VISUALISASI SERANGAN *MALWARE SPYWARE* DENGAN MENGGUNAKAN METODE *RANDOM FOREST*

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

OLEH:

Amelia Pramudita
09011381924085

JURUSAN SISTEM KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA

2023

LEMBAR PENGESAHAN

VISUALISASI SERANGAN MALWARE SPYWARE DENGAN MENGGUNAKAN METODE *RANDOM FOREST*

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh

Amelia Pramudita
09011381924085

Palembang, 16 Juni 2023

Mengetahui,

Ketua Jurusan Sistem Komputer	**Pembimbing Tugas Akhir**
Dr.Ir. Sukemi, M.T.	**Deris Stiawan, M.T., Ph.D.**
NIP. 196612032006041001	**NIP. 197806172006041002**

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada

Hari : Kamis

Tanggal : 25 Mei 2023

Tim Penguji

1. Ketua : Ahmad Fali Oklilas, M.T
2. Sekretaris : Nurul Afifah, M.Kom
3. Pembimbing : Deris Stiawan, M.T., Ph.D
4. Penguji : Ahmad Heryanto, M.T

Mengetahui, 16/6/23

Ketua Jurusan Sistem Komputer

Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini

Nama : Amelia Pramudita

NIM : 09011381924085

Judul : Visualisasi Serangan Malware *Spyware* Dengan Menggunakan Metode *Random Forest*

Hasil Pengecekan Plagiat/Turnitin: 8%

Menyatakan bahwa laporan tugas akhir ini merupakan hasil karya saya sendiri dan tidak mengandung unsur penjiplakan atau plagiat. Apabila terbukti adanya penjiplakan atau plagiat dalam laporan tugas akhir ini, maka saya siap menerima sanksi akademik dari Universitas Sriwijaya.

Demikian pernyataan ini saya buat dengan kesadaran penuh dan tanpa adanya paksaan dari pihak manapun.

Palembang, Juni 2023

Yang menyatakan

Amelia Pramudita

NIM. 09011381924085

HALAMAN PERSEMBAHAN

MOTTO

"Sesungguhnya sesudah kesulitan itu ada kemudahan"

(Q.S Al-Insyirah : 6)

"Nobody else can change your life unless yourself"

Tidak akan ada yang bisa mengerti perjuangan atas masa sulitnya kita, yang mereka ketahui hanyalah bagian suksesnya. Berjuanglah untuk diri sendiri agar kelak di masa depan akan sangat bangga dengan apa yang telah kita perjuangkan.

"Tidak ada kesuksesan tanpa kerja keras. Tidak ada keberhasilan tanpa kebersamaan. Tidak ada kemudahan tanpa doa."

–– Ridwan Kamil ––

Dengan mengucapkan syukur Alhamdulillah atas rahmat Allah SWT,
skripsi ini saya persembahkan untuk :

Orang tua saya tercinta yang telah memberikan dukungan baik moril maupun materil dan selalu memanjatkan doa yang luar biasa untuk anaknya selama ini. Terima kasih kepada keluarga, saudara serta nenek saya yang selalu mendukung dan bekerja keras dalam mendidik saya. Saya sangat berterima kasih kepada kalian semua yang telah mendorong saya dengan motivasi dan juga semangat yang diberikan hingga saya berada di titik ini.

Last but not least, I wanna thanks me, I wanna thanks me for believing in me, I wanna thanks me for doing all this hard work, I wanna thanks me for having no days off, I wanna thanks me for never quitting, for just me at all times.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh,

Puji dan syukur Alhamdullilah penulis ucapkan kehadirat Allah SWT karena rahmat dan karunia-Nya penulis dapat menyelesaikan penulisan Proposal Tugas Akhir ini yang berjudul **"Visualisasi Serangan Malware Spyware Dengan Menggunakan Metode *Random Forest*"**.

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

1. Allah SWT yang telah memberikan berkah serta nikmat kesehatan dan kesempatan sehingga saya dapat menyelesaikan penulisan Proposal Tugas Akhir ini dengan baik dan lancar.
2. Kedua Orang tua dan keluarga yang sangat saya sayangi, yang telah membesarkan, mendukung, dan mendidik saya dengan kasih sayang. Terima kasih untuk segala do'a, motivasi dan dukungannya baik moril, materil maupun spiritual selama ini.
3. Bapak Jaidan Jauhari, S.Pd., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Dr. Ir. H. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Bapak Deris Stiawan, M.T., Ph.D., IPU., ASEAN ENG. selaku Dosen Pembimbing Tugas Akhir serta Dosen Pembimbing Akademik yang telah berkenan meluangkan waktunya untuk membimbing, memberikan saran dan motivasi serta bimbingan terbaik untuk penulis dalam menyelesaikan Tugas Akhir ini.
6. Mbak Nurul Afifah M.Kom yang telah memberikan bimbingan dan saran selama penulis menyelesaikan Tugas Akhir ini.

7. Mbak Sari selaku admin Jurusan Sistem Komputer yang telah membantu mengurus seluruh berkas.
8. Sartika dan Zarah Fitriani yang memberikan dukungan dan saran selama proses penyusunan laporan skripsi.
9. Grup riset COMNETS.

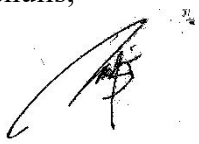
Penulis menyadari bahwa laporan tugas akhir ini masih sangat jauh dari kata sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan penulis agar penulisan laporan ini dapat menjadi lebih baik lagi dan dapat dijadikan sumber referensi yang bermanfaat dan berguna untuk khalayak.

Akhir kata penulis mengharapkan agar laporan tugas akhir ini dapat menghasilkan sesuatu yang bermanfaat, khususnya bagi Fakultas Ilmu Komputer Universitas Sriwijaya secara langsung ataupun tidak langsung sebagai sumbangan pikiran dalam peningkatan mutu pembelajaran dan penelitian.

Wassalamu'alaikum Warahmatullah Wabarakatuh.

Palembang, Juni 2023

Penulis,



Amelia Pramudita

NIM. 09011381924085

VISUALISASI SERANGAN MALWARE SPYWARE DENGAN MENGGUNAKAN METODE RANDOM FOREST

AMELIA PRAMUDITA (09011381924085)
Jurusan Sistem Komputer, Fakultas Ilmu Komputer. Universitas Sriwijaya
Email : ameliadita69@gmail.com

ABSTRAK

Spyware merupakan salah satu dari jenis *malware* yang bertujuan untuk mengumpulkan informasi dan data penting seperti informasi keuangan dan kata sandi tanpa izin dan mengirimkan mereka kepada penyerang. Teknik visualisasi diperlukan untuk dapat mempermudah dalam menganalisa pola serangan dan karakteristik dari spyware. Penelitian ini menggunakan algoritma *Random Forest*. Dataset yang digunakan berasal dari CIC-MalMem2022 dengan jenis data benign dan *Spyware Gator* dalam bentuk .csv. Penelitian ini menerapkan seleksi fitur yang bertujuan untuk menemukan fitur – fitur relevan dan mengurangi fitur yang tidak relevan dengan menggunakan *Correlation Based Feature Selection.* Proses seleksi fitur ini menghasilkan 7 fitur, 16 fitur dan 31 fitur relevan yang akan di visualisasikan dengan menggunakan diagram garis *parallel coordinates*. Hasil validasi dari ketiga jumlah fitur yang berbeda dengan menggunakan *stratified k-fold* menghasilkan akurasi terbaik untuk 7 fitur pada 6-Fold sebesar 99,94%. Hasil akurasi terbaik untuk 16 fitur terdapat pada 4-Fold yaitu sebesar 99,97% dan hasil akurasi terbaik untuk 31 fitur terdapat pada 9-Fold yaitu sebesar 99,98%.

Kata Kunci : *Spyware, Random Forest,* Visualisasi, *RandomizedSearchCV, Stratified K-Fold, Confusion Matrix*

VISUALIZATION OF SPYWARE MALWARE ATTACKS USING RANDOM FOREST METHOD

AMELIA PRAMUDITA (09011381924085)

Department of Computer Systems, Computer Science Faculty,

Sriwijaya University

Email : ameliadita69@gmail.com

ABSTRACT

Spyware is a type of malware that aims to collect important information and data such as financial information and passwords without permission and send them to the attacker. Visualization techniques are needed to make it easier to analyze attack patterns and characteristics of spyware. This study used the Random Forest algorithm. The dataset is from CIC-MalMem2022 with benign data types and Spyware Gator in .csv form. In this study applied feature selection to find relevant features and reduce irrelevant features by using Correlation Based Feature Selection. This feature selection process produces 7 features, 16 features and 31 relevant features that will be visualized using parallel coordinates line diagrams. The validation results of the three different number of features using stratified k-fold produce the best accuracy for 7 features at 6-Fold is 99.94%. The best accuracy results for 16 features are found at 4-Fold that is 99.97% and the best accuracy results for 31 features are found at 9-Fold that is 99.98%.

Kata Kunci : *Spyware, Random Forest, Visualization, RandomizedSearchCV, Stratified K-Fold, Confusion Matrix*

DAFTAR ISI

Halaman

DAFTAR GAMBAR

DAFTAR TABEL

BAB I
PENDAHULUAN

1.1 Latar Belakang

Malware atau *malicious software* merupakan suatu program berbahaya yang berjalan didalam sebuah sistem komputer tanpa izin yang bertujuan untuk mencuri data pribadi dan juga dapat merusak sistem[1]. Pada penelitian[2] menyebutkan bahwa terlepas dari peningkatan yang signifikan dari mekanisme *cyber security* dan evolusi berkelanjutan, malware masih menjadi salah satu ancaman paling efektif di *cyber spac*e. Dalam penelitian [1] mengatakan bahwa malware dapat dikategorikan ke dalam beberapa jenis, seperti virus, *worm, root-kit, trojan horses, backdoor, spyware, adware, botnet* dan lainnya berdasarkan perilaku dan pola penyebaranya.

Spyware adalah program perangkat lunak berbahaya yang bertujuan untuk mengumpulkan data penting dan berharga seperti kata sandi dan keuangan informasi tanpa izin dari pemilik dan mengirimkan mereka kepada penyerang[3]. Spyware juga merupakan langkah "pengintaian" awal untuk serangan canggih berturut-turut. Menurut penelitian[4], *spyware* sering dikaitkan dengan ancaman yang lebih berbahaya seperti penyadapan dunia maya yang menargetkan eksekutif puncak, serangan yang dapat mengambil sejumlah besar uang dari perusahaan, dan pencurian data pribadi untuk tujuan pemerasan.

Penelitian ini menggunakan metode *Random Forest* sebagai algoritma pengklasifikasian. *Random forest* adalah pengklasifikasi dengan keputusan berganda pohon. Ini memiliki model yang fleksibel dan pelatihan yang cepat. Itu bisa memecahkan kesalahan klasifikasi yang disebabkan oleh *unbalanced data* [5]. *Random Forest* merupakan suatu algoritma yang memakai teknik pembagian kelas biner berulang demi mendapat titik final di dalam struktur *decision tree* berdasar dari klasifikasi serta regresi *tree* [6]. Model ini memiliki beberapa kelebihan yaitu dapat memberikan hasil error yang rendah, hasil performa yang

bagus dalam proses klasifikasi, dan mampu mengatasi data *training* secara efektif dengan jumlah yang besar, serta metode efektif untuk memprediksi *missing data* [7].

Dari penelitian yang sebelumnya [8], dilakukan identifikasi *Malicious Web* menggunakan metode Random Forest. Hasil yang di dapat dari penelitian ini menghasilkan nilai precision 94%, nilai recall 94%, nilai F1-score adalah 93%, dan memiliki nilai *support* nya adalah 97%. Pada penelitian yang dilakukan [4], membahas bagaimana mengklasifikasikan malware *spyware* menggunakan beberapa algoritma. Dalam proses klasifikasinya menggunakan 10-fold *Cross Validation* yang diulang sebanyak 30 kali memberikan hasil bahwa algoritma Random Forest memiliki nilai skor F1 terbaik yaitu 97.6%, disusul olah 97.1% untuk Decision Tree, 86.9% untuk algoritma KNN, 78.2% untuk Naïve Bayes, dan 42.3% untuk algoritma SVM.

Pada penelitian [9], algoritma Random Forest digunakan untuk mendeteksi malware *Ransomware*. Hasil akurasi yang di dapat dari penelitian ini cukup tinggi yaitu 97.74%. Dalam penelitian [10], dilakukan klasifikasi binary dengan pendekatan big data untuk deteksi malware menggunakan analisis memori. Dalam penelitian ini menggunakan dataset CIC-MalMem-2022 yang berasal dari *Canadian Institute for Cybersecurity* pada tahun 2022. Penelitian tersebut menggunakan *machine learning* dan *deep learning* untuk menganalisis memori dan mendeteksi malware. Hasil dari model tersebut terlihat bahwa semua model mencapai performa tinggi dalam klasifikasi malware. Di antara algoritma tersebut, Regresi Logistik adalah algoritma dengan kinerja terbaik dengan akurasi 99,98%. Kemudian diikuti dengan metode Random Forest yang memiliki akurasi sebesar 99,97%.

Berdasarkan latar belakang dan ulasan penelitian sebelumnya yang telah disebutkan diatas, maka pada tugas akhir ini penulis akan membahas tentang bagaimana visualisasi serangan malware *spyware*, yang mana penelitian ini diberi judul “Visualisasi Serangan *Malware* S*pyware* dengan Menggunakan Metode

Random Forest". Penggunaan metode Random Forest ini akan berguna untuk klasifikasi serangan malware spyware Gator dan memberikan visual dari pola serangannya.

1.2 Rumusan Masalah

Perumusan masalah dari penulisan Proposal Tugas Akhir ini adalah Penelitian ini akan membahas

1. Bagaimana cara mengklasifikasikan yang mana data *spyware Gator* dan yang mana data *benign*?
2. Bagaimana cara untuk memilih fitur terbaik yang dapat digunakan untuk membuat proses komputasi lebih cepat?
3. Bagaimana memvisualisasikan serangan *spyware Gator* berdasarkan fitur dan atribut agar lebih mudah di pahami?

1.3 Tujuan

Tujuan yang akan dicapai dalam penelitian ini adalah sebagai berikut :

1. Menerapkan metode Random Forest untuk mengklasifikasi malware *spyware Gator* dan data benign pada penelitian ini.
2. Menggunakan *feature selection* dalam proses klasifikasi malware *spyware Gator*.
3. Melakukan visualisasi serangan malware *spyware Gator* ke dalam bentuk diagram garis *parallel coordinates.*

1.4 Manfaat

1. Dapat mengklasifikasikan data *spyware Gator* dan data *benign* dengan algoritma *Random Forest*.
2. *Feature selection* berguna agar proses komputasi lebih cepat.
3. Dapat memberikan informasi yang mudah dipahami dengan adanya visualisasi serangan malware *spyware Gator* .

1.5 Batasan Masalah

Adapun batasan masalah pada penulisan Proposal Tugas Akhir ini adalah sebagai berikut :

1. Algoritma yang digunakan dalam penelitian yaitu algoritma *Random Forest*.
2. Dataset yang digunakan berasal dari *Canadian Institute for Cybersecurity (CIC)* yaitu CICMalMem2022 dengan jenis *spyware Gator* dan benign.
3. Malware yang digunakan adalah malware jenis *spyware Gator*.
4. Dalam penelitian ini tidak membahas tentang bagaimana pencegahan terhadap malware *spyware Gator*.

1.6 Metodologi Penelitian

Metodologi yang digunakan dalam penelitian Tugas Akhir ini akan melewati beberapa tahapan sebagai berikut:

1. Studi Pustaka/ Literatur

Tahap ini dilakukan untuk mencari masalah yang akan dibahas sesuai dan relevan dengan tujuan untuk dijadikan sebagai penelitian. Setelah menemukan masalah tersebut dilanjutkan dengan mencari referensi yang berhubungan dengan penelitian seperti artikel, jurnal, buku, dan lainnya.

2. Pengolahan Data

Pada tahap ini membahas bagaimana proses mengolah suatu dataset yang masih mentah menjadi siap olah, menerapkan metode *Random Forest* pada tugas akhir, dan melakukan visualisasi terhadap data yang telah diolah.

3. Visualisasi Data

Pada tahapan inilah proses visualisasi data spyware dan data normal dilakukan dengan menggunakan algortima *Random Forest*.

4. Analisa

Setelah memperoleh data yang didapat pada tahapan visualisasi, maka langkah selanjutnya adalah membust sebuah analisis dari hasil yang telah diperoleh sebelumnya yang kemudian akan memberikan hasil yang objektif.

5. Kesimpulan dan Saran

Tahap terakhir yaitu membuat kesimpulan berdasarkan dari permasalahan yang ada, studi pustaka, metodelogi, dan analisa hasil visualisasi. Selain itu, saran yang diberikan dapat menjadi acuan untuk penelitian selanjutnya.

1.7 Sistematikan Penulisan

Adapun sistematika penulisan dalam penelitian Tugas Akhir ini adalah sebagai berikut:

BAB I. PENDAHULUAN

Bab ini berisikan latar belakang, tujuan dan manfaat, rumusan masalah, batasan masalah, metodologi penelitian serta sistematika penulisan.

BAB II. TINJAUAN PUSTAKA

Pada bab ini berisi literature review yang berkaitan dengan masalah visualisasi *spyware* dengan menggunakan algoritma *Random Forest* yang mengacu terhadap penelitian terdahulu.

BAB III. METODOLOGI PENELITIAN

Bab ini menjelaskan proses penelitian secara sistematis. Penjelasan yang ada di bab ini berisi tahapan serta persiapan data *spyware* yang akan digunakan dan diterapkan kedalam algoritma *Random Forest.*

BAB IV. HASIL DAN ANALISIS

Pada bab ini akan menjelaskan proses dan analisa yang akan ditampilkan dalam bentuk visual dari setiap data yang diperoleh dari hasil pengujian motode sistem.

BAB V. KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang telah didapat dari penelitian yang telah dilakukan, dan jawaban yang diperoleh dari tujuan yang hendak dicapai, serta memberikan saran untuk penelitan selanjutnya.

DAFTAR PUSTAKA

[1] A. Mishra, S. Mishra, and P. Jain, “Malware Category Prediction Using,” vol. 10, no. 02, pp. 787–797, 2019.

[2] D. Ucci, L. Aniello, and R. Baldoni, “Survey of machine learning techniques for malware analysis,” *Comput. Secur.*, vol. 81, pp. 123–147, 2019, doi: 10.1016/j.cose.2018.11.001.

[3] D. Javaheri, M. Hosseinzadeh, and A. M. Rahmani, “Detection and elimination of spyware and ransomware by intercepting kernel-level system routines,” *IEEE Access*, vol. 6, pp. 78321–78332, 2018, doi: 10.1109/ACCESS.2018.2884964.

[4] F. Pierazzi, G. Mezzour, Q. Han, M. Colajanni, and V. S. Subrahmanian, “A Data-driven Characterization of Modern Android Spyware,” *ACM Trans. Manag. Inf. Syst.*, vol. 11, no. 1, 2020, doi: 10.1145/3382158.

[5] 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*, no. Iccece, pp. 542–545, 2021, doi: 10.1109/ICCECE51280.2021.9342376.

[6] T. Zulhaq Jasman, E. Hasmin, C. Susanto, and W. Musu, “Perbandingan Logistic Regression, Random Forest, dan Perceptron pada Klasifikasi Pasien Gagal Jantung,” *CRSID J.*, vol. 14, no. 3, pp. 271–286, 2022, [Online]. Available: https://www.doi.org/10.22303/csrid.14.3.2022.271-286.

[7] L. BREIMAN, “Random Forests,” 2001, doi: 10.1109/ICCECE51280.2021.9342376.

[8] C. Science, V. P. Perdana, and N. H. Octavya, “Identifikasi Malicious Web Menggunakan Metode Random Forest,” vol. 4, no. 1, pp. 978–979, 2018.

[9] B. M. Khammas, “Ransomware Detection using Random Forest Technique,” *ICT Express*, vol. 6, no. 4, pp. 325–331, 2020, doi: 10.1016/j.icte.2020.11.001.

[10] M. Dener, G. Ok, and A. Orman, “Malware Detection Using Memory Analysis Data in Big Data Environment,” *Appl. Sci.*, vol. 12, no. 17, 2022, doi: 10.3390/app12178604.

[11] J. L. Alvares, “Malware Classification with BERT,” 2021, [Online]. Available: https://scholarworks.sjsu.edu/etd_projects/998.

[12] A. R. Yogaswara, “Klasifikasi Malware Family menggunakan Metode k-Nearest Neighbor (k-NN),” *J. Repos.*, vol. 3, no. 3, pp. 319–323, 2021, doi: 10.22219/repositor.v2i3.1313.

[13] G. Sun and Q. Qian, “Deep Learning and Visualization for Identifying Malware Families,” *IEEE Trans. Dependable Secur. Comput.*, vol. 18, no. 1, pp. 283–295, 2021, doi: 10.1109/TDSC.2018.2884928.

[14] K. Sethi, R. Kumar, L. Sethi, P. Bera, and P. K. Patra, “A novel machine learning based malware detection and classification framework,” *2019 Int. Conf. Cyber Secur. Prot. Digit. Serv. Cyber Secur. 2019*, pp. 1–4, 2019, doi: 10.1109/CyberSecPODS.2019.8885196.

[15] R. Sihwail, K. Omar, and K. A. Z. Ariffin, “An Effective Memory Analysis for Malware Detection and Classification,” *Comput. Mater. Contin.*, vol. 67, no. 2, pp. 2301–2320, 2021, doi: 10.32604/cmc.2021.014510.

[16] O. Aslan and A. A. Yilmaz, “A New Malware Classification Framework Based on Deep Learning Algorithms,” *IEEE Access*, vol. 9, pp. 87936–87951, 2021, doi: 10.1109/ACCESS.2021.3089586.

[17] T. Carrier, P. Victor, A. Tekeoglu, and A. Lashkari, “Detecting Obfuscated Malware using Memory Feature Engineering,” no. Icissp, pp. 177–188,

2022, doi: 10.5220/0010908200003120.

[18] A. Tripathi, N. Bhoj, M. Khari, and B. Pandey, “Feature Selection and Scaling for Random Forest Powered Malware Detection System,” pp. 1–14, 2021.

[19] E. Tansen and D. W. Nurdiarto, “Analisis dan Deteksi Malware dengan Metode Hybrid Analysis Menggunakan Framework MOBSF,” *J. Teknol. Inf.*, vol. 4, no. 2, pp. 191–201, 2020, doi: 10.36294/jurti.v4i2.1338.

[20] B. Cakir and E. Dogdu, “Malware classification using deep learning methods,” *Proc. ACMSE 2018 Conf.*, vol. 2018-Janua, no. March, 2018, doi: 10.1145/3190645.3190692.

[21] M. Wazid *et al.*, “A framework for detection and prevention of novel keylogger spyware attacks,” *7th Int. Conf. Intell. Syst. Control. ISCO 2013*, pp. 433–438, 2013, doi: 10.1109/ISCO.2013.6481194.

[22] T. P. Setia, A. P. Aldya, and N. Widiyasono, “Reverse Engineering untuk Analisis Malware Remote Access Trojan,” *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 1, p. 40, 2019, doi: 10.26418/jp.v5i1.28214.

[23] A. Razzaq, M. Aditya, A. Widya, O. Kuncoro, D. Lesmana, and P. Widodo, “Serangan Hacking Tools sebagai Ancaman Siber dalam Sistem Pertahanan Negara (Studi Kasus : Predator),” vol. 6, no. April, pp. 35–46, 2022, doi: 10.34010/gpsjournal.v6i1.

[24] M. Boldt, A. Jacobsson, and B. Carlsson, “Exploring spyware effects,” *Privacy-Invasive Softw.*, no. May 2014, 2010, [Online]. Available: http://www.researchgate.net/profile/Martin_Boldt/publication/42756579_Privacy-Invasive_Software/links/00b4952a8786a07545000000.pdf#page=87.

[25] S. Assitant and C. Science, “Detection of Spyware in Software,” no. Icoei, pp. 1138–1142, 2019.

[26] Y. Lozanov, S. Tzvetkova, and A. Petleshkov, “Use of machine learning techniques for classification of thermographic images,” *2020 12th Electr. Eng. Fac. Conf. BulEF 2020*, pp. 4–7, 2020, doi: 10.1109/BulEF51036.2020.9326046.

[27] N. Thamaraikannan and S. Manju, “Review on Image Classification Techniques in Machine Learning for Satellite Imagery,” *Proc. - Int. Conf. Artif. Intell. Smart Syst. ICAIS 2021*, pp. 144–149, 2021, doi: 10.1109/ICAIS50930.2021.9395808.

[28] Di. C. Nguyen *et al.*, “Enabling AI in Future Wireless Networks: A Data Life Cycle Perspective,” *IEEE Commun. Surv. Tutorials*, vol. 23, no. 1, pp. 553–595, 2021, doi: 10.1109/COMST.2020.3024783.

[29] S. M. Ahamed and V. N. Sharma, “Malware Detection using Optimized Random Forest Classifier within Mobile Devices,” vol. 3, no. 5, pp. 90–99, 2016.

[30] T. Carneiro, R. V. M. Da Nobrega, T. Nepomuceno, G. Bin Bian, V. H. C. De Albuquerque, and P. P. R. Filho, “Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications,” *IEEE Access*, vol. 6, pp. 61677–61685, 2018, doi: 10.1109/ACCESS.2018.2874767.

[31] O. Tomic, “Using machine learning and Repeated Elastic Net Technique for identification of biomarkers of early Alzheimer’s disease,” 2021, [Online]. Available: https://nmbu.brage.unit.no/nmbu-xmlui/handle/11250/2980497.

[32] M. Chen-Wishart, *python machine learning, third edition*, no. January 2010. 2014.

[33] S. Widodo, H. Brawijaya, and S. Samudi, “Stratified K-fold cross validation optimization on machine learning for prediction,” *Sinkron*, vol. 7,

no. 4, pp. 2407–2414, 2022, doi: 10.33395/sinkron.v7i4.11792.

[34] Y. T. Bau, T. Sasidaran, and C. Le Goh, “Improving Machine Learning Algorithms for Breast Cancer Prediction,” *J. Syst. Manag. Sci.*, vol. 12, no. 4, pp. 251–266, 2022, doi: 10.33168/JSMS.2022.0416.

[35] A. Mallak and M. Fathi, “A Hybrid Approach: Dynamic Diagnostic Rules for Sensor Systems in Industry 4.0 Generated by Online Hyperparameter Tuned Random Forest,” *Sci*, vol. 2, no. 4, pp. 75–70, 2020, doi: 10.3390/sci2040075.

[36] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization,” *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, 2012.

[37] P. Probst, M. N. Wright, and A. L. Boulesteix, “Hyperparameters and tuning strategies for random forest,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 9, no. 3, pp. 1–15, 2019, doi: 10.1002/widm.1301.

[38] M. H. L. Louk and B. A. Tama, “Tree-Based Classifier Ensembles for PE Malware Analysis: A Performance Revisit,” *Algorithms*, vol. 15, no. 9, pp. 1–15, 2022, doi: 10.3390/a15090332.

[39] A. Paul, D. P. Mukherjee, P. Das, A. Gangopadhyay, A. R. Chintha, and S. Kundu, “Improved Random Forest for Classification,” *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4012–4024, 2018, doi: 10.1109/TIP.2018.2834830.