KLASIFIKASI PDF MALWARE PADA GARBA RUJUKAN DIGITAL (GARUDA) KEMDIKBUD DIKTI MENGGUNAKAN METODE RECURRENT NEURAL NETWORK

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat Memperoleh

Gelar Sarjana Komputer

OLEH :

Indah Ria Andina

09011382025161

JURUSAN SISTEM KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2025

LEMBAR PENGESAHAN

KLASIFIKASI PDF MALWARE PADA GARBA RUJUKAN DIGITAL (GARUDA) KEMDIKBUD DIKTI MENGGUNAKAN METODE RECURRENT NEURAL NETWORK

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer (S1)

Program Studi Sistem Komputer
Jenjang S1
Oleh :

INDAH RIA ANDINA
09011382025161

Palembang, 17 Januari 2025

Pembimbing I Tugas Akhir

Prof. Ir. Deris Stiawan, M.T., Ph.D
NIP. 197806172006041002

Pembimbing II Tugas Akhir

Nurul Afifah, M.Kom
NIP. 199211102023212049

Mengetahui,
Ketua Jurusan Sistem Komputer

Dr. Ir. Sukemi, M.T.
NIP. 196612032006041001

AUTHENTICATION PAGE

PDF CLASSIFICATION OF MALWARE ON DIGITAL REFERENCE GARUDA (GARUDA) OF MINISTRY OF EDUCATION, CULTURE AND HIGHER EDUCATION USING THE RECURRENT NEURAL NETWORK METHOD

SKRIPSI

Submitted To Complete One Of The Requirements For Obtaining A Bachelor's Degree in Computer Science

By:

INDAH RIA ANDINA

Palembang, 2 Januari 2025

Final Project Advisor I	Final Project Advisor II
Prof. Ir. Deris Stiawan, M.T., Ph.D	**Nurul Afifah, M.Kom**
NIP. 197806172006041002	**NIP. 199211102023212049**

Acknowledge.

Head Of Computer Science Departement

Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

LEMBAR PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Senin

Tanggal : 23 Desember 2024

Tim Penguji :

1. **Ketua : Kemahyanto Exaudi, S.Kom., M.T.**
2. **Penguji : Huda Ubaya, M.T.**
3. **Pembimbing I : Prof. Ir. Deris Stiawan, M.T., Ph.D.**
4. **Pembimbing II : Nurul Afifah, M.Kom**

Mengetahui,

Ketua Jurusan Sistem Komputer

Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Indah Ria Andina

NIM : 09011382025161

Judul : KLASIFIKASI PDF MALWARE PADA GARBA RUJUKAN DIGITAL (GARUDA) KEMDIKBUD DIKTI MENGGUNAKAN METODE RECURRENT NEURAL NETWORK

Hasil Pengecekan Software *iThenticatel Turnitin* : 11%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil pengjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan tugas akhir ini, maka saya menerima sanksi akademik dari Universitas Sriwijaya.

Demikian, pernyataan ini saya buat dalam keadaan sadar dan tanpa paksaan dari siapapun.

Palembang, Januari 2025

Yang Menyatakan

Indah Ria Andina

NIM. 09011382025161

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh,

Puji dan syukur penulis haturkan atas kehadirat Allah SWT, yang telah memberikan rahmat dan karunia-Nya berupa akal pikiran, ilmu pengetahuan kesehatan dan kekuatan sehingga penulis dapat menyelesaikan penyusunan tugas akhir ini dengan judul **Klasifikasi PDF Malware Pada Garba Rujukan Digital (GARUDA) Kemdikbud Dikti Menggunakan Metode Recurrent Neural Network.**

Pada penyusunan tugas akhir ini, tidak lepas dari motivasi, semangat, bimbingan dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis mengucapkan rasa syukur dan terima kasih kepada:

1. Allah Subhanahu Wata'ala yang telah memberikan rahmat dan karunia-Nya kepada penulis dalam penyusunan tugas akhir ini.
2. Orangtua tercinta (Hera Maya) yang selalu memberikan dukungan baik moral maupun finansial, semangat serta do'a yang tiada hentinya.
3. Keluarga besar penulis yang tersayang. Terima kasih atas semua kebaikan dan dukungan yang diberikan.
4. Adik satu-satunya (Chealsea Andira) yang selalu memberi semangat dan do'a.
5. Bapak Dr. Erwin, M.Si selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
6. Bapak Dr. Ir. H. Sukemi, M.T. selaku Ketua Jurusan Sistem Komputer Universitas Sriwijaya.
7. Bapak Deris Stiawan, M.T., Ph.D., IPU., ASEAN-Eng selaku Pembimbing I Tugas Akhir penulis di Jurusan Sistem Komputer yang telah meluangkan untuk membimbing dan memberikan motivasi selama kuliah dan pengerjaan Tugas Akhir.
8. Mbak Nurul Afifah, M.Kom. selaku pembimbing II yang telah membimbing penulis dalam pengerjaan Tugas Akhir dari awal pembuatan *dataset* sampai pembenaran dalam penulisan laporan Tugas Akhir.

9. Mbak Sari Anhar selaku Admin Jurusan Sistem Komputer yang baik dan ramah dalam membantu administrasi Tugas Akhir.
10. Teman- teman satu kelompok riset yang selalu memberi solusi dan semangat Viginita Putri Lestari, Riski Wahyuni, Muhammad Ramadhanil.
11. Teman-temanku tersayang yang jadi teman terdekat selama di perkuliahan, terima kasih yaa M. Rizky Juliansyah, Krisna Agustini, Chyntia Angraini, Siti Khairunnisa.
12. Kakak-kakak tingkat yang menjadi panutan, teman-teman seperjuangan Jurusan Sistem Komputer Angkatan 2020 terkhusus kelas B, serta semua orang baik yang sempat hadir dalam kehidupan penulis yang tidak dapat penulis cantumkan satu persatu.
13. Civitas Akademika Fakultas Ilmu Komputer Universitas Sriwijaya.
14. Almamater Universitas Sriwijaya.

Penulis menyadari bahwa masih ada banyak kekurangan dalam penulisan laporan tugas akhir ini. Mengingat kurangnya pengetahuan dan pengalaman penulis dalam hal ini. Oleh karena itu kritik dan saran yang mendukung sangat penting bagi penulis.

Akhir kata dengan segala keterbatasan, penulis berharap semoga laporan ini menghasilkan sesuatu yang bermanfaat bagi kita semua khususnya bagi mahasiswa Fakultas Ilmu Komputer Universitas Sriwijaya secara langsung ataupun tidak langsung sebagai sumbangan pikiran dalam peningkatanmatu pembelajaran.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Palembang, Januari 2025

Penulis

Indah Ria Andina

NIM. 09011382025161

KLASIFIKASI PDF MALWARE PADA GARBA RUJUKAN DIGITAL (GARUDA) KEMDIKBUD DIKTI MENGGUNAKAN METODE RECURRENT NEURAL NETWORK

Indah Ria Andina (09011382025161)

Jurusan sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email: ndahhyaa2002@gmail.com

ABSTRAK

Garba Rujukan Digital (GARUDA) merupakan sumber informasi yang mencakup berbagai bidang antara lain sains, psikologi perilaku, ilmu kemoralan, math, dan computer, serta publikasi ilmiah lain yang dikelola oleh Kementerian Pendidikan dan Kebudayaan serta Pendidikan Tinggi. Masalah malware lebih berbahaya bagi lembaga dan organisasi pemerintah dibandingkan bagi pengguna pribadi. Berbagai niat dilakukan oleh para penggelap. untuk melakukan kegiatan berbahaya yang dapat merugikan orang lain, seperti penyadapan atau mendapatkan hak akses tanpa sepengetahuan dan izin pemiliknya. Teknik yang digunakan dalam klasifikasi yaitu Recurrent Neural Network. Dataset yang digunakan sebanyak 10.000 dengan 21 variabel prediksi dan mendapat akurasi sebesar 99%.

Kata Kunci : Klasifikasi, Recurrent Neural Network, PDF Malware, Imbalance Dataset

PDF CLASSIFICATION OF MALWARE ON DIGITAL REFERENCE GARUDA (GARUDA) OF MINISTRY OF EDUCATION, CULTURE AND HIGHER EDUCATION USING THE RECURRENT NEURAL NETWORK METHOD

Indah Ria Andina (09011382025161)

Department of Computer Systems, Faculty of Computer Science Sriwijaya University

Email: ndahhyaa2002@gmail.com

ABSTRACT

Garba Rujukan Digital (GARUDA) is a source of information covering various fields including science, behavioral psychology, moral science, math, and computers, as well as other scientific publications managed by the Ministry of Education and Culture and Higher Education. Malware problems are more dangerous for government institutions and organizations than for private users. Various intentions are carried out by embezzlers. to carry out dangerous activities that can harm others, such as wiretapping or gaining access rights without the knowledge and permission of the owner. The technique used in the classification is Recurrent Neural Network. The dataset used is 10,000 with 21 prediction variables and gets an accuracy of 99%.

Keywords: Classification, Recurrent Neural Network, PDF Malware, Imbalance Dataset

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB I
PENDAHULUAN

1.1 Latar Belakang

Internet User semakin banyak melakukan kejahatan dengan memanfaatkan teknologi karena banyaknya kejahatan dunia maya. Kejahatan dunia maya yang digunakan oleh para penyerang adalah perangkat lunak berbahaya, yang juga disebut malware. Malware adalah program jahat yang dirancang untuk merusak atau menghancurkan perangkat lunak atau system operasi, melakukan penyadapan, memperoleh hak akses computer tanpa izin pemiliknya, manipulasi transaksi perbankan untuk mendapatkan keuntungan, mencuri informasi pribadi, menyebabkan kerugian finansial, dan merusak reputasi organisasi [2].

File PDF dapat disimpan di penyimpanan cloud yang mungkin memiliki celah yang dapat dieksploitasi oleh peretas untuk melakukan kejahatan seperti mengunduh dokumen atau program berbahaya, melakukan serangan terarah, dan mengirim email kepada orang yang tidak bertanggung jawab. File PDF dapat dieksploitasi oleh peretas dengan tujuan jahat untuk menghasilkan malware yang berisi kode mencurigakan yang disertakan dalam file tersebut. File PDF terdiri dari berbagai komponen utama seperti header, body, cross-reference table (CRT), dan trailer. Oleh karena itu, peretas tertarik untuk memasukkan berbagai jenis konten malware ke dalam file PDF [4].

Garba Rujukan Digital (GARUDA) merupakan sumber informasi yang mencakup berbagai bidang antara lain sains, psikologi perilaku, ilmu kemoralan, math, dan computer, serta publikasi ilmiah lain yang dikelola oleh Kementerian Pendidikan dan Kebudayaan serta Pendidikan Tinggi. Masalah malware lebih berbahaya bagi lembaga dan organisasi pemerintah dibandingkan bagi pengguna pribadi. Berbagai niat dilakukan oleh para penggelap untuk melakukan kegiatan berbahaya yang dapat merugikan orang lain, seperti penyadapan atau mendapatkan hak akses tanpa sepengetahuan dan izin pemiliknya[6].

Recurrent Neural Network (RNN) merupakan metode yang mampu mengetahui pola data dan membuat prediksi yang tepat untuk mengklasifikasikan jenis malware dan non-malware. RNN juga merupakan salah satu aplikasi jaringan saraf berulang yang paling banyak digunakan, metode ini dapat memproses data berurutan dan mengklasifikasikan atau memprediksi berdasarkan pengetahuan sebelumnya seperti Pengenalan teks, analisis sentimen, pengenalan ucapan [7].

Penelitian Diva [8] dalam penelitiannya mengenai "Klasifikasi Gender Berdasarkan Suara Menggunakan RNN mendapatkan hasil Dengan tingkat akurasi 90% pada data uji dan 95% pada data latih, perkembangan informasi di bidang teknologi terus berkembang pesat. Perkembangan teknologi tidak dapat dilepaskan dari beberapa faktor, antara lain sentuhan, penglihatan, dan suara. Setiap orang memiliki karakteristik yang berbeda dengan orang lain yang dapat dikenali dari suaranya. Konsep pemrosesan suara sangat berguna bagi semua sistem yang membutuhkan interaksi manusia dalam kehidupan sehari-hari. Salah satu metode penyampaian linguistik adalah klasifikasi yang berdampak langsung pada sistem penyampaian linguistik. SimpleRNN dan LSTM merupakan model deep learning yang dapat digunakan untuk mengklasifikasikan sentimen. Metode ini mampu memproses data seperti suara, video, dan teks secara berurutan.

Penelitian Asri [9] dalam penelitiannya mengenai "Prediksi Harga Saham menggunakan Metode Recurrent Neural Network" Memprediksi Akurasi 94% untuk data pelatihan dan 55% untuk data uji dicapai setelah pelatihan menggunakan RNN untuk memprediksi tujuh fitur variabel. Studi ini memprediksi harga saham dengan menggunakan analisis riwayat harga saham dan RNN. Harga terendah, tertinggi, harga buka, harga terdekat, volume, harga rata-rata, dan pergerakan adalah semua fitur yang diidentifikasi. Teknologi pemrosesan prediktif dan pembelajaran mesin saat ini dapat secara otomatis mengidentifikasi prediksi harga saham.

Penelitian Zhibin Guan [10] dalam penelitiannya mengenai “Easy Data Augmentation for Improved Malware Detection: A Comparative Study” , hasil ketepatan metode mendapatkan akurasi sebesar 94% dalam melakukan deteksi

kode berbahaya Web. Kode berbahaya dapat disematkan ke dalam aplikasi Web dengan berbagai cara, yang akan menyebabkan serangan Web berbahaya yang sering. Studi ini menunjukkan bahwa kinerja metode deteksi kode berbahaya Web berbasis RNN sangat dipengaruhi oleh cara preprocessing sumber kode.

Resampling adalah metode yang digunakan untuk mengatasi permasalahan Kumpulan data imbalance dengan menambahkan jumlah sampel kelas minor dari data duplikat hasil sampel yang sudah ada. Dengan menggunakan teknik ini, model pembelajaran mesin memiliki kondisi di mana jumlah sampel dalam sebuah dataset tidak seimbang di antara berbagai kelas atau kategori, seperti data medis, keamanan, dan deteksi anomali, di mana jumlah sampel dari satu kelas mungkin jauh lebih sedikit daripada jumlah sampel dari kelas lainnya. Ketidakseimbangan ini dapat menyebabkan model yang dilatih bias terhadap kelas mayoritas, yang menghasilkan klasifikasi kelas minoritas yang buruk [11].

Pada tugas akhir ini, dilakukan penelitian mengenai PDF Malware Garuda menggunakan metode Recurrent Neural Network untuk mengetahui malware PDF serta sumber yang dianalisis untuk digunakan dalam klasifikasi tugas akhir ini yang berjudul "**KLASIFIKASI PDF *MALWARE* PADA GARUDA KEMDIKBUD DIKTI MENGGUNAKAN METODE *RECURRENT NEURAL NETWORK*".**

1.2 Rumusan Masalah

Beberapa Rumusan masalah dari penelitian Tugas Akhir ini yaitu:

1. Bagaimana cara mengatasi dataset imbalance menggunakan Teknik Resampling untuk meningkatkan kinerja model?
2. Bagaimana metode Recurrent Neural Network dalam mengklasifikasi jenis malware benign, mal-html, dan mal-pdf dengan akurat?
3. Bagaimana performa evaluasi model Recurrent Neural Network dalam melakukan klasifikasi PDF *Malware* GARUDA?

1.3 Batasan Masalah

Untuk mencapai paparan yang diharapkan, perlu dilakukan betasan masalah sebagai berikut:

1. *Dataset* yang digunakan untuk penelitian tugas akhir adalah pdf yang berasal dari repositori GARUDA.
2. Melakukan Klasifikasi PDF benign, mal-html dan mal-pdf dengan menggunakan program python dan algoritma Recurrent Neural Network.
3. *Recurrent Neural Network* merupakan metode yang digunakan untuk mengklasifikasikan Anomali PDF Malware.

1.4 Tujuan

Beberapa Tujuan dari penelitian Tugas Akhir yaitu:

1. Menerapkan tenik Resampling pada dataset imbalance untuk meningkatkan kinerja model terbaik.
2. Menerapkan metode Recurrent Neural Network untuk mengklasifikasi jenis malware benign, mal-html, dan mal-pdf dengan akurat.
3. Menerapkan model Recurrent Neural Network untuk mendapatkan performa klasifikasi terbaik.

1.5 Manfaat

Beberapa Manfaat dari Penelitian Tugas Akhir yaitu:

1. Dapat meningkatkan jumlah sampel pada kelas minoritas menggunakan Teknik Resampling.
2. Meningkatkan akurasi dalam mengklasifikasikan Benign, Mal-html, dan Mal-pdf dalam metode Recurrent Neural Network.
3. Dapat mengetahui hasil kinerja dari klasifikasi PDF malware menggunakan metode Recurrent Neural Network.

1.6 Metodologi Penelitian

Tahap ini merupakan Metodologi penelitian yang digunakan:

1. Analisi Literatur

   Metode ini dilakukan dengan cara mencari dan mengumpulkan referensi yang berupa literature yang terdapat pada buku dan internet mengenai PDF Malware, Recurrent Neural Network, *dataset* imbalance, resampling, dan hal-hal yang dibutuhkan dalam penelitian.

2. Metode Diskusi

   Metode ini melakukan konsultasi kepada pihak-pihak yang memiliki pengetahuan serta wawasan yang baik dalam mengatasi permasalahan yang ditemui dalam penulisan tugas akhir.

3. Metode Pendataan

   Metode ini dilakukan dengan mengumpulkan dataset, yang mana dalam penelitian ini digunakan dataset dari Garba Rujukan Digital (GARUDA) Kemdikbud Dikti.

4. Analisis Data

   Metode ini dilakukan dengan menganalisis file PDF terbelih dahulu sehingga menjadi dataset yang siap diolah. Dataset yang diperoleh merupakan dataset imbalance sehingga diperlukan tahap resampling. Pada penelitian ini tahap resampling dilakukan dengan menggabungkan antara teknik oversampling dan undersampling.

5. Metode Analisis

   Metode ini menganalisis hasil pemrosesan data dan kemudian memvalidasinya untuk menarik kesimpulan.

6. Metode Kesimpulan dan Saran

   Metode ini adalah metode terakhir yang digunakan setelah mendapatkan poin penting. Metode ini selanjutnya dimaksudkan sebagai kesimpulan dari penelitian akhir ini dan dengan demikian dapat digunakan sebagai refrensi.

1.7 Sistematika Penulisan

Metode Penulisan merupakan pendekatan yang digunakan untuk menjabarkan tugas akhir:

BAB I. PENDAHULUAN

Bagian I terdapat Latar Belakang penelitian yang dilakukan, Rumusan Masalah, Batasan Masalah, Tujuan Penelitan, Manfaat Penelitian, Metodologi Penelitian dan Sistematika Penelitian.

BAB II. TINJAUAN PUSTAKA

Bagian II membahas landasan teori malware PDF, proses analisis dataset, proses resampling menggunakan resampling, klasifikasi dengan Metode RNN dan yang relevan dengan penelitian.

BAB III. METODOLOGI

Bagian III akan memaparkan Metodologi dan diagram alir pada setiap tahapan perancangan sistem dalam tugas akhir.

BAB IV. ANALISA DAN PEMBAHASAN

Bagian IV akan memaparkan hasil dari proses pengolahan data. Dari hasil tersebut akan dianalisis sehingga diperoleh data yang akurat. Analisis diproses secara manual dengan menghitung nilai-nilai pada matriks konfusi.

BAB V. KESIMPULAN DAN TINDAK LANJUT

Bab ini akan menjelaskan tentang kesimpulan yang didapat dari data penelitian yang telah dilakukan. Dan saran yang diharapkan dapat membuat penelitian ini dikembangkan lebih baik.

DAFTAR PUSTAKA

DAFTAR PUSTAKA

[1] "Malware Menggunakan Metode Malware Analisis Dinamis dan Malware Analisis Statis".

[2] T. Tsafrir, A. Cohen, E. Nir, and N. Nissim, "Efficient feature extraction methodologies for unknown MP4-Malware detection using Machine learning algorithms," *Expert Syst. Appl.*, vol. 219, no. August 2020, p. 119615, 2023, doi: 10.1016/j.eswa.2023.119615.

[3] M. Kuribayashi and K. S. Wong, "StealthPDF: Data hiding method for PDF file with no visual degradation," *J. Inf. Secur. Appl.*, vol. 61, no. June, p. 102875, 2021, doi: 10.1016/j.jisa.2021.102875.

[4] D. E. dan F. Panjaitan, "mengklasifikasi jenis malware pada metode RNN," *EAI/Springer Innov. Commun. Comput.*, vol. 23, no. 3, pp. 53–61, 2021, doi: 10.1007/978-3-030-57077-4_7.

[5] "garba rujukan digital".

[6] Y. Li, Y. Wang, Y. Wang, L. Ke, and Y. an Tan, "A feature-vector generative adversarial network for evading PDF malware classifiers," *Inf. Sci. (Ny).*, vol. 523, pp. 38–48, 2020, doi: 10.1016/j.ins.2020.02.075.

[7] B. Ghojogh and A. Ghodsi, "Recurrent Neural Networks and Long Short-Term Memory Networks: Tutorial and Survey," 2023.

[8] D. T. Adherda and M. Hikmatyar, "klasifikasi gender berdasarkan suara menggunakan RNN," vol. 17, no. 1, pp. 111–122, 2023.

[9] M. A. D. Suyudi, E. C. Djamal, and A. Maspupah, "Prediksi Harga Saham menggunakan Metode Recurrent Neural Network," *Semin. Nas. Apl. Teknol. Inf.*, pp. 1907–5022, 2019.

[10] F. W. R. Wkh *et al.*, "Easy Data Augmentation for Improved Malware Detection: A Comparative Study".

[11] Y. Pang, Z. Chen, L. Peng, K. Ma, C. Zhao, and K. Ji, "A signature-based assistant random oversampling method for malware detection," *Proc. - 2019 18th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. IEEE Int. Conf. Big Data Sci. Eng. Trust. 2019*, pp. 256–263, 2019, doi: 10.1109/TrustCom/BigDataSE.2019.00042.

[12] W. Qiang, L. Yang, and H. Jin, “Efficient and Robust Malware Detection Based on Control Flow Traces Using Deep Neural Networks,” *Computers and Security*, vol. 122. 2022. doi: 10.1016/j.cose.2022.102871.

[13] X. Wu, J. Shen, W. Zheng, L. Lin, Y. Sui, and A. O. A. Semasaba, “https://www.semanticscholar.org/paper/Klasifikasi Recurrent Neural Network-Sample RNN,” *Knowledge-Based Syst.*, vol. 279, p. 110955, 2023, doi: 10.1016/j.knosys.2023.110955.

[14] “Malware classification with fine-tune convolution neural.”

[15] J. Singh and J. Singh, “A survey on machine learning-based malware detection in executable files,” *J. Syst. Archit.*, vol. 112, no. August 2020, p. 101861, 2021, doi: 10.1016/j.sysarc.2020.101861.

[16] J. Stiborek, T. Pevný, and M. Rehák, “Multiple instance learning for malware classification,” *Expert Syst. Appl.*, vol. 93, pp. 346–357, 2018, doi: 10.1016/j.eswa.2017.10.036.

[17] N. Fleury, T. Dubrunquez, and I. Alouani, “PDF-Malware: An Overview on Threats, Detection and Evasion Attacks,” 2021.

[18] M. Cova, C. Kruegel, and G. Vigna, “Detection and analysis of drive-by-download attacks and malicious JavaScript code,” *Proc. 19th Int. Conf. World Wide Web, WWW ’10*, pp. 281–290, 2010, doi: 10.1145/1772690.1772720.

[19] C. Smutz, “Department of Computer Science Malicious PDF Detection Using Metadata and Structural Features,” 2012.

[20] A. Charim, S. Basuki, and D. R. Akbi, “Detect Malware in Portable Document Format Files (PDF) Using Support Vector Machine and Random Decision Forest,” *J. Online Inform.*, vol. 3, no. 2, p. 99, 2019, doi: 10.15575/join.v3i2.196.

[21] W. Xia, W. Zhu, B. Liao, M. Chen, L. Cai, and L. Huang, “Novel architecture for long short-term memory used in question classification,” *Neurocomputing*, vol. 299, pp. 20–31, 2018, doi: 10.1016/j.neucom.2018.03.020.

[22] R. Neural and N. Rnns, “Chapter 4,” vol. 197, pp. 117–138.

[23] W. Zhou, C. Zhu, and J. Ma, “Single-layer folded RNN for time series

prediction and classification under a non-Von Neumann architecture," *Digit. Signal Process. A Rev. J.*, vol. 147, no. February, p. 104415, 2024, doi: 10.1016/j.dsp.2024.104415.