JIKO (Jurnal Informatika dan Komputer) Accredited KEMENDIKTI SAINSTEK, No.0173/C3/DT.05.00/2025
Vol. 8, No. 2, August 2025, pp. 79-84 p-ISSN: 2614-8897
DOI: 10.33387/jiko.v8i2.9811 e-ISSN: 2656-1948

COMBINATION OF MULTI-VIEW LEARNING AND DEEP REINFORCEMENT LEARNING TO IMPROVE WEBSITE PHISING DETECTION

Muhamad Hasbia[1*], Wina Witanti[2], Gunawan Abdillah[3]

[1,2,3] Informatics, Faculty of Science and Informatics, Jendral Achmad Yani University, Kota Cimahi, Indonesia
Email: *[1]mhasbia21@if.unjani.ac.id, [2]witanti@gmail.com , [3]Gunawanabdillah03@gmail.com

(Received: 06 May 2025, Revised: 22 May 2025, Accepted: 10 June 2025)

Abstract

Phishing is one of the most common and dangerous forms of cyberattacks, where perpetrators attempt to obtain sensitive information by masquerading as trustworthy entities. Traditional detection methods often fail to anticipate new attacks due to the dynamic nature of phishing. This research proposes an adaptive phishing detection system that combines Multi-Kernel Learning (MKL) and Deep Q-Network (DQN) approaches. MKL is utilized to integrate features from URL structure, domain metadata, and webpage content into a rich multi-view representation, while DQN enhances the model's adaptability through a reward-based learning mechanism. This combination was chosen because MKL effectively captures feature variations from different sources, while DQN excels at handling rapidly changing attack patterns. The dataset consists of 11,056 entries with 32 features, divided in an 80:20 ratio for training and testing. Moreover, evaluation is performed using a 5-Fold Cross Validation method to ensure result stability, and hyperparameter exploration is conducted to obtain the best configuration. Evaluation results show that the system achieves an accuracy of 96.34%, precision of 95.8%, recall of 97.85%, F1-score of 96.73%, and AUC of 0.98. These results demonstrate that the MKL-DQN approach is highly effective in accurately and adaptively detecting phishing.

Keywords: *Multi-Kernel Learning, Deep Q-Network, Phishing Detection, Reinforcement Learning, Multi-View Learning.*

**Corresponding Author: Muhammad Hasbia*

1. INTRODUCTION

Over the last twenty years, the rapid growth of information and communication technologies has significantly impacted various domains of daily life, including digital communication, online financial transactions, and personal data management. The internet now serves as the central infrastructure for global socio-economic operations. However, this technological advancement has been accompanied by a surge in cybersecurity threats among which phishing remains one of the most commonly exploited techniques.

Phishing is a type of social engineering attack that manipulates users into revealing confidential information via emails, counterfeit websites, or messaging platforms. According to the Anti-Phishing Working Group (APWG), over 1.3 million phishing attempts were reported in the first quarter of 2023 [1], illustrating the growing frequency and sophistication of such attacks. Phishing continues to thrive due to its cost-effectiveness, simplicity in execution, and high success rate in deceiving users.

Conventional detection methods such as blacklists and signature-based approaches have notable shortcomings. Blacklists are limited to identifying previously known malicious domains and are ineffective against newly launched phishing websites [2]. Similarly, signature-based techniques often fail to detect modified phishing schemes that maintain malicious intent while altering content patterns. These reactive strategies contribute to detection delays and system vulnerability.

In response to these challenges, researchers have increasingly turned to machine learning (ML) and deep learning (DL) techniques for automated phishing detection. These techniques enable systems to analyze past data and identify phishing characteristics without relying on explicitly programmed rules [3]. Multi-Kernel Learning (MKL), in particular, has shown promise in enhancing classification accuracy by combining diverse feature sets extracted from different

website components. Supporting this, Firmansyah and Setiawan (2023) demonstrated that hybrid deep learning models offer substantial improvements in phishing URL detection performance over traditional methods [16].

Moreover, the Multi-View Learning (MVL) framework further improves detection accuracy by integrating multiple feature perspectives, such as URL characteristics, domain metadata, and page content [4]. Each view contributes distinct information that helps differentiate phishing websites from legitimate ones, enhancing the model's overall predictive power [5].

Despite these advancements, many existing detection models still rely on static learning, which restricts their ability to adapt to emerging attack strategies. To overcome this limitation, researchers have started adopting Deep Reinforcement Learning (DRL) methods. For example, Ridho et al. (2024) found that the Deep Q-Network (DQN), a type of DRL algorithm, increases system robustness by continuously adjusting decisions through experience-based rewards [6].

In parallel, the work of Al Ghifari et al. (2022) reaffirmed the importance of URL-based features in ML-driven phishing detection systems, even as phishing tactics evolve [7]. Motivated by these insights, this research proposes a combined approach that leverages the multi-perspective representation capability of MKL and the adaptive learning strength of DQN to address the complex and evolving nature of phishing threats.

This study, therefore, aims to design and develop a phishing detection model that utilizes MKL for integrating heterogeneous feature views namely, URL structure, domain metadata, and webpage content into a unified representation. The DQN component subsequently enhances the model's adaptability to changes in attack patterns by applying reinforcement learning. The expected outcome is a system with superior accuracy, improved adaptability, and lower rates of false positives and false negatives..

Ultimately, this research contributes to the advancement of adaptive phishing detection methodologies by combining MKL and DQN to enhance accuracy, resilience, and generalizability in dynamic threat environments.

2. RESEARCH METHOD.

This research falls under the category of applied studies, aiming to design and implement a practical solution for detecting phishing websites in an adaptive manner. The approach integrates Multi-View Learning via Multi-Kernel Learning (MKL) and Deep Reinforcement Learning with the Deep Q-Network (DQN) algorithm to develop a detection system capable of accurately identifying phishing threats while adapting to their evolving nature.

A quantitative experimental framework is adopted to assess the system's performance. Several key evaluation metrics are employed, including accuracy, precision, recall, F1-score, and Area Under the Curve (AUC), calculated using a distinct test set. This evaluation aims to determine the comparative advantage of the MKL-DQN approach over conventional phishing detection methods. detection compared to traditional approaches.

The research process is divided into several core phases: data acquisition, data preprocessing, model construction, training and testing, and performance analysis using classification-based statistical methods.

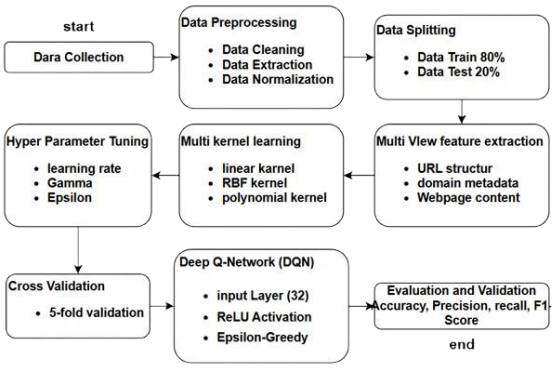

Figure 1. Research Flow

2.1 Data collection

The dataset used was sourced from credible repositories such as PhishTank, OpenPhish, and verified prior studies. It contains 11,056 records, each described by 32 features reflecting different website characteristics such as URL length, HTTPS usage, presence of special characters, whether an IP address appears in the URL, domain age, and HTML content features. The classification label assigned is 1 for legitimate sites and -1 for phishing.

2.2 Data Preprocessing

To prepare the dataset for model training, a series of preprocessing steps were performed. These include removing duplicate records, correcting incomplete data, and addressing missing values by imputing the mean for numerical attributes and the mode for categorical ones. To avoid dominance by larger numerical ranges, Min-Max Scaling is applied, transforming all numeric values into the [0, 1] interval. Categorical features are transformed into numerical format using encoding techniques. Finally, the dataset is partitioned into training (80%) and testing (20%) subsets to evaluate generalization on unseen data.

2.3 Multi-View Feature Extraction

The Multi-View Learning strategy enables the model to harness heterogeneous data sources. In this study, feature extraction is divided into three distinct "views." The first view focuses on URL-based indicators such as length, suspicious character counts, and HTTPS usage. The second view relates to domain metadata, including domain registration age, SSL certificate presence, and registrant details. The third view captures page-level features like the number of input forms, hidden scripts, and manipulated HTML

elements. Each view is preserved as an independent feature set to maintain its unique information. These are later merged using the Multi-Kernel Learning method, allowing different kernels to be combined to yield a more expressive feature representation.

2.4 Implementation of Multi-Kernel Learning (MKL)

Each view is transformed into a specific kernel to capture its data distribution. A linear kernel is used for URL-based features, a Radial Basis Function (RBF) kernel for metadata, and a polynomial kernel for page content. These kernels are fused using a convex optimization technique that dynamically assigns weights to generate an optimal composite kernel K(x,x′)K(x, x')K(x,x′)[9]. This enriched feature space enhances the model's ability to learn complex phishing patterns, leading to better classification outcomes.

2.5 Implementation of Deep Q-Network (DQN)

The kernel fusion output from MKL serves as the input state for the Deep Q-Network model. The DQN architecture includes an input layer of 32 neurons, followed by two hidden layers with 64 and 32 neurons, respectively, each activated using ReLU, and an output layer consisting of two neurons that represent the legitimate and phishing classes. The training process follows the Q-learning paradigm, rewarding correct predictions and penalizing incorrect ones. To manage the exploration-exploitation trade-off, an epsilon-greedy approach is applied.

2.6 MKL-DQN System Integration

The combined MKL-DQN framework leverages the feature-rich representation from MKL and the adaptive decision-making capability of DQN. Once MKL generates the feature embeddings, they are fed directly into the DQN training pipeline. This cohesive integration results in a phishing detection system that adapts better to novel attack strategies and demonstrates improved performance metrics compared to static models.

2.7 Hyperparameter Tuning

To fine-tune model behavior, hyperparameters such as the learning rate, discount factor (gamma), and exploration rate (epsilon) are systematically adjusted. Learning rates tested include 0.001, 0.01, and 0.1. Gamma is explored across values of 0.9, 0.95, and 0.99, while epsilon is gradually decayed from 1.0 to 0.1. The configuration yielding the best empirical results is selected to optimize learning stability and classification performance.

2.8 Validation with Cross Validation

The model's ability to generalize is validated using 5-Fold Cross-Validation. The dataset is partitioned into five segments, where each fold takes a turn as the test set while the remaining four serve as training data. This rotation ensures each data point is tested exactly once. Accuracy for each fold is calculated as:

$$Fold\ accuracy = \frac{1}{5}\sum_{i=1}^{5} \frac{TP_i+TN_i}{TP_i+TN_i+FP_i+FN_i} \quad (1)$$

where TP represents True Positives, TN represents True Negatives, FP represents False Positives, and FN represents False Negatives. The final accuracy is the average of the accuracies across all folds, ensuring that the model performs consistently on different subsets of the data.

2.9 Performance Evaluation

The effectiveness of the phishing detection model is measured using common binary classification metrics: accuracy, precision, recall, F1-score, and AUC (Area Under the ROC Curve)[11]. Accuracy measures the overall percentage of correct predictions, while precision evaluates the accuracy of predictions against phishing sites. Recall measures the model's ability to capture all existing phishing cases, and F1-score is used to balance precision and recall, especially in imbalanced datasets. AUC is used to assess the model's ability to distinguish between phishing and legitimate classes at various prediction thresholds. This evaluation is performed on test data that was not used in the training process, to ensure that the model is able to generalize well to new data.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (2)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (3)$$

$$Recall = \frac{\mathrm{TP}}{(\mathrm{TP}+\mathrm{FN})} \quad (4)$$

$$F1-Score = 2 \times \frac{(\mathrm{Precision} \times \mathrm{Recoll})}{(\mathrm{Precision}+\mathrm{Recall})} \quad (5)$$

The AUC is derived from the Receiver Operating Characteristic (ROC) curve, which assesses the model's ability to distinguish between phishing and legitimate websites at various thresholds. These metrics collectively ensure that the model is both accurate and reliable in detecting phishing attempts.

3. RESULT AND DISCUSSION

3.1 System Implementation

The phishing detection system was developed using Python, incorporating libraries such as Pandas and NumPy for data preprocessing and manipulation, and Scikit-learn for feature extraction and kernel construction. Each feature view underwent a separate kernel transformation as part of the Multi-Kernel Learning (MKL) phase before being combined. The Deep Q-Network (DQN) component was built using TensorFlow and Keras, and results were visualized with Matplotlib and Seaborn.

The architecture of the DQN model comprises an input layer with 32 neurons each representing a feature followed by two hidden layers with 64 and 32 neurons, respectively, both activated using the ReLU function.

The output layer contains two neurons corresponding to the phishing and legitimate classes. The model was trained over 100 epochs using a batch size of 32. An epsilon-greedy strategy was applied to balance the exploration of new actions with the exploitation of learned patterns.

To provide deeper insights into the training dynamics, the key hyperparameters are detailed as follows: the learning rate was fixed at 0.001 to ensure a stable and efficient convergence process. The discount factor (gamma) was set to 0.95, prioritizing long-term reward accumulation over immediate feedback. The exploration rate (epsilon) was linearly decayed from 1.0 to 0.1 across epochs, allowing the model to start with high exploration and progressively focus on exploiting the best-known actions. These parameters collectively ensured the model's robust learning behavior.

3.2 Experimental Result

The experimental evaluation was aimed at measuring the impact of combining MKL and DQN in detecting phishing threats. Initial preprocessing steps included data cleaning (e.g., removing duplicates), handling missing values, and applying Min-Max normalization to ensure consistent feature scaling and stability during model training.

The dataset comprised 11,056 samples with an even distribution between phishing and legitimate labels, divided into training (80%) and testing (20%) subsets to avoid overfitting and enable unbiased generalization assessment.

Feature views were transformed using three distinct kernel types: linear for URL-based features, RBF for metadata, and polynomial for content-related attributes. These kernels were integrated using convex optimization, resulting in a more expressive and diverse feature representation.

The combined kernel outputs were used to train the DQN model over 100 epochs. The training loss curve (see Figure 2) showed a steady decline, indicating successful convergence. Minor oscillations in the loss were observed, which are typical in reinforcement learning contexts.

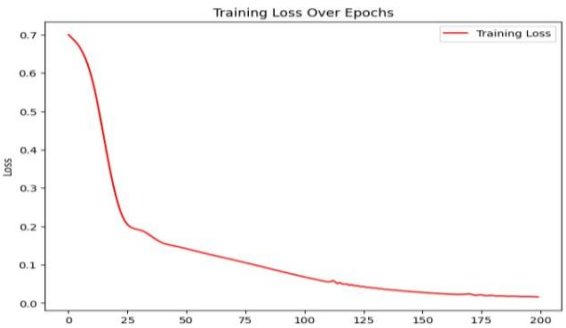

Figure 2. Training Loss Graph per Epoch

Figure 2 shows the loss curve over 100 epochs. The graph indicates a gradual decrease in loss with slight fluctuations, which is typical in reinforcement learning training. The overall downward trend suggests that the model effectively avoided overfitting and learned the underlying patterns in the data.After training, model performance was evaluated on the test set using a confusion matrix and standard classification metrics.

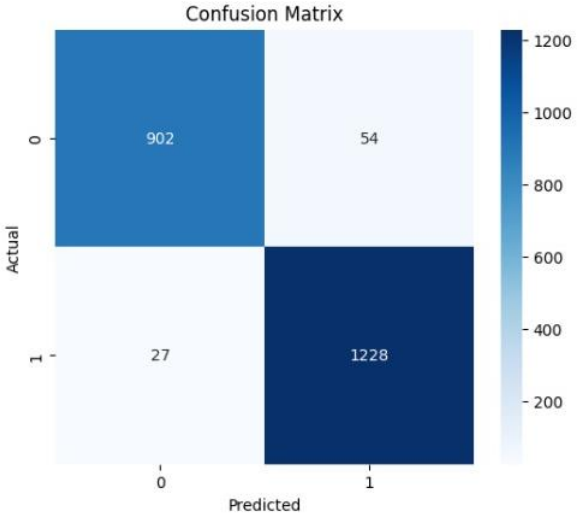

Figure 3. Confusion Matrix

According to the matrix, the model correctly identified 902 legitimate instances (True Negatives) and 1,228 phishing instances (True Positives). It misclassified 54 legitimate entries as phishing (False Positives) and 27 phishing entries as legitimate (False Negatives). These results indicate strong classification performance, with a low rate of misclassification and a balanced distribution of errors.

Performance was further assessed using standard binary classification metrics: accuracy, precision, recall, F1-score, and AUC. The estimated AUC, based on the high and balanced precision and recall values, is approximately 0.98.

Table 1 Evaluation Matrix Results

Metrik Evaluasi	Nilai
Akurasi	96,34%
Presisi	95,8%
Recall	97,85%
F1-Score	96,73%
AUC	0,98

These metrics suggest that the system is highly accurate in distinguishing between phishing and legitimate websites. The high precision shows minimal false positives, while the strong recall indicates nearly all phishing sites were correctly detected. The high F1-score confirms a well-balanced performance, and the AUC close to 1.0 suggests excellent discriminatory capability.

Table 2 Comparison of Studies and Previous Research

Researcher	Method	Accuracy	Recall	Precision	F1-Score	AUC
Fauzan et al. (2021)	SVM	91,7%	90,2%	89,5%	90,9%	0.93
Al Ghifari et al. (2022)	RF	93,8%	92,1%	91,5%	92,9%	0,95
Fauzan et al. (2021)	MKL	94,1%	95%	93,5%	94,5%	0,96
Proposed Method	MKL + DQN	96,34%	97,85%	95,5%	96,73%	0,98

3.3 Comparison of Studies and Previous Research

To highlight the advantages of the proposed MKL-DQN model, its performance is compared with other commonly used methods in phishing detection. Table 2 presents a comparative analysis of accuracy, recall, precision, F1-score, and AUC, along with relevant references.

The proposed MKL-DQN method demonstrates superior performance across all evaluated metrics. It outperforms traditional models like SVM and RF, which lack adaptive mechanisms. While MKL alone improves feature representation, integrating it with DQN introduces adaptability, making the system more resilient to new attack patterns.

3.4 Discussion

The experimental outcomes clearly show that the hybrid MKL-DQN model performs exceptionally well in identifying phishing threats. Achieving 96.34% accuracy suggests the system can reliably make correct predictions, outperforming several existing machine learning approaches.

The precision value (95.8%) is particularly important, as it reflects the model's ability to minimize false positives critical in maintaining user trust in real-world deployments. Similarly, the recall of 97.85% demonstrates high sensitivity, capturing nearly all phishing attempts. This surpasses the results of earlier MKL-based studies such as Fauzan et al. (2021), which reported a recall of approximately 95% [12].

With an F1-score of 96.73%, the model exhibits a strong balance between precision and recall. This finding aligns with Tukino & Fifi (2024), who emphasized the effectiveness of multi-source feature integration in enhancing classification consistency. Moreover, the AUC value of 0.98 supports results by Lestari (2022), who showed that DQN-based systems are well-suited to dynamic cybersecurity environments [13], [14].

In terms of method contribution, the integration of MKL and DQN is proven to provide complementary advantages: (1) MKL enriches the feature representation by combining kernels from various information sources (URL structure, domain metadata, and page content); (2) DQN provides an adaptive mechanism to continuously improve classification decisions based on feedback (rewards) from previous prediction results.

Compared to conventional supervised learning models, this integrated framework is more robust when facing the evolving nature of phishing techniques. This is consistent with findings by Pratama (2024), who explored DRL-based approaches in dynamic domains [15].

Nonetheless, there are challenges. The model's complexity can lead to longer training durations and sensitivity to hyperparameter choices, which may impact performance stability. Future enhancements could involve using automated hyperparameter tuning and lightweight architectures to improve computational efficiency.

In conclusion, this study demonstrates that combining MKL's ability to extract rich feature representations with the adaptability of DQN creates a robust and accurate phishing detection system. The proposed approach shows strong potential for wider application in dynamic cybersecurity scenarios.

4. CONCLUSION

This study has successfully introduced an adaptive phishing detection system by leveraging the strengths of Multi-Kernel Learning (MKL) and Deep Q-Network (DQN). Through the fusion of multi-view feature representations from multiple data sources and the adaptive learning capabilities offered by reinforcement-based strategies, the model delivered strong classification performance achieving 96.34% accuracy, 95.8% precision, 97.85% recall, an F1-score of 96.73%, and an AUC value of 0.98.

These findings indicate that the MKL-DQN hybrid framework effectively overcomes the constraints of traditional phishing detection techniques, especially in addressing the shifting and increasingly sophisticated nature of phishing attacks. MKL enhances the learning process by combining diverse feature perspectives, while DQN provides adaptability by continuously refining predictions based on experience-driven feedback.

When compared to earlier methods, the proposed approach consistently yields superior outcomes, particularly in terms of sensitivity and predictive accuracy. Nonetheless, the model's structural complexity and its reliance on finely tuned hyperparameters present certain limitations, suggesting areas for future improvement particularly in optimizing scalability and computational efficiency.

In summary, the combination of MKL and DQN offers a powerful and reliable approach for strengthening phishing detection capabilities. With its high adaptability and robust performance, this method holds great promise for deployment across a wide range of cybersecurity applications.

5. REFERENCE

[1] A. P. Working Group, "Phishing Activity Trends Report," 2023. [Online]. Available: https://apwg.org/trendsreports/

[2] M. S. F. Purwani, "Analisis peran dan

penanggulangan kejahatan siber: Studi kasus spearphishing," *Restor. J. Indones. Probat. Parol. Syst*., vol. 1, no. 1, pp. 33–45, 2023, doi: 10.59653/restor.v1i1.56.

[3] I. H. R. Hatta *et al.*, *Kecerdasan Buatan*. Cendikia Mulia Mandiri, 2024.

[4] C. Xu, J. Si, Z. Guan, W. Zhao, Y. Wu, and X. Gao, "Reliable Conflictive Multi-View Learning," *Proc. AAAI Conf. Artif. Intell.*, vol. 38, no. 14, pp. 16129–16137, 2024, doi: 10.1609/aaai.v38i14.29546.

[5] Y. Wang, W. Ma, H. Xu, Y. Liu, and P. Yin, "A lightweight multi-view learning approach for phishing attack detection using transformer with mixture of experts," *Appl. Sci.*, vol. 13, no. 13, p. 7429, 2023, doi: 10.3390/app13137429.

[6] M. R. Ridho, N. Fajrah, and F. Fifi, "Literatur review: Penerapan deep reinforcement learning dalam business intelligence," *J. Desain Dan Anal. Teknol*., vol. 3, no. 2, pp. 96–103, 2024, doi: 10.58291/jdat.v3i2.379.

[7] M. A. G. Al Ghifari, B. Hananto, and B. T. Wahyono, "Implementasi ekstensi Google Chrome dalam mendeteksi situs web phishing menggunakan algoritma Random Forest," *in Proc. Semnas Mahasiswa Bidang Ilmu Komputer dan Aplikasinya,* 2022, pp. 640–649.

[8] R. Rahmadani, A. Rahim, and R. Rudiman, "Analisis Sentimen Ulasan 'Ojol The Game' Di Google Play Store Menggunakan Algoritma Naive Bayes Dan Model Ekstraksi Fitur Tf-Idf Untuk Meningkatkan Kualitas Game," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024.

[9] E. N. Yudistira and S. Kom, *Deep Learning: Teori, Contoh Perhitungan, dan Implementasi*. Deepublish, 2024.

[10] R. Y. Putra and F. T. E. D. I. Cerdas, "Perencanaan Gerakan pada Mobil Otonom di Jalan Raya Menggunakan Quantile Regression Deep Q Network," Institut Teknologi Sepuluh Nopember, 2021.

[11] Z. A. Dwiyanti and C. Prianto, "Prediksi cuaca kota Jakarta menggunakan metode Random Forest," *J. Tekno Insentif*, vol. 17, no. 2, pp. 127–137, 2023, doi: 10.36787/jti.v17i2.1201.

[12] R. Fauzan, A. V. Vitianingsih, D. Cahyono, A. L. Maukar, and Y. A. B. Suprio, "Penerapan algoritma klasifikasi pada machine learning untuk deteksi phishing: Application of classification algorithms in machine learning for phishing detection," *MALCOM Indonesia. J. Mach. Learn. Comput. Sci*., vol. 5, no. 2, pp. 531–540, 2025, doi: 10.33050/malcom.v5i2.4126.

[13] T. Tukino and F. Fifi, "Penerapan Support Vector Machine untuk analisis sentimen pada layanan ojek online,"*J. Desain Dan Anal. Teknol.*, vol. 3, no.2, pp. 104–113, 2024, doi: 10.58291/jdat.v3i2.380.

[14] W. S. Lestari, "Deteksi serangan DDoS menggunakan Q-learning," *JATISI J. Tek. Inform. dan Sist. Informasi*, vol. 9, no. 1, pp. 648–658, 2022, doi: 10.35957/jatisi.v9i1.772..

[15] D. A. S. Pratama, "Pengembangan kontrol adaptif untuk kendaraan otonom dengan studi kasus pada mobil elektrik berbasis deep reinforcement learning," *M.S. thesis, Dept. Electrical Eng., Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*, 2024.

[16] R. Firmansyah and A. Setiawan, "Hybrid deep learning untuk deteksi phishing URL," *JIKO J. Inform. Dan Komput*., vol. 5, no. 2, pp. 122–131, 2023, doi: 10.33387/jiko.v5i2.6724.