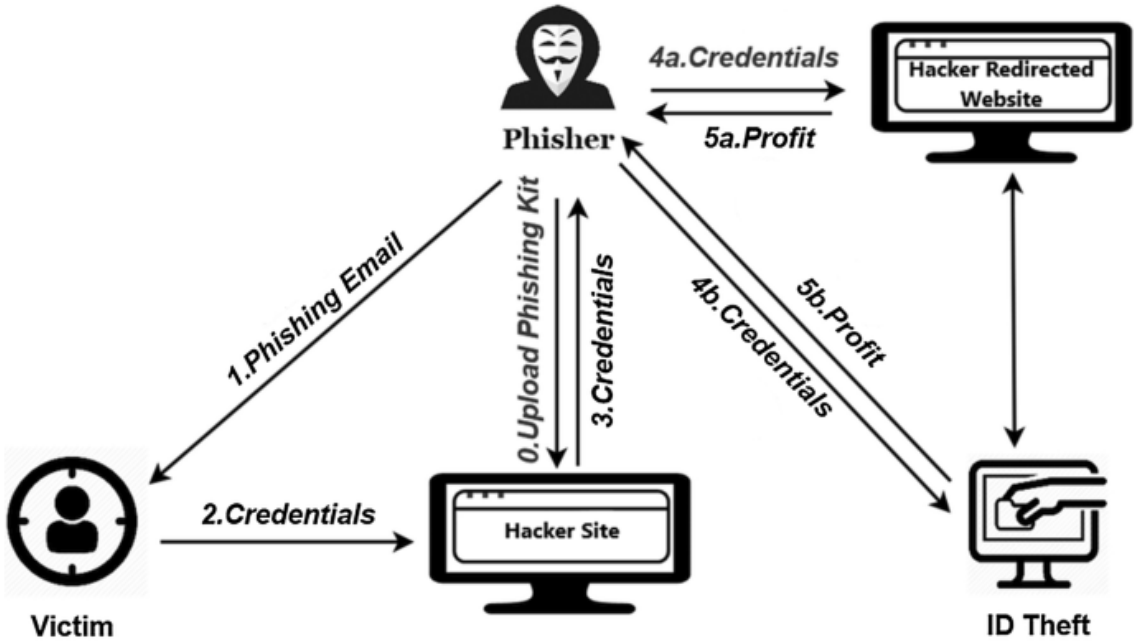
Developing Machine Learning Models for Automated Phishing Detection in Cloud-Based Platforms Phishing attacks pose a grave threat to both individuals and organizations as they evolve in prevalence, utilizing social engineering tactics to deceive individuals into revealing sensitive information or engaging with harmful links. Traditional email filtering methods often prove inadequate against these dynamic threats, particularly as cloud-based platforms like Google Workspace and Microsoft 365 become more prevalent. [1] As phishing attacks continue to grow in sophistication and prevalence, individuals and organizations face the threat of divulging private information or falling victim to financial losses and data breaches through deceptive tactics. The shift towards cloud-based email platforms like Google Workspace and Microsoft 365 has further exacerbated this issue, with hackers targeting these centralized communication hubs due to their increased usage. The vast volume of emails exchanged on these platforms presents a challenge for manual security teams, struggling to keep pace with the influx of phishing attempts. To address this pressing concern, the research aims to develop a scalable machine learning system specifically designed for automated phishing detection within cloud-based email environments. By harnessing features from URLs, sender details, and email content, efficient machine learning models can be trained to accurately identify and classify phishing emails, fortifying defenses against evolving

threats. [2] In conclusion, the escalating sophistication and prevalence of phishing attacks pose a significant risk to both individuals and organizations, necessitating advanced defense mechanisms. By leveraging machine learning to enhance automated phishing detection in cloud-based email platforms, the research endeavors to combat these evolving threats effectively. Through the adept utilization of pertinent features and robust models, a scalable and reliable system can be established to safeguard against the deceptive tactics utilized in phishing attempts.

PHISHING THREATS IN CLOUD-BASED EMAIL PLATFORMS:

Phishing attacks have evolved significantly over the years, adapting to technological advancements and targeting vulnerabilities in cloud-based email platforms. The evolution includes Social Engineering Techniques: Early phishing attacks relied heavily on basic social engineering, exploiting trust and manipulating users into divulging sensitive information. Email Spoofing: Attackers progressed to email spoofing, creating deceptive emails that appeared legitimate to trick users into clicking malicious links or providing confidential data. Spear Phishing: With increased sophistication, attackers began tailoring their attacks for specific individuals or organizations, leveraging information about the target to enhance the believability of phishing attempts. Clone Phishing: Phishers started duplicating legitimate emails, modifying content to include malicious links or attachments, and fooling users who were expecting communication from trusted sources. [3]Whaling Attacks: Targeting high-profile individuals, such as executives or key decision-makers, became prevalent, and often led to more severe consequences due to the potential access to sensitive corporate information.

Figure 1 - Phishing Attack



MACHINE LEARNING FOR PHISHING DETECTION:

Traditional Methods vs. Machine Learning Approaches:

Traditional Methods: Rule-Based Systems: Early phishing detection systems relied on predefined rules to identify phishing patterns. [4] These rules often struggled to adapt to the evolving nature of phishing

attacks. Heuristics and Signature-Based Detection: Static heuristics and signature-based approaches were employed to match incoming emails against known phishing signatures, but these methods were limited in their ability to detect new, previously unseen attacks. Machine Learning Approaches: Advantages of Machine Learning: ML models offer adaptability and the ability to learn from data, enabling them to identify patterns and anomalies associated with phishing attacks. Feature Learning: ML models can automatically identify relevant features and patterns, reducing the reliance on predefined rules and signatures. Feature Extraction and Selection: Feature Extraction: Effective phishing detection relies on extracting informative features from emails. Common features include sender reputation, URL characteristics, email content analysis, and lexical analysis. Natural Language Processing (NLP): NLP techniques are employed to analyse the textual content of emails, extracting semantic meaning and identifying suspicious language patterns. Image and Attachment Analysis: Phishing attacks often involve malicious attachments or images. ML models can analyse these components, detecting anomalies and potential threats. Behavioral Features: Monitoring user behaviour, such as email interaction patterns and click rates, can be crucial for identifying deviations indicative of phishing attempts. Supervised Learning: Classification Models: Supervised learning algorithms, such as Support Vector Machines (SVM), Decision Trees, and Random Forests, are trained on labeled datasets to classify emails as phishing or legitimate based on predefined features. Ensemble Methods: Combining multiple models through ensemble techniques improves overall detection accuracy and robustness. Unsupervised Learning: Clustering Algorithms: Unsupervised learning, particularly clustering algorithms like K-means or hierarchical clustering, can identify patterns within data without labeled training sets. Anomaly Detection: Unsupervised learning is effective for anomaly detection, identifying emails that deviate from the norm and may indicate phishing attempts. Deep Learning Techniques: Neural Networks: Deep learning models, especially neural networks, can automatically learn hierarchical features from raw data, enabling them to capture complex patterns associated with phishing attacks. Convolutional Neural Networks (CNN): CNNs are effective for image-based phishing detection, analyzing email content and embedded images for malicious elements. Recurrent Neural Networks (RNN): RNNs are suitable for sequential data, making them valuable for analyzing the temporal nature of email communication and detecting phishing patterns. [5]Transfer Learning: Leveraging pre-trained models and fine-tuning them for phishing detection tasks accelerates model training and enhances performance, especially when dealing with limited labeled data.

LITERATURE REVIEW IN USING MACHINE LEARNING ALGORITHMS

Table 1- Literature Review

Techniques	Year	Advantages	Disadvantages	Objectives	Research
Neural network	2020	Using data analysis check attacks NN used to resolve attack issues	Cyber-attack issue Run time error, Malicious security attack	Consumers face many risks and security issues and loss of sensitive data and other details	Still in research
Random Forest	2020	Use random forests to detect phishing data. Resolve threat and attack issues	Not secure data. Phishing attacks and authentication issues	The main role is to secure the client's login and other data using the algorithm	Still in research

				to resolve the threat and attack issue	
Neural Network	2019	Website phishing attacks, improve performance, reduce middle attack	Malicious attack Links attack in secure data	The objective is to maintain the system daily and check attacks and apply a decision tree to maintain the data	Still in research
Decision Tree	2017	DDoS attack issues using decision tree improve security problems, the ML algorithm used to improve performance	Run time error, huge problem in execution	To maintain the system daily check attacks and apply a decision tree to maintain data	Middle term issue
Naïve Bayes, Random Forest	2020	AI procedure, Naïve Bayes, and random forest to detect malicious and website attacks, ML used to identify Phishing attacks	Loss of personal data in secure sites, data not secure	NB algorithm is used to identify the attack and resolve it	Long term issue
SVM, Naïve Bayes	2018	Resolve issues using SVM ML strategies to remove attacks and threats improve and secure data	The network is not secure and loses personal data	SVM used to identify and detect the attack	Still in research
GWO-Bert	2024	effectiveness with more training data	need for fine-tuning	to develop a 2 stage hybrid deep learning model which can accurately detect emails as spam or ham	Still in research
Artificial Neural Network	2023	This signifies its high efficacy in detecting and classifying malicious emails	greater computational burden, proneness to overfitting	To identify between a safe and spam email	Steal in research

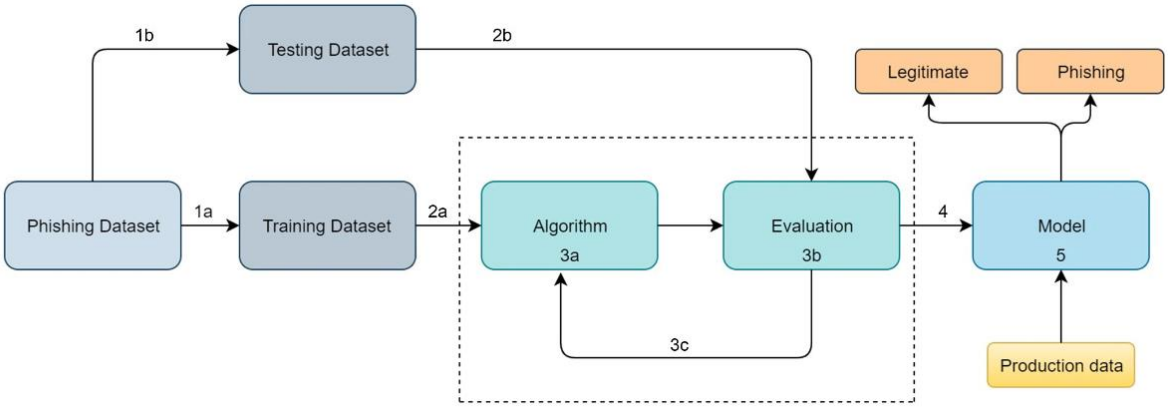
Table 2- Different Algorithm results

	Logistic Regression	Random Forest	Decision Tree	Naïve Bayes	Support vector machine	XG boost	light GBM	MLP Neural Network
Training score	0.802889	0.993755	0.993755	0.607787	0.719646	0.948903	0.917354	0.750434
Testing score	0.806298	0.908578	0.858415	0.611509	0.718567	0.900977	0.894463	0.752009
Classifier auc score	0.802966	0.906394	0.857589	0.600391	0.706047	0.898735	0.891750	0.743340
Estimator auc score	0.857625	0.96677	0.859618	0.673123	0.775768	0.961408	0.956050	0.796522

DATASET CONSIDERATIONS:

Phishing detection involves using various datasets, [6] including public and private ones, to identify and detect phishing emails. Public datasets like Phishtank and Enron Email Dataset provide real-world examples, while private datasets can be curated by organizations. Imbalanced datasets are crucial to avoid bias and temporal datasets provide a realistic representation of the threat landscape. Challenges in dataset selection include labeling accuracy, representativeness, privacy concerns, and scalability. Synthetic data generation techniques, such as GANs, can help overcome limitations and improve model robustness. Transfer learning with synthetic data can enhance model performance by combining synthetic and real datasets. However, challenges include maintaining data realism, avoiding overfitting, and ensuring samples cover real-world phishing scenarios.

Figure 2 - Machine earning for phishing attack Detection



EVALUATION METRICS AND BENCHMARKS:

Phishing detection performance is evaluated using metrics such as False Positive Rate (FPR), False Negative Rate (FNR), Accuracy, Precision, Recall, F1 Score, and Area Under the ROC Curve (AUC-ROC). Cloud-powered machine learning systems are benchmarked for their processing speed, scalability, resource utilization, reliability, and availability. Red team testing, adversarial testing, user feedback, incident response simulation, and long-term performance are used to assess the effectiveness of the

system. [6]A lower FPR indicates better precision in avoiding false alarms, while a lower FNR ensures real threats are not overlooked. The model's reliability and availability are also assessed through real-world testing and simulations. The system's performance over an extended period is evaluated to assess its ability to adapt to evolving phishing techniques and maintain effectiveness.

CASE STUDIES AND IMPLEMENTATIONS:

Cloud-powered machine learning (ML) systems have been successfully implemented in various industries, including financial, healthcare, technology, and educational sectors. These systems have significantly reduced the number of successful phishing attacks, improved detection accuracy, and reduced disruptions to critical operations. However, challenges include integration complexity, adversarial tactics, user awareness and education, and regulatory compliance. To overcome these, it is essential to prioritize thorough compatibility testing, implement regular model retraining, stay vigilant to emerging phishing techniques, and prioritize a privacy-by-design approach. Additionally, combining technology solutions with ongoing user awareness programs can create a multi-layered defense strategy. Overall, ML systems offer valuable insights for businesses and organizations to protect sensitive data and maintain business continuity.

FUTURE DIRECTIONS AND EMERGING TRENDS:

Future machine learning (ML) algorithms for phishing detection are expected to prioritize explainability, ensemble techniques, online learning, federated learning, real-time threat intelligence integration, automated threat response, cross-platform collaboration, differential privacy, ethical AI practices, user consent and control, interactive training platforms, behavioral analysis for education, and simulated phishing exercises. These advancements will enhance detection accuracy, robustness, and user awareness while ensuring transparency and avoiding discriminatory outcomes. Ethical AI practices will be crucial, and user consent mechanisms will be enhanced. Additionally, interactive training platforms will be developed to educate users about phishing tactics and improve their ability to identify and report suspicious emails.

SECURITY AND PRIVACY CONCERNS:

Cloud-based systems require strong encryption, robust access controls, and an understanding of data residency and jurisdiction to prevent unauthorized access to sensitive information. Incident response planning and balancing security needs with user privacy are crucial. Email content analysis and user consent can help balance security needs with privacy. Data minimization and transparency can reduce the impact on user privacy. Adhering to data protection regulations, cross-border data transfers, and user notification and rights are essential. Regular security audits and certifications demonstrate an organization's commitment to maintaining a secure cloud-based email security system. By implementing these measures, organizations can ensure the security of their cloud-based email systems.

CONCLUSION:

The study delves into the advancements of a cloud-based machine learning system for automated phishing detection and email security, emphasizing the evolution of phishing tactics, the pivotal role of machine learning in bolstering detection capabilities, and the essentiality of robust cloud security frameworks Phishing attacks continue to evolve, becoming more sophisticated and challenging to detect; here, machine

learning plays a crucial role in bolstering defense mechanisms. The implementation of cloud security architectures ensures the protection of sensitive data within email platforms. Consideration of robust datasets and performance evaluation metrics is pivotal for enhancing detection accuracy. Successful integration in diverse industries underscores the tangible advantages of cloud-powered machine learning in combatting phishing. Looking ahead, emphasis on innovation, collaboration, user-centric design, and ethical frameworks will drive continual progress in this dynamic landscape, necessitating unwavering commitment to outmanoeuvre cyber threats. The future of cloud-powered machine learning for phishing detection is promising, demanding continuous evolution and collaboration. Embracing user-centric approaches and ethical considerations will be paramount in staying ahead of cyber adversaries. The dynamic nature of this field necessitates perpetual adaptability and collective effort to fortify email security in the face of evolving threats.

REFERENCES

1. C. U. P. Bowl., "https://it.cornell.edu/phish-bowl," 3 January 2021.
2. C. P. E. E. Dataset, "http://www.cs.cmu.edu/~enron/," 16 January 2021.
3. A. Das, S. Baki, A. El Aassal, R. Verma and A. S. Dunbar, "A Comprehensive Reexamination of Phishing Research From the Security Perspective.," *Commun. Surv. Tuts,* 2020.
4. P. Muncaster, "COVID19 Drives Phishing Emails Up 667% in Under a Month," *https://www.infosecurity-magazine.com/news/covid19-drive-phishing-emails-667?utm_source=twitterfeed&utm_medium=twitter,* 2020.
5. Abdul Basit, Maham Zaphar,Xuan Liu, "Springler Linker," 23 October 2020. [Online]. Available: https://link.springer.com/article/10.1007/s11235-020-00733-2.
6. Basit A., Zaphar m., Liu X., et al. A comprehensive survey of AI-enabled phishing attacks detection techniques, "Sprinkler Link," 23 October 2020. [Online]. Available: https://rdcu.be/dxBKb. [Accessed 9 October 2020].
7. A. Muneer, R. F. Ali, A. A. Al-Sharai and S. M. Fati,, pp. 1-6, " "A Survey on Phishing Emails Detection Techniques,"," in *2021 International Conference on Innovative Computing (ICIC),*, Lahore, Pakistan, 2021.
8. Bin Sulaiman, R., Schetinin, V. & Sant, P. , "Review of Machine Learning Approach on Credit Card Fraud Detection. Hum-Cent Intell Syst 2, 55–68 (2022).," in *Springer Nature*, 5 May 2022.
9. Mughaid, A., AlZu'bi, S., Hnaif, A. et al.
10. Mughaid, A., AlZu'bi, S., Hnaif, A. et al. , "An intelligent cyber security phishing detection system using deep learning techniques. Cluster Comput 25, 3819–3828 (2022).," Mughaid, A., AlZu'bi, S., Hnaif, A. et al., 2022.
11. I. Sarker, "Machine Learning for Intelligent Data Analysis and Automation in Cybersecurity: Current and Future Prospects. Ann. Data. Sci. 10,," 19 September 2022. [Online]. Available: https://rdcu.be/dxB5y. [Accessed December 2023].
12. Ravi, Roshan,Shillare, Abhishek Arvind,Prakash Bhoir, Prathamesh,Charumathi, K S, "URL-based Email Phishing Detection Application," *International Research Journal of Engineering and Technology,* 2021.
13. Valenzuela, Chelsea,Sturman, Daniel,Plate, Oliver,Tanvir, Tazin,Auton, Jaime C.Bayl-Smith, Piers, "The role of cue utilization in the detection of phishing emails," *Applied Ergonomics*, p. 106, 2023.

14. C. Thapa, J. Tang, A. Abuadbba, Y. Gao, S. Camtepe, S. Nepal, M. Almashor and Y. Zheng, "Evaluation of Federated Learning in Phishing Email Detection.," no. ttps://doi.org/10.3390/s23094346, 2023.
15. T. J. A. A. G. Y. C. S. N. S. A. M. Z. Y. Thapa C, "valuation of Federated Learning in Phishing Email Detection. Sensors.," no. https://doi.org/10.3390/s23094346, 2023.
16. C. J. W. T. A. A. Y. G. S. C. S. N. M. A. a. Y. Z. Thapa, ""Evaluation of Federated Learning in Phishing Email Detection" Sensors 23," no. https://doi.org/10.3390/s23094346.
17. M. N. N. B. M. a. A. H. S. M. S. Rana, ""Deepfake Detection: A Systematic Literature Review,"," vol. vol 10, pp. pp. 25494-25513, 2022.
18. S. D. J. B. K. Sergi D Bray, "Testing human ability to detect 'deepfake' images of human faces,," *Journal of Cybersecurity*, vol. Volume 9, no. Issue 1, 2023.
19. N. Shoham, T. Avidor, A. Keren, N. Israel, D. Benditkis, L. Mor-Yosef and I. Zeitak, "Overcoming Forgetting in Federated Learning on Non-IID Data.," *Google Scholar*, 2019.
20. B. McMahan, E. Moore, D. Ramage, S. Hampson and B. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data.," 2017.
21. P. Mohassel and P. A. Rindal, "A Mixed Protocol Framework for Machine Learning," pp. pg 35 - 52, 15–19 October 2018.
22. S. Wagh, D. Gupta and N. Chandran, "SecureNN: 3-Party Secure Computation for Neural Network Training," 2019.
23. A. Cidon, L. Gavish, I. Bleier, N. Korshun, M. Schweighauser and A. Tsitkin, "High precision detection of business email compromise," 14–16 August 2019.
24. H. Gascon, S. Ullrich, B. Stritter and K. Rieck, "Reading Between the Lines: Content-Agnostic Detection of Spear-Phishing Emails.," *In Proceedings of the RAID 2018*, 10–12 September 2018;.
25. M. Nguyen, T. Nguyen and T. Nguyen, " A deep learning model with hierarchical lstms and supervised attention for anti-phishing," *In Proceedings of the 1st AntiPhishing Shared Pilot at 4th ACM IWSPA, Tempe, AZ, USA*, 21 March 2018.
26. J. Zhang and X. Li, "Phishing detection method based on borderline-smote deep belief network.," *In Proceedings of the International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage, Guangzhou, China*, , pp. pg 45-53, 12–15 December 2017.