Ceddi Journal of Information System and Technology (JST)
Vol. 3 No. 2 December (2024)
https://doi.org/10.56134/jst.v3i2.100

ISSN: 2829-808X (print)
ISSN: 2829-6575 (online)

Evaluating Deep Learning Models for Website Phishing Attack Detection: A Comparative Analysis

Abdullahi Raji Egigogo[1*], Ismaila Idris[2], Morufu Olalere[3], Abisoye Opeyemi Aderiike[4]

[1-4]Department of Cyber Security Science, Federal University of Technology Minna, Minna, Nigeria
[3]Department of Cybersecurity, National Open University, Abuja, Nigeria

Abstract

Phishing attacks remain a significant security threat in cyberspace, targeting individuals and businesses to steal confidential information. Traditional detection methods often struggle to identify newly created or altered phishing sites, highlighting the need for more adaptive solutions. This study evaluates the performance of various deep learning (DL) models for detecting online phishing attacks. A comparative analysis of single and hybrid DL models, including CNN, LSTM, BiGRU, and their combinations, is conducted. The evaluation is based on metrics such as accuracy, precision, recall, and F1-score, derived from 17 peer-reviewed publications published between 2019 and 2024. Results indicate that hybrid models, particularly ODAE-WPDC, exhibit superior performance, achieving accuracy rates of up to 99.28% and robust results across all metrics. Single models, such as CNN and BiGRU, also demonstrate strong performance, with accuracy ranging from 97% to 99.5%. This research underscores the efficacy of deep learning architectures in phishing detection and offers practical guidance for selecting optimal models based on specific requirements.

Keywords: Deep learning models; Hybrid architectures; Phishing detection; Performance evaluation; Cybersecurity threats.

Received: 15 March 2024 Revised: 20 May 2024 Accepted: 25 November 2024

Introduction

Phishing attacks involve building bogus websites that mimic legitimate ones to steal sensitive information such as usernames, passwords, and financial details (Manoj et al., 2021; Alabdan, 2020; Ali & Mohd Zaharon, 2024; Alkhalil et al., 2021; Chaudhary, 2012; Gupta et al., 2017; James, 2005; Mohammad et al., 2015). Notwithstanding enormous technology improvements, phishing remains a continuous concern owing to attackers' adaptable and dynamic features (Do et al., 2022). Traditional approaches to phishing detection, such as blacklist-based systems and heuristic methods, are generally useless against new or significantly updated phishing websites (Somesha et al., 2020; Eka Purwiantono & Tjahyanto, 2017; Hidayanto et al., 2022; Irawan et al., 2021; Mahmud & Wirawan, 2024).

The deployment of Deep Learning (DL) models has been increasingly popular in recent years due to their capacity to learn complicated representations and their resilience to developing attack techniques (Alzubaidi et al., 2021; Admin, 2024; Afinda, 2024; Amazon, 2024; Nursyafitri, 2023). DL models can examine enormous datasets to uncover patterns that automatically distinguish phishing websites from real ones (Citra, 2024; Hazzataqiy, 2024; Nada, 2019; Pilo, 2023; Puskomedia, 2024). This makes them well-suited to handle the difficulty of identifying phishing assaults, notably when coupled with feature extraction techniques that allow models to focus on essential website attributes like URLs, HTML content, and domain registration information (Tesfom et al., 2023; Almomani et al., 2022; Kara et al., 2022; Opara et al., 2020). Numerous deep learning models are available for identifying phishing websites; however, the effectiveness of each model may vary significantly based on the type of attack, feature selection, and data preprocessing practices (Iswahyudi et al., 2023; Santoso, 2023; Wahyuni et al., 2024). A careful examination of these models is essential to develop the best-performing solutions for real-world applications.

With the wide range of deep learning models available, it is essential to comprehend the advantages and disadvantages of various phishing detection architectures to maximize efficiency and reduce false positives or negatives. This article

*Corresponding author.
E-mail address: rajiaegigogo@auk.edu.ng (Abdullahi Raji Egigogo)

analyzes various deep learning architectures for phishing detection and investigates their effectiveness in identifying phishing websites.

Various scholars have proposed several deep learning methods to detect online phishing attacks. (Alshingiti et al., 2023) introduce three deep learning algorithms, CNN, LSTM, and a hybrid LSTM-CNN, to detect phishing websites. Their experiments show that CNN is the most effective model in phishing detection, surpassing the other models with an accuracy of 99.2%. (Adebowale et al., 2023; Adebowale et al., 2023) introduce the Intelligent Phishing Detection System (IPDS), a hybrid model merging CNN and LSTM, which leverages URLs and website content, including visuals and text. The model achieves 93.28% accuracy with an average detection time of 25 seconds, balancing precision and speed. (Wu et al., 2022) provide a BiGRU model modified with an attention mechanism and dropout regularization to prevent overfitting.

This technique performs well in detecting malicious URLs 97.92 and gives substantial insights for practical applications. (Wang & Chen., 2022) create TCURL, a hybrid network that integrates convolution and transformer branches to handle both local and global URL correlations. It gets up to 99.77% accuracy, outperforming previous techniques. (Assefa & Katarya., 2022) offer a neural network-based autoencoder that uses outlier analysis to detect phishing websites, including zero-hour phishing attacks. This technique provides a solution for real-time phishing detection. Zheng et al. (2022) offer an HDP-CNN model that mixes character-level and word-level URL representations with an accuracy of 98.30%. It excels at handling imbalanced datasets and obtaining detailed information from URLs.A multi-feature extraction technique utilizing MLP, CNN, and RNN models is presented by Yu et al. (2022). This method can handle challenging phishing detection tasks, as seen by its 97.75% accuracy rate. A deep autoencoder-based model (ODAE-WPDC) with a 99.28% accuracy is provided by (Alqahtani et al., 2022) using artificial algae and invasive weed optimization for feature selection, suggesting its applicability in phishing detection. (Tang & Mahmoud., 2022) provide a browser plug-in for real-time phishing detection that incorporates RNN-GRU, blacklist interception, and whitelist filtering. The approach obtained 99.18% accuracy with low false positives, indicating its utility.

(Mourtaji et al., 2021) offer a hybrid rule-based approach using numerous machine learning models and deep learning approaches such as CNN and MLP. CNN outperforms other models with a 97.945% accuracy, suggesting the use of such hybrid methods for efficient phishing detection. Using sensitive word segmentation, (Zhang et al., 2021) propose a CNN-BiLSTM hybrid model for phishing detection. The model accurately captures long-distance URL dependencies, obtaining good accuracy and recall. (Ozcan et al., 2021) develop hybrid models utilizing LSTM and DNN, emphasizing NLP-based feature extraction. Their algorithms surpass existing strategies in terms of accuracy, demonstrating promise for phishing URL identification. (Feng & Yue., 2020) demonstrate RNN-based models leveraging URL lexical information to detect phishing attempts with more than 99% accuracy.

Their work also introduces unique visualization tools to evaluate model activity and increase phishing detection accuracy. (Somesha et al., 2020b) focus on three models (DNN, LSTM, CNN) for phishing detection, reaching great accuracy (99.57% for LSTM). Relying on a few third-party services, these architectures are known for their speed and robustness. The character-level CNN paradigm proposed by (Aljofey et al., 2020) eschews third-party services and content retrieval. It yields 98.58% accuracy, making it a speedy and efficient URL-based detection technique. (Wang et al., 2019) propose PDRCNN, a fast phishing detection model using bi-directional LSTM and CNN to identify URLs. Without depending on-page content or other services, it achieves 97% accuracy and excellent processing speed. Table 1 displays the literature summary of the selected papers regarding performance.

Method

In Table 1, a comparative analysis of the deep learning model for website phishing detection is carried out, based on a review of existing research rather than new experiments, seventeen studies were selected from peer-reviewed journals and conference proceedings published between 2019 and 2024, focusing on deep learning models. Searches were conducted in IEEE Xplore, SpringerLink, and Scopus using the keywords "deep learning for website phishing detection," "website phishing detection approaches," and "hybrid deep learning models for website phishing detection." Only papers that provided detailed performance metrics, such as accuracy, precision, recall, F1-score, or detection time, were included. The models were then categorized into Single and Hybrid Models, and their performance was evaluated based on the standard metrics shown in Figure 1.

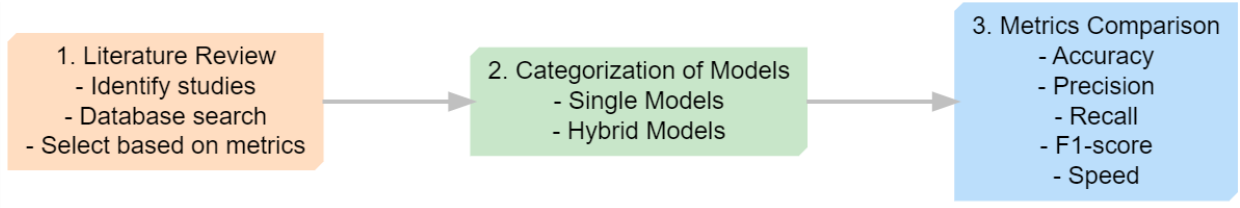

Figure 1. Methodology

Table 1. Summary of the review of the selected studies based on performance

Authors/Year	Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)	Training Time (s)	Test Time (s)
(Alshingiti et al., 2023)	LSTM–CNN	97.6	96.9	98.2	97.6	Nil	Nil	Nil
(Adebowale et al., 2023)	IPDS **(CNN+ LSTM)**	93.28	93.28	93.28	93.28	Nil	Nil	Nil
(Wu et al., 2022)	DA-BiGRU	0.9792	0.9691	0.9834	0.9553	Nil	Nil	Nil
(Wang & Chen, 2022)	TCURL	0.9692	0.9718	0.9664	0.9691	Nil	Nil	Nil
(Assefa & Katarya, 2022)	Autoencod Er	91.24	Nil	Nil	Nil	Nil	Nil	Nil
(Yu et al., 2022)	multi-feature (MLP, CNN & RNN)	97.75	96.65	99.01	97.82	Nil	Nil	Nil
(Zheng et al., 2022)	HDP-CNN	98.30	nil	Nil	94.95	Nil	Nil	Nil
(Alqahtani et al., 2022)	ODAE-WPDC	99.28	99.29	99.24	99.27	Nil	Nil	Nil
(Tang & Mahmoud, 2022)	RNN-GRU	99.18	98.6	Nil	99.15	Nil	Nil	Nil
(Ozcan et al., 2021)	DNN+BiLSTM	98.79	Nil	Nil	98.81	98.78	Nil	Nil
(Mourtaji et al., 2021)	CNN	97.945	Nil	Nil	98.591	Nil	Nil	Nil
(Zhang et al., 2021)	CNN-BiLSTM	98.84	99.71	98.04	98.87	Nil	Nil	Nil
(Somesha et al., 2020b)	CNN	99.43	Nil	Nil	Nil	Nil	Nil	Nil
(Feng & Yue, 2020)	BiGRU	99.5%	97.1%	96.4%	96.8%	Nil	Nil	Nil
(Aljofey et al., 2020)	CNN	98.58	98.55	98.62	98.56	98.58	5281.81	32.70
(W. Wang et al., 2019)	PDRCNN	95.6	97.33	93.78	95.52	98.96	4426.15	40.66

Table 1 shows the performance of the literature across different measures, including accuracy, precision, recall, f1 score, AUC, training, and test time.

Results and Discussion

Result

Assessing the effectiveness of each deep learning model is crucial to comprehending how well it can detect phishing attempts. The result of every single model (CNN and BiGRU) is presented in Tables 2 and 2 and visualized in Figures 2 and 2, respectively, focusing on their detection accuracy, precision, recall, f1 score, and model adopted.

Table 2 Categorization of single CNN models performance

Authors/Year	Model	Accuracy (%)	Precision (%)	Recall	F1-score
Alshingiti et al. (2023)	CNN	99.2	99	99.2	99.2
Mourtaji et al. (2021)	CNN	97.94	Nil	Nil	Nil
Somesha et al. (2020)	CNN	99.43			
Aljofey et al. (2020)	CNN	98.58	98.55	98.62	98.56

Table 2 shows the performance of CNN models from various studies, all showing high accuracy between 97.94% and 99.43%. (Alshingiti et al., 2023; Aljofey et al., 2020) provide comprehensive evaluations with precision, recall, and F1-scores closely matching their accuracy, indicating well-balanced and reliable performance. However, (Mourtaji et al., 2021; Somesha et al., 2020) report only accuracy, making it harder to assess the effectiveness of their models fully. While all models demonstrate substantial accuracy, the detailed results from Alshingiti and Aljofey offer more precise insights into their consistency and robustness. Figure 2 shows the categorization of the performance of single models.

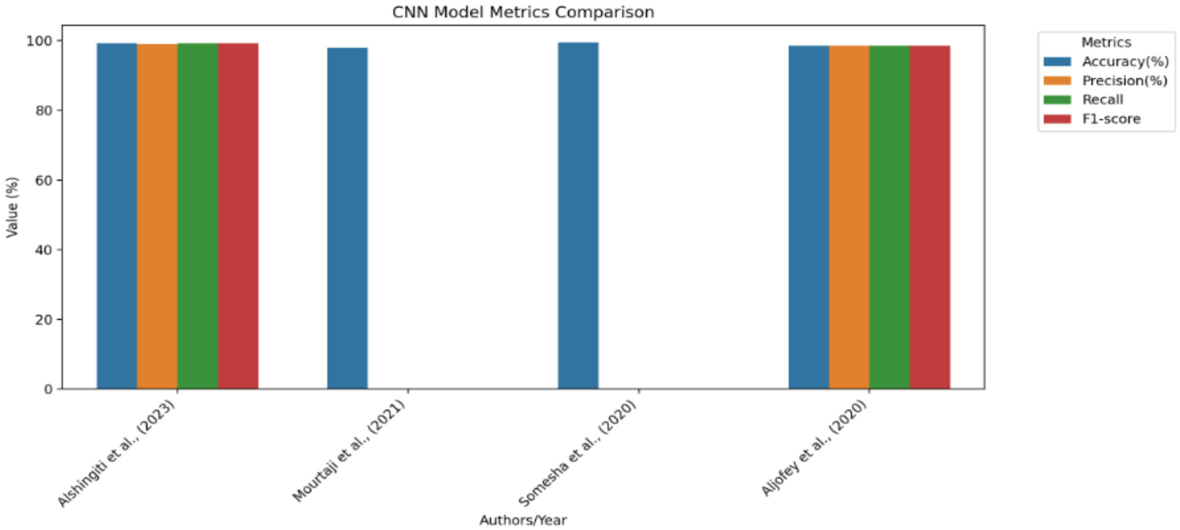

Figure 2. Categorization of single model's performance

Table 3 Categorization of single (BiGRU) models performance

Authors/Year	**Model**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Feng & Yue, (2020)	BiGRU	99.5	97.1	96.4	96.8

Table 3 shows the performance of the BiGRU model by (Feng & Yue., 2020) with an accuracy of 99.5%, indicating it makes nearly flawless predictions. While precision (97.1%) and recall (96.4%) are slightly lower, the model is highly effective at minimizing false positives, though it occasionally misses a few positive cases. The F1-score (96.8%) reflects a strong balance between precision and recall. Despite minor variations in the metrics, the model demonstrates high effectiveness and reliability across all key performance indicators. Figure 3 depicts the categorization of a single (BiGRU) model.

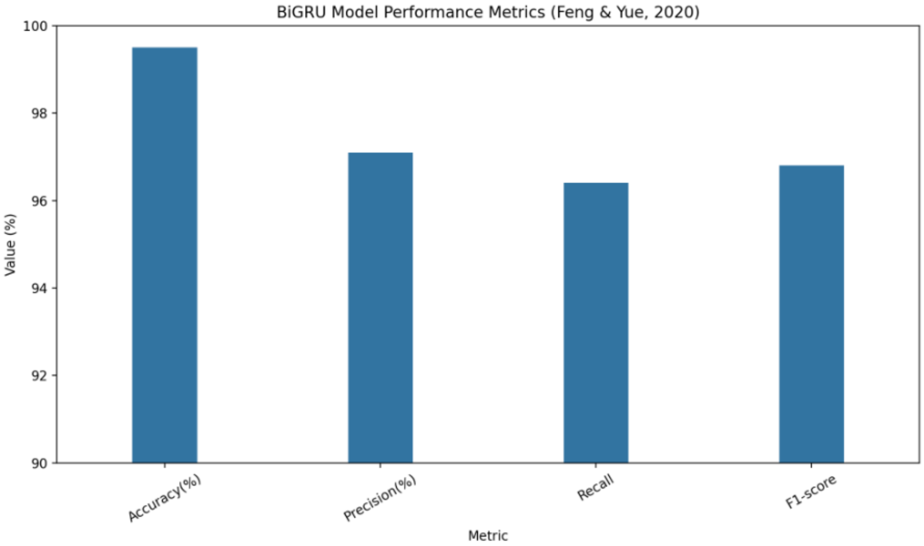

Figure 3. Categorization of single models performance

The research results are presented in full and in accordance with the scope of the study. The results of the research can be completed with tables, graphs (images), and/or charts. Tables and figures are numbered and titled. The results of the data analysis were interpreted correctly.

Table 4. Categorization of hybrid model's performance

Auths/Year	Method	Accuracy(%)	Precision (%)	Recall(%)	F1-Score(%)
(Alshingiti et al., 2023)	LSTM–CNN	97.6	96.9	98.2	97.6
(Adebowale et al., 2023)	IPDS **(CNN+ LSTM)**	93.28	93.28	93.28	93.28
(Wu et al., 2022)	DA-BiGRU	0.9792	0.9691	0.9834	0.9553
(Wang & Chen, 2022)	TCURL	0.9692	0.9718	0.9664	0.9691
(Yu et al., 2022)	multi-feature (MLP, CNN & RNN)	97.75	96.65	99.01	97.82
(Zheng et al., 2022)	HDP-CNN	98.30	nil	Nil	94.95
(Alqahtani et al., 2022)	ODAE-WPDC	99.28	99.29	99.24	99.27
(Tang & Mahmoud, 2022)	RNN-GRU	99.18	98.6	Nil	99.15
(Ozcan et al., 2021)	DNN+BiLSTM	98.79	Nil	Nil	98.81
(Zhang et al., 2021)	CNN-BiLSTM	98.84	99.71	98.04	98.87
(W. Wang et al., 2019)	PDRCNN	95.6	97.33	93.78	95.52

Table 4 displays multiple hybrid models according to accuracy, precision, recall, and F1-score. The ODAE-WPDC model (Alqahtani et al., 2022) obtains the highest accuracy at 99.28%, with comparably good results across other criteria. The RNN-GRU (Tang & Mahmoud., 2022) and CNN-BiLSTM (Zhang et al., 2021) models work as well, topping 99% accuracy. In comparison, models like IPDS (Adebowale et al., 2023) and PDRCNN (W. Wang et al., 2019) show somewhat lower accuracy but retain equal effectiveness across all metrics. Overall, each model has various characteristics, with some excelling in certain areas while others give more uniform performance. Figure 4 displays the Categorization of Hybrid Models Performance.

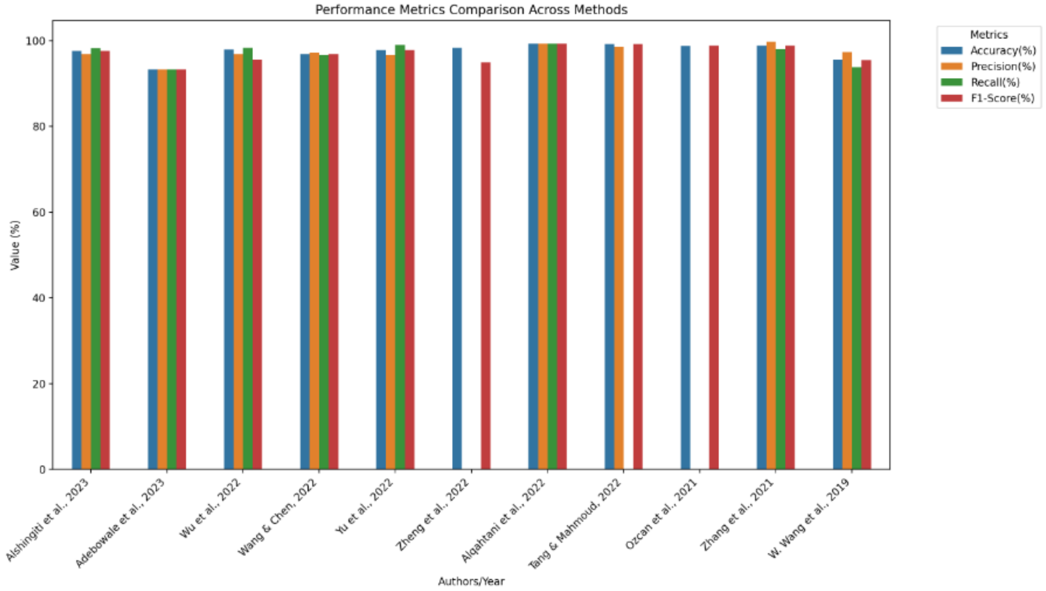

Figure 4. Categorization of hybrid models' performance

Table 5. Comparison of the performance of the single models

Authors/Year	Method	Accuracy(%)	Precision (%)	Recall(%)	F1-Score(%)
Alshingiti et al. (2023)	CNN	99.2	99	99.2	99.2
Mourtaji et al.(2021)	CNN	97.94	Nil	Nil	Nil
Somesha et al. (2020)	CNN	99.43	Nil	Nil	Nil
Aljofey et al. (2020)	CNN	98.58	98.55	98.62	98.56
Feng & Yue (2020)	BiGRU	99.5	97.1	96.4	96.8

Table 5 illustrates and Figure 5 compares the performance of single models for detecting website phishing attacks. Table 5 assesses the performance of single models, primarily CNN and BiGRU, based on accuracy, precision, recall, and F1-score. CNN models consistently deliver strong results, with accuracy ranging from 97.94% to 99.43%. The highest accuracy (99.43%) is reported by Somesha et al. (2020), although precision, recall, and F1-score values are not provided. (Alshingiti et al., 2023) present a well-rounded CNN model with 99.2% accuracy and high precision, recall, and F1-score. Similarly, (Aljofey et al., 2020) show a high-performing CNN with 98.58% accuracy and balanced metrics. On the other hand, the BiGRU model from (Feng & Yue., 2020) achieves slightly higher accuracy (99.5%) but exhibits lower precision, recall, and F1-scores, suggesting that while it excels in accuracy, its performance in other areas is less consistent. Overall, CNN models demonstrate robust, consistent performance, while the BiGRU model stands out in accuracy but is less balanced across different metrics.

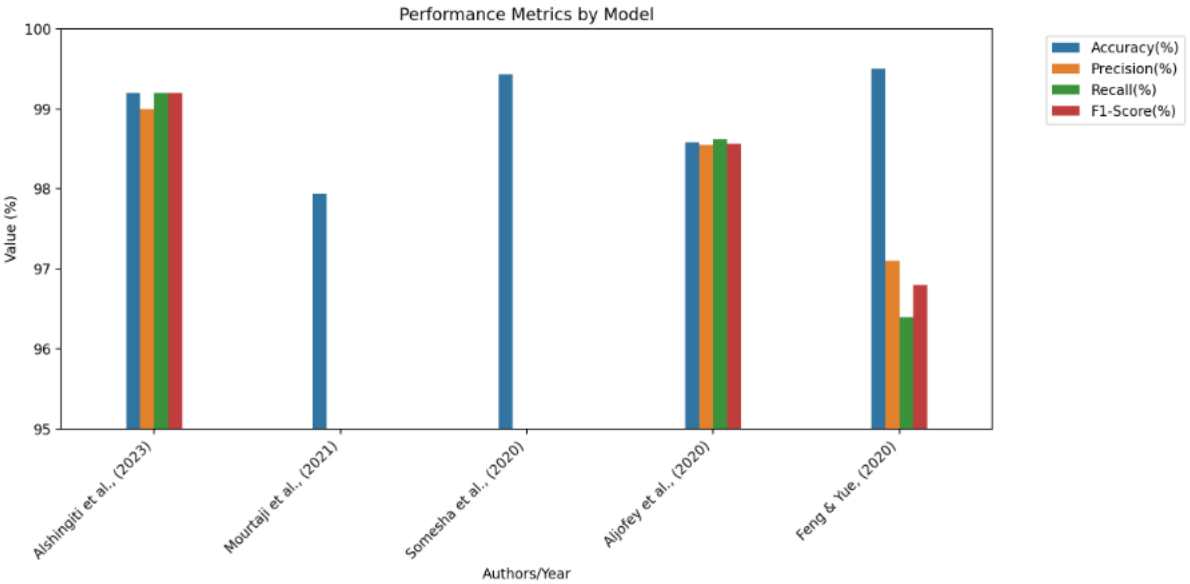

Figure 5. Comparison of the performance of the single models

Table 6. Comparison of the performance of the hybrid models

Authors/Year	Method	Accuracy(%)	Precision (%)	Recall(%)	F1-Score(%)
(Alshingiti et al., 2023)	LSTM–CNN	97.6	96.9	98.2	97.6
(Adebowale et al., 2023)	IPDS **(CNN+ LSTM)**	93.28	93.28	93.28	93.28
(Wu et al., 2022)	DA-BiGRU	0.9792	0.9691	0.9834	0.9553
(Wang & Chen, 2022)	TCURL	0.9692	0.9718	0.9664	0.9691
(Yu et al., 2022)	multi-feature (MLP, CNN & RNN)	97.75	96.65	99.01	97.82
(Zheng et al., 2022)	HDP-CNN	98.30	Nil	Nil	94.95
(Alqahtani et al., 2022)	ODAE-WPDC	99.28	99.29	99.24	99.27
(Tang & Mahmoud, 2022)	RNN-GRU	99.18	98.6	Nil	99.15
(Ozcan et al., 2021)	DNN+BiLSTM	98.79	Nil	Nil	98.81
(Zhang et al., 2021)	CNN-BiLSTM	98.84	99.71	98.04	98.87
(W. Wang et al., 2019)	PDRCNN	95.6	97.33	93.78	95.52

Table 6 illustrates and Figure 6 depicts the comparisons of the performance of single models for the detection of website phishing attacks. Table 6 examines the effectiveness of multiple hybrid models with regard to accuracy, precision, recall, and F1 score. The ODAE-WPDC model (Alqahtani et al., 2022) scores top with 99.28% accuracy and near-perfect precision, recall, and F1-score, suggesting remarkable overall efficiency. Other effective models are RNN-GRU (Tang & Mahmoud., 2022) and CNN-BiLSTM (Zhang et al., 2021), with both having accuracy surpassing 99%, however, some measures are lacking. LSTM-CNN (Alshingiti et al., 2023) and the multi-feature model (Yu et al., 2022) also perform well, with an accuracy of roughly 97.6-97.75%. In comparison, IPDS (Adebowale et al., 2023) and PDRCNN (W. Wang et al., 2019) show lower accuracy, with IPDS maintaining consistent scores across all metrics, while PDRCNN displays more variability. DA-BiGRU (Wu et al., 2022) and TCURL (Wang & Chen., 2022) achieve around 97% accuracy with balanced performance. The top models deliver high accuracy and balanced metrics; others show inconsistencies or missing data.

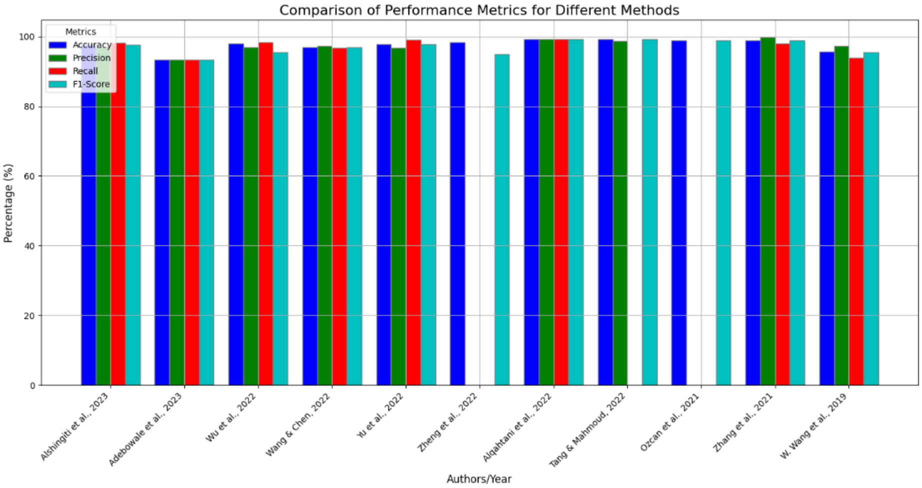

Figure 6. Comparison of the perfirmance of the singel models

Table 7. Comparison of training timeacross models

Authors/Year	Method	Training Time (s)	Test Time (s)
(Aljofey et al., 2020)	CNN	5281.81	32.70
(W. Wang et al., 2019)	PDRCNN	4426.15	40.66

Table 7 compares the training and test times of two models: CNN (Aljofey et al., 2020) and PDRCNN (W. Wang et al., 2019). The CNN model takes longer to train, requiring 5281.81 seconds, while PDRCNN completes training in 4426.15 seconds, making it quicker in that phase. However, CNN compensates with a faster test time of 32.70 seconds, compared to PDRCNN's 40.66 seconds. This indicates that CNN requires more training time and is more efficient during testing, making it potentially more suitable for applications where rapid prediction is essential.

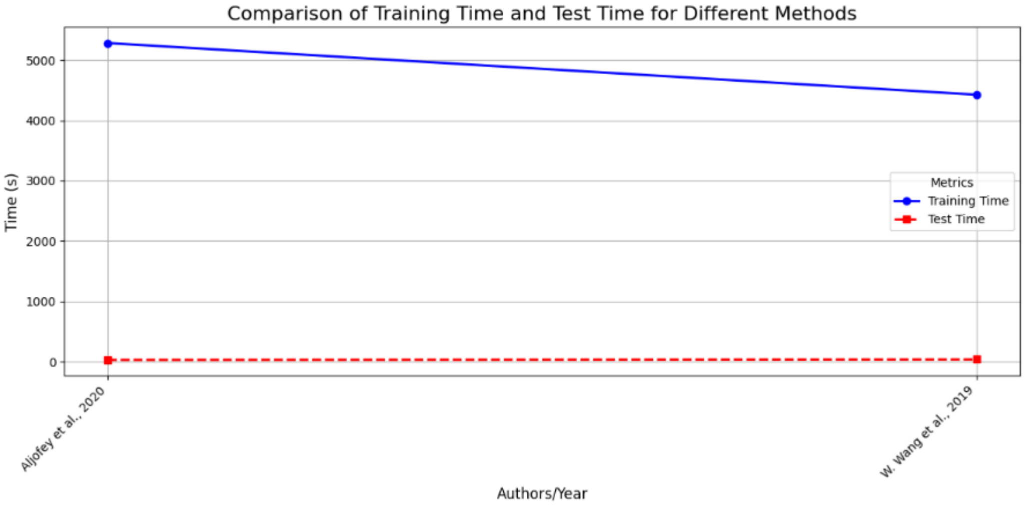

Figure 7. Comparison of the performance of the single models

Discussion

The paper titled "Evaluating Deep Learning Models for Website Phishing Attack Detection" analyzes the usefulness of deep learning (DL) models, including CNN, RNN, LSTM, and hybrid models, in detecting phishing websites. It highlights the versatility of phishing attacks, which undermine conventional detection methods, and proposes DL as a robust option due to its pattern recognition capabilities. Key findings indicate that hybrid models, such as ODAE-WPDC and RNN-GRU, achieve excellent accuracy and balanced performance across metrics like precision and recall. CNN models also provide consistent and reliable results, while BiGRU records the highest accuracy among single models. However, the study also identifies several challenges, including long training times, the need for real-time detection improvements, and issues arising from imbalanced datasets.

The research recommends exploring efficient hybrid architectures, advancing real-time detection capabilities, and implementing explainable AI to enhance transparency and trust. It emphasizes the critical role of DL in advancing phishing detection and cybersecurity. This study aligns with recent research exploring deep learning models for phishing website detection, demonstrating their effectiveness in this domain. A hybrid CNN-Attention-LSTM model achieved 97% accuracy by leveraging URL features and semantic dependencies (Sultana et al., 2023).

Another study combining CNN and LSTM achieved 93.28% accuracy using URL, image, and frame element data (Adebowale et al., 2019). An empirical analysis of various deep learning algorithms, including DNN, CNN, LSTM, and GRU, revealed that no single algorithm excelled across all performance metrics, underscoring the need for tailored solutions (Do et al., 2021). A comparative study of deep learning and machine learning algorithms found that a combined CNN-LSTM model outperformed other approaches with 93.1% accuracy (Tesfom et al., 2023). These findings highlight the potential of deep learning techniques in phishing detection, particularly hybrid models while emphasizing the importance of considering specific application requirements when selecting an approach.

Conclusions and Suggestions

Conclusions

The comparable examination of deep learning models for phishing website detection indicates that though single models such as CNN and BiGRU exhibit excellent accuracy and reliable achievement, hybrid models, specifically ODAE-WPDC, show superior accuracy, precision, and recall efficacy. This study underscores the significance of blending diverse deep learning architectures to tackle the dynamic nature of phishing attempts, giving useful insights for creating more efficient detection systems. The findings also underline the significance of addressing accuracy and computational efficiency, with models like CNN offering higher test speeds despite longer lengthy training durations. Future studies ought to focus on enhancing hybrid models and testing them in real-world circumstances to ensure practical applicability.

Suggestions

The recommendations from this study for future development are as follows:

1. Further develop hybrid models such as ODAE-WPDC by incorporating new deep learning methods or more efficient optimization algorithms.
2. Use larger and more diverse datasets from various languages, regions, or phishing patterns to enhance the model's generalization.
3. Test the integration of the model with existing security systems to evaluate its effectiveness and scalability in complex operational environments.

Contribution: Abdullahi Raji Egigogo: Content development, Formal analysis, investigation, methodology, writing-original draft, review & editing. Prof. Ismaila Idris: Conceptualisation, quality assessment, review & editing, supervision. Prof. Morufu Olalere: Visualisation, quality assessment, review & editing, supervision. Dr. Mrs. Abisoye: Resources, quality assessment, supervision, review & editing.ss

References

Adebowale, M. A., Lwin, K. T., & Hossain, M. A. (2019). Deep Learning with Convolutional Neural Network and Long Short-Term Memory for Phishing Detection. *2019 13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, 1–8. https://api.semanticscholar.org/CorpusID:211058598

Adebowale, M. A., Lwin, K. T., & Hossain, M. A. (2023). Intelligent phishing detection scheme using deep learning algorithms. *Journal of Enterprise Information Management*, *36*(3), 747–766. https://doi.org/https://doi.org/10.1108/JEIM-01-2020-0036

Admin. (2024). *Teknologi Deep Learning: Mendorong Batasan Inovasi di Berbagai Industri*. Pusdasi.Uma.Ac.Id. https://pusdasi.uma.ac.id/teknologi-deep-learning-mendorong-batasan-inovasi-di-berbagai-industri/

Afinda, A. M. (2024). *Neural Network: Cikal Bakal Revolusi Deep Learning*. Www.Dicoding.Com. https://www.dicoding.com/blog/neural-network-cikal-bakal-revolusi-deep-learning/

Alabdan, R. (2020). Phishing attacks survey: Types, vectors, and technical approaches. *Future Internet*, *12*(10), 168. https://doi.org/https://doi.org/10.3390/fi12100168

Ali, M. M., & Mohd Zaharon, N. F. (2024). Phishing A cyber fraud: The types, implications and governance. *International Journal of Educational Reform*, *33*(1), 101–121. https://doi.org/https://doi.org/10.1177/10567879221082966

Aljofey, A., Jiang, Q., Qu, Q., & Huang, M. (2020). *An Effective Phishing Detection Model Based on Character Level Convolutional Neural Network from URL*. https://doi.org/10.3390/electronics9091514

Alkhalil, Z., Hewage, C., Nawaf, L., & Khan, I. (2021). Phishing attacks: A recent comprehensive study and a new anatomy. *Frontiers in Computer Science*, *3*, 563060. https://doi.org/https://doi.org/10.3389/fcomp.2021.563060

Almomani, A., Alauthman, M., Shatnawi, M. T., Alweshah, M., Alrosan, A., Alomoush, W., & Gupta, B. B. (2022). Phishing Website Detection With Semantic Features Based on Machine Learning Classifiers: A Comparative Study. *Int. J. Semantic Web Inf. Syst.*, *18*, 1–24. https://api.semanticscholar.org/CorpusID:246920682

Alqahtani, H., Alotaibi, S. S., Alrayes, F. S., Al-Turaiki, I., Alissa, K. A., Aziz, A. S. A., Maray, M., & Al Duhayyim, M. (2022). Evolutionary Algorithm with Deep Auto Encoder Network Based Website Phishing Detection and Classification. *Applied Sciences (Switzerland)*, *12*(15). https://doi.org/10.3390/app12157441

Alshingiti, Z., Alaqel, R., Al-Muhtadi, J., Haq, Q. E. U., Saleem, K., & Faheem, M. H. (2023). A Deep Learning-Based Phishing Detection System Using CNN, LSTM, and LSTM-CNN. *Electronics (Switzerland)*, *12*(1). https://doi.org/10.3390/electronics12010232

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1). https://doi.org/10.1186/s40537-021-00444-8

Amazon. (2024). *Apa itu Deep Learning?* Aws.Amazon.Com. https://aws.amazon.com/id/what-is/deep-learning/

Assefa, A., & Katarya, R. (2022). Intelligent Phishing Website Detection Using Deep Learning. *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, *1*, 1741–1745. https://doi.org/10.1109/ICACCS54159.2022.9785003

Chaudhary, S. (2012). *Recognition of phishing attacks utilizing anomalies in phishing websites*.

Citra. (2024). *Deep Learning*. Wangs.Id. https://www.wangs.id/literasi-bersama/apa-itu-deep-learning/

Do, N. Q., Selamat, A., Krejcar, O., Herrera-Viedma, E., & Fujita, H. (2022). Deep Learning for Phishing Detection: Taxonomy, Current Challenges and Future Directions. *IEEE Access*, *10*, 36429–36463. https://doi.org/10.1109/ACCESS.2022.3151903

Do, N. Q., Selamat, A., Krejcar, O., Yokoi, T., & Fujita, H. (2021). Phishing Webpage Classification via Deep Learning-Based Algorithms: An Empirical Study. *Applied Sciences*. https://api.semanticscholar.org/CorpusID:244611674

Eka Purwiantono, F., & Tjahyanto, A. (2017). Model Klasifikasi untuk Deteksi Situs Phising di Indonesia. *Institut Teknologi Sepuluh Nompember Surabaya*, 156. https://doi.org/10.13140/RG.2.2.29627.52003

Feng, T., & Yue, C. (2020). Visualizing and interpreting RNN Models in URL-based phishing detection. *Proceedings of ACM Symposium on Access Control Models and Technologies, SACMAT*, 13–24. https://doi.org/10.1145/3381991.3395602

Gupta, B. B., Tewari, A., Jain, A. K., & Agrawal, D. P. (2017). Fighting against phishing attacks: state of the art and future challenges. *Neural Computing and Applications*, *28*, 3629–3654. https://doi.org/https://doi.org/10.1007/s00521-016-2275-y

Hazzataqiy, H. (2024). *Penerapan Deep Learning dalam Pengenalan Pola dan Analisis Data*. Kompasiana.Com. https://www.kompasiana.com/hazzataqiyhiroshi4541/674586d5ed641506835ac272/penerapan-deep-learning-dalam-pengenalan-pola-dan-analisis-data

Hidayanto, A. C., Gamaliel, Y. Y., & Hutagalung, M. (2022). *Pengembangan Deep Learning untuk Mendeteksi Situs Phising dengan Menggunakan Convolutional Neural Network*. Institut Teknologi Harapan Bangsa.

https://repository.ithb.ac.id/id/eprint/56/9/1318013_Paper-TA.pdf

Irawan, A. S. Y., Heryana, N., Hopipah, H. S., & Rahma, D. (2021). Identifikasi Website Phishing dengan Perbandingan Algoritma Klasifikasi. *Syntax : Jurnal Informatika*, *10*(01), 57–67. https://doi.org/10.35706/SYJI.V10I01.5292

Iswahyudi, M. S., Irmawati, I., Widians, J. A., Mahendra, G. S., Pratiwi, M., Hayati, N., Pomalingo, S., Miranda, E., Waryono, W., & Yanuarsyah, H. I. (2023). *Aplikasi Machine Learning di Berbagai Bidang: Solusi Cerdas Untuk Masa Depan*. PT. Sonpedia Publishing Indonesia.

James, L. (2005). *Phishing exposed*. Elsevier.

Kara, I., Ok, M., & Ozaday, A. (2022). Characteristics of Understanding URLs and Domain Names Features: The Detection of Phishing Websites With Machine Learning Methods. *IEEE Access*, *10*, 124420–124428. https://api.semanticscholar.org/CorpusID:253660812

Mahmud, A. F., & Wirawan, S. (2024). Deteksi Phishing Website menggunakan Machine Learning Metode Klasifikasi. *Jurnal Sistem Informasi*, *13*, 1368–1380. https://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/viewFile/3456/781

Manoj, P., Bhuvan Kumar, Y., Rakshitha, D., & Megha, G. (2021). Detection and classification of phishing websites. *Trends in Computer Science and Information Technology*, 053–059. https://doi.org/10.17352/tcsit.000040

Mohammad, R. M., Thabtah, F., & McCluskey, L. (2015). Tutorial and critical analysis of phishing websites methods. *Computer Science Review*, *17*, 1–24. https://doi.org/https://doi.org/10.1016/j.cosrev.2015.04.001

Mourtaji, Y., Bouhorma, M., Alghazzawi, D., Aldabbagh, G., & Alghamdi, A. (2021). Hybrid Rule-Based Solution for Phishing URL Detection Using Convolutional Neural Network. *Wireless Communications and Mobile Computing*, *2021*. https://doi.org/10.1155/2021/8241104

Nada, M. (2019). *Penerapan Deep Learning Menggunakan Convolutional Neural Network (CNN)*. Medium.Com. https://medium.com/@mukhlishatunnada02/penerapan-deep-learning-menggunakan-convolutional-neural-network-cnn-d02dc6532f5b

Nursyafitri, G. D. (2023). *Memahami Deep Learning, Bagian Machine Learning*. Dqlab.Id. https://dqlab.id/memahami-deep-learning-bagian-machine-learning

Opara, C. C., Chen, Y., & Bo.wei. (2020). Look Before You Leap: Detecting Phishing Web Pages by Exploiting Raw URL And HTML Characteristics. *Expert Syst. Appl.*, *236*, 121183. https://api.semanticscholar.org/CorpusID:226282035

Ozcan, A., Catal, C., Donmez, E., & Senturk, B. (2021). A hybrid DNN–LSTM model for detecting phishing URLs. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-021-06401-z

Pilo, R. (2023). *Deep Learning: Model AI yang Terinspirasi dari Otak Manusia*. Phintraco.Com. https://phintraco.com/deep-learning/

Puskomedia. (2024). *Machine Learning dan Deep Learning: Menambah Kehebatan Komputasi dan Analisis Data*. Www.Puskomedia.Id. https://www.puskomedia.id/blog/machine-learning-dan-deep-learning-menambah-kehebatan-komputasi-dan-analisis-data/

Santoso, J. T. (2023). Teknologi Keamanan Siber (Cyber Security). *Penerbit Yayasan Prima Agus Teknik*, 1–173.

Somesha, M., Pais, A. R., Rao, R. S., & Rathour, V. S. (2020a). Efficient deep learning techniques for the detection of phishing websites. *Sadhana - Academy Proceedings in Engineering Sciences*, *45*(1). https://doi.org/10.1007/s12046-020-01392-4

Somesha, M., Pais, A. R., Rao, R. S., & Rathour, V. S. (2020b). *Efficient deep learning techniques for the detection of phishing websites*. *45*(1).

Sultana, R., Rahman, M. A., & Khan, M. I. (2023). Hybrid Model Based Phishing Websites Detection Using Deep Learning Technique. *2023 26th International Conference on Computer and Information Technology (ICCIT)*, 1–6. https://api.semanticscholar.org/CorpusID:268044677

Tang, L., & Mahmoud, Q. H. (2022). A Deep Learning-Based Framework for Phishing Website Detection. *IEEE Access*, *10*, 1509–1521. https://doi.org/10.1109/ACCESS.2021.3137636

Tesfom, B., Belay, F., Daniel, S., Salem, R., & Otoum, S. (2023). Phishing Detection Using Deep Learning and Machine Learning Algorithms: Comparative Analysis. *2023 IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Cloud and Big Data Computing, International Conference on Cyber Science and Tec*, 684–689. https://doi.org/10.1109/DASC/PiCom/CBDCom/Cy59711.2023.10361457

Wahyuni, S., Darnila, E., Gustiana, Z., Prayoga, J., Saffiera, C. A., Eka, M., Fadhilah, C., & others. (2024). *Data*

Science. Serasi Media Teknologi.

Wang, C., & Chen, Y. (2022). TCURL: Exploring hybrid transformer and convolutional neural network on phishing URL detection. *Knowledge-Based Systems*, *258*. https://doi.org/10.1016/j.knosys.2022.109955

Wang, W., Zhang, F., Luo, X., & Zhang, S. (2019). PDRCNN: Precise Phishing Detection with Recurrent Convolutional Neural Networks. *Security and Communication Networks*, *2019*. https://doi.org/10.1155/2019/2595794

Wu, T., Wang, M., Xi, Y., & Zhao, Z. (2022). Malicious URL Detection Model Based on Bidirectional Gated Recurrent Unit and Attention Mechanism. *Applied Sciences (Switzerland)*, *12*(23). https://doi.org/10.3390/app122312367

Yu, S., An, C., Yu, T., Zhao, Z., Li, T., & Wang, J. (2022). Phishing Detection Based on Multi-Feature Neural Network. *2022 IEEE International Performance, Computing, and Communications Conference (IPCCC)*, 73–79. https://doi.org/10.1109/IPCCC55026.2022.9894337

Zhang, Q., Bu, Y., Chen, B., Zhang, S., & Lu, X. (2021). Research on phishing webpage detection technology based on CNN-BiLSTM algorithm. *Journal of Physics: Conference Series*, *1738*(1). https://doi.org/10.1088/1742-6596/1738/1/012131

Zheng, F., Yan, Q., Leung, V. C. M., Yu, F. R., & Ming, Z. (2022). HDP-CNN: Highway deep pyramid convolution neural network combining word-level and character-level representations for phishing website detection. *COMPUTERS & SECURITY*, *114*. https://doi.org/10.1016/j.cose.2021.102584