PENERAPAAN RANDOM FOREST CLASSIFIER UNTUK DETEKSI PDF MALWARE PADA LAYANAN AGREGATOR GARBA RUJUKAN DIGITAL (GARUDA) KEMENDIKBUD DIKTI

SKRIPSI

OLEH:

ALFIAH NUR FATMAWATI

09011181823131

JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA

2024

Penerapan *Random Forest Classifier* untuk Deteksi PDF *Malware* pada Layanan Agregator Garba Rujukan Digital (GARUDA) Kemendikbud Dikti

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

OLEH
Alfiah Nur Fatmawati
09011181823131

JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA

2024

LEMBAR PENGESAHAN

Penerapan *Random Forest Classifier* untuk Deteksi PDF *Malware* pada Layanan Agregator Garba Rujukan Digital (GARUDA) Kemendikbud Dikti

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Satu Syarat

Memperoleh Sarjana Komputer

OLEH

Alfiah Nur Fatmawati

09011181823131

Indralaya, 27 Januari 2024

Pembimbing 1 Tugas Akhir

Prof. 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

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : **Jum'at**

Tanggal : **12 Januari 2024**

Tim Penguji :

1. **Ketua** : **Dr. Firdaus, M.Kom.**
2. **Sekretaris** : **Muhammad Ali Buchari, S.Kom., M.T.**
3. **Penguji** : **Rossi Passarella, M.Eng.**
4. **Pembimbing I** : **Prof. Deris Stiawan, M.T., Ph.D.**
5. **Pembimbing II** : **Nurul Afifah, M.Kom.**

Mengetahui,

Ketua Jurusan Sistem Komputer

Dr.Ir. Sukemi,.M.T

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Alfiah Nur Fatmawati

NIM : 09011181823131

Judul : Penerapan *Random Forest Classifier* untuk Deteksi PDF *Malware* pada Layanan Agregator Garba Rujukan Digital (GARUDA) Kemendikbud Dikti

Hasil Pengecekan Software *iThentiate/Turnitin* : 1%

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

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

Indralava, Januari 2024

Alfiah Nur Fatmawati

NIM. 09011181823131

HALAMAN PERSEMBAHAN

Bismillahirahmanirahim Tugas Akhir ini saya persembahkan untuk Bangsa dan Negara, kedua untuk Universitas Sriwijaya selaku rumah bagi saya dalam menimba ilmu. Saya ucapkan rasa Puji Syukur yang teramat besar kepada Allah swt, yang telah memberikan saya rezeki sehingga saya dapat berkuliah di Universitas Sriwijaya, saya bisa menjadi bagian dari salah satu manusia yang diberkati dari banyaknya orang yang ingin berkuliah disini, Ketiga saya ucapkan terimakaksih kepada kedua orang tua saya yang telah medoakan saya dan juga sabar atas segala hal yang saya lakukan sampai sejauh ini. Keluhan rasa sedih cinta dan doa semua membaur menjadi satu.

Hasbunallah Wa ni'mal Wakil

"Cukuplah Allah menjadi Penolong kami dan Allah adalah sebaik-baik Pelindung"

[Ali 'Imran, 3: 173]

"Optimis yakinkan diri pasti bisa"

(Alfiah Nur Fatmawati)

KATA PENGANTAR

Assalamu'alaikum Wr.Wb.

Pertama-tama kami panjatkan rasa puja dan puji syukur Alhamdulillah penulis panjatkan atas kehadirat Allah SWT yang telah memberikan karunia dan rahmat-Nya, sehingga penulis dapat menyelesaikan penulisan Tugas Akhir ini yang berjudul **"Penerapan *Random Forest Classifier* untuk deteksi PDF *Malware* pada Layanan Agregrator Garba Rujukan Digital (GARUDA) Kemendikbud Dikti)".**

Dalam laporan bertujuan untuk mengetahui hasil Performasi dan Evaluasi pada PDF Malware. Dalam penelitian ini kami menggunakan metode *Randome Forest Classifier* sebagai tools nya. Dimana hal ini meminimalisir terjadinya infeksi Malware. Tentang karakteristik malware berdasarkan hasil Analisa bahwa terdapat beberapa signature yang dapat di simpulkan. Penulis dengan ini sangat penuh harap agar tulisan ini dapat bermanfaat bagi banyak orang.

Pada kesempatan ini penulis ingin mengucapkan terima kasih kepada beberapa pihak atas ide dan saran serta bantuannya dalam menyelesaikan penulisan Tugas Akhir ini. Oleh karena itu, penulis ingin mengucapkan rasa syukurkepada Allah SWT dan terimakasih kepada yang terhormat :

1. Allah SWT, yang telah memberikan rahmat dan karunia-Nya sehingga saya dapat menyelesaikan penulisan Proposal Tugas Akhir ini dengan baik dan lancar.
2. Orang tua saya tercinta (Yon Sahono dan Nur Asiah) yang telah membesarkan saya dengan cinta dan kasih saying, dukungan dan doa serta keberkahan yang begitu sangat luar biasa sehingga berada di titik saat ini.
3. Kepada keluarga penulis, yang tersayang Nenek (Mujinah) dan kakek (Muhawi) saya yang telah wafat semoga allah tempatkan di sisi terbaiknya
4. Adik saya tercinta (Fachri Jamil) kebahagiaan menyertaimu terimakasih atas doa dan dukunganya

5. Prof. DR.Erwin, S.Si.,M.Si. selaku Dekan Fakultas Ilmu KomputerUniversitas Sriwijaya.
6. Bapak Dr. Ir. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
7. Prof. Deris Stiawan, M.T., PH.D., IPU,.ASEAN-Eng selaku Dosen Pembimbing Tugas Akhir 1 yang telah berkenan meluangkan waktunya guna membimbing, memberikan saran dan motivasi serta bimbingan terbaik untuk penulis dalam menyelesaikan Tugas Akhir ini.
8. Mbak Nurul Afifah, M.Kom. selaku Pembimbing Tugas Akhir II saya sangat berterimakasih dan bersyukur atas segala kebaikan dan bimbingan, support dari mba nurul dan berkenan membimbing saya.
9. Bapak Rossi Passarella, M.eng. selaku Pembimbing Akademik Jurusan Sistem Komputer.
10 Kak Yopi selaku admin Jurusan Sistem Komputer yang telah membantu mengurus berkas-berkas yang di perlukan.
11. Teman saya Nata Arista yang selalu mendukung saya
12. Menyadari bahwa penulis laporan ini masih sangat jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangatlah diharapkan penulis. Akhir kata penulis berharap, semoga proposal tugas akhir ini bermanfaat dan berguna bagikhalayak.

Wassalamu'alaikum Wr. Wb.

Indralaya, Januari 2024
Penulis,

Alfiah Nur Fatmawati
NIM. 09011181823131

PENERAPAN *RANDOM FOREST* CLASSIFIER UNTUK DETEKSI PDF *MALWARE* PADA LAYANAN AGREGATOR GARBA RUJUKAN DIGITAL (GARUDA) KEMENDIKBUD DIKTI

ALFIAH NUR FATMAWATI (09011181823131)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email : alfiahnurfatmawati27@gmail.com

ABSTRAK

Garba Rujukan Digital (GARUDA) merupakan salah satu Penyedia layanan yang memberikan akses informasi dan sarana pengetahuan yang di hasilkan dari akademisi penelitian yang ada di Indonesia, pada saat ini Layanan GARUDA tersebut memberikan sebuah data berbentuk file *Portable Document Format* (PDF) yang mana PDF tersebut umunya di gunakan untuk layanan berbagi informasi secara cepat dan mudah, namun banyak oknum yang tidak bertanggung jawab justru memanfaatkan hal tersebut untuk melakukan kejahatan digital *Cyber Crime*. GARUDA memberikan data sebesar 10000 PDF Malware yang akan di gunakan sebagai *Dataset* dalam penelitian ini. Dari dataset yang di miliki maka di peroleh data benign 8900, data non-malpdf 196, dan data mal-pdf 6. Virus Total dan PDFiD di gunakan Pada untuk Analisis Statis pada dataset tersebut. Penelitian ini bertujuan untuk menghasilkan Performasi dan evaluasi dari kenerja Algoritma Random Forest dalam dataset Imbalance pada dataset PDF malware GARUDA. Pada proses penyeimbangan data di lakukan dengan dua pendekatan yaitu menggunakan *Over-sampling* dengan teknik SMOTE dan Under-sampling serta penggunaan *K-Fold*. Penerapan *Random Forest Classifier* di hasilkan akurasi dengan tingkat 86,22%, Precision dengan 79,55% , Recall dengan 78,98% dan F1-Score dengan 79,26%.

Kata Kunci : *Random Forest Classifier,* PDF *Malware, Virus Total,* PDFiD, *Synthetic Minority Over-sampling Technique SMOTE*

Pembimbing I

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

Pembimbing II

Nurul Afifah, M.Kom
NIP. 199211102023212049

Ketua Jurusan Sistem Komputer

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

APPLICATION RANDOM FOREST CLASSIFIER FOR DETECTION PDF MALWARE ON AGREGATOR GARBA RUJUKAN DIGITAL SERVICE (GARUDA) KEMENDIKBUD DIKTI

ALFIAH NUR FATMAWATI (09011181823131)

Computer Engineering Department, Computer Science Faculty, Sriwijaya University

Email : alfiahnurfatmawati27@gmail.com

ABSTRACT

Garba Rujukan Digital (GARUDA) is a Service Provider whice provides information and means whice has been produced by research academics in Indonesia. Currently the service (GARUDA) Provides data in file room . Portable Document Format (PDF). Whice PDF is ussually for fast and easy information sharing service. Howaver many people are irresponsible and actually tak and advantage of this to commit digital crime Cyber Crime GARUDA have data a big and give a 10000 PDF Malware is used for dataset in research. Dataset in mine then get it data benign 8900, data non-pdf 196, and data mal-pdf 6. Virus Total and PDFiD in used for analysis statis to dataset. This research arms to produce performance and evaluation of the performance of the Random Forest algorithm in the Imbalance dataset on the GARUDA malware PDF dataset. The data Balancing process is carried out using two approaches, namely using Over-sampling with the SMOTE technique and Under-sampling as well as using K-Fold. The application of the Random Forest Classifier resulted in an accuracy rate of 86,22%, precision with 79,55%, recall with 78,98% and F1-Score with 79,26%

Key Word *: Random Forest Classifier , PDF Malware, Virus Total, PDFiD, Synthetic Minority Over-sampling Technique (SMOTE)*

First Supervisor

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

Second Supervisor

Nurul Afifah, M.Kom
NIP. 199211102023212049

Head Of Computer Engineering Department

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

DAFTAR ISI

Halaman

DAFTAR GAMBAR

Halaman

DAFTAR TABEL

Halaman

BAB I

PENDAHULUAN

1.1 Latar Belakang

Portable Document Format yang biasanya di kenal PDF yaitu sistem format berkas yang di buat oleh adobe sytem pada tahun 1993 yang di gunakan sebagai pertukaran dokumen digital. Format PDF di manfaatkan untuk mempresentasikan dokumen dua dimensi yang meliputi grafik vektor dua dimensi ,huruf,teks,citra. PDF di kenalkan pertama kali di publikasikan pada tahun 1993, pada masa itu penggunaan format PDF relatif masih rendah.[1] Tidak dapat dipungkiri bahwa, seiring dengan kemajuan teknologi, masyarakat kini mengandalkan teknologi untuk menjalankan tugas sehari-hari maupun untuk terlibat dalam bidang politik, sosial, dan akademik [1]. Teknologi dimanfaatkan dalam berbagai macam hal seperti pada dunia Pendidikan, teknologi digunakan sebagai media belajar, untuk memahami perkembangan dunia digital penggunaan nya dengan cara memanfaatkan *e-book* , *e-learning* maupun pemanfaatan *Website* pemanfaatan dari teknologi yang berkembang ini tentunya memberikan keuntungan bagi kita yang mana hidup pada zaman modern yang begitu banyak memanfaatkan sarana teknologi digital [2].

Malicious Software atau yang biasa di sebut *malware* yaitu perangkat lunak secara eksplisit yang di desain untuk menjalankan berbagai macam perusak maupun aktifivas yang berbahaya bagi perangkat lunak lainya seperti *spyware, trojan, eksploit maupun virus*[3].

Oleh karena itu, untuk memastikan apakah aplikasi yang terdeteksi adalah malware dan untuk mengidentifikasi jenis malware, analisis dan deteksi merupakan langkah penting dalam menentukan potensi konsekuensi dari eksekusi sistem berbahaya. Mengenai hal-hal kebenaran yang dapat kita ketahui dari Malware yang telah di jelaskan pentingnya di lakukan metode Analisa agar mengetahui jenis-jenis dari serangan malware agar kita dapat mengetahui ciri atas malware itu sendiri.[2]

Garba Rujukan Digital yang di singkat GARUDA, merupakan wadah yang di gunakan sebagai tempat berkumpulnya informasi dan sarana pengetahuan yang menaungi sumber informasi yang melingkupi banyak aspek karya ilmiah yang ada di negara Indonesia, yang mana hal tersebut bertujuan untuk mengelolah dan mengakses karya ilmiah dengan lengkap dan mudah, aspek tersebut meliputi komputer, matematika dan prilaku[3].

Dataset *Imbalance* memungkinkan akan mengalami penurunan kesetabilan. Perolehan hasil yang di dapatkan lebih dominan akan memberikan lebih banyak pada mayoritas kelas. Dalam beberapa hal seperti pada *Multiclass classification*. Imbalance data akan memberikan representasi data sehingga mengakibatkan data yang lemah akan di acuhkan. Near Miss dan SMOTE yang merupakan kepanjangan dari *Syntetic Minorty Oversampling Technique* adalah metode undersampling dan Oversampling di gunakan oleh banyak orang untuk menghadapi masalah seperti dataset imbalance.

Randome Forest Classifier[3] adalah metode asembling yang di gunakan sebagai klasifikasi dan regresi dan tugas lainya yang beroperasi dengan membangun keputusan untuk pelatih dalam klasifikasi yang beroperasi dengan membangun banyak pohon keputusan pada waktu pelatihan untuk tugas klasifikasi hasil hutan acak untuk regresi prediksi rata-rata dari masing-masing pohon atau rata-rata dari masing-masing regresi untuk regresi masing-masing pohon di kembalikan [1][3][4].

Kinerja algoritma *Random Forest* di bandingkan dengan *Naive Bayes dan K-Nerest* setelah melalukan proses klasifikasi dengan data set yang telah di kumpulkan menunjukkan bahawa *Random Forest* memiliki akurasi dan daya ingat yang lebih tinggi di antara metode yang lain. Hal ini menunjukkan lompatan akurasi dibandingkan penelitian

Penelitian dalam hal ini akan berkonsentrasi pada sejumlah contoh yang telah kami kumpulkan, seperti file malware PDF. Dari 10.000 kumpulan data yang diambil, hanya 197 yang berbahaya; fitur kumpulan data ini belum diketahui saat ini, oleh karena itu diperlukan penyelidikan lebih lanjut.

Analisis dan deteksi ini dilakukan untuk mencegah berbagai ancaman dan serangan yang dapat merugikan korbannya. Kerugian yang dapat ditimbulkan antara lain pencurian informasi secara kasar, peretasan, dan masuknya virus berbahaya ke

dalam perangkat pengguna.

Dengan memberikan konteks, menjadi jelas bahwa selain manfaat yang bisa kita rasakan, ada juga kelemahan yang harus kita waspadai. Dengan pengetahuan yang kita miliki, kita dapat menangkis serangan virus dengan lebih efektif. Oleh karena itu, kami mempunyai harapan yang besar agar penelitian ini dapat membantu masyarakat dan memberikan pengaruh yang baik.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah di jabarkan maka di perolehlah perumusan sebagai berikut:

1. Bagaimana teknik yang digunakan dalam mengekstrak Raw PDF dari PDF Malware Layanan Agregator GARUDA menjadi dataset.
2. Bagaimana cara penerapan untuk mengklasifikasi PDF Malware berupa *mal-pdf,benign, mal-html* menggunakan Algoritma *Random Forest Classifier*
3. Bagaimana pengaruh hasil performasi dan evaluasi dari kinerja algoritma *Random Forest Classifier* dalam dataset Imbalance pada dataset PDF Malware Layanan Agregator GARUDA.

1.3. Batasan Masalah

Batasan masalah tersebut antara lain merupakan batasan yang penulis miliki saat ini, yang bertujuan agar pembahasan tetap pada topik.

1. Data set untuk penelitian ini berasal dari PDF Malware di agregator garba rujukan digital (GARUDA) kemdikbud Dikti berjumlah 10.000

2. Menganalisa karakteristik PDF Malware hanya menggunakan metode yang diusulkan yaitu metode *Random Forest Classifier* pada PDF malware di agregator GARUDA kemdikbud Dikti
3. Tidak membahas bagaimana cara masuk dan mencegah serangan Malware pada file PDF.

1.4 Tujuan

Berdasarkan hasil penelitian ini tujuan yang akan dicapai adalah sebagai berikut :

1. Memudahkan dalam mengekstrak Raw data pada PDF malware di agregator Garba Rujukan Digital (GARUDA) kemdikbud Dikti menjadi data yang di olah agar menjadi dataset.
2. Penerapan Random Forest Classifier untuk klasifikasi PDF malware berupa *mal-pdf,benign, mal-html.*
3. Pengaruh hasil performasi dan evaluasi dari kinerja algoritma Random Forest dalam dataset Imbalance pada dataset PDF malware GARUDA.

1.5 Manfaat

Adapun manfaatnya yang dapat di peroleh dari penelitian Skripsi ini sebagai berikut;

1. Kumpulan data Raw PDF malware yang telah di ekstraksi akan di ubah menjadi dataset
2. Memahami dan mampu mengidentifikasi malware berdasarkan kategorisasi PDF Malware Referensi Digital Kementerian Pendidikan dan Kebudayaan. Agregator GARUDA Garba menggunakan teknik Random Forest Classifier
3. Memperoleh hasil yang tepat dan optimal dalam mendekteksi PDF malware

1.6. Manfaat bagi kampus

Adapun manfaatnya yang dapat di peroleh dari penelitian Skripsi ini sebagai berikut;

1. Memahami permasalahan mengenai penyerangan Malware.
2. Menganalisa karakteristik malware yang ada pada PDF Malware di agregator GARUDA kemdikbud Dikti dengan Metode *Random Forest Classifier*
3. Menambah wawasan dan pengetahuan yang bisa digunakan sebagai acuan dalam Deteksi Malware pada suatu instansi, serta kampus yang mana bisa bermanfaat sebagai bahan ajar agar ilmu secara teori.
4. Tidak membahas tentang bagaimana mencegah serangan Malware.

1.7 Metode Penelitian

Metedologi yang akan digunakan dalam tugas akhir ini akan melewati tahapan sebagai berikut :

1. Metode studi pustaka/literatur

   Pada tahap ini melakukan literatur riview yang berkaitan tentang Malware menggunakan metode Random Forest Classifier, serta untuk menyelesaikannya dibantu melalui jurnal internasional, buku dan internet yang berkaitan dengan tugas akhir.

2. Metode Konsultasi

   Metode konsultasi pada penelitian ini melakukan konsultasi kepada Dosen pembimbing serta semua orang yang mempunyai pengetahuan dan wawasan pada saat terdapat permasalahan dalam melakukan tugas akhir.

3. Metode Pengumpulan Data

   Pada metode kali ini, database yang digunakan Dataset PDF Malware GARUDA sebanyak 10.000 pdf yang di ektraksi menggunakan Virus taotal.

4. Metode Observasi

   Pada metode kali ini melakukan pengamatan dan pencatatan dan pengumpulan terhadap data-data yang diproleh.

5. Metode Perancangana Software

Dilakukan perancangan pada tahap ini agar dapat memudahkan menganalisa serta deteksi malware pada PDF malware.

6. Metode Pengujian

Tahapan selanjutnya adalah pengujian pada Malware apakah malware tersebut berbahaya atau tidak karena pada malware yang telah di ekstraksi hanya terdapat sedikit malware yang sesungguhnya

7. Metode Analisa dan Kesimpulan

Menganalisa hasil dari pengujian metode sebelumnya telah di lakukan untuk dapat memahami karakteristik serta yang di timbulkan dari malware agar dilakukannya pengembangan untuk penelitian kedepannya..

1.8 Lingkungan perangkat keras dan perangkat lunak

Dalam tugas akhir ini perangkat lunak yang digunakan adalah Kali LINUX, Virtual Box dan Virus Total. Sedangkan perangkat keras yang digunakan adalah pc atau laptop.

1.9 Sistematika Penulisan

Sistematika pada penulisan yang digunakan dalam tugas akhir ini adalah sebagai berikut yang mana tinjauan ini di gunakan untuk mendeskripsikan subbab-bab yang tersusun. Sebagaimana berikut susunannya.

BAB I. PENDAHULUAN

Pada bab I ini akan berisi latar belakang masalah, tujuan dan manfaat serta metodologi penelitian dan sistematika penulisan mengenai Malware.

BAB II. TINJAUAN PUSTAKA

Pada bab II akan berisi dasar teori dan literatur review malware yang diteliti oleh peneliti lain dengan menggunakan metode yang beragam. Pada bab ini juga akan menjelaskan kelemahan dari metode yang digunakan pada penelitian lain.

BAB III. METODOLOGI PENELITIAN

Pada bab III akan membahas Analisis dan Deteksi Malware menggunakan metode *Random Forest Classifier*.

BAB IV. HASIL DAN ANALISA

Pada bab IV membahas proses yang serta hasil Analisa dan Pendeteksian Malware perangkat lunak menggunakan metode *Random Forest Classifier* untuk mendapatkan hasil Malware yang di butuhkan.

BAB V. KESIMPULAN DAN SARAN

Pada bab V nantinya akan berisi kesimpulan yang diambil dari bab sebelumnya mengenai hasil Analisa deteksi PDF Malware. Dan juga berisi saran yang dapat digunakan untuk perhatian pada penelitian selanjutnya.

DAFTAR PUSTAKA

[1] W. Deng, Z. Huang, J. Zhang, and J. Xu, "A Data Mining Based System for Transaction Fraud Detection," *2021 IEEE Int. Conf. Consum. Electron. Comput. Eng. ICCECE 2021*, pp. 542–545, 2021, doi: 10.1109/ICCECE51280.2021.9342376.

[2] S. D. S. K. Virgiawan A. Manoppo, Arie S. M. Lumenta, "Analisa Malware Menggunakan Metode Dynamic Analysis Pada Jaringan Universitas Sam Ratulangi," *J. Tek. Elektro Dan Komput.*, vol. 9, no. 3, pp. 181–188, 2020.

[3] F. C. C. Garcia, "Random Forest for Malware Classification," pp. 1–4.

[4] N. Afifah, D. Stiawan, and S. Nurmaini, "The Implementation of Deep Neural Networks Algorithm for Malware Classification," *Comput. Eng. Appl.*, vol. 8, no. 3, pp. 189–202, 2019.

[5] T. A. Cahyanto, V. Wahanggara, and D. Ramadana, "Analisis dan Deteksi Malware Menggunakan Metode Malware Analisis Dinamis dan Malware Analisis Statis," *Justindo, J. Sist. Teknol. Inf. Indones.*, vol. 2, no. 1, pp. 19–30, 2017, [Online]. Available: http://jurnal.unmuhjember.ac.id/index.php/JUSTINDO/article/view/1037

[6] S. Madan, S. Sofat, and D. Bansal, "Tools and Techniques for Collection and Analysis of Internet-of-Things malware: A systematic state-of-art review," *J. King Saud Univ. - Comput. Inf. Sci.*, no. 2022, doi: 10.1016/j.jksuci.2021.12.016.

[7] A. Beaudet, C. Escudero, and É. Zamaï, "Malicious anomaly detection approaches robustness in manufacturing ICSs," *IFAC-PapersOnLine*, vol. 54, no. 1, pp. 146–151, 2021, doi: 10.1016/j.ifacol.2021.08.016.

[8] A. S. Rusdi, N. Widiyasono, and H. Sulastri, "Analisis Infeksi Malware Pada Perangkat Android Dengan Metode Hybrid Analysis," *Jl. Siliwangi No*, vol. 46115, no. 24, 2019.

[9] Z. Shen, M. U. Rehman, W. Chen, Y. Liu, J. Liu, and T. Zhong, "A Method

based on Modified PageRank-Algorithm for Measuring and Rating Android Malwares," *Procedia Comput. Sci.*, vol. 174, pp. 252–255, 2020, doi: 10.1016/j.procs.2020.06.081.

[10] R. Blanquero, E. Carrizosa, P. Ramírez-Cobo, and M. R. Sillero-Denamiel, "Variable selection for Naïve Bayes classification," *Comput. Oper. Res.*, vol. 135, p. 105456, 2021, doi: 10.1016/j.cor.2021.105456.

[11] N. Bala, A. Ahmar, W. Li, F. Tovar, A. Battu, and P. Bambarkar, "DroidEnemy: battling adversarial example attacks for Android malware detection," *Digit. Commun. Networks*, 2021, doi: 10.1016/j.dcan.2021.11.001.

[12] T. Olsson, M. Ericsson, and A. Wingkvist, "To automatically map source code entities to architectural modules with Naive Bayes," *J. Syst. Softw.*, vol. 183, p. 111095, 2022, doi: 10.1016/j.jss.2021.111095.

[13] M. H. Junejo, A. A. H. Ab Rahman, R. A. Shaikh, K. M. Yusof, D. Kumar, and I. Memon, "Lightweight Trust Model with Machine Learning scheme for secure privacy in VANET," *Procedia Comput. Sci.*, vol. 194, pp. 45–59, 2021, doi: 10.1016/j.procs.2021.10.058.

[14] R. Upadhyay, U. R. Bhatt, and H. Tripathi, "DDOS Attack Aware DSR Routing Protocol in WSN," *Phys. Procedia*, vol. 78, no. December 2015, pp. 68–74, 2016, doi: 10.1016/j.procs.2016.02.012.

[15] J. Kim and P. R. Kumar, "Security of control systems with erroneous observations," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 2225–2230, 2020, doi: 10.1016/j.ifacol.2020.12.008.

[16] X. Xie, Q. Lu, A. K. Parlikad, and J. M. Schooling, "Digital twin enabled asset anomaly detection for building facility management," *IFAC-PapersOnLine*, vol. 53, no. 3, pp. 380–385, 2020, doi: 10.1016/j.ifacol.2020.11.061.

[17] V. Syrris and D. Geneiatakis, "On machine learning effectiveness for malware detection in Android OS using static analysis data," *J. Inf. Secur. Appl.*, vol. 59, no. May, p. 102794, 2021, doi: 10.1016/j.jisa.2021.102794.

[18] S. Farhana, "Classification of Academic Performance for University Research Evaluation by Implementing Modified Naive Bayes Algorithm," *Procedia*

Comput. Sci., vol. 194, pp. 224–228, 2021, doi: 10.1016/j.procs.2021.10.077.

[19] S. R. T. Mat, M. F. A. Razak, M. N. M. Kahar, J. M. Arif, and A. Firdaus, "A Bayesian probability model for Android malware detection," *ICT Express*, no. xxxx, 2022, doi: 10.1016/j.icte.2021.09.003.

[20] M. K. Ishak and F. K. Khan, "Unique message authentication security approach based controller area network (can) for anti-lock braking system (abs) in vehicle network," *Procedia Comput. Sci.*, vol. 160, pp. 93–100, 2019, doi: 10.1016/j.procs.2019.09.448.

[21] Hubert, P. Phoenix, R. Sudaryono, and D. Suhartono, "Classifying Promotion Images Using Optical Character Recognition and Naïve Bayes Classifier," *Procedia Comput. Sci.*, vol. 179, no. 2020, pp. 498–506, 2021, doi: 10.1016/j.procs.2021.01.033.

[22] J. Yuste, E. G. Pardo, and J. Tapiador, "Optimization of code caves in malware binaries to evade Machine Learning detectors," *Comput. Secur.*, p. 102643, 2022, doi: 10.1016/j.cose.2022.102643.

[23] N. Zakeya, K. Ségla, T. Chamseddine, and B. B. Alvine, "Probing AndroVul dataset for studies on Android malware classification," *J. King Saud Univ. - Comput. Inf. Sci.*, no. 2021, doi: 10.1016/j.jksuci.2021.08.033.

[24] W. Casbolt, I. Esnaola, and B. Jones, "Denial of service attacks on control systems with packet loss," *IFAC-PapersOnLine*, vol. 53, pp. 3488–3495, 2020, doi: 10.1016/j.ifacol.2020.12.1699.

[25] N. Al Sarah, F. Y. Rifat, M. S. Hossain, and H. S. Narman, "An Efficient Android Malware Prediction Using Ensemble machine learning algorithms," *Procedia Comput. Sci.*, vol. 191, no. 2019, pp. 184–191, 2021, doi: 10.1016/j.procs.2021.07.023.

[26] A. Yudhana, D. Sulistyo, and I. Mufandi, "GIS-based and Naïve Bayes for nitrogen soil mapping in Lendah, Indonesia," *Sens. Bio-Sensing Res.*, vol. 33, p. 100435, 2021, doi: 10.1016/j.sbsr.2021.100435.

[27] P. Bhat and K. Dutta, "A multi-tiered feature selection model for android malware detection based on Feature discrimination and Information Gain," *J.*

King Saud Univ. - Comput. Inf. Sci., 2021, doi: 10.1016/j.jksuci.2021.11.004.

[28] Y. Cahyaningrum, I. R. Widiasari, and J. O. Notohamidjojo, "Analisis Performa Container Berplatform Docker atas Serangan Malicious Software (Malware)," pp. 47–54, 2020.

[29] A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Creat. Inf. Technol. J.*, vol. 2, no. 3, pp. 207–217, 2015.

[30] G. Fiore, E. De Santis, and M. D. Di Benedetto, "Secure Mode Distinguishability for Switching Systems Subject to Sparse Attacks," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 9361–9366, 2017, doi: 10.1016/j.ifacol.2017.08.1442.