Bridge : Jurnal Publikasi Sistem Informasi dan Telekomunikasi
Volume. 3 Nomor. 3 Agustus 2025
e-ISSN : 3046-725X; p-ISSN : 3046-7268, Hal. 15-37
DOI: https://doi.org/10.62951/bridge.v3i3.485
Available online at: https://journal.aptii.or.id/index.php/Bridge

Evaluasi Kinerja AI berbasis *Recurrent Neural Network* (RNN) dalam Mengidentifikasi Ancaman *Phising* pada URL *Website*

Nailah Azzahra[1*], Merry Dwi Handayani[2], Awwaliyah Aliyah[3]
[1-3]Departemen Sistem Informasi, Fakultas Teknik Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember, Indonesia
nailahazhr@gmail.com[1], dwimerry0603@gmail.com[2], aliyahawwaliyah@gmail.com[3]

*Korespondensi Penulis: nailahazhr@gmail.com**

Abstract. *Phishing is an evolving form of cybercrime that targets users' sensitive information through URL manipulation. Conventional detection methods such as blacklists and signature-based approaches have become increasingly inadequate in addressing the dynamic variations of modern phishing attacks. This study evaluates the effectiveness of Recurrent Neural Network (RNN) variants, such Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), and Gated Recurrent Unit (GRU), in detecting phishing threats based on URL data. The methodology involves a Systematic Literature Review (SLR) of scholarly publications from the past ten years, complemented by experimental implementation of the models using a public dataset from Kaggle. Literature findings show that Bi-LSTM consistently achieves the highest accuracy, up to 99%, while GRU stands out for its computational efficiency. Experimental results support these findings, with Bi-LSTM achieving an accuracy of 96.22%, GRU 96.29%, and LSTM 95.43%. Classification metrics indicate that RNN-based models perform very well in detecting benign and defacement URLs, although their performance on phishing URLs remains challenged, particularly in terms of recall. These results confirm that RNNs remain a promising approach for phishing detection systems, especially when integrated into hybrid models with complementary architectures. This study is expected to provide a foundation for developing precise and adaptive AI systems to combat increasingly sophisticated phishing threats.*

Keywords: *Phishing, URL, RNN, LSTM, Bi-LSTM, GRU*

Abstrak. Phishing merupakan bentuk kejahatan siber yang terus berkembang, menargetkan informasi sensitif pengguna melalui manipulasi URL. Metode deteksi konvensional seperti blacklist dan signature-based semakin terbukti tidak memadai menghadapi variasi phishing modern yang dinamis. Penelitian ini mengevaluasi efektivitas varian Recurrent Neural Network (RNN) , yakni Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), dan Gated Recurrent Unit (GRU), dalam mendeteksi ancaman phishing berbasis URL. Pendekatan yang digunakan adalah Systematic Literature Review (SLR) atas publikasi ilmiah dari 10 tahun terakhir serta eksperimen langsung terhadap model-model tersebut menggunakan dataset publik dari Kaggle. Hasil studi literatur menunjukkan bahwa Bi-LSTM secara konsisten mencatat akurasi tertinggi hingga 99%, sementara GRU unggul dalam efisiensi komputasi. Hasil eksperimen memperkuat temuan tersebut dengan Bi-LSTM mencatat akurasi 96,22%, GRU 96,29%, dan LSTM 95,43%. Evaluasi metrik klasifikasi menunjukkan bahwa model RNN sangat baik dalam mendeteksi URL benign dan defacement, namun performa pada phishing URL masih menghadapi tantangan dari segi recall. Temuan ini menegaskan bahwa RNN tetap menjadi pendekatan prospektif dalam sistem deteksi phishing, terutama jika dikombinasikan dengan arsitektur lain dalam model hibrida. Studi ini diharapkan dapat menjadi dasar bagi pengembangan sistem AI yang presisi dan adaptif dalam menghadapi ancaman phishing yang semakin kompleks.

Kata Kunci: *Phishing*, URL, RNN, LSTM, Bi-LSTM, GRU

1. PENDAHULUAN

Dalam era digital, phishing menjadi salah satu bentuk serangan siber paling umum dan merugikan. Teknik ini menipu korban agar menyerahkan informasi sensitif dengan menyamar sebagai entitas tepercaya, tidak hanya melalui email dan situs web, tetapi juga aplikasi mobile seiring meningkatnya penggunaan perangkat pintar. Ancaman phishing menunjukkan lonjakan signifikan. Menurut Anti-Phishing Working Group (APWG), serangan meningkat dari 779.000

Received: Mei 26, 2025; Revised: Juni 06, 2025; Accepted: Juni 20, 2025; Online Available: Juni 23, 2025

kasus pada 2019 menjadi 4,99 juta pada 2023 (Husain, 2025). Sepanjang 2024, rata-rata insiden bulanan mencapai lebih dari 300.000 kasus (APWG, 2025). Sekitar 57% organisasi global bahkan mengalami serangan ini setiap minggu. Fakta tersebut menegaskan bahwa phishing kini merupakan risiko siber harian yang sistematis, mendorong kebutuhan terhadap sistem deteksi yang lebih adaptif dan presisi tinggi.

Berbagai metode telah dikembangkan, mulai dari blacklist, signature-based detection, hingga pendekatan machine learning (Balogun et al., 2021; Bharath, 2025). Meski masih digunakan luas, metode konvensional tidak cukup tangguh menghadapi varian phishing baru yang kerap lolos deteksi. Machine learning pun menghadapi tantangan seperti rendahnya akurasi pada kelas minoritas dan concept drift akibat perubahan karakteristik data (Singh & Meenu, 2020; Nezhab & Langarib, 2025).

Melihat kebutuhan akan pendekatan yang lebih responsif terhadap pola dinamis, Recurrent Neural Network (RNN) dinilai memiliki potensi kuat karena kemampuannya memproses data sekuensial seperti struktur URL atau konten situs yang telah di-obfuscate (Lamina et al., 2024). Oleh karena itu, penelitian ini bertujuan mengkaji performa model RNN dan variannya (LSTM, GRU) dalam mendeteksi phishing berbasis web. Peninjauan ini juga merangkum temuan-temuan penting dari literatur untuk mengevaluasi keefektifan RNN, tantangan yang dihadapi, dan arah riset ke depan.

2. TINJAUAN PUSTAKA

RNN

RNN adalah arsitektur deep learning untuk data sekuensial, dengan koneksi rekursif yang menyimpan informasi historis dalam state internal (Ghojogh & Ghodsi, 2023). Namun, RNN standar rentan terhadap *vanishing gradient*, sehingga dikembangkan varian seperti LSTM dan GRU dengan *gating mechanism* untuk mempertahankan informasi penting (Lipton et al., 2015). Dalam deteksi phishing, Bi-LSTM menunjukkan akurasi hingga 99%, unggul atas model lain (Roy et al., 2022), karena mampu memahami pola dua arah dalam struktur URL. Senouci & Benaouda (2025) juga menunjukkan efektivitas RNN-LSTM dalam lingkungan cloud (akurasi 98,88%). Selain itu, Halgas et al. (2019) menemukan bahwa RNN dapat mengidentifikasi pola semantik pada teks phishing. Meskipun model seperti CNN dan transformer juga digunakan, RNN tetap unggul dalam menangani urutan, meski kurang efisien dalam pelatihan dan panjang input, menjadikannya cocok dalam model hybrid atau ensemble.

Phishing

Phishing adalah serangan berbasis rekayasa sosial yang meniru entitas resmi guna mencuri informasi pribadi. Deteksi phishing telah banyak dilakukan dengan pendekatan machine learning seperti SVM, Random Forest, dan kNN, yang dilaporkan efektif secara akurasi (Jadhav & Chandre, 2022). Namun, pendekatan tersebut kurang tepat untuk pola berurutan yang kompleks. Oleh sebab itu, model deep learning seperti LSTM-CNN mulai digunakan. Alshingti et al. (2023) mencatat bahwa sistem ini mencapai akurasi 97,4% dalam pengujian berbasis URL phishing. Tantangan tetap ada, terutama terkait keterbatasan dataset yang tidak mencerminkan kompleksitas serangan dunia nyata (Mohammad et al., 2015).

URL (Uniform Resource Locator)

URL menjadi komponen utama dalam deteksi phishing karena mengandung informasi terstruktur yang dapat dianalisis untuk menemukan anomali. Struktur URL yang terdiri dari protokol, domain, path, dan query string sering dimanipulasi oleh pelaku phishing. Model seperti PhiUSIIL mendeteksi serangan homograph, bitsquatting, dan combosquatting dengan memanfaatkan indeks kemiripan dan pembelajaran inkremental (Prasad & Chandra, 2024). Rashid et al. (2024) menunjukkan bahwa teknik unsupervised domain adaptation berhasil meningkatkan generalisasi model, dengan rata-rata peningkatan F1-score sebesar 0,06 hingga 0,2. Analisis URL juga mencakup fitur seperti panjang URL, keberadaan IP address, dan jumlah parameter yang mencurigakan.

Website

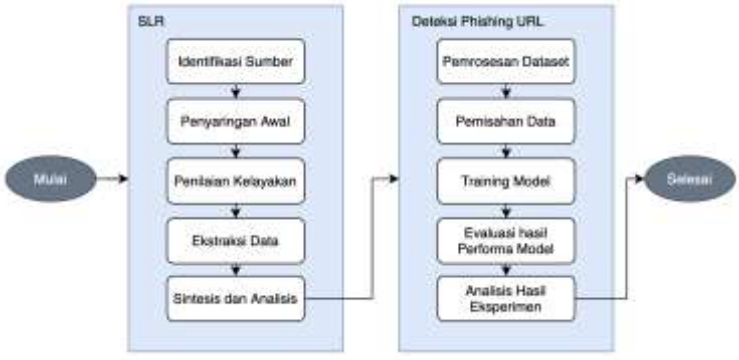
Website menjadi target utama phishing karena mudah ditiru tampilannya. Teknik seperti typosquatting, homograph attacks, dan pengalihan JavaScript sering digunakan untuk menipu pengguna (Vijayalakshmi et al., 2020). Oleh karena itu, deteksi berbasis machine learning mulai diterapkan. Mahmood & Tang (2021) menunjukkan bahwa algoritma seperti SVM dan Random Forest dapat mengklasifikasikan situs phishing secara akurat berdasarkan fitur seperti panjang URL dan jumlah subdomain. Namun, pendekatan klasik ini masih kurang efektif dalam menangkap dinamika temporal dan struktur data yang kompleks, sehingga mendorong penggunaan model deep learning.

Artificial Intelligence

Artificial Intelligence (AI) berperan penting dalam deteksi phishing modern karena mampu mengenali pola kompleks dari data besar. Model-model seperti RNN, LSTM, dan CNN-RNN digunakan untuk memproses input berurutan seperti URL. Gupta et al. (2024) mencatat peningkatan akurasi deteksi phishing melalui optimasi hyperparameter dengan *whale optimization algorithm*. Meta-analisis oleh Catal et al. (2022) juga menyimpulkan bahwa model deep learning lebih unggul dibanding algoritma tradisional karena tidak memerlukan rekayasa fitur manual. Survei oleh Safi & Singh (2023) menunjukkan bahwa CNN dapat mencapai akurasi hingga 99,98% dalam klasifikasi phishing. Studi oleh Asiri et al. (2024) juga menunjukkan bahwa sistem AI mampu memblokir URL berbahaya sebelum pengguna sempat mengaksesnya, yang menunjukkan pentingnya AI dalam perlindungan real-time terhadap serangan siber.

3. METODE PENELITIAN

Penelitian ini mengadopsi pendekatan Systematic Literature Review (SLR) untuk mengumpulkan dan menganalisis studi terkait kinerja model Recurrent Neural Network (RNN) dalam mendeteksi phishing pada website (Thakur et al., 2023). SLR dipilih karena menawarkan tinjauan literatur yang komprehensif, sistematis, dan minim bias, sekaligus membantu mengidentifikasi tren, celah penelitian, dan arah studi mendatang. Berikutnya, dilakukan eksperimen pembanding terhadap performa model RNN (LSTM, GRU, Bi-LSTM) serta model lain (1D CNN dan MLP) dalam mendeteksi phishing URL. Diagram metodologi penelitian ditunjukkan pada Gambar 3.1.



Gambar 1. Tahapan Metodologi

***Systematic Literature Review* (SLR)**

Proses Systematic Literature Review (SLR) dalam penelitian ini mengikuti pedoman Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) untuk menjamin transparansi dan replikasi hasil (Kyaw, Gutierrez, & Ghobakhlou, 2024). Berdasarkan pendekatan dari Thakur et al. (2023), lima tahapan utama dilakukan sebagai berikut:

- Identifikasi Sumber

  Pencarian sistematis dilakukan pada basis data IEEE Xplore, ACM Digital Library, ScienceDirect, dan SpringerLink, menggunakan kata kunci seperti *"Deep Learning"*, *"Phishing Detection"*, *"RNN"*, *"LSTM"*, *"GRU"*, dan sejenisnya. Literatur dibatasi pada 5–10 tahun terakhir dan berfokus pada penerapan RNN untuk deteksi phishing pada website atau URL.
- Penyaringan Awal

  Duplikasi dihapus, kemudian setiap artikel diseleksi melalui peninjauan judul dan abstrak. Artikel yang tidak relevan dengan topik langsung dikeluarkan.
- Penilaian Kelayakan

  Artikel yang lolos diseleksi kembali melalui pembacaan teks penuh untuk memastikan kesesuaian dengan kriteria inklusi: penerapan RNN dalam deteksi phishing dan publikasi dalam jurnal ilmiah.
- Ekstraksi Data

  Informasi yang dikumpulkan mencakup: jenis arsitektur RNN (LSTM, Bi-LSTM, GRU, atau hibrida), teknik pra-pemrosesan, sumber dan karakteristik dataset, metrik evaluasi (akurasi, presisi, recall, F1-score), serta kelebihan dan keterbatasan tiap studi.
- Sintesis dan Analisis

  Data yang telah diekstrak dianalisis secara kualitatif dan kuantitatif untuk mengidentifikasi pola, membandingkan performa antar model, dan menemukan celah penelitian guna merumuskan arah studi berikutnya.

Pendeteksian Phishing URL

Langkah-langkah eksperimen mengikuti lima tahap utama, seperti digambarkan dalam diagram alir berikut:

- Akuisisi dan Pra-Pemrosesan Dataset

  Dataset diperoleh dari Kaggle dengan nama file *malicious_phish.csv*, berisi URL dalam

empat kategori: benign (0), defacement (1), phishing (2), dan malware (3). Pra-pemrosesan mencakup pembersihan URL, encoding label, ekstraksi fitur, tokenisasi dan padding untuk model sekuensial, serta penskalaan fitur numerik.

- Pemisahan Data Latih dan Uji

  Dataset dibagi menjadi 80% data latih dan 20% data uji menggunakan train_test_split, dengan random_state=42 dan stratify=y untuk menjaga proporsi kelas. Tujuannya adalah menguji generalisasi model pada data yang belum pernah dilihat sebelumnya.
- Implementasi dan Pelatihan Model Neural Network

  Lima arsitektur diimplementasikan: LSTM, GRU, Bi-LSTM, 1D CNN, dan MLP. Model sekuensial menggunakan lapisan Embedding, sedangkan MLP menerima input fitur numerik langsung. Semua model dikompilasi dengan fungsi loss dan optimizer yang sama, menggunakan Dropout untuk regularisasi dan dilatih dengan batch size 64 serta validation split 10%.
- Evaluasi Performa Model

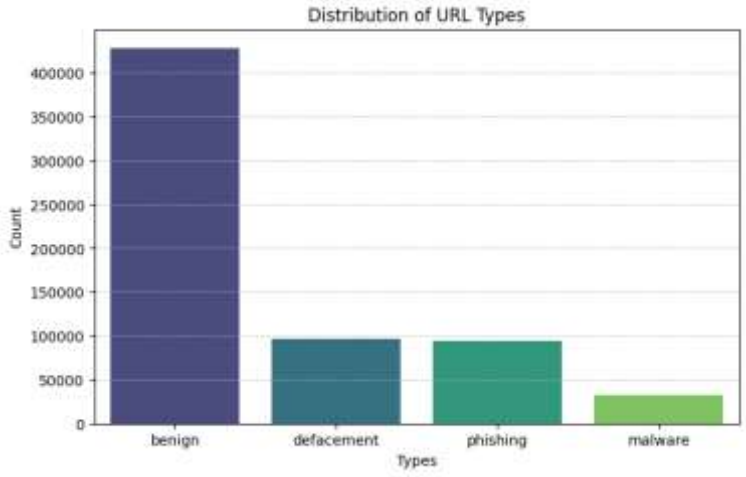
  Evaluasi dilakukan pada data uji menggunakan metrik akurasi, confusion matrix, dan classification report(presisi, recall, F1-score), untuk mengukur kinerja klasifikasi tiap kelas, khususnya deteksi terhadap kelas berbahaya.
- Analisis Hasil dan Perbandingan Model

  Performa kelima model dibandingkan berdasarkan metrik evaluasi dan hasil analisis confusion matrix, dengan penekanan pada kemampuan mendeteksi phishing, malware, dan defacement. Temuan ini dianalisis dalam konteks studi-studi literatur yang telah dikaji melalui proses SLR sebelumnya untuk menilai efektivitas relatif setiap model.

4. PEMBAHASAN

Dataset Penelitian Deteksi Phishing URL

Dataset digunakan dari Kaggle, terdiri atas 653.000 URL dalam empat kategori keamanan: Benign (0), Defacement (1), Phishing (2), dan Malware (3). Distribusi kelas ditunjukkan pada Tabel 1, dengan kelas Benign mendominasi 65%(428.103 URL), sedangkan Defacement (96.457), Phishing (94.111), dan Malware (32.520) jauh lebih sedikit. Komposisi visual data ditampilkan pada Gambar 4.1.



Gambar 2. Distribusi Tipe URL dalam Dataset

Distribusi yang timpang ini mencerminkan ketidakseimbangan kelas yang signifikan, di mana model cenderung fokus pada kelas mayoritas (Benign), sehingga risiko bias terhadap kelas minoritas meningkat, terutama Phishing dan Malware. Akibatnya, metrik seperti recall dan F1-score per kelas bisa menurun meskipun akurasi keseluruhan tinggi. Studi sebelumnya (Thakur et al., 2023; Kyaw et al., 2024) juga menegaskan bahwa data tidak seimbang dapat menyebabkan overfitting terhadap kelas dominan dan mengurangi kemampuan generalisasi. Oleh karena itu, penelitian ini menekankan evaluasi performa model per kelas untuk mengidentifikasi potensi kelemahan dalam mendeteksi ancaman aktual secara lebih akurat.

Pra-proses Data URL

Pra-pemrosesan merupakan tahap penting dalam membangun model deep learning untuk deteksi phishing berbasis URL. Karena URL bersifat sekuensial dan terdiri atas karakter yang kompleks, data perlu dikonversi ke bentuk numerik agar dapat dipahami oleh jaringan saraf. Umumnya, URL diproses sebagai rangkaian karakter, yang kemudian diubah menjadi urutan angka melalui tokenisasi tingkat karakter (Roy et al., 2022). Metode encoding yang umum digunakan antara lain one-hot encoding atau pendekatan dari Natural Language Processing (NLP) seperti Word2Vec dan embedding layer (Asiri et al., 2024).

Agar model dapat menerima input dengan dimensi tetap, padding dan truncation dilakukan untuk menyamakan panjang URL (L). Jika URL terlalu panjang, karakter sisanya dipotong; jika terlalu pendek, akan ditambahkan nilai nol agar panjangnya sesuai (Roy et al., 2022). Setelah itu, urutan token dikonversi menjadi embedding vektor berdimensi tetap, yang menjadi input utama bagi model deep learning.

Dalam eksperimen ini, dataset dari Kaggle (kaggle.com/datasets/malicious-urls-dataset) digunakan sebagai sumber data. Setiap URL ditokenisasi pada tingkat karakter,

diproses dengan padding/truncation hingga panjang 200 karakter, lalu dikonversi ke dalam embedding vektor berdimensi 128. Langkah ini dilakukan untuk memastikan konsistensi format input antar model dan membantu menjaga stabilitas pelatihan, khususnya pada arsitektur RNN seperti LSTM dan GRU.

Arsitektur dan Konfigurasi Model

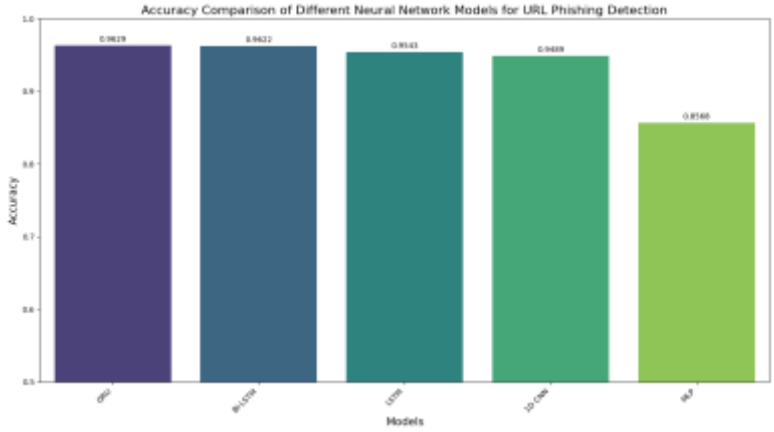
Penelitian ini mengimplementasikan lima arsitektur jaringan saraf untuk deteksi phishing URL, yang masing-masing dirancang untuk menangani karakteristik data sekuensial. Pemilihan model didasarkan pada pemahaman bahwa URL merupakan urutan karakter, sehingga arsitektur yang mampu mengenali pola berurutan dan kontekstual menjadi penting.

- LSTM (Long Short-Term Memory): LSTM dirancang untuk menangani dependensi jangka panjang dan mengatasi masalah vanishing gradient pada RNN tradisional. Model terdiri dari lapisan Embedding, LSTM (128 unit), Dropout (0,5), dan Dense softmax output.
- GRU (Gated Recurrent Unit): GRU memiliki struktur lebih sederhana dari LSTM, dengan dua gerbang utama: reset dan update. Arsitektur GRU mirip dengan LSTM, namun menawarkan efisiensi komputasi lebih baik dengan performa serupa.
- Bi-LSTM (Bidirectional LSTM): Bi-LSTM memproses data dari dua arah (maju dan mundur), sehingga lebih mampu memahami konteks menyeluruh dalam URL. Arsitekturnya identik dengan LSTM, tetapi dibungkus dalam lapisan Bidirectional.
- 1D CNN: CNN satu dimensi mengekstraksi pola lokal, seperti n-gram atau substring, melalui Conv1D (128 filter, kernel size 5), GlobalMaxPooling1D, dan dua lapisan Dense, dengan Dropout (0,5) sebagai regularisasi. Meski efektif untuk pola lokal, CNN kurang tepat untuk menangani hubungan jangka panjang.
- MLP (Multilayer Perceptron): MLP berfungsi sebagai baseline. Model terdiri dari tiga lapisan Dense (256–128–64) dengan aktivasi ReLU dan Dropout, serta menerima fitur numerik datar hasil transformasi URL. MLP tidak memproses urutan, sehingga informasi sekuensial dapat hilang.

Dalam tahap konfigurasi pelatihan, seluruh model dikompilasi dengan optimizer Adam, categorical cross entropy sebagai fungsi loss, dan akurasi sebagai metrik. Pelatihan dilakukan selama 5 epoch, dengan batch size 64 dan validation split 10%. Dimensi embedding ditetapkan sebesar 100 untuk semua model sekuensial.

Perbandingan Akurasi Seluruh Model

Perbandingan akurasi uji dari kelima model disajikan pada Gambar 4.2, yang menunjukkan bahwa model berbasis Recurrent Neural Network (RNN), yakni GRU, Bi-LSTM, dan LSTM, secara konsisten mengungguli model pembanding 1D CNN dan MLP. GRU mencatat akurasi tertinggi sebesar 96,29%, diikuti oleh Bi-LSTM (96,22%), dan LSTM (95,43%), sementara 1D CNN mencatat 94,89% dan MLP tertinggal jauh di 85,68%.



Gambar 3. Perbandingan Akurasi Model RNN dan pembanding

Keunggulan model RNN terletak pada kemampuannya dalam mempelajari dependensi temporal dalam data sekuensial seperti URL. Arsitektur seperti LSTM dan GRU mampu mempertahankan informasi penting dari elemen sebelumnya, sedangkan Bi-LSTM lebih lanjut memperluas konteks dengan membaca data dari dua arah, meningkatkan sensitivitas terhadap pola phishing yang tersebar (Thakur et al., 2023). Sebaliknya, 1D CNN lebih terbatas pada pola lokal, dan MLP, yang hanya menerima input numerik datar, tidak dapat menangkap informasi struktural atau urutan sama sekali.

Ringkasan hasil klasifikasi dari semua model, termasuk akurasi uji keseluruhan serta rata-rata presisi, recall, dan F1-score (makro dan tertimbang), disajikan pada Tabel 4.1 berikut.

Tabel 1. Kinerja Model RNN dan Variannya dalam Deteksi Phishing URL

Model RNN	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)
GRU	96.29	89-99	86-99	87-99
Bi-LSTM	96.22	88-99	86-100	87-99
LSTM	95.43	85-99	86-98	85-99
1D CNN	94.89	90-99	76-100	82-98
MLP	85.68	81-94	18-99	29-92

Temuan ini sejalan dengan studi sebelumnya. Roy et al. (2022) melaporkan Bi-LSTM mencapai akurasi hingga 99%, dengan presisi dan recall yang tinggi pada dua kelas utama URL. GRU juga mencatat akurasi 97,5%, dengan efisiensi lebih baik karena struktur gerbang

yang lebih ringan (Roy et al., 2022; Asiri et al., 2024). Penelitian oleh Kyaw et al. (2024) bahkan melaporkan akurasi hingga 99,1%, meskipun variasi preprocessing dan dataset mempengaruhi hasil. Kinerja model dalam tinjauan literatur dirangkum ke dalam tabel perbandingan berikut.

Tabel 2. Kinerja Model RNN berdasarkan Literatur

Arsitektur Model	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	Dataset yang Digunakan
LSTM	97.0	94-99	96-99	96-97	Kaggle Dataset (450K URLs)
Bi-LSTM	99.0	99	99	99	
GRU	97.5	95-99	97-99	97-98	
Bi-LSTM with Attention	99	99	99	99	Tidak spesifik
CNN-BiLSTM-Attention (CCBLA)	99.85	N/A	N/A	N/A	N/A
RNN-GRU	99.18	N/A	N/A	N/A	N/A

Analisis Kinerja Model RNN

LSTM

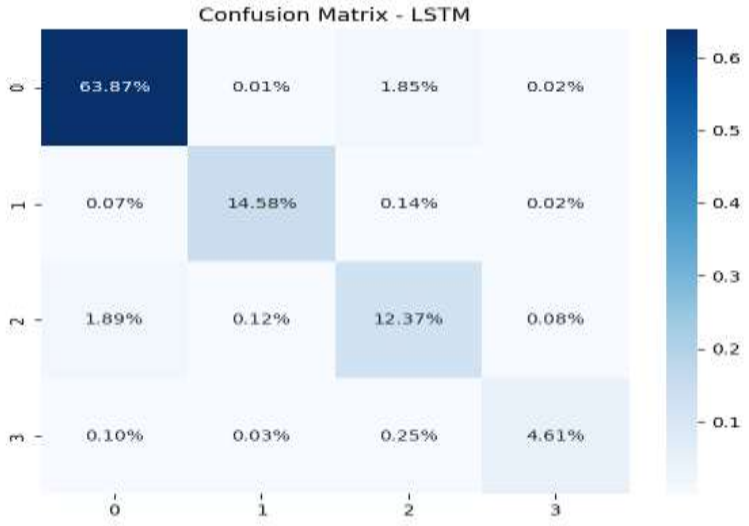
```
LSTM Test Accuracy: 95.43%
4070/4070 ━━━━━━━━━━━━━━━━━━━━ 15s 4ms/step
              precision    recall  f1-score   support

           0       0.97      0.97      0.97     85621
           1       0.99      0.98      0.99     19292
           2       0.85      0.86      0.85     18822
           3       0.98      0.92      0.95      6504

    accuracy                           0.95    130239
   macro avg       0.95      0.93      0.94    130239
weighted avg       0.95      0.95      0.95    130239
```

Gambar 4. Classification Report LSTM

Model LSTM menghasilkan akurasi uji sebesar 95,43%. Pada Kelas 2 (Phishing), LSTM mencatat nilai presisi 0,85, recall 0,86, dan F1-score 0,85. Nilai-nilai ini menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengidentifikasi URL phishing, meskipun kinerjanya masih di bawah kelas mayoritas. Pada Kelas 0 (Benign), presisi dan recall masing-masing sebesar 0,97, dengan F1-score 0,97. Ini mengindikasikan bahwa model sangat efektif dalam mengenali URL yang aman, yang memang mendominasi dataset. Untuk Kelas 1 (Defacement) dan Kelas 3 (Malware), nilai F1-score masing-masing adalah 0,99 dan 0,95.



Gambar 5. Confusion Matrix LSTM

Berdasarkan confusion matrix di atas, ditemukan bahwa 1,89% URL phishing diklasifikasikan secara keliru sebagai benign, dan 0,10% malware juga diklasifikasikan sebagai benign. Kesalahan ini mencerminkan kecenderungan model untuk bias terhadap kelas mayoritas, terutama dalam dataset yang tidak seimbang. Meskipun demikian, struktur gerbang internal pada LSTM (input, forget, output) berperan penting dalam menjaga konteks sekuensial, yang menjadikannya tetap efektif dalam mengenali pola kompleks pada struktur URL phishing.

GRU

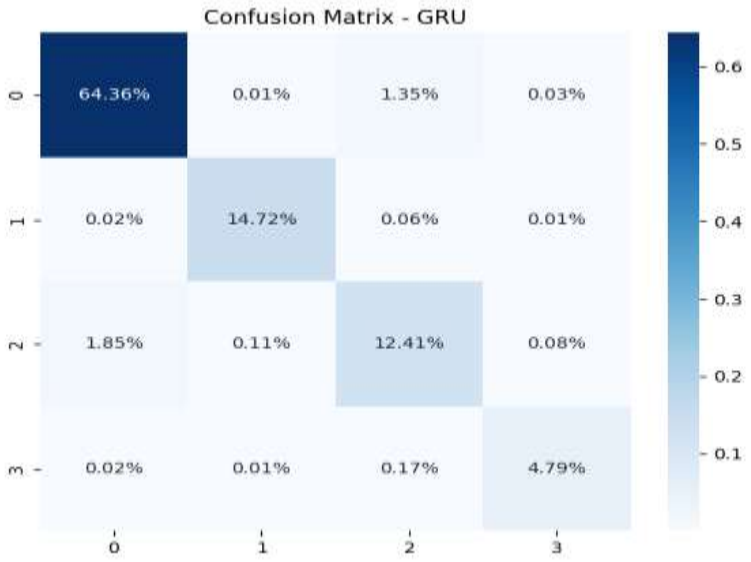
GRU Test Accuracy: 96.29%

4070/4070 ━━━━━━━━━━━━━━━━━━━━ 16s 4ms/step

	precision	recall	f1-score	support
0	0.97	0.98	0.98	85621
1	0.99	0.99	0.99	19292
2	0.89	0.86	0.87	18822
3	0.98	0.96	0.97	6504
accuracy			0.96	130239
macro avg	0.96	0.95	0.95	130239
weighted avg	0.96	0.96	0.96	130239

Gambar 6. Classification Report GRU

Model GRU mencatat akurasinya tertinggi, yaitu sebesar 96,29%, sekaligus memberikan performa terbaik dalam mendeteksi phishing. Pada Kelas 2 (Phishing), GRU mencatat presisi 0,89, recall 0,86, dan F1-score 0,87. Nilai presisi yang lebih tinggi dibandingkan LSTM mengindikasikan bahwa GRU lebih tepat dalam mengenali phishing URL tanpa banyak kesalahan positif. Untuk Kelas 0 (Benign), model meraih presisi 0,97 dan recall 0,98. Pada Kelas 1 (Defacement), presisi dan recall mencapai 0,99, dan pada Kelas 3 (Malware), recall meningkat menjadi 0,96, lebih tinggi dibandingkan LSTM.



Gambar 7. Confusion Matrix GRU

Confusion matrix di atas menunjukkan bahwa 1,85% phishing URL diklasifikasikan sebagai benign, dan 1,35% benign URL diklasifikasikan sebagai phishing. Angka ini sedikit lebih baik dibandingkan LSTM, terutama dalam mengurangi false positive dan false negative pada kelas phishing. Dengan arsitektur yang lebih ringan, hanya terdiri dari dua gerbang (reset dan update), GRU terbukti mampu mencapai efisiensi komputasi tanpa mengorbankan akurasi, bahkan menunjukkan kinerja yang lebih unggul dalam konteks deteksi phishing URL.

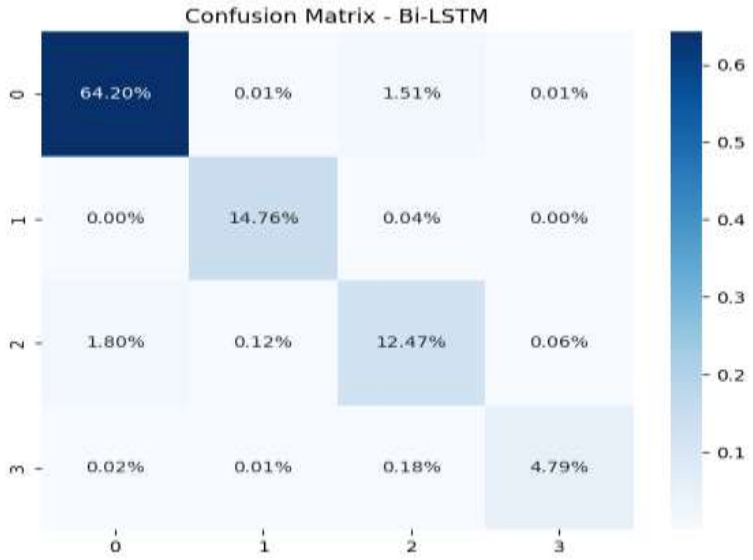
Bi-LSTM

Bi-LSTM Test Accuracy: 96.22%
4070/4070 ━━━━━━━━━━━━━━━━━━━━ 25s 6ms/step

	precision	recall	f1-score	support
0	0.97	0.98	0.97	85621
1	0.99	1.00	0.99	19292
2	0.88	0.86	0.87	18822
3	0.98	0.96	0.97	6504
accuracy			0.96	130239
macro avg	0.96	0.95	0.95	130239
weighted avg	0.96	0.96	0.96	130239

Gambar 8. Classification Report Bi-LSTM

Model Bi-LSTM mencatat akurasi uji sebesar 96,22%, yang sangat kompetitif dan mendekati GRU. Untuk Kelas 2 (Phishing), Bi-LSTM mencatat presisi 0,88, recall 0,86, dan F1-score 0,87, sebanding dengan GRU, serta lebih tinggi dibandingkan LSTM. Pada Kelas 0 (Benign), model mencatat presisi 0,97 dan recall 0,98, dengan F1-score 0,97. Pada Kelas 1 (Defacement), model menunjukkan recall sempurna sebesar 1,00 dengan presisi 0,99, dan F1-score 0,99. Sementara itu, Kelas 3 (Malware) mencapai F1-score 0,97, sama seperti GRU.



Gambar 9. Confusion Matrix Bi-LSTM

Berdasarkan confusion matrix di atas, model Bi-LSTM menunjukkan tingkat kesalahan terendah dalam mengklasifikasikan phishing URL sebagai benign, yaitu 1,80%. Selain itu, 1,51% benign URL diklasifikasikan sebagai phishing, menunjukkan keseimbangan yang cukup baik antara false positive dan false negative. Keunggulan utama Bi-LSTM terletak pada arsitektur dua arah yang memproses urutan input dari sisi kiri dan kanan secara bersamaan. Hal ini membuat Bi-LSTM lebih peka terhadap pola phishing yang tersebar di awal maupun akhir URL, sehingga memberikan konteks yang lebih kaya dalam proses klasifikasi.

Analisis Kinerja Model Pembanding

1D CNN

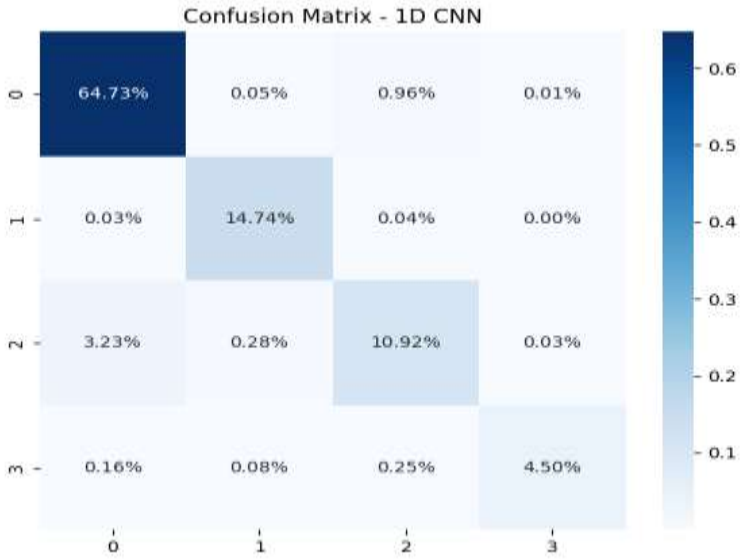
```
1D CNN Test Accuracy: 94.89%
4070/4070 ━━━━━━━━━━━━━━━━━━━━ 5s 1ms/step
              precision    recall  f1-score   support

           0       0.95      0.98      0.97     85621
           1       0.97      1.00      0.98     19292
           2       0.90      0.76      0.82     18822
           3       0.99      0.90      0.94      6504

    accuracy                           0.95    130239
   macro avg       0.95      0.91      0.93    130239
weighted avg       0.95      0.95      0.95    130239
```

Gambar 10. Classification Report 1D CNN

Model 1D CNN mencatat akurasi uji sebesar 94,89%. Untuk deteksi phishing (Kelas 2), model menunjukkan presisi 0,90dan recall 0,76, menghasilkan F1-score 0,82. Meskipun presisi tergolong tinggi, nilai recall yang rendah menunjukkan banyak URL phishing yang tidak berhasil dikenali. Pada Kelas 0 (Benign) dan Kelas 1 (Defacement), model mencatat F1-score masing-masing 0,97 dan 0,98, sementara pada Kelas 3 (Malware), F1-score mencapai 0,94.



Gambar 11. Confusion Matrix 1D CNN

Confusion matrix di atas menunjukkan bahwa 3,23% benign URL diklasifikasikan sebagai phishing, dan true positive phishing hanya 10,92%, lebih rendah dibandingkan model RNN. Hal ini mengindikasikan keterbatasan 1D CNN dalam mendeteksi phishing, khususnya karena CNN hanya mengekstraksi pola lokal (n-gram), tanpa menangkap dependensi karakter yang tersebar luas dalam struktur URL. Meski demikian, kinerja keseluruhan 1D CNN masih kompetitif dan lebih baik dibandingkan MLP.

MLP

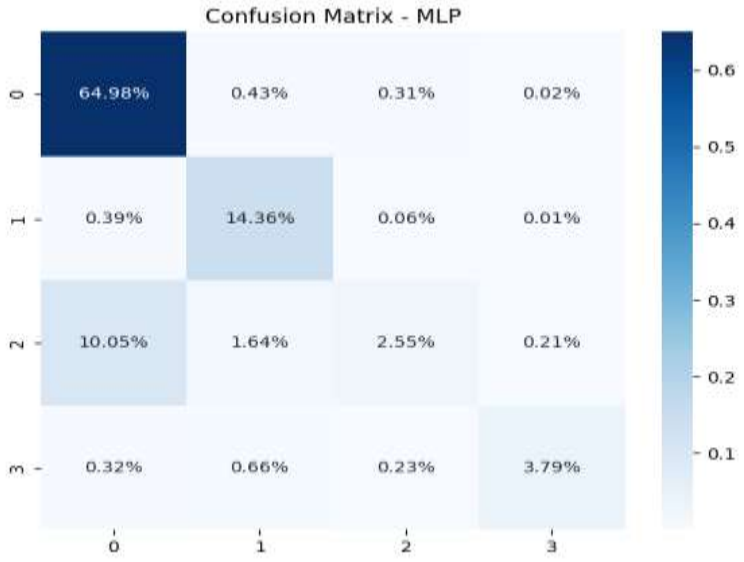
```
MLP Test Accuracy: 85.68%
4070/4070 ━━━━━━━━━━━━━━━━━━━━ 5s 1ms/step
              precision    recall  f1-score   support

           0       0.86      0.99      0.92     85621
           1       0.84      0.97      0.90     19292
           2       0.81      0.18      0.29     18822
           3       0.94      0.76      0.84      6504

    accuracy                           0.86    130239
   macro avg       0.86      0.72      0.74    130239
weighted avg       0.85      0.86      0.82    130239
```

Gambar 12. Classification Report MLP

Model MLP memiliki akurasi uji terendah, yaitu 85,68%. Untuk phishing (Kelas 2), kinerjanya sangat rendah dengan presisi 0,81, recall hanya 0,18, dan F1-score 0,29. Nilai recall yang sangat kecil menunjukkan bahwa sebagian besar URL phishing gagal dikenali. Pada kelas benign, meskipun recall tinggi (0,99), presisi rendah (0,86) menyebabkan banyak false positive dari kelas lain yang salah diklasifikasikan sebagai benign. F1-score tertinggi dicapai pada Kelas 0 (0,92) dan Kelas 1 (0,90), sedangkan Malware (Kelas 3) memiliki F1-score 0,84.



Gambar 13. Confusion Matrix MLP

Confusion matrix memperlihatkan true positive phishing hanya 2,55%, dengan false positive benign sebesar 10,05%, yang jauh lebih tinggi dibandingkan model lainnya. Hal ini mencerminkan ketidakmampuan MLP dalam membedakan URL phishing dan benign secara akurat. Karena MLP tidak memproses data secara sekuensial, seluruh informasi URL harus diubah menjadi fitur vektor datar terlebih dahulu, yang menyebabkan hilangnya konteks penting seperti urutan karakter dan struktur domain. Kelemahan ini menjadikan MLP tidak cocok untuk tugas deteksi phishing URL, khususnya tanpa rekayasa fitur lanjutan.

Keunggulan dan Keterbatasan Model RNN

Keunggulan

Penggunaan model Recurrent Neural Network (RNN) dan variannya seperti Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), serta Gated Recurrent Unit (GRU) memiliki keunggulan signifikan dalam menangani data sekuensial seperti URL, yang strukturnya kompleks dan sangat bergantung pada urutan karakter. LSTM dirancang khusus untuk mengatasi masalah vanishing gradient yang sering terjadi pada RNN konvensional, sehingga mampu mempertahankan informasi dari karakter awal hingga akhir URL. Mekanisme ini memungkinkan LSTM untuk secara efektif menangkap ketergantungan jangka panjang dalam struktur URL, menjadikannya sangat cocok untuk mendeteksi pola phishing yang tersembunyi dalam urutan karakter tersebut (Chen, Zhang, & Su, 2018).

Mengembangkan lebih jauh kemampuan LSTM, arsitektur dua arah seperti Bi-LSTM menawarkan peningkatan performa dengan memproses data secara simultan dari dua arah sekaligus (*forward* dan *backward*). Pendekatan ini memperluas pemahaman konteks sekuensial

karena model dapat mengenali pola penting yang mungkin tersembunyi di bagian tengah URL, seperti penggunaan kata-kata mengecoh atau simbol-simbol khusus yang sering dijumpai dalam serangan phishing. Studi oleh Roy *et al.* (2022) menunjukkan bahwa penggunaan Bi-LSTM dalam klasifikasi URL phising mampu menghasilkan akurasi hingga sekitar 99%, menjadikannya lebih unggul dibandingkan model LSTM satu arah dalam mengidentifikasi URL berbahaya secara presisi. Selain itu, penelitian terbaru oleh Baskota (2025) juga melaporkan bahwa arsitektur ini mencapai akurasi hingga 97 % pada dataset besar yang memuat 650 ribu URL, menegaskan efektivitas Bi-LSTM dalam skenario berskala besar.

Selain ketepatan prediksi, efisiensi komputasi juga menjadi keunggulan penting dari arsitektur RNN modern. Rangapur *et al.* (2021) dalam studi Phish-Defence menunjukkan bahwa GRU, sebagai alternatif yang lebih ringan dibanding LSTM, dirancang dengan struktur lebih sederhana sehingga memerlukan jumlah parameter yang lebih sedikit dan memberikan latensi inferensi yang lebih cepat (hanya 0,53–0,8 detik) pada perangkat seperti Raspberry Pi 4, namun tetap mempertahankan kinerja yang baik dalam memahami urutan data sekuensial. Eksperimen juga menunjukkan bahwa GRU mampu mencapai akurasi hingga 98,51%, presisi 99,08%, serta f-score dan recall yang tinggi, menjadikannya pilihan efisien untuk mendeteksi URL phishing secara real-time pada perangkat kecil seperti mobile dan Raspberry Pi.

Keterbatasan

Meskipun model RNN dan variannya menunjukkan performa tinggi dalam mendeteksi URL phishing, efektivitas model tetap sangat bergantung pada kualitas dan keragaman dataset yang digunakan selama pelatihan. Safi dan Singh (2023) dalam tinjauan literatur sistematisnya menekankan bahwa sebagian besar pendekatan phishing detection masih mengandalkan dataset dari sumber terbatas seperti PhishTank dan Alexa, yang dapat membatasi representasi variasi URL di dunia nyata. Oleh karena itu, penting untuk mempertimbangkan cakupan data pelatihan dalam mengevaluasi kemampuan model secara menyeluruh.

Di sisi lain, meskipun model RNN mampu mencapai tingkat akurasi tinggi (lebih dari 95%), tantangan dalam bentuk kesalahan klasifikasi tetap muncul. Kesalahan ini biasanya berupa false positives dan false negatives, yang umumnya terjadi ketika URL phising dirancang secara cermat agar menyerupai URL yang sah (*legitimate*), baik dalam struktur domain, subdomain, maupun penggunaan karakter simbolik. Roy *et al.* (2022) melaporkan bahwa model seperti LSTM dan Bi-LSTM masih kesulitan membedakan pola semacam itu, terutama ketika variasi phishing tersebut tidak tercakup dalam data pelatihan. Permasalahan ini semakin kompleks karena phishing modern seringkali menggunakan pendekatan yang sangat subtil dan

menyamarkan ciri-ciri khas phishing, sehingga model sekuensial seperti RNN membutuhkan augmentasi tambahan untuk dapat mengatasi jenis serangan yang semakin kreatif (Safi & Singh, 2023).

Dalam konteks tersebut, salah satu tantangan yang juga perlu diperhatikan adalah kemampuan model dalam mengenali pola lokal penting pada URL, seperti substring "login", "secure", atau "verify", yang sering kali menjadi indikator kuat adanya phishing. Studi oleh Opara *et al.* (2024) menunjukkan bahwa pendekatan berbasis Convolutional Neural Network (CNN) yang diterapkan dalam model WebPhish lebih efektif dalam mengekstraksi pola lokal tersebut secara efisien, bahkan tanpa memerlukan rekayasa fitur manual. Sejalan dengan itu, Bozkir *et al.* (2023) mengembangkan model GramBeddings yang memanfaatkan embedding karakter berbasis n-gram dan menggabungkannya dengan CNN, LSTM, dan attention mechanism. Model ini terbukti meningkatkan akurasi deteksi phishing hingga 98,27% pada kumpulan data berskala besar, menunjukkan bahwa penggunaan embedding n-gram secara signifikan memperkuat sensitivitas terhadap pola-pola lokal dalam URL. Temuan-temuan ini membuka peluang untuk merancang arsitektur hibrida yang menggabungkan keunggulan CNN dan model sekuensial seperti RNN guna meningkatkan efektivitas sistem deteksi phishing terhadap berbagai jenis pola, baik lokal maupun kontekstual.

Perbandingan Kinerja RNN dengan Model Deep Learning Lain

RNN vs CNN

Recurrent Neural Network (RNN), khususnya variannya seperti Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), dan Bidirectional LSTM (Bi-LSTM), dirancang untuk mengatasi permasalahan vanishing gradient dan menangkap ketergantungan jangka panjang dalam data sekuensial seperti struktur karakter URL. Keunggulan utama RNN terletak pada kemampuannya memahami konteks historis dalam data yang berurutan, sehingga menjadikannya cocok untuk mendeteksi pola-pola phishing yang tersembunyi dalam susunan karakter URL.

Meski demikian, pendekatan berbasis Convolutional Neural Network (CNN) juga menunjukkan performa yang kompetitif dalam deteksi phishing berbasis URL. CNN lebih unggul dalam mengekstraksi fitur lokal yang khas secara paralel melalui penggunaan kernel konvolusi, yang memungkinkan identifikasi elemen mencurigakan dalam URL (seperti kata "login", "secur", atau "verify") tanpa perlu memahami konteks urutan secara menyeluruh.

Dalam eksperimen yang dilakukan pada penelitian ini, model RNN mencatat akurasi diatas 95%. Sementara itu, model 1D CNN sebagai pembanding mencatat akurasi sebesar

94.58%. Meskipun sedikit lebih rendah dari varian RNN, hasil ini menunjukkan bahwa CNN tetap merupakan pilihan yang kuat, khususnya dalam mengekstraksi pola lokal dari URL yang kompleks.

Namun, hasil ini tidak sepenuhnya selaras dengan beberapa temuan dalam studi sebelumnya. Misalnya, dalam penelitian oleh Alshingiti *et al.* (2023), perbandingan antara arsitektur CNN, LSTM, dan kombinasi LSTM–CNN menunjukkan bahwa CNN justru mencatat akurasi tertinggi sebesar 99,2%, melampaui performa LSTM–CNN (97,6%) dan LSTM murni (96,8%). Perbedaan ini kemungkinan disebabkan oleh variasi dalam dataset, metode preprocessing, konfigurasi model, atau teknik optimasi yang digunakan.

Dengan demikian, baik RNN maupun CNN memiliki kekuatan masing-masing, dan efektivitasnya sangat bergantung pada konteks implementasi serta karakteristik data yang digunakan. Hal ini membuka peluang untuk eksplorasi model hibrida yang dapat menggabungkan kekuatan pemahaman konteks dari RNN dan keunggulan ekstraksi pola lokal dari CNN.

RNN vs Transformer/Self-Attention

Dalam beberapa tahun terakhir, arsitektur berbasis self-attention seperti Transformer telah merevolusi berbagai aplikasi pemrosesan bahasa alami, termasuk deteksi phishing URL. Berbeda dengan RNN yang memproses data secara berurutan dan cenderung kesulitan menangani urutan panjang, Transformer mampu memahami konteks secara global dalam satu langkah melalui mekanisme attention yang memperhitungkan hubungan antar semua elemen input secara langsung. Kemampuan ini memungkinkan model mengenali korelasi kompleks antar bagian URL yang berjauhan tanpa batasan urutan.

Xu (2021) melaporkan bahwa model Transformer yang dirancang khusus untuk mendeteksi phishing URL mencapai akurasi 97,3%, melampaui pendekatan tradisional berbasis RNN. Sementara itu, URLTran yang dikembangkan oleh Maneriker *et al.* (2021) menunjukkan keunggulan signifikan: pada false positive rate (FPR) rendah sebesar 0,01%, model ini mencapai true positive rate (TPR) hingga 86,8%, jauh di atas baseline RNN/CNN yang hanya mencapai 71,2%. Hasil ini menunjukkan bahwa pendekatan self-attention dan Transformer tidak hanya unggul dalam akurasi, tetapi juga dalam ketahanan terhadap kesalahan klasifikasi pada tingkat deteksi yang ketat.

RNN vs Hybrid models (CNN+Bi-LSTM+Attention)

Model hybrid, yang menggabungkan berbagai jenis arsitektur deep learning, menjadi pendekatan yang sangat menarik dan semakin banyak digunakan dalam deteksi phishing karena mampu mengkombinasikan keunggulan masing-masing komponen. CNN mampu mengekstraksi pola lokal dari string URL, sedangkan RNN seperti Bi-LSTM menangkap konteks sekuensial, dan mekanisme attention membantu menyoroti bagian-bagian penting dari input. Misalnya, studi oleh Sangeetha *et al.* (2025) menunjukkan bahwa arsitektur hybrid GRU–CNN berhasil mencapai akurasi sebesar 99%, lebih tinggi dibandingkan penggunaan CNN (98%) atau GRU (97,8%) secara terpisah.

Sebagai tambahan, Asiri *et al.* (2024), dalam pengembangan PhishTransformer yang merupakan gabungan antara CNN dan Transformer Encoder, berhasil mencapai F1-score sebesar 99%, menunjukkan efektivitas pendekatan hybrid bahkan di lingkungan deteksi yang kompleks. Temuan-temuan ini membuktikan bahwa integrasi fitur lokal, konteks global, dan mekanisme attention secara sinergis tidak hanya mampu meningkatkan akurasi, tetapi juga memperkuat ketahanan model terhadap beragam variasi serangan phishing.

Untuk memberikan gambaran yang lebih jelas, perbandingan kinerja model RNN, CNN, Transformer, dan hybrid dalam mendeteksi phishing URL dirangkum dalam tabel berikut.

Tabel 3. Kinerja Model RNN, CNN, Transformer, dan Hybrid dalam Deteksi Phishing URL

Perbandingan	Kekurangan Model RNN	Kelebihan Model Non-RNN	Kelebihan Hybrid
LSTM vs CNN	Akurasi ~96–97 %, perlu interpretasi konteks	Akurasi ~99,2 %, cepat ekstraksi lokal	-
LSTM vs Transformer	Memerlukan interpretasi sekuensial berlapis	TPR tinggi dalam FPR rendah, 97–99 % akurasi	-
GRU vs GRU–CNN	GRU terkadang kurang peka pola lokal	GRU–CNN 99 %, GRU 97,8 %	Hybrid menangkap local+context

5. KESIMPULAN

Penelitian ini mengevaluasi efektivitas model RNN, yakni LSTM, GRU, dan Bi-LSTM, dalam mendeteksi URL phishing melalui pendekatan Systematic Literature Review dan eksperimen komputasional. Hasil menunjukkan bahwa seluruh model RNN secara konsisten mengungguli model pembanding (1D CNN dan MLP), khususnya dalam mendeteksi kelas phishing. GRU mencatat akurasi tertinggi sebesar 96,29% dengan efisiensi komputasi yang

baik, sementara Bi-LSTM menampilkan stabilitas klasifikasi terbaik. LSTM juga menunjukkan performa kuat dalam mengenali pola jangka panjang, meskipun sedikit di bawah GRU dan Bi-LSTM. Sebaliknya, 1D CNN unggul dalam mengekstraksi pola lokal namun kurang mampu menangkap dependensi sekuensial penuh, dan MLP menunjukkan keterbatasan paling besar karena tidak dirancang untuk memproses urutan karakter.

Kinerja model sangat dipengaruhi oleh pra-pemrosesan URL, arsitektur yang digunakan, dan distribusi data yang timpang. Ketidakseimbangan kelas menjadi tantangan utama karena menyebabkan model bias terhadap kelas mayoritas dan mengurangi akurasi deteksi pada kelas minoritas seperti phishing dan malware. Oleh karena itu, penelitian selanjutnya disarankan untuk mengembangkan atau menggunakan dataset phishing yang lebih representatif dan bervariasi, serta menerapkan teknik penyeimbangan kelas seperti oversampling atau synthetic minority over-sampling (SMOTE). Selain itu, eksplorasi dapat diperluas pada model hybrid seperti CNN-BiLSTM atau arsitektur berbasis attention untuk meningkatkan pemahaman kontekstual URL. Implementasi sistem deteksi phishing secara real-time serta pengujian model di lingkungan operasional juga menjadi arah penting untuk menjamin efektivitas model dalam praktik nyata. Dengan demikian, pendekatan berbasis RNN tetap menjadi solusi yang relevan dan adaptif dalam menghadapi evolusi serangan phishing berbasis URL.

DAFTAR PUSTAKA

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*, *12*(1), 232. https://doi.org/10.3390/electronics12010232

Anti-Phishing Working Group. (2025, Maret 19). *Phishing Activity Trends Report: 4th Quarter 2024*. https://docs.apwg.org/reports/apwg_trends_report_q4_2024.pdf

Asiri, S., Xiao, Y., & Li, T. (2024). PhishTransformer: A Novel Approach to Detect Phishing Attacks Using URL Collection and Transformer. *Electronics*, *13*(1), 30. https://doi.org/10.3390/electronics13010030

Asiri, S., Xiao, Y., Alzahrani, S., Ju, M. (2024). PhishingRTDS: A Real-time Detection System for Phishing Attacks Using a Deep Learning Model. *Computers and Security, 141*, 103843. https://doi.org/10.1016/j.cose.2024.103843

Balogun, A. O., Adewole, K. S., & Raheem, M. O., Akande, O. N., Usman-Hamza, F. E., Mabayoje, M. A., Akintola, A. G., Asaju-Gbolagade, A. W., Jimoh, M. K., Jimoh, R. G., & Adeyemo, V. E. (2021). Improving the phishing website detection using empirical analysis of Function Tree and its variants. Heliyon, 7(7), e07437. https://doi.org/10.1016/j.heliyon.2021.e07437

Baskota, S. (2025). Phishing URL Detection using Bi-LSTM. *arXiv*. https://doi.org/10.48550/arXiv.2504.21049

Bharath, G. (2025). *A Comparative Study of Traditional and AI-Based Phishing Detection Techniques*. Insights2TechInfo. https://insights2techinfo.com/a-comparative-study-of-traditional-and-ai-based-phishing-detection-techniques/

Bozkir, A. S., Dalgic, F. C., & Aydos, M. (2023). GramBeddings: A New Neural Network for URL Based Identification of Phishing Web Pages Through N-gram Embeddings. *Computers & Security*, *124*, 102964. https://doi.org/10.1016/j.cose.2022.102964

Catal, C., Giray, G., Tekinerdogan, B., Kumar, S., & Shukla, S. (2022). Applications of deep learning for phishing detection: A systematic literature review. *Knowledge and Information Systems*, *64*(6), 1457–1500. https://doi.org/10.1007/s10115-022-01675-0

Chen, W., Zhang, W., Su, Y. (2018). Phishing Detection Research Based on LSTM Recurrent Neural Network. In: Zhou, Q., Gan, Y., Jing, W., Song, X., Wang, Y., Lu, Z. (eds) Data Science. ICPCSEE 2018. *Communications in Computer and Information Science*, *vol 901*. Springer, Singapore. https://doi.org/10.1007/978-981-13-2203-7_52

Ghojogh, B., & Ghodsi, A. (2023). Recurrent Neural Networks and Long Short-Term Memory Networks: Tutorial and Survey. arXiv preprint arXiv:2304.11461. https://arxiv.org/abs/2304.11461

Gupta, B. B., Gaurav, A., Attar, R. W., Arya, V., Alhomoud, A., & Chui, K. T. (2024). Optimized phishing detection with recurrent neural network and whale optimizer algorithm. *Computers, Materials & Continua*, *80*(3), 4895–4916. https://doi.org/10.32604/cmc.2024.050815

Halgas, L., Agrafiotis, I., & Nurse, J. R. C. (2019). Catching the Phish: Detecting Phishing Attacks using Recurrent Neural Networks (RNNs). arXiv preprint arXiv:1908.03640. https://arxiv.org/abs/1908.03640

Husain, O. (2025, Februari 6). *99 Global Phishing Statistics & Industry Trends (2023–2025)*. Control D. https://controld.com/blog/phishing-statistics-industry-trends/

Kyaw, P. H., Gutierrez, J., & Ghobakhlou, A. (2024). A Systematic Review of Deep Learning Techniques for Phishing Email Detection. *Electronics*, *13*(19), 3823. https://doi.org/10.3390/electronics13193823

Lamina, O. A., Ayuba, W. A., Adebiyi, O. E., Michael, G. E., Samuel, O. D., & Samuel, K. O. (2024). AI-Powered Phishing Detection and Prevention. Path of Science, 10(12), 112–117. https://doi.org/10.22178/pos.112-7

Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A Critical Review of Recurrent Neural Networks for Sequence Learning. arXiv preprint arXiv:1506.00019. https://arxiv.org/abs/1506.00019

Maneriker, P., Stokes, J. W., Garcia Lazo, E., Carutasu, D., Tajaddodianfar, F., & Gururajan, A. (2021). URLTran: Improving Phishing URL Detection Using Transformers. *arXiv*. https://doi.org/10.48550/arXiv.2106.05256

Mohammad, R. M., Thabtah, F., & McCluskey, L. (2015). Predicting phishing websites based on self-structuring neural network. Neural Computing and Applications, 25(2), 443–458. https://doi.org/10.1007/s00521-013-1490-z

Nezhab, M. A., & Langarib, N. (2025). Phishing Detection Techniques: A review. Journal of Computing and Applied Informatics (JoCAI), 9(1), 32-46. https://doi.org/10.32734/jocai.v9.i1-19904

Opara, C., Chen, Y., & Wei, B. (2024). Look before you leap: Detecting phishing web pages by exploiting raw URL and HTML characteristics. *Expert Systems with Applications*, *236*, 121183. https://doi.org/10.1016/j.eswa.2023.121183

Prasad, A., & Chandra, S. (2024). PhiUSIIL: A diverse security profile empowered phishing URL detection framework based on similarity index and incremental learning. *Computers & Security*, *136*, 103545. https://doi.org/10.1016/j.cose.2023.103545

Rangapur, A., Kanakam, T., & Dhanvanthini, P. (2021). Phish-Defence: Phishing Detection Using Deep Recurrent Neural Networks (Version 4). *arXiv*. https://doi.org/10.48550/arXiv.2110.13424

Rashid, F., Doyle, B., Han, S. C., & Seneviratne, S. (2024). Phishing URL detection generalisation using unsupervised domain adaptation. *Computer Networks*, *245*, 110398. https://doi.org/10.1016/j.comnet.2024.110398

Roy, S. S., Awad, A. I., Amare, L. A., Erkihun, M. T., & Anas, M. (2022). Multimodel Phishing URL Detection Using LSTM, Bidirectional LSTM, and GRU Models. *Future Internet*, *14*(11), 340. https://doi.org/10.3390/fi14110340

Safi, A., & Singh, S. (2023). A systematic literature review on phishing website detection techniques. *Journal of King Saud University - Computer and Information Sciences*, *35*(2), 590–611. https://doi.org/10.1016/j.jksuci.2023.01.004

Sangeetha, M., Navaz, K., Ravva, S. K., Roopa, R., Balaji, P., & Kumar, R. T. (2025). Enhanced Phishing URL Detection Using a Novel GRU-CNN Hybrid Approach. *Journal of Machine and Computing*, 5(1). https://doi.org/10.53759/7669/jmc202505007

Senouci, O., & Benaouda, N. (2025). Enhancing Phishing Detection in Cloud Environments Using RNN-LSTM in a Deep Learning Framework. Journal of Telecommunications and Information Technology. https://doi.org/10.26636/jtit.2025.1.1916

Singh, C., & Meenu. (2020). Phishing Website Detection Based on Machine Learning: A Survey. 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE. http://doi.org/10.1109/ICACCS48705.2020.9074400

Thakur, K., Ali, M. L., Obaidat, M. A., & Kamruzzaman, A. (2023). A Systematic Review on Deep-Learning-Based Phishing Email Detection. *Electronics*, *12*(21), 4545. https://doi.org/10.3390/electronics12214545

Vijayalakshmi, M., Shalinie, S. M., Yang, M. H., & Meenakshi, R. U. (2020). Web phishing detection techniques: A survey on the state-of-the-art, taxonomy and future directions.

IET Networks, 9(5), 235–246. https://doi.org/10.1049/iet-net.2020.0078

Xu, P. (2021). A Transformer-based Model to Detect Phishing URLs (arXiv:2109.02138). *arXiv*. https://doi.org/10.48550/arXiv.2109.02138