ELKOM (Jurnal Elektronika dan Komputer)
p-ISSN : 1907-0012 (print) e-ISSN : 2714-5417 (online)
Vol. 19, No. 1, Juli 2026
https://journal.stekom.ac.id/index.php/elkom

Perbandingan Metode *Random Forest* dan *Convolutional Neural Network* dalam Deteksi *Website Phishing* pada Lingkungan Universitas XYZ

Elang Prasakti Ghani[1*], Ria Putri Sunaryo[2]
[1] Department of Electrical Engineering, Universitas Diponegoro, Semarang, Indonesia
[2] Department of Electrical Engineering, Universitas Darma Persada, Jakarta Timur, Indonesia
Email author: elangpg01@gmail.com, riaputri949@gmail.com

Article Info

Article history:

Received June 23, 2026
Revised June 30, 2026
Accepted July 7, 2026
Available: July 13, 2026
Published: July 30, 2026

Keywords:

Phishing Detection;
Random Forest;;
Convolutional Neural Network;
Machine Learning;
Engineering

ABSTRACT (10 PT)

Phishing attacks are a cybersecurity threat often used to steal sensitive user information through fake websites that resemble legitimate sites. Therefore, this study aims to analyze and compare the performance of the Random Forest and Convolutional Neural Network (CNN) algorithms in detecting phishing websites based on Uniform Resource Locator (URL) features. The dataset used was obtained from Web Application Firewall (WAF) security logs on the network infrastructure at XYZ University, which record URL access activities on the web system. The data was then processed and labeled into two categories: phishing and legitimate websites. The dataset used in this study consists of 549,346 URL records. The research stages include data exploration (Exploratory Data Analysis / EDA), URL text preprocessing, character-based feature extraction and URL tokenization, and model training using the Random Forest algorithm and a 1D Convolutional Neural Network (1D-CNN) architecture. Model evaluation was conducted using accuracy, precision, recall, and F1-score metrics as well as confusion matrix analysis. The results showed that the Random Forest model achieved an accuracy of 82.69%, while the 1D-CNN model achieved a higher accuracy of 95.94%. Furthermore, the CNN training process demonstrated a steady increase in accuracy and a decrease in loss values in each epoch. Based on these results, it can be concluded that the deep learning approach using CNN outperforms the Random Forest method in detecting URL-based phishing websites.

Corresponding Author:

Elang Pasakti Ghani,
Universitas Diponegoro
Jl. Prof. Soedarto, Tembalang, Semarang
Email: elangpg01@gmail.com

1. INTRODUCTION

Perkembangan teknologi informasi yang pesat memberikan kemudahan dalam berbagai aktivitas digital, namun juga meningkatkan risiko kejahatan siber seperti phishing. Phishing merupakan

serangan yang bertujuan menipu pengguna agar memberikan informasi sensitif melalui email atau website palsu yang menyerupai situs resmi sehingga korban tidak menyadari bahwa informasi yang diberikan digunakan untuk tujuan kejahatan [1]. Serangan ini terus berkembang dengan teknik yang semakin kompleks sehingga diperlukan metode deteksi yang mampu mengidentifikasi website phishing secara otomatis. Salah satu pendekatan yang banyak digunakan adalah metode machine learning yang mampu menganalisis karakteristik URL dan fitur website untuk membedakan situs phishing dan situs yang sah. Penelitian sebelumnya menunjukkan bahwa algoritma klasifikasi seperti Decision Tree maupun Random Forest dapat digunakan untuk mendeteksi website phishing dengan tingkat akurasi yang cukup baik dalam proses klasifikasi [2].

Pendekatan berbasis machine learning dalam mendeteksi serangan siber semakin berkembang karena kemampuannya dalam menganalisis pola data jaringan secara otomatis dan menghasilkan model klasifikasi yang adaptif terhadap berbagai jenis ancaman. Algoritma seperti Random Forest dan Decision Tree banyak digunakan dalam penelitian keamanan siber karena mampu melakukan proses klasifikasi data dengan tingkat akurasi yang tinggi serta mampu menangani dataset yang kompleks. Metode tersebut memanfaatkan karakteristik data seperti struktur URL, fitur halaman web, maupun pola lalu lintas jaringan untuk membedakan aktivitas normal dan aktivitas berbahaya sehingga dapat membantu proses identifikasi serangan secara lebih efektif [3]. Selain itu, penerapan metode machine learning juga memungkinkan sistem deteksi dapat bekerja secara otomatis dalam menganalisis data dalam jumlah besar sehingga proses deteksi ancaman siber dapat dilakukan secara lebih cepat dan efisien [4]. Berbagai penelitian juga menunjukkan bahwa penggunaan teknik klasifikasi berbasis data mampu meningkatkan kemampuan sistem dalam mengenali pola serangan yang terus berkembang pada lingkungan jaringan modern [5].

Meskipun berbagai pendekatan deteksi telah dikembangkan, sistem keamanan jaringan masih menghadapi berbagai permasalahan dalam mendeteksi serangan siber secara efektif. Banyak sistem keamanan masih bergantung pada pendekatan berbasis signature yang hanya mampu mengenali serangan yang telah diketahui sebelumnya sehingga kurang efektif dalam mendeteksi serangan baru atau pola serangan yang telah dimodifikasi. Selain itu, pertumbuhan jumlah pengguna internet serta meningkatnya kompleksitas lalu lintas jaringan menyebabkan volume data yang harus dianalisis menjadi semakin besar sehingga proses identifikasi ancaman secara manual menjadi sulit dilakukan [6]. Kondisi tersebut menyebabkan berbagai serangan seperti phishing, malware, maupun Distributed Denial of Service masih berpotensi lolos dari sistem keamanan apabila tidak didukung oleh metode analisis berbasis data yang lebih cerdas dan adaptif [7]. Oleh karena itu, diperlukan pendekatan deteksi yang mampu menganalisis pola data jaringan secara lebih mendalam agar ancaman siber dapat diidentifikasi secara lebih akurat [8].

Berbagai penelitian sebelumnya telah mengkaji penerapan algoritma machine learning untuk meningkatkan efektivitas deteksi serangan siber. Salah satu pendekatan yang digunakan adalah algoritma Decision Tree yang mampu melakukan klasifikasi data berdasarkan struktur pohon keputusan sehingga dapat digunakan untuk mendeteksi situs phishing dengan memanfaatkan fitur URL maupun karakteristik halaman web [9]. Penelitian lain menunjukkan bahwa algoritma Random Forest mampu memberikan performa klasifikasi yang lebih stabil karena menggunakan pendekatan ensemble learning yang menggabungkan beberapa decision tree dalam proses klasifikasi data [10]. Selain itu, metode klasifikasi berbasis data juga terbukti mampu meningkatkan kemampuan sistem dalam mengenali pola serangan jaringan yang kompleks sehingga dapat digunakan sebagai dasar dalam pengembangan sistem deteksi intrusi yang lebih efektif [11].

Berdasarkan berbagai penelitian yang telah dilakukan, masih terdapat kebutuhan untuk mengembangkan pendekatan deteksi serangan siber yang mampu menganalisis karakteristik data jaringan secara lebih komprehensif. Penelitian ini mengusulkan pendekatan deteksi serangan berbasis machine learning yang memanfaatkan algoritma klasifikasi untuk menganalisis pola lalu lintas jaringan dalam mengidentifikasi aktivitas yang berpotensi sebagai serangan siber. Pendekatan ini diharapkan mampu meningkatkan kemampuan sistem dalam membedakan aktivitas jaringan normal dan aktivitas berbahaya berdasarkan karakteristik data yang dianalisis [12]. Selain itu, metode yang digunakan dalam penelitian ini diharapkan dapat memberikan pendekatan analisis yang lebih sistematis dalam proses identifikasi ancaman sehingga dapat meningkatkan efektivitas sistem keamanan jaringan [13]. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan metode deteksi serangan siber berbasis data yang lebih akurat dan adaptif [14].

Melalui pendekatan yang diusulkan, penelitian ini diharapkan mampu menghasilkan sistem deteksi serangan siber yang lebih efektif dalam mengidentifikasi berbagai jenis ancaman pada jaringan komputer. Implementasi metode machine learning pada penelitian ini diharapkan dapat membantu proses analisis data jaringan secara otomatis sehingga aktivitas mencurigakan dapat dideteksi lebih awal sebelum menimbulkan dampak yang lebih besar terhadap system. Selain itu, sistem yang dihasilkan diharapkan mampu memberikan tingkat akurasi deteksi yang lebih tinggi dalam mengidentifikasi berbagai jenis serangan jaringan yang semakin kompleks [15]. Dengan adanya sistem deteksi yang lebih adaptif, keamanan infrastruktur teknologi informasi diharapkan dapat ditingkatkan sehingga mampu meminimalkan risiko kerugian akibat serangan siber.

2. METHOD

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk menganalisis dan membandingkan kinerja algoritma Random Forest dan Convolutional Neural Network (CNN) dalam mendeteksi website phishing berbasis URL. Data penelitian diperoleh dari log keamanan Web Application Firewall (WAF) pada infrastruktur jaringan Universitas XYZ yang merekam aktivitas akses URL pada sistem web. Dataset yang digunakan berjumlah 549.346 data URL yang kemudian diproses dan dilabeli menjadi dua kategori yaitu phishing dan legitimate website. Proses penelitian meliputi tahapan Exploratory Data Analysis (EDA), preprocessing teks URL, ekstraksi fitur berbasis karakter dan token URL, serta pelatihan model menggunakan algoritma Random Forest dan arsitektur 1D Convolutional Neural Network. Evaluasi kinerja model dilakukan menggunakan metrik accuracy, precision, recall, dan f1-score serta analisis confusion matrix untuk mengukur kemampuan masing-masing metode dalam mengklasifikasikan URL berbahaya dan URL yang sah. Pendekatan ini memungkinkan pengukuran performa model secara objektif berdasarkan hasil perhitungan statistik sehingga dapat diketahui metode yang paling efektif dalam mendeteksi serangan phishing pada lingkungan sistem web [16].

2.1. Tahap Penelitian

Tahapan penelitian pada studi ini dilakukan secara sistematis untuk membangun dan mengevaluasi model deteksi website phishing berbasis URL menggunakan algoritma Random Forest dan Convolutional Neural Network (CNN). Proses penelitian dimulai dari tahap pengumpulan data hingga evaluasi performa model yang dihasilkan. Dataset yang digunakan berasal dari log keamanan Web Application Firewall (WAF) yang mencatat aktivitas akses URL pada sistem web. Data tersebut kemudian diproses melalui beberapa tahapan seperti Exploratory Data Analysis (EDA), preprocessing data, ekstraksi fitur, pelatihan model, serta evaluasi model untuk mengetahui kinerja algoritma yang digunakan dalam mendeteksi website phishing. Alur tahapan penelitian yang dilakukan dalam studi ini ditunjukkan pada Gambar 1.

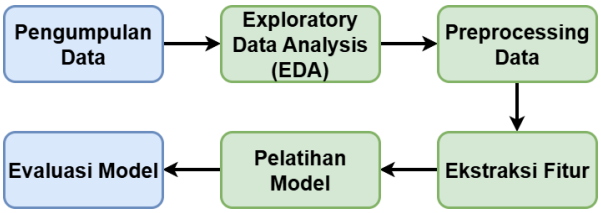

Figure 1. Alur Tahapan Penelitian Deteksi Website Phishing Berbasis URL

2.2. Pengumpulan Data

Pada penelitian ini, data yang digunakan berupa dataset alamat situs web yang digunakan untuk mengidentifikasi website phishing dan website yang sah (legitimate). Dataset diperoleh dari log

aktivitas akses web yang kemudian dikompilasi dalam bentuk dataset klasifikasi phishing. Dataset tersebut berisi 549.346 data yang terdiri dari dua atribut utama yaitu Uniform Resource Locator (URL) dan label. Atribut URL berisi alamat halaman web yang diakses oleh pengguna, sedangkan atribut label menunjukkan kategori dari alamat tersebut, yaitu bad untuk website phishing dan good untuk website legitimate. Dataset ini dipilih karena memiliki jumlah data yang besar sehingga dapat merepresentasikan berbagai pola URL yang berkaitan dengan aktivitas phishing maupun akses web yang normal.

Contoh sebagian data dari dataset yang digunakan dalam penelitian ini ditunjukkan pada Gambar 2. Gambar tersebut menampilkan beberapa sampel data yang terdiri dari atribut *index*, *URL*, dan *label*. Kolom *URL* berisi alamat halaman web yang dianalisis, sedangkan kolom *label* menunjukkan kategori dari alamat web tersebut. Data yang ditampilkan pada gambar hanya merupakan sebagian kecil dari keseluruhan dataset yang digunakan dalam penelitian ini dan ditampilkan sebagai ilustrasi struktur data sebelum dilakukan proses analisis lebih lanjut.

index	URL	Label
0	nobell.it/70ffb52d079109dca5664cce6f317373782/login.SkyPe.com/en/cgi-bin/verification/login/70ffb52d079109dca5664cce6f317373/index.php?cmd=_profile-ach&outdated_page_tmpl=p/gen/failed-to-load&nav=0.5.1&login_access=1322408526	bad
1	www.dghjdgf.com/paypal.co.uk/cycgi-bin/webscrcmd=_home-customer&nav=1/loading.php	bad
2	serviciosbys.com/paypal.cgi.bin.get-into.herf.secure.dispatch35463256rzr321654641dsf654321874/href/href/href/secure/center/update/limit/seccure/4d7a1ff5c55825a2e632a679c2fd5353/	bad
3	mail.printakid.com/www.online.americanexpress.com/index.html	bad
4	thewhiskeydregs.com/wp-content/themes/widescreen/includes/temp/promocoessmiles/?84784787824HDJNDJDSJSHD//2724782784/	bad

Show 25 per page

Figure 2. Contoh Sebagian Dataset Uniform Resource Locator (URL) untuk Deteksi Website Phishing

2.3. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) dilakukan untuk memahami karakteristik awal dataset sebelum dilakukan proses pemodelan. Tahap ini bertujuan untuk mengidentifikasi pola data, distribusi kelas, serta karakteristik URL yang dapat menjadi indikator dalam mendeteksi website phishing. Melalui analisis ini, peneliti dapat memperoleh gambaran awal mengenai perbedaan karakteristik antara URL phishing dan URL legitimate.

Distribusi kelas pada dataset ditunjukkan pada Gambar 3. Berdasarkan grafik tersebut dapat diketahui bahwa jumlah data dengan label good (legitimate) lebih banyak dibandingkan dengan data berlabel bad (phishing). Ketidakseimbangan jumlah data ini menunjukkan bahwa sebagian besar URL dalam dataset merupakan URL legitimate, sedangkan URL phishing memiliki jumlah yang lebih sedikit. Informasi ini penting untuk memahami kondisi dataset sebelum dilakukan proses pelatihan model.

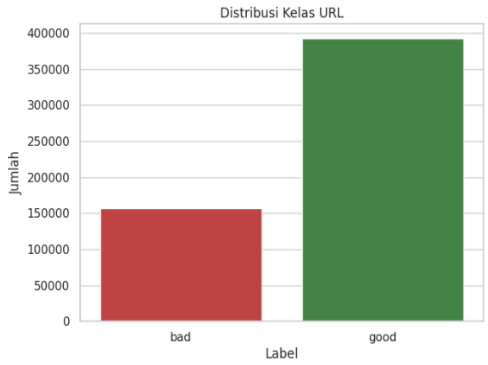

Figure 3. Distribusi Kelas pada Dataset Uniform Resource Locator (URL)

Selanjutnya dilakukan analisis terhadap panjang URL yang ditunjukkan pada Gambar 4. Grafik distribusi tersebut memperlihatkan bahwa sebagian besar URL memiliki panjang antara sekitar 20 hingga 60 karakter. Selain itu terlihat bahwa URL phishing cenderung memiliki variasi panjang yang lebih besar dibandingkan URL legitimate. Hal ini menunjukkan bahwa panjang URL dapat menjadi salah satu indikator dalam membedakan URL phishing dengan URL legitimate.

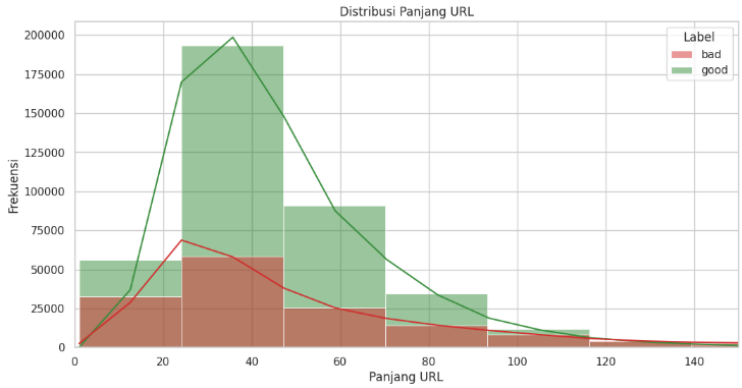

Figure 4. Distribusi Panjang Uniform Resource Locator (URL)

Analisis berikutnya dilakukan terhadap jumlah karakter khusus yang terdapat dalam URL, seperti tanda titik (.), tanda hubung (-), garis miring (/), simbol sama dengan (=), dan simbol persen (%). Hasil analisis yang ditunjukkan pada Gambar 5 menunjukkan bahwa URL *phishing* umumnya memiliki rata-rata jumlah karakter khusus yang lebih tinggi dibandingkan URL *legitimate*. Karakter khusus tersebut sering digunakan dalam struktur URL *phishing* untuk menyamarkan alamat situs agar menyerupai situs asli.

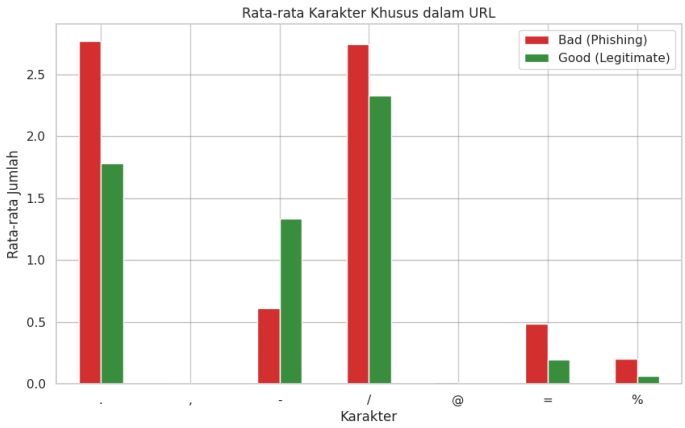

Figure 5. Rata-rata Karakter Khusus pada Uniform Resource Locator (URL)

Selain itu, dilakukan analisis terhadap distribusi Top Level Domain (TLD) pada dataset seperti yang ditunjukkan pada Gambar 6. Hasil analisis menunjukkan bahwa domain .com merupakan domain yang paling banyak digunakan baik pada URL phishing maupun URL legitimate. Selain .com, domain lain yang cukup sering muncul antara lain .html, .org, .htm, dan .php. Informasi ini menunjukkan bahwa penyerang sering memanfaatkan domain umum agar URL phishing terlihat lebih meyakinkan bagi pengguna.

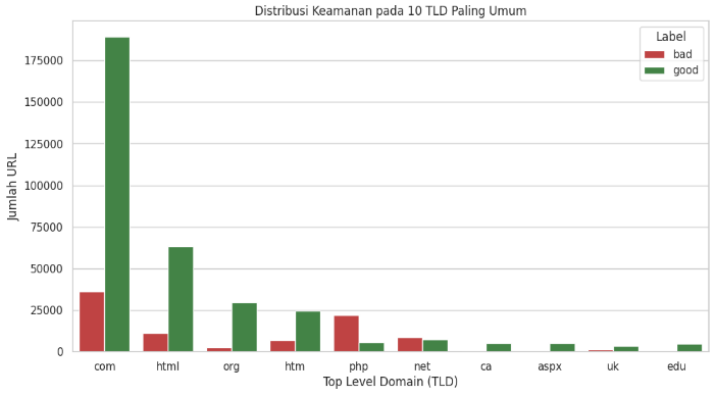

Figure 6. Distribusi Top Level Domain (TLD) pada Dataset URL

Analisis terakhir dilakukan terhadap kata-kata yang paling sering muncul dalam URL phishing dan URL legitimate seperti yang ditunjukkan pada Gambar 7. Pada URL phishing ditemukan beberapa kata yang sering muncul seperti login, paypal, images, content, dan secure. Kata-kata tersebut umumnya digunakan untuk meniru halaman autentikasi atau halaman layanan dari suatu situs resmi. Sementara itu pada URL legitimate ditemukan kata-kata seperti wiki, wikipedia, youtube, facebook, dan news yang lebih berkaitan dengan layanan atau platform website yang umum digunakan oleh pengguna.

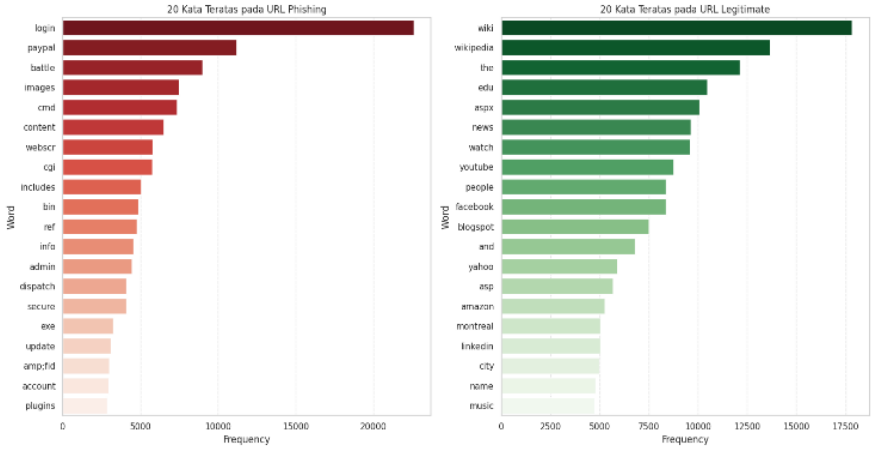

Figure 7. Kata yang Paling Sering Muncul pada URL Phishing dan URL Legitimate

Hasil dari tahap EDA ini memberikan gambaran awal mengenai pola-pola yang terdapat pada dataset URL. Informasi tersebut kemudian digunakan sebagai dasar dalam proses preprocessing data dan ekstraksi fitur sebelum dilakukan pelatihan model deteksi website phishing.

2.4. Preprocessing Data

Tahap preprocessing data dilakukan untuk menyiapkan dataset agar dapat digunakan dalam proses analisis dan pemodelan machine learning. Proses ini bertujuan untuk meningkatkan kualitas data dengan menghilangkan data yang tidak relevan serta melakukan normalisasi terhadap teks pada Uniform Resource Locator (URL). Dataset yang telah dikumpulkan pada tahap sebelumnya kemudian diproses sehingga dapat digunakan secara optimal dalam proses ekstraksi fitur dan pelatihan model.

Pada tahap ini dilakukan beberapa proses utama, yaitu pembersihan data (data cleaning), normalisasi teks URL, serta penghapusan data duplikat. Proses data cleaning dilakukan dengan memeriksa apakah terdapat nilai kosong atau data yang tidak valid dalam dataset. Selanjutnya dilakukan normalisasi teks dengan mengubah seluruh karakter URL menjadi huruf kecil (lowercase) agar tidak terjadi perbedaan representasi karakter selama proses analisis. Selain itu, dilakukan proses penghapusan data duplikat untuk menghindari pengaruh data yang sama terhadap proses pelatihan model.

Setelah proses pembersihan dan normalisasi data dilakukan, dataset kemudian disiapkan untuk tahap berikutnya yaitu ekstraksi fitur (feature extraction). Tahap ini bertujuan untuk mengubah data URL yang bersifat teks menjadi representasi fitur yang dapat diproses oleh algoritma machine learning seperti Random Forest maupun arsitektur Convolutional Neural Network (CNN).

2.5. Ekstraksi Fitur

Tahap ekstraksi fitur dilakukan untuk mengubah data Uniform Resource Locator (URL) yang berupa teks menjadi fitur numerik yang dapat diproses oleh algoritma machine learning. Proses ini bertujuan untuk mengambil karakteristik penting dari struktur URL yang berpotensi menjadi indikator adanya aktivitas phishing. Fitur-fitur tersebut kemudian digunakan sebagai input pada proses pelatihan model klasifikasi.

Pada penelitian ini beberapa fitur diekstraksi dari setiap URL, seperti panjang URL, jumlah karakter titik (.), jumlah tanda hubung (-), jumlah garis miring (/), jumlah angka pada URL, serta jumlah kata yang dicurigai (suspicious words) yang sering muncul pada URL phishing. Fitur-fitur tersebut dipilih karena URL phishing umumnya memiliki struktur yang lebih kompleks dan sering menggunakan karakter atau kata tertentu untuk meniru alamat situs resmi.

Hasil dari proses ekstraksi fitur ditunjukkan pada Gambar 8. Tabel tersebut menampilkan contoh sebagian data setelah dilakukan proses ekstraksi fitur. Kolom URL berisi alamat situs web yang dianalisis, sedangkan kolom label menunjukkan kategori URL yaitu bad untuk website phishing dan good untuk website legitimate. Selain itu terdapat beberapa kolom fitur numerik seperti length yang menunjukkan panjang URL, dot_count yang menunjukkan jumlah karakter titik dalam URL, hyphen_count yang menunjukkan jumlah tanda hubung, slash_count yang menunjukkan jumlah garis miring, digit_count yang menunjukkan jumlah angka pada URL, serta suspicious_word_count yang menunjukkan jumlah kata yang sering muncul pada URL phishing.

	URL	Label	length	dot_count	hyphen_count	slash_count	digit_count	suspicious_word_count
0	nobell.it/70ffb52d079109dca5664cce6f317373782/...	bad	225	6	4	10	58	1
1	www.dghjdgf.com/paypal.co.uk/cycgi-bin/webscrc...	bad	81	5	2	4	1	0
2	serviciosbys.com/paypal.cgi.bin.get-into.herf....	bad	177	7	1	11	47	1
3	mail.printakid.com/www.online.americanexpress....	bad	60	6	0	2	0	0
4	thewhiskeydregs.com/wp-content/themes/widescre...	bad	116	1	1	10	21	0

Figure 8. Contoh Hasil Ekstraksi Fitur pada Dataset Uniform Resource Locator (URL)

2.6. Pemodelan dan Perbandingan Model

Tahap pemodelan dilakukan untuk membangun sistem klasifikasi yang mampu membedakan website phishing dan website legitimate berdasarkan karakteristik Uniform Resource Locator (URL) yang telah diproses pada tahap sebelumnya. Dataset yang telah melalui proses preprocessing dan ekstraksi fitur kemudian digunakan sebagai data pelatihan untuk membangun model klasifikasi. Pada penelitian ini digunakan dua pendekatan metode yang berbeda yaitu algoritma Random Forest dan Convolutional Neural Network (CNN) [17].

Algoritma Random Forest digunakan untuk memodelkan hubungan antara fitur fitur numerik yang dihasilkan dari proses ekstraksi fitur URL dengan label klasifikasi. Metode ini bekerja dengan membangun sejumlah pohon keputusan (decision tree) dan menggabungkan hasil prediksi dari setiap pohon untuk menghasilkan keputusan akhir yang lebih stabil. Pendekatan ini banyak digunakan dalam klasifikasi karena mampu menangani data dengan jumlah fitur yang cukup besar serta mengurangi risiko overfitting.

Selain itu, penelitian ini juga menggunakan model Convolutional Neural Network (CNN) untuk mempelajari pola karakter yang terdapat pada struktur URL [18]. Model CNN yang digunakan berupa arsitektur 1D Convolutional Neural Network yang dirancang untuk memproses data berbentuk urutan karakter atau teks. Melalui proses konvolusi, model CNN dapat mengekstraksi pola tertentu dari urutan karakter URL yang berpotensi menjadi indikator adanya aktivitas phishing.

Kedua model tersebut kemudian dilatih menggunakan dataset yang telah dipersiapkan pada tahap sebelumnya sehingga sistem mampu mempelajari pola perbedaan antara URL phishing dan URL yang bersifat aman.

2.7. Evaluasi Model

Tahap evaluasi model dilakukan untuk menilai kemampuan sistem dalam mengklasifikasikan website phishing dan website legitimate berdasarkan Uniform Resource Locator (URL) yang dianalisis. Evaluasi ini bertujuan untuk mengukur sejauh mana model mampu melakukan prediksi secara tepat terhadap data yang belum pernah digunakan pada proses pelatihan sebelumnya. Proses pengujian dilakukan dengan menggunakan data uji yang diperoleh dari pemisahan dataset pada tahap sebelumnya sehingga performa model dapat diukur secara objektif.

Penilaian kinerja model dilakukan menggunakan beberapa metrik evaluasi klasifikasi yang umum digunakan dalam penelitian machine learning, yaitu accuracy, precision, recall, dan f1-score. Selain itu, analisis juga dilakukan menggunakan confusion matrix untuk melihat distribusi prediksi model terhadap kelas phishing dan kelas legitimate [19]. Melalui evaluasi ini dapat diketahui kemampuan model dalam mengidentifikasi URL berbahaya serta meminimalkan kesalahan klasifikasi terhadap URL yang bersifat aman.

Selain menggunakan metrik evaluasi, pengujian juga dilakukan dengan menganalisis beberapa contoh URL pada skenario pengujian nyata untuk melihat bagaimana sistem memberikan prediksi terhadap URL yang dianalisis. Tahap ini bertujuan untuk menguji konsistensi model dalam mendeteksi pola URL yang mencurigakan maupun URL yang tergolong aman. Hasil evaluasi tersebut kemudian akan digunakan untuk menganalisis performa masing masing model pada bagian hasil dan pembahasan penelitian.

3. RESULT DAN ANALISIS

Bagian ini menyajikan hasil eksperimen serta analisis performa model Random Forest dan Convolutional Neural Network (CNN) dalam mendeteksi website phishing berdasarkan fitur URL yang telah diproses.

3.1. Hasil Klasifikasi Menggunakan Random Forest

Pada penelitian ini, algoritma Random Forest digunakan untuk melakukan klasifikasi terhadap data Uniform Resource Locator (URL) yang telah melalui proses preprocessing dan ekstraksi fitur. Random Forest merupakan metode ensemble learning yang bekerja dengan membangun sejumlah pohon keputusan (decision tree) secara acak dan kemudian menggabungkan hasil prediksi dari setiap pohon untuk menentukan kelas akhir. Prediksi akhir pada Random Forest diperoleh melalui mekanisme majority voting, yaitu kelas yang paling banyak diprediksi oleh seluruh pohon keputusan akan menjadi hasil klasifikasi akhir.

Secara matematis, proses klasifikasi pada Random Forest dapat dinyatakan sebagai berikut (1):

$$\hat{y} = mode\{h_1(x), h_2(x), \cdots h_n(x)\} \quad (1)$$

Di mana $h_i(x)$merupakan prediksi dari pohon keputusan ke-i, nmenyatakan jumlah pohon keputusan dalam model Random Forest, dan $\hat{y}$ merupakan kelas hasil prediksi akhir yang diperoleh dari nilai mayoritas (majority voting) seluruh pohon keputusan.

Persamaan tersebut menunjukkan bahwa hasil klasifikasi diperoleh dari nilai mayoritas (mode) dari seluruh prediksi pohon keputusan terhadap data masukan x.

Hasil evaluasi model Random Forest ditunjukkan pada Tabel 1. Berdasarkan hasil tersebut, model memperoleh nilai accuracy sebesar 0.83. Untuk kelas good (legitimate website), model

menghasilkan nilai precision sebesar 0.89, recall sebesar 0.86, dan f1-score sebesar 0.88. Sementara itu, pada kelas bad (phishing), model memperoleh nilai precision sebesar 0.68, recall sebesar 0.75, dan f1-score sebesar 0.71.

Table 1. Hasil Evaluasi Model Random Forest

Class	Precision	Recall	F1-Score	Support
Good	89	86	88	78585
Bad	68	75	71	31285
Accuracy			83	109870
Macro Avg	79	80	79	109870
Weighted Avg	83	83	83	109870

Selain itu, analisis performa model juga dilakukan menggunakan confusion matrix yang ditunjukkan pada Gambar 9. Berdasarkan matriks tersebut, sebanyak 67.489 data legitimate website berhasil diklasifikasikan dengan benar sebagai kelas good, sedangkan 23.316 data phishing berhasil terdeteksi sebagai kelas bad. Namun demikian, masih terdapat beberapa kesalahan klasifikasi, yaitu 11.096 data legitimate website yang diprediksi sebagai phishing dan 7.969 data phishing yang diprediksi sebagai legitimate website. Hasil ini menunjukkan bahwa model Random Forest mampu mengenali pola pada struktur URL dengan cukup baik, meskipun masih terdapat beberapa kesalahan klasifikasi pada sebagian data phishing.

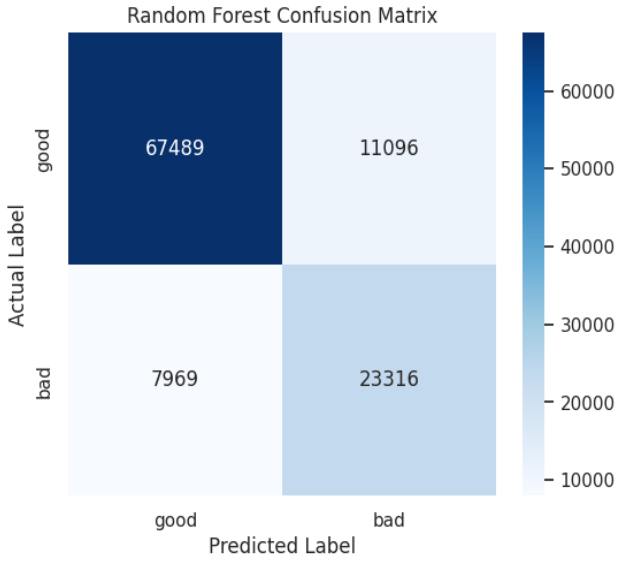

Figure 9. Confusion Matrix Model Random Forest

3.2. Pembangunan Model Convolutional Neural Network (CNN)

Pada penelitian ini, pendekatan deep learning menggunakan arsitektur Convolutional Neural Network (CNN) diterapkan untuk melakukan klasifikasi terhadap data Uniform Resource Locator (URL). Model CNN digunakan untuk menangkap pola karakter dan struktur pada URL yang berpotensi menunjukkan indikasi phishing [21]. Arsitektur CNN yang digunakan terdiri dari beberapa lapisan utama yaitu Embedding layer, Conv1D layer, Global Max Pooling layer, Dense layer, serta Dropout layer. Lapisan embedding digunakan untuk mengubah token URL menjadi representasi vektor numerik sehingga dapat diproses oleh jaringan saraf.

Konfigurasi parameter model CNN yang digunakan dalam penelitian ini ditunjukkan pada Tabel 2. Model memiliki total 23.427 parameter yang seluruhnya bersifat trainable. Lapisan Conv1D digunakan untuk mengekstraksi pola lokal pada urutan karakter URL, sedangkan Global Max Pooling

berfungsi untuk mereduksi dimensi fitur yang dihasilkan. Selanjutnya, lapisan Dense digunakan untuk melakukan proses klasifikasi terhadap fitur yang telah dipelajari oleh jaringan.

Table 2. Hasil Arsitektur dan Parameter Model CNN

Layer (type)	Output Shape	Param #
Embedding (Embedding)	(None, 100, 50)	11,650
Conv1d (Conv1D)	(None, 98, 64)	9,664
Global_max_pooling1d (globalmaxpooling1d)	(None, 64)	0
Dense (Dense)	(None, 32)	2,080
Dropout (Dropout)	(None, 32)	0
Dense_1 (dense)	(None, 1)	33

Secara matematis, operasi konvolusi pada CNN dapat dinyatakan sebagai berikut (2):

$$y_i = f\left(\sum_{k=1}^{n} w_k \cdot x_{i+k-1} + b\right) \quad (2)$$

Di mana x merupakan input fitur, w_k adalah bobot kernel konvolusi, $b$merupakan bias, dan f adalah fungsi aktivasi yang digunakan dalam jaringan saraf. Persamaan ini menunjukkan bahwa setiap neuron pada lapisan konvolusi menghitung kombinasi linier dari input dengan bobot tertentu yang kemudian dilewatkan ke fungsi aktivasi.

Proses pelatihan model dilakukan selama beberapa epoch untuk mempelajari pola dari dataset URL. Hasil proses pelatihan ditunjukkan pada Gambar 10, yang memperlihatkan perkembangan nilai accuracy dan loss selama proses training. Berdasarkan hasil tersebut, nilai accuracy model menunjukkan peningkatan secara bertahap pada setiap epoch, sementara nilai loss mengalami penurunan. Hal ini menunjukkan bahwa model CNN mampu mempelajari pola karakteristik URL secara efektif selama proses pelatihan.

```
Total params: 23,427 (91.51 KB)
Trainable params: 23,427 (91.51 KB)
Non-trainable params: 0 (0.00 B)
Epoch 1/5
6867/6867 ━━━━━━━━━━━━━━━━━━━━ 132s 19ms/step - accuracy: 0.8910 - loss: 0.3903 - val_accuracy: 0.9491 - val_loss: 0.2916
Epoch 2/5
6867/6867 ━━━━━━━━━━━━━━━━━━━━ 130s 19ms/step - accuracy: 0.9429 - loss: 0.3142 - val_accuracy: 0.9504 - val_loss: 0.2870
Epoch 3/5
6867/6867 ━━━━━━━━━━━━━━━━━━━━ 130s 19ms/step - accuracy: 0.9480 - loss: 0.3060 - val_accuracy: 0.9546 - val_loss: 0.2814
Epoch 4/5
6867/6867 ━━━━━━━━━━━━━━━━━━━━ 130s 19ms/step - accuracy: 0.9509 - loss: 0.3014 - val_accuracy: 0.9575 - val_loss: 0.2777
Epoch 5/5
6867/6867 ━━━━━━━━━━━━━━━━━━━━ 128s 19ms/step - accuracy: 0.9528 - loss: 0.2986 - val_accuracy: 0.9585 - val_loss: 0.2763
```

Figure 10. Proses Training Model CNN pada Setiap Epoch

Selanjutnya, visualisasi performa model selama proses pelatihan ditampilkan melalui grafik model accuracy dan model loss yang ditunjukkan pada Gambar 11. Grafik tersebut memperlihatkan perbandingan antara nilai training accuracy dan validation accuracy, serta nilai training loss dan validation loss. Berdasarkan grafik tersebut dapat diamati bahwa nilai accuracy pada data pelatihan maupun validasi menunjukkan tren peningkatan yang stabil, sementara nilai loss mengalami penurunan seiring bertambahnya jumlah epoch. Hal ini menunjukkan bahwa model CNN memiliki kemampuan yang baik dalam mempelajari pola dari data URL tanpa mengalami overfitting yang signifikan.

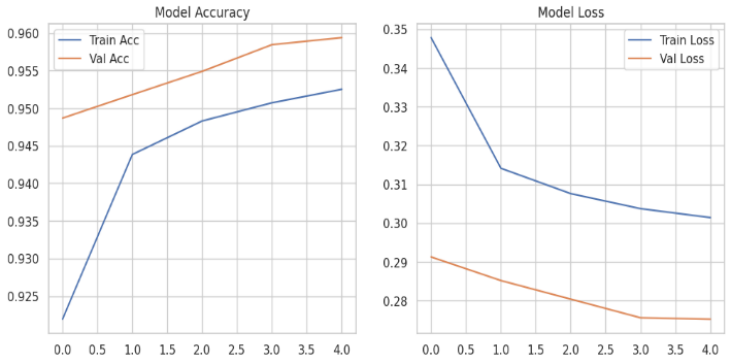

Figure 11. Grafik Model Accuracy dan Model Loss pada CNN

3.3. Hasil Perbandingan Kinerja Model Random Forest dan CNN

Pada tahap akhir penelitian ini dilakukan analisis perbandingan performa antara dua model klasifikasi yang digunakan, yaitu Random Forest dan Convolutional Neural Network (CNN). Tujuan dari tahap ini adalah untuk mengetahui model mana yang memiliki kemampuan lebih baik dalam mendeteksi URL phishing berdasarkan dataset yang telah diproses pada tahap sebelumnya. Evaluasi dilakukan menggunakan metrik akurasi yang diperoleh dari hasil pengujian model terhadap data uji.

Hasil perbandingan akurasi kedua model ditunjukkan pada Gambar 12. Berdasarkan hasil tersebut, model Random Forest memperoleh nilai akurasi sebesar 82,69%, sedangkan model 1D-CNN memperoleh akurasi yang lebih tinggi yaitu sebesar 95,94%. Perbedaan ini menunjukkan bahwa model CNN memiliki kemampuan yang lebih baik dalam mempelajari pola karakter dan struktur URL dibandingkan metode berbasis ensemble learning seperti Random Forest.

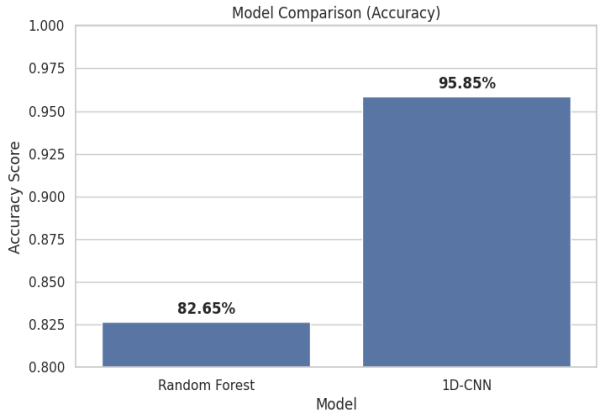

Figure 12. Perbandingan Akurasi Model Random Forest dan CNN

Secara umum, model CNN mampu mengekstraksi pola kompleks pada urutan karakter URL melalui proses konvolusi dan pembelajaran representasi fitur secara otomatis. Hal ini memungkinkan model untuk mengidentifikasi karakteristik URL phishing dengan lebih akurat. Oleh karena itu, berdasarkan hasil evaluasi yang diperoleh, model CNN menunjukkan performa yang lebih unggul dibandingkan Random Forest dalam tugas klasifikasi URL phishing pada penelitian ini.

3.4. Rekomendasi dan Arah Pengembangan Riset

Berdasarkan hasil penelitian yang telah dilakukan, penggunaan pendekatan machine learning dan deep learning menunjukkan potensi yang signifikan dalam mendeteksi URL phishing secara otomatis. Model CNN terbukti mampu memberikan performa klasifikasi yang lebih baik dibandingkan

Random Forest dalam mengidentifikasi pola karakteristik URL berbahaya. Untuk pengembangan penelitian selanjutnya, disarankan agar dilakukan eksplorasi metode deep learning yang lebih kompleks seperti Long Short-Term Memory (LSTM) atau Transformer yang dapat menangkap pola urutan karakter URL secara lebih mendalam. Selain itu, penelitian berikutnya juga dapat mempertimbangkan penggunaan dataset yang lebih beragam serta integrasi fitur tambahan seperti informasi domain, sertifikat keamanan, atau analisis konten halaman web guna meningkatkan akurasi deteksi dan memperkuat sistem klasifikasi phishing secara lebih komprehensif.

4. DISCUSSION/CONCLUSION

Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan URL phishing menggunakan pendekatan *machine learning* dan *deep learning* berdasarkan dataset URL yang diperoleh dari sumber publik serta didukung oleh pola serangan yang umumnya terdeteksi pada log *Web Application Firewall* (WAF). Proses penelitian meliputi tahap pengumpulan data, *exploratory data analysis* (EDA), *preprocessing*, ekstraksi fitur URL, pembangunan model, serta evaluasi kinerja model. Dua algoritma klasifikasi digunakan dalam penelitian ini yaitu *Random Forest* dan *Convolutional Neural Network* (CNN). Hasil eksperimen menunjukkan bahwa model CNN mampu memberikan performa yang lebih baik dengan tingkat akurasi sebesar 95,94%, dibandingkan dengan Random Forest yang memperoleh akurasi sebesar 82,69%. Hal ini menunjukkan bahwa pendekatan *deep learning* memiliki kemampuan yang lebih efektif dalam mempelajari pola kompleks pada struktur URL yang berkaitan dengan aktivitas *phishing*. Dengan demikian, metode yang diusulkan dapat menjadi pendekatan yang potensial untuk mendukung sistem deteksi *phishing* secara otomatis serta membantu meningkatkan mekanisme perlindungan keamanan pada aplikasi *website* dan sistem berbasis jaringan.

REFERENCES

[1] M. R. Fatiha, I. Setiawan, A. N. Ikhsan, and I. R. Yunita, "Optimisasi sistem deteksi phishing berbasis web menggunakan algoritma decision tree," Jurnal Ilmiah IT CIDA, vol. 10, no. 2, 2024.
[2] D. B. Suwarno and M. Hardjianto, "Deteksi website phishing dari analisis URL menggunakan algoritma Random Forest," BIT (Fakultas Teknologi Informasi Universitas Budi Luhur), vol. 21, no. 2, 2024.
[3] S. Silcilia, N. P. Salsabila, and T. Apriani, "Analisis performa Random Forest, Decision Tree, dan Naive Bayes untuk deteksi link phishing berbasis fitur URL," ROUTERS: Jurnal Sistem dan Teknologi Informasi, vol. 4, no. 1, pp. 13–20, 2026.
[4] A. Mughaid, S. AlZu'bi, A. Hnaif, S. Taamneh, A. Alnajjar, and E. A. Elsoud, "An intelligent cyber security phishing detection system using deep learning techniques," Cluster Computing, 2022, doi: 10.1007/s10586-022-03604-4.
[5] M. Aljabri and A. Alharthi, "Machine learning based phishing detection using URL features," IEEE Access, vol. 10, pp. 105127–105139, 2022.
[6] M. A. Alshamrani, A. Chowdhary, S. Pisharody, and D. Huang, "Phishing detection using machine learning: A systematic literature review," IEEE Access, vol. 10, pp. 103564–103586, 2022.
[7] E. P. Ghani, A. Sofwan, and M. Somantri, "AI-driven network security: Detecting and mitigating DDoS, malware, and backdoor attacks with isolation and Random Forest algorithm," in Proceedings of the International Conference on Smart Computing IoT and Machine Learning, 2025, pp. 1–6.
[8] U. Haresh and P. R. Kumari, "Advanced machine learning framework for robust phishing website identification," International Journal of Scientific Research and Engineering Trends, vol. 11, no. 3, 2025.
[9] A. Jadhao, L. Mahindre, K. Rahangdale, V. Singh, R. Shipurkar, and U. Kosarkar, "Phish guard: Phishing website detection using machine learning algorithms," International Journal of Trend in Scientific Research and Development, 2024.
[10] M. S. Kumari, C. K. Priya, G. Bhavya, H. Neha, M. Awasthi, and S. Tripathi, "Viable detection of URL phishing using machine learning approach," E3S Web of Conferences, vol. 430, 2023, doi: 10.1051/e3sconf/202343001037.
[11] M. Fahri, "Penerapan algoritma Random Forest untuk deteksi phishing pada website," Jurnal Ilmiah Teknologi Sistem Informasi, 2025.
[12] S. D. Gupta, K. T. Shahriar, H. Alqahtani, D. Alsalman, and I. H. Sarker, "Modeling hybrid feature-based phishing websites detection using machine learning techniques," Annals of Data Science, 2022, doi: 10.1007/s40745-022-00379-8.
[13] Z. Alshingiti, R. Alaqel, J. Al-Muhtadi, Q. E. Ul Haq, K. Saleem, and M. H. Faheem, "A deep learning-based phishing detection system using CNN, LSTM, and LSTM-CNN," Electronics, vol. 12, 2023.
[14] H. Abutair and A. Belghith, "Using machine learning to detect phishing attacks in URLs," Computers & Security, vol. 118, p. 102696, 2022.
[15] M. Zamir, H. Khan, M. I. Khan, M. S. Nisar, M. A. Khan, and S. Kadry, "Phishing website detection using machine learning and deep learning models," IEEE Access, vol. 8, pp. 132544–132554, 2020.
[16] A. Wijoyo, A. Saputra, A. Aditia, M. R. A. Pratama, and R. Rahman, "Analisis serangan phishing dan strategi deteksinya," JRIIN: Jurnal Riset Informatika dan Inovasi, vol. 1, no. 4, 2023.
[17] V. Indriani, H. Listiyono, and Saefurrahman, "Deteksi phishing website menggunakan support vector machine, gradient boosting, dan neural networks," JATI (Jurnal Mahasiswa Teknik Informatika), vol. 9, no. 2, pp. 3107–3113, 2025.
[18] Lukito and W. B. T. Handaya, "Deteksi website phishing menggunakan teknik machine learning," Jurnal Informatika Atma Jogja, vol. 6, no. 1, pp. 69–80, 2025.
[19] V. A. Windarni, A. F. Nugraha, S. T. A. Ramadhani, D. A. Istiqomah, F. M. Puri, and A. Setiawan, "Deteksi website phishing menggunakan teknik filter pada model machine learning," Information Systems Journal, vol. 6, no. 1, pp. 39–43, 2023.
[20] N. Azzahra, M. D. Handayani, and A. Aliyah, "Evaluasi kinerja AI berbasis recurrent neural network (RNN) dalam mengidentifikasi ancaman phishing pada URL website," Bridge: Jurnal Publikasi Sistem Informasi dan Telekomunikasi, vol. 3, no. 3, pp. 15–37, 2025.

[21] E. A. Aldakheel, M. Zakariah, G. A. Gashgari, F. A. Almarshad, and A. I. Alzahrani, "A deep learning-based innovative technique for phishing detection in modern security with uniform resource locators," Sensors, vol. 23, no. 9, p. 4403, 2023.