KLASIFIKASI MALICIOUS URL PADA FILE BERBASIS HOST-BASED FEATURE EXTRACTION MENGGUNAKAN METODE BI-DIRECTIONAL LSTM

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Sistem Komputer

OLEH :

MUHAMMAD ANDIKO PUTRA
09011281823048

JURUSAN SISTEM KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2025

HALAMAN PENGESAHAN

SKRIPSI

KLASIFIKASI MALICIOUS URL PADA FILE BERBASIS HOST-BASED FEATURE EXTRACTION MENGGUNAKAN METODE BI-DIRECTIONAL LSTM

Sebagai salah satu syarat untuk penyelesaian studi di

Program Studi S1 Sistem Komputer

Oleh:

MUHAMMAD ANDIKO PUTRA

09011281823048

Pembimbing 1 : **Dr. Ir. Ahmad Heryanto, M.T.**
NIP. 198701222015041002

Mengetahui

Ketua Jurusan Sistem Komputer

Dr. Ir. Sukemi, M.T

196612032006041001

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada

Hari : Jum'at

Tanggal : 14 Maret 2025

Tim penguji :

1. Ketua : Aditya Putra Perdana Prasetyo, M.T.
2. Penguji : Dr. Ahmad Zarkasi, M.T.
3. Pembimbing I : Dr. Ir. Ahmad Heryanto, M.T.

Mengetahui, 7/6/25

Ketua Jurusan Sistem Komputer

Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Muhammad Andiko Putra

NIM : 09011281823048

Judul : Klasifikasi Malicious URL Pada File Berbasis Host-Based Feature Extraction Menggunakan Metode Bi-Directional LSTM

Hasil Pengecekan Software *iThenticate/Turnitin* : 10%

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

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

Indralaya, Juni 2025

Muhammad Andiko Putra

NIM, 09011281823048

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Alhamdulillahirobbil'alamin, Segala puji dan syukur atas kehadirat ALLAH SWT Tuhan semesta alam, karena berkat Rahmat dan Karunia-Nya lah sehingga penulis dapat menyelesaikan penyusunan Skripsi ini yang berjudul **"Klasifikasi Malicious URL Pada File Berbasis Host-Based Feature Extraction Menggunakan Metode *Bi-Directional LSTM"*.**

Dalam penyusunan skripsi ini penulis menampilkan penjelasan tentang proses daripada klasifikasi serangan Botnet hingga menampilkan hasil dan data akhir yang didapatkan. Penulis berharap agar skripsi ini dapat menjadi manfaat untuk orang banyak serta menjadi menjadi bacaan yang menarik sehingga menjadi sumber referensi penelitian lain yang mengambil tema penelitian bidang Networking.

Penulisan skripsi ini merupakan syarat akhir untuk memenuhi sebagian kurikulum dan syarat kelulusan pada Jurusan Sistem Komputer, Universitas Sriwijaya.

Pada kesempatan ini penulis mengucapkan terima kasih kepada banyak pihak yang telah memberikan bantuan, dorongan, motivasi, semangat dan bimbingan dalam penyusunan Skripsi ini :

1. ALLAH SWT Tuhan semesta alam.
2. Ayah dan Ibu yang telah sangat membantu dalam hal moral dan materil.
3. Kakak saya Sindi Ramadita yang selalu memberikan semangat dan membantu dalam hal materil.
4. Prof. Dr. Taufiq Marwa, S.E., M.Si. selaku Rektor Universitas Sriwijaya
5. Bapak Prof. Dr. Erwin, S.Si., 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 Prof. Deris Stiawan, Ph.D., IPU., ASEAN Eng. selaku Dosen Pembimbing Akademik

8. Bapak Dr. Ir. Ahmad Heryanto, M.T. selaku Dosen Pembimbing Skripsi yang telah berkenan memberikan saran dan masukan selama pembuatan skripsi ini.
9. Agung yang telah membantu dalam pembuatan program Skripsi.
10. Sahabat – sahabat Aan, Irey, Kyay, Ibob, Ale, Alung, Julian, Dharby, Rezi, Iqbal yang selalu memberikan support dan selalu ada dalam suka dan duka.
11. Teman-teman Sistem Komputer B 2018 Indralaya.

Dalam penyusunan Skripsi ini penulis menyadari sepenuhnya bahwa skripsi ini belum masuk dalam kata sempurna, oleh karena itu penulis mengharapkan saran dan kritik dari semua pihak yang berkenan demi laporan yang lebih baik lagi.

Akhir kata penulis harap semoga Skripsi ini dapat bermanfaat serta dapat memberikan pengetahuan dan wawasan bagi semua pihak yang membutuhkannya.

Palembang, Juni 2025

Muhammad Andiko Putra
NIM.09011281823048

KLASIFIKASI MALICIOUS URL PADA FILE BERBASIS HOST-BASED FEATURE EXTRACTION MENGGUNAKAN METODE BI-DIRECTIONAL LSTM

Muhammad Andiko Putra (09011281823048)
Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya
Palembang, Indonesia
Email : muhammadandikoputra@gmail.com

ABSTRAK

Perkembangan teknologi yang semakin maju, membuat serangan atau ancaman bagi pengguna internet semakin banyak jenisnya. Serangan berupa phising, malware, spyware, dan ransomware merupakan jenis serangan yang biasa menyerang pengguna internet saat ini. Serangan yang efektif saat ini yaitu dengan menggunakan URL. URL (Uniform Resource Locator) adalah sebuah alamat yang digunakan untuk menemukan lokasi dari sebuah file yang berada di Internet. Hal ini membuat URL digunakan sebagai salah satu metode untuk melakukan serangan siber disebut sebagai Malicious URL. Malicious URL atau situs berbahaya di internet memuat banyak konten berupa spam, phising, yang digunakan untuk memulai serangan. Pada penelitian ini menggunakan file PDF Garba Rujukan Digital dengan mengambil Alamat URL yang ada pada setiap file PDF dan kemudian di parsing agar mendapatkan dataset berupa data benign dan malicious. Dataset yang di dapat akan di klasifikasi menggunakan Bi-Directional LSTM (Long Short Term Memory) dengan fitur ektraksi Host-Based. Training data menggunakan perbandingan 50:50, 40:60, 30:70, 20:80 dan 10:90. Menerapkan tuning Hyperparameter saat proses training data dengan perbandingan data 50:50. Dari hasil performa klasifikasi Malicious URL terbukti berjalan dengan baik, dengan menghasilkan akurasi 93.35%, presisi 96.79%, recall 89.67%, dan spesifitas sebesar 97.03%.

Kata Kunci : URL, Uniform Resource Locator, Malicious URL, Host – Based Feature Extraction, Tuning Hyperparameter, Bi-Directional LSTM

CLASSIFICATION MALICIOUS URL ON FILE BASED ON HOST-BASED FEATURE EXTRACTION USING BI-DIRECTIONAL LSTM METHOD

Muhammad Andiko Putra (09011281823048)
Department of Computer System, Faculty of Computer Science, University of Sriwijaya
Palembang, Indonesia
Email : muhammadandikoputra@gmail.com

ABSTRACT

The advancement of technology had led to an increasing variety of attacks or threats targeting internet users. Attacks such as phishing, malware, spyware, and ransomware were the common types that targeted internet users. One of the most effective attack methods at the time was through the use of URLs. A URL (Uniform Resource Locator) was an address used to locate a file on the Internet. This made URLs one of the methods used to carry out cyberattacks, known as Malicious URLs. Malicious URLs or harmful websites on the internet contained various types of content such as spam and phishing, which were used to initiate attacks. In this study, PDF files from Garba Rujukan Digital were used by extracting the URLs contained in each PDF file, which were then parsed to create a dataset consisting of benign and malicious data. The resulting dataset was classified using a Bi-Directional LSTM (Long Short Term Memory) with host-based feature extraction. The training data was divided using various ratios: 50:50, 40:60, 30:70, 20:80, and 10:90. Hyperparameter tuning was applied during the data training process, particularly with the 50:50 data ratio. The classification performance of the Malicious URL model proved to be effective, achieving an accuracy of 93.35%, a precision of 96.79%, a recall of 89.67%, and a specificity of 97.03%.

Keyword : URL, Uniform Resource Locator, Malicious URL, Host – Based Feature Extraction, Tuning Hyperparameter, Bi-Directional LSTM

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB I
PENDAHULUAN

1.1 Latar Belakang

Dalam era teknologi yang semakin berkembang pesat saat ini, komputer digunakan untuk memudahkan pekerjaan manusia, dalam pengoperasiannya ada *software* yang berjalan diatas sistem operasi, dan sangat berperan penting dalam melakukan tugas-tugas yang dikerjakan oleh pengguna. Karena melalui *software* inilah suatu komputer dapat menjalankan perintah sehingga membantu pengguna dalam menyelesaikan pekerjaannya. Namun tidak semua *software* dapat membantu dan memudahkan manusia dalam melakukan pekerjaannya, ada pula jenis *software* yang diciptakan untuk melakukan perusakan atau tindak kejahatan yang dapat merugikan orang lain, software tersebut dikategorikan sebagai *Malicous Software*[1].

Malicous Software atau yang lebih dikenal sebagai *Malware* merupakan perangkat lunak yang secara eksplisit didesain untuk melakukan aktifitas berbahaya atau perusak perangkat lunak lainnya seperti *Trojan, Virus, Spyware* dan Exploit[2][3].

Hal ini berdampak pada perkembangan dalam dunia penyebaran informasi yang begitu masif mendorong setiap media untuk menggunakan internet dalam membagikan informasi kepada masyarakat melalui layanan berupa *website*. Sehingga membuat perkembangan *website* pada layanan internet berkembang dengan cepat. Dengan meningkatkanya layanan penyedia *website hosting* ini, dapat menciptakan celah untuk kejahatan *cyber* yang menyerang penggunanya melalui *URL* atau halaman *web* yang telah di rancang khusus, untuk menjebak pengguna internet dan memliki ciri masa hidup yang pendek. *URL* berbahaya ini, disebarkan dapat melalui email, pesan, facebook, twitter, iklan pada *website*, *file* berbentuk *docx, .pdf, dan .txt* dan media lainnya. Pada *URL* atau *website* tersebut terkandung berupa *malware, virus (ransomware),* dan juga *keylogger* yang digunakan untuk membuka pintu untuk serangan ke komputer pengguna dan kemudian mencuri data dari penggunanya.

URL (Uniform Resource Locator) digunakan untuk merujuk sebuah alamat pada halaman di internet. *URL* disajikan dalam bentuk dan karakteristik yang terbagi dalam dua komponen dasar yaitu sebagai pengidentifikasi protokol yang menunjukkan protokol apa yang sedang digunakan dan *resource name* yang digunakan dalam menentukan alamat IP atau domain dari website tersebut berada[4].

Dengan sulitnya dalam melakukan analisis, identifikasi, klasifikasi dan deteksi pada *malicious URL* dan *benign URL*, membuat beberapa penyerang mengevolusi *malware* yang mereka buat agar terhindar dari sistem pendeteksi yang menyebabkan kesulitan bagi pengguna untuk mengetahui apakah *device* mereka terinfeksi *malware* atau tidak[5]. Belum lagi dokumen-dokumen seperti PDF atau word yang didalamnya terdapat sebuah *URL* yang mengandung *malware* yang tanpa sengaja terbuka oleh pengguna dan dapat menyebabkan *device* mereka terinfeksi virus yang dapat mengambil data pribadi atau yang lebih berbahaya lagi, sampai bisa menyebarkan dan mengambil uang yang ada pada *device* yang terinfeksi *malware*[2]. Berdasarkan data pada 5 tahun terakhir ini, *malicious URL* terus mengalami peningkatan dengan pertambahan sebanyak 90% dari pengguna yang terdampak dari serangan ini[6]

Pada penelitian dengan judul *Shades of Grey: On the effectiveness of reputation-based blacklist*, penelitian ini menggunakan metode *blacklisting*, didapat hasil yang baik dengan tingkat false negatif yang tinggi dibandingkan dengan tingkat ekspetasi dari tingkat false positif dan dengan error 5%. Tetapi metode ini juga memiliki kelemahan dalam melakukan deteksi terhadap *URL* yang baru muncul diluar dari *blacklist* yang ada, selain itu juga *blacklisting* tidak terlalu mampu dalam mengatasi *malicious URL* dalam jumlah yang besar. Dan penyerang juga dapat dengan mudah memanipulasi system dengan sedikit melakukan perubahan dari satu atau lebih komponen string dari *URL* dan *URL* yang tidak ada dalam *blacklist*, maka sistem pendeteksi tidak mencurigai itu sebagai *malicious*[7].

Upaya yang dilakukan untuk mengatasi kekurangan dari metode sebelumnya adalah dengan mengembangkan pendekatan menggunakan algoritma *machine learning* dan *Deep Learning*. Ada beberapa metode *machine learning* dan *deep learning* yang dapat digunakan untuk melakukan deteksi, klasifikasi, dan

clustering, yaitu *Support Vector Machine, Decission Tree, K-nearest Neighbor, K-Means,* dan *Long Short-Term Memory*[8][9]. Metode-metode tersebut dapat memberikan kemudahan dalam melakukan fungsi prediksi yang digunakan untuk mengelompokkan URL sebagai malicious dan benign. Dan dengan pendekatan machine learning dan deep learning, dapat digunakan untuk menganalisis informasi dari sebuah URL dengan melakukan ekstraksi fitur seperti host based-feature extraction, lexical based feature extraction, dan content-based feature extraction yang menghasilkan fitur masukan yang akan digunakan dalam analisis, identifikasi, klasifikasi, dan deteksi pada malicious URL dan benign URL, Seperti URL length, domain name length, IP address, host-name URL. Berikut ini merupakan hasil penelitian menggunakan machine learning dan deep learning dengan feature extraction yang digunakan untuk mengidentifikasi malicious URL yang menjadi acuan utama dalam penelitian ini.

Pada penelitian[9] dengan judul *Empirical study on malicious URL detection using machine learning*, penelitian ini menggunakan fitur ekstraksi yaitu lexical feature dan host-based feature extraction. Pada penelitian ini menggunakan dua metode machine learning yaitu SVM (Support Vector Machine) dan Random Forest. Penelitian tersebut dilakukan dengan menggunakan 10 iterasi dan menggunkana tiga pembagian rasio 60:40, 70:30, dan 80:20. Hasil yang didapat dalam penilitian ini yaitu tingkat akurasi pada SVM lebih kecil dibandingkan Random Forest dan hasil plot dari Random Forest dengan variasi model yang dipakai 82 – 90%.

Pada penelitian[10] dengan judul menggunakan metode hybrid deep-learning yang dinamakan URLdeepDetect yang digunakan untuk melakukan analisis dan klasifikasi untuk medeteksi malicious URL. URLdeepdetect menggunakan mekanisme supervised dan unsupervised yaitu LSTM (Long Short-Term Memory) dan K-Means Clustering untuk mengklasifikasi URL. Pada penelitian ini mendapatkan hasil yang baik yaitu dengan 98.3% akurasi untuk LSTM dan 99.7% untuk KMeans Clustering. Dalam penelitian ini, hasil yang didapatkan baik, tetapi pada metode K-means Clustering masih perlu ditingkatan kembali pada bagian presisi dan recall yang masih rendah dibandingkan metode LSTM.

Dari penelitian terdahulu yang telah dijelaskan pada pembahasan sebelumnya dengan performa dan hasil dari metode yang telah digunakan, maka penulis mengangkat judul Klasifikasi Malicious *URL* Pada *File* Berbasis *Host-Based Feature Extraction* Menggunakan Metode *Bi-Directional LSTM* dengan menggunakan dataset pada file pdf garuda yang telah diparser untuk mengambil URL yang ada didalam file pdf.

1.2 Perumusan Masalah

Rumusan masalah dari pembuatan Skripsi ini antara lain :

1. Bagaimana perbedaan karakteristik yang terdapat pada benign URL atau malicious URL berdasarkan fitur-fitur yang digunakan untuk mengekstraksi URL dengan Host-Based Feature Extraction?
2. Bagaimana cara penerapan klasifikasi URL berbahaya berupa benign dan malicious URL menggunakan algoritma Bi-Directional LSTM?
3. Bagaimana pengaruh hasil akurasi, presisi, recall, dan spesifitas dari benign URL dan malicious URL ?

1.3 Batasan Masalah

Adapun batasan masalah dari pembuatan Skripsi ini antara lain :

1. Penelitian yang dilakukan berfokus untuk klasifikasi serangan malicious URL dan benign URL.
2. Menerapkan Host-Based Feature Extraction untuk Ekstraksi Fitur URL.
3. Klasifikasi dilakukan dengan menggunakan Teknik dari metode Bi-Directional LSTM.
4. Menggunakan dataset GARUDA PDF.

1.4 Tujuan

1.4.1 Tujuan Umum

Adapun tujuan umum dari pembuatan skripsi ini antara lain :

1. Guna memenuhi syarat nilai untuk lulus mata kuliah Skripsi sebanyak 6 sks bagi mahasiswa Sistem Komputer, Fakultas Ilmu Komputer.
2. Sebagai salah satu syarat lulus untuk memenuhi sebagian kurikulum dan syarat kelulusan jenjang pendidikan Strata 1 (S1) pada jurusan Sistem Komputer.
3. Menerapkan ilmu yang telah dipakai selama menjalani proses pembelajaran menjadi mahasiswa Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

1.4.2 Tujuan Khusus

Adapun tujuan khusus yang akan dicapai dari pembuatan skripsi ini antara lain:

1. Menerapkan seleksi fitur Host-Based agar mendapatkan informasi sesuai dengan fitur yang digunakan untuk mengoptimalkan proses klasifikasi serangan Malicious URL dan Benign URL.
2. Menerapkan metode Bi-Directional LSTM dalam klasifikasikan dari serangan Malicious URL.
3. Mengetahui hasil dari performa klasifikasi serangan Malicious URL dilihat dari hasil akurasi, spesifitas, recall, presisi.

1.5 Manfaat

Adapun manfaat dari pembuatan skripsi ini antara lain :

1. Menerapkan metode Bi-Directional Long Short-Term Memory untuk klasifikasi serangan Malicious URL.
2. Mengoptimalkan metode Bi-Directional Long Short Term Memory sehingga mendapatkan nilai akurasi yang tinggi.
3. Mengetahui performa hasil Bi-Directional Long Short-Term Memory untuk mengklasifikasi serangan Malicious URL.

1.6 Metedologi Penelitian

Adapun Metodologi yang akan digunakan pada pembuatan skripsi ini antara lain:

1. Tahap Pertama (Persiapan data)

   Pada tahap ini melakukan pengumpulan dataset yang akan digunakan kemudian melakukan pembelajaran dan pemahaman terhadap data yang akan diolah sehingga kebutuhan untuk topik penelitian dapat terpenuhi.

2. Tahap Kedua (Studi Pustaka dan Literatur)

   Pada tahap ini penulis akan mencari informasi-informasi dengan mengumpulkan jurnal, paper, maupun pencarian internet yang membahas tentang skema serangan Malicious URL dan penjelasan Bi-Directional LSTM yang berkenaan tentang pembuatan skripsi ini.

3. Tahap Ketiga (Pengujian)

   Pada tahap ini membangun rancangan model yang terstruktur untuk tahapan klasifikasi serangan Malicious URL dengan mentraining dataset guna mendapatkan hasil yang diharapkan.

4. Tahap Keempat (Analisis dan Kesimpulan)

   Pada tahap terakhir ini setelah mendapatkan hasil dari klasifikasi serangan Malicious URL kemudian melakukan analisis pada klasifikasi yang telah dilakukan sebelumnya dan menarik kesimpulan pada pembuatan skrisi ini.

1.7 Sistematika Penulisan

Adapun Sistematika penulisan pada skripsi ini untuk menjelaskan isi dari setiap sub bab antara lain :

BAB I. PENDAHULUAN

Dalam bab I , menjelaskan tentang latar belakang, perumusan masalah, batasan masalah, tujuan dan manfaat serta sistematika penulisan dari pembahasan topik skripsi ini yaitu Klasifikasi serangan Malicious URL dengan metode Bi-Directional LSTM.

BAB II. TINJAUAN PUSTAKA

Dalam bab II, menampilkan literature riview yang berhubungan tentang pembahasan teori serangan Malicious URL, metode Bi-Directional LSTM dan teori-teori lainnya yang berkaitan dengan skripsi ini.

BAB III. METODOLOGI

Dalam bab III, menjelaskan tahapan proses penelitian yang dilakukan secara terstruktur dengan menampilkan tahapan-tahapan pada persiapan dataset Malicious URL, kemudian penerapan metode Bi-Directional LSTM guna memenuhi tujuan dari pembuatan sripsi ini.

BAB IV, HASIL DAN ANALISIS

Dalam bab IV, menampilkan hasil yang telah didapatkan dari tahapan yang telah dilakukan, serta utuk melihat performa sistem kemudian melakukan analisis dari metode Bi-Directional LSTM.

DAFTAR PUSTAKA

[1] 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 (N Y)*, vol. 523, pp. 38–48, 2020, doi: 10.1016/j.ins.2020.02.075.

[2] G. Meng, M. Patrick, Y. Xue, Y. Liu, and J. Zhang, "Securing Android App Markets via Modeling and Predicting Malware Spread between Markets," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 7, pp. 1944–1959, 2019, doi: 10.1109/TIFS.2018.2889924.

[3] S. Rameem Zahra, M. Ahsan Chishti, A. Iqbal Baba, and F. Wu, "Detecting Covid-19 chaos driven phishing/malicious URL attacks by a fuzzy logic and data mining based intelligence system," *Egyptian Informatics Journal*, no. xxxx, 2021, doi: 10.1016/j.eij.2021.12.003.

[4] C. Do Xuan, H. D. Nguyen, and T. V. Nikolaevich, "Malicious URL detection based on machine learning," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 1, pp. 148–153, 2020, doi: 10.14569/ijacsa.2020.0110119.

[5] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, and S. Venkatraman, "Robust Intelligent Malware Detection Using Deep Learning," *IEEE Access*, vol. 7, pp. 46717–46738, 2019, doi: 10.1109/ACCESS.2019.2906934.

[6] T. Manyumwa, P. F. Chapita, H. Wu, and S. Ji, "Towards Fighting Cybercrime: Malicious URL Attack Type Detection using Multiclass Classification," in *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 1813–1822. doi: 10.1109/BigData50022.2020.9378029.

[7] J. Ispahany and R. Islam, "Detecting malicious COVID-19 URLs using machine learning techniques," in *2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events, PerCom Workshops 2021*, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 718–723. doi: 10.1109/PerComWorkshops51409.2021.9431064.

[8] R. Patgiri, H. Katari, R. Kumar, and D. Sharma, "Empirical study on malicious URL detection using machine learning," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in*

Bioinformatics), Springer Verlag, 2019, pp. 380–388. doi: 10.1007/978-3-030-05366-6_31.

[9] S. Afzal, M. Asim, A. R. Javed, M. O. Beg, and T. Baker, "URLdeepDetect: A Deep Learning Approach for Detecting Malicious URLs Using Semantic Vector Models," *Journal of Network and Systems Management*, vol. 29, no. 3, Jul. 2021, doi: 10.1007/s10922-021-09587-8.

[10] G. Palaniappan, S. Sangeetha, B. Rajendran, Sanjay, S. Goyal, and B. S. Bindhumadhava, "Malicious Domain Detection Using Machine Learning on Domain Name Features, Host-Based Features and Web-Based Features," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 654–661. doi: 10.1016/j.procs.2020.04.071.

[11] S. Aarthi, N. V. Kishan, V. Surya Teja, N. V Harsha, and V. Gupta, "Classification of Phishing Website Based on URL Features," *International Journal of Emerging Technologies in Engineering Research (IJETER)*, vol. 7, no. 5, 2019, [Online]. Available: www.ijeter.everscience.org

[12] S. Kumi, C. Lim, and S. G. Lee, "Malicious url detection based on associative classification," *Entropy*, vol. 23, no. 2, pp. 1–12, Feb. 2021, doi: 10.3390/e23020182.

[13] M. Aljabri *et al.*, "An Assessment of Lexical, Network, and Content-Based Features for Detecting Malicious URLs Using Machine Learning and Deep Learning Models," *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/3241216.

[14] J. S. Ambata, J. Gaurana, D. Jacinto, and J. De Goma, "Malicious URL Classification Using Extracted Features, Feature Selection Algorithm, and Machine Learning Techniques."

[15] John. Elder, ACM Digital Library., Association for Computing Machinery. Special Interest Group on Knowledge Discovery & Data Mining., and Association for Computing Machinery. Special Interest Group on Management of Data., *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 2009.

[16] Q. Wang, L. Li, B. Jiang, Z. Lu, J. Liu, and S. Jian, "Malicious domain detection based on k-means and smote," in *Lecture Notes in Computer Science (including*

subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Science and Business Media Deutschland GmbH, 2020, pp. 468–481. doi: 10.1007/978-3-030-50417-5_35.

[17] D. Vaishnavi, S. Suwetha, Y. B. Jinila, R. Subhashini, and S. P. Shyry, "A comparative analysis of machine learning algorithms on malicious URL prediction," in *Proceedings - 5th International Conference on Intelligent Computing and Control Systems, ICICCS 2021*, Institute of Electrical and Electronics Engineers Inc., May 2021, pp. 1398–1402. doi: 10.1109/ICICCS51141.2021.9432138.

[18] V. Vundavalli, F. Barsha, M. Masum, H. Shahriar, and H. Haddad, "Malicious URL Detection Using Supervised Machine Learning Techniques," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Nov. 2020. doi: 10.1145/3433174.3433592.

[19] M. Aldwairi and R. Alsalman, "MALURLs: Malicious URLs Classification System." [Online]. Available: http://www.mywot.com/en/download.

[20] K. D. Vara, V. S. Dimble, M. M. Yadav, and A. A. Thorat, "Based on URL Feature Extraction Identify Malicious Website Using Machine Learning Techniques," *International Research Journal of Innovations in Engineering and Technology (IRJIET)*, vol. 6, no. 3, pp. 144–148, 2022, doi: 10.47001/IRJIET/2022.603019.

[21] D. Chiba, K. Tobe, T. Mori, and S. Goto, "Detecting malicious websites by learning IP address features," in *Proceedings - 2012 IEEE/IPSJ 12th International Symposium on Applications and the Internet, SAINT 2012*, 2012, pp. 29–39. doi: 10.1109/SAINT.2012.14.

[22] A. K. Jain and B. B. Gupta, "A machine learning based approach for phishing detection using hyperlinks information," *J Ambient Intell Humaniz Comput*, vol. 10, no. 5, pp. 2015–2028, May 2019, doi: 10.1007/s12652-018-0798-z.

[23] T. Manyumwa, P. F. Chapita, H. Wu, and S. Ji, "Towards Fighting Cybercrime: Malicious URL Attack Type Detection using Multiclass Classification," in *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 1813–1822. doi: 10.1109/BigData50022.2020.9378029.

[24] F. O. Catak, K. Sahinbas, and V. Dörtkardeş, “Malicious URL Detection Using Machine Learning,” pp. 160–180, 2020, doi: 10.4018/978-1-7998-5101-1.ch008.

[25] M. Chakraborty, M. Singh, V. E. Balas, and I. Mukhopadhyay, “Lecture Notes in Networks and Systems 163 The ‘Essence’ of Network Security: An End-to-End Panorama.” [Online]. Available: http://www.springer.com/series/15179

[26] C. A. Germain, “URLs: Uniform resource locators or unreliable resource locators.”

[27] A. B. Sayamber and A. M. Dixit, “Malicious URL Detection and Identification,” 2014. [Online]. Available: https://mail.google.com/mail/#inboxIt

[28] B. Cui, S. He, X. Yao, and P. Shi, “Biographical notes: Baojiang Cui received his BS in the Hebei University of Technology, China, in 1994, MS in the Harbin Institute of Technology, China, in 1998 and PhD in Control Theory and,” 2018.

[29] E. Aghaei and G. Serpen, “Host-based anomaly detection using Eigentraces feature extraction and one-class classification on system call trace data.” [Online]. Available: https://www.researchgate.net/publication/337560479

[30] Han’guk T’ongsin Hakhoe, IEEE Communications Society, Denshi Jōhō Tsūshin Gakkai (Japan). Tsūshin Sosaieti, and Institute of Electrical and Electronics Engineers, *RNN-Based Node Selection for Sensor Networks with Energy Harvesting*.

[31] P. Zhao and X. Yang, “Opportunistic routing for bandwidth-sensitive traffic in wireless networks with lossy links,” *Journal of Communications and Networks*, vol. 18, no. 5, pp. 806–817, Oct. 2016, doi: 10.1109/JCN.2016.000109.

[32] A. Sherstinsky, “Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network,” *Physica D*, vol. 404, Mar. 2020, doi: 10.1016/j.physd.2019.132306.

[33] N. Gupta, V. Jindal, and P. Bedi, “LIO-IDS: Handling class imbalance using LSTM and improved one-vs-one technique in intrusion detection system,” *Computer Networks*, vol. 192, Jun. 2021, doi: 10.1016/j.comnet.2021.108076.

[34] M. Gerald Rizky and J. Jusak, “Analisis Perbandingan Metode Lstm Dan Bilstm Untuk Klasifikasi Sinyal Jantung Phonocardiogram,” 2021. [Online]. Available: http://jurnal.dinamika.ac.id/index.php/jcone

[35] A. Peimankar and S. Puthusserypady, “DENS-ECG: A deep learning approach for ECG signal delineation,” *Expert Syst Appl*, vol. 165, Mar. 2021, doi: 10.1016/j.eswa.2020.113911.

[36] Y. Karyadi and H. Santoso, “Prediksi Kualitas Udara Dengan Metoda LSTM, Bidirectional LSTM, dan GRU”.