KLASIFIKASI SERANGAN *SPYWARE* DENGAN MENGGUNAKAN METODE *K-NEAREST NEIGHBORS* (KNN)

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer

OLEH:

MULKI PEDERSON

09011282025086

JURUSAN SISTEM KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2024

LEMBAR PENGESAHAN

KLASIFIKASI SERANGAN *SPYWARE* DENGAN MENGGUNAKAN METODE *K-NEAREST NEIGHBORS* (KNN)

SKRIPSI

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

MULKI PEDERSON
09011282025086

Indralaya, 11 Juni 2024

Mengetahui,

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

AUTHENTICATION PAGE

CLASSIFICATION OF SPYWARE ATTACK USING *K-NEAREST NEIGHBORS* (KNN) METHOD

THESIS

Departement of Computer System

Bachelor's Degree

By:

MULKI PEDERSON

09011282025086

Indralaya, 11 June 2024

Aknowlodge,

Head of Computer System Departement,	Final Project Advisor,
Dr. Ir. Sukemi, M. T.	Prof. Deris Stiawan, M. T., Ph. D.
NIP. 196612032006041001	NIP. 197806172006041002

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : Rabu

Tanggal : 22 Mei 2024

Tim Penguji

1. Ketua : **Kemahyanto Exaudi, M. T.**

2. Sekretatis : **Abdurahman, S. Kom., M. Han**

3. Penguji : **Ahmad Heryanto, M. T.**

4. Pendamping : **Prof. Deris Stiawan, M. T., Ph. D**

Mengetahui,
Ketua Jurusan Sistem Komputer

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

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama	**:**	**Mulki Pederson**
NIM	**:**	**09011282025086**
Judul	**:**	**Klasifikasi Serangan *Spyware* Dengan Menggunakan Metode *K -Nearest Neighbors* (KNN)**

Hasil Pengecekan *Software iThenticate* / Turnitin : 9%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam 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 2024

Mulki Pederson
NIM. 09011282025086

KATA PENGANTAR

Assalamu'alaikum Warrahmatullahi Wabarakatuh.

Alhamdulilahirabbil'alamin, puji dan syukur penulis panjatkan atas kehadirat Allah SWT yang telah melimpahkan nikmat, taufik, dan hidayah-Nya yang sangat besar dan tidak pernah berhenti kepada penulis sehingga penulis dapat menyelesaikan penulisan Tugas Akhir ini yang berjudul **Klasifikasi Serangan *Spyware* Dengan Menggunakan Metode *K – Nearest Neighbors (*KNN)** untuk memenuhi salah satu syarat memperoleh gelar Sarjana Komputer (Strata 1) pada Jurusan Komputer di Universitas Sriwijaya.

Dalam lembaran ini, dengan segala kerendahan hati penulis menyampaikan ucapan terima kasih kepada setiap pihak atas bantuan, bimbingan, dan saran yang telah diberikan dalam menyelesaikan penulisan Tugas Akhir ini. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Allah SWT, yang telah memberikan nikmat dan kesempatan, terutama nikmat iman dan kesehatan sehingga penulis dapat menyusun serta menyelesaikan penulisan tugas akhir ini.
2. Orang Tua dan Keluarga yang selalu memberikan doa, motivasi dan dukungannya baik secara moril, materil maupun spiritual hingga saat ini.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Dr. Ir. Sukemi, M.T. selaku Ketua Jurusan Sistem Komputer Universitas Sriwijaya.
5. Bapak Prof. Deris Stiawan, M.T., Ph.D, selaku Dosen Pembimbing Tugas Akhir yang telah berkenan meluangkan waktunya untuk membimbing, memberikan saran, dan motivasi kepada penulis dalam penyusunan tugas akhir.
6. Bapak Muhammad Ali Buchari, S.Kom., M.T. selaku Pembimbing Akademik.
7. Staff Fakultas Ilmu Komputer Univesitas Sriwijaya yang telah membantu dalam proses penyelesaian laporan tugas akhir.

8. Rekan – rekan penulis yang senantiasa membantu dan memberikan saran kepada penulis selama penyusunan tugas akhir.
9. Teman-teman seperjuangan Jurusan Sistem Komputer Angkatan 2020.
10. Grup riset COMNETS.
11. Seluruh pihak yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan semangat serta do'a.

Akhir kata, penulis menyadari bahwa penulisan tugas akhir ini masih jauh dari kata sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan, semoga tugas akhir ini dapat bermanfaat bagi semua pihak.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Indralaya, Juni 2024
Penulis,

Mulki Pederson
NIM. 09011282025086

CLASSIFICATION OF SPYWARE ATTACK USING K-NEAREST NEIGHBORS (KNN) METHOD

MULKI PEDERSON (09011282025086)

Department of Computer System, Faculty of Computer Science,

Sriwijaya University

E-mail: mlkypdrsn17@gmail.com

ABSTRACT

Spyware is one type of malware that threatens computer systems because it can steal users' personal information and sensitive data without their knowledge. Spyware can monitor user activities and steal data such as visited websites, email addresses, and even record keyboard and screen activities. This research aims to classify spyware attacks using the K-Nearest Neighbors (KNN) algorithm. The research dataset consists of spyware malware data and benign data available in the CICMalMem2022 dataset. In this study, data splitting is performed with Stratified K-Fold to obtain the optimal number of folds that can achieve more optimal classification results for each parameter k.The research results indicate that the KNN algorithm is highly accurate in classifying spyware attack data, achieving the highest results with 20 best features using k=3 and fold=4, reaching an accuracy of 99.91%. With such results, it can be said that the use of the KNN algorithm is effective in identifying spyware attacks with a high level of accuracy.. With these results, it can be said that the use of the KNN algorithm is effective in identifying spyware attacks with a high level of accuracy.

Keywords: *Malware Classification, Spyware, K-Nearest Neighbors, Stratified K-Fold.*

Aknowlodge,

Head of Computer System Departement, 9/6/24	**Final Project Advisor,**
Dr. Ir. Sukemi, M. T.	**Prof. Deris Stiawan, M. T., Ph. D.**
NIP. 196612032006041001	**NIP. 197806172006041002**

KLASIFIKASI SERANGAN *SPYWARE* DENGAN MENGGUNAKAN METODE *K-NEAREST NEIGHBORS* (KNN)

MULKI PEDERSON (09011282025086)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

E-mail: mlkypdrsn17@gmail.com

ABSTRAK

Spyware merupakan salah satu jenis malware yang mengancam sistem komputer karena dapat mencuri informasi pribadi dan data sensitif pengguna tanpa sepengetahuannya. *Spyware* dapat memantau aktivitas pengguna dan mencuri data seperti, daftar situs web yang dikunjungi, alamat email, dan bahkan dapat merekam aktivitas *keybord* dan layar. Penelitian ini bertujuan untuk melakukan klasifikasi serangan *spyware* menggunakan algoritma *K-Nearest Neighbors* (KNN). Dataset penelitian terdiri dari data malware jenis *spyware* dan data *benign* yang terdapat pada dataset CICMalMem2022. Dalam penelitian ini dilakukan *split* data dengan *Stratified K-Fold* untuk mendapatkan jumlah *fold* terbaik yang dapat memperoleh hasil klasifikasi yang lebih optimal pada setiap parameter *k*. Hasil penelitian menunjukkan bahwa algoritma KNN sangat akurat dalam melakukan klasifikasi data serangan *spyware* dengan memperoleh hasil tertinggi pada 20 fitur terbaik dengan *k*=3 dan menggunakan *fold*=4 yang mencapai akurasi sebesar 99,91%. Dengan hasil seperti ini dapat dikatakan bahwa penggunaan algoritma KNN efektif dalam mengidentifikasi serangan *spyware* dengan tingkat akurasi yang tinggi.

Kata Kunci: Klasifikasi Malware, *Spyware*, *K-Nearest Neighbors*, *Stratified K-Fold.*

Mengetahui,

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

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB 1
PENDAHULUAN

1.1 Latar Belakang

Salah satu permasalahan yang sering dihadapi pengguna internet dari dulu hingga saat ini adalah malware. Malware (*Malicious Software*) merupakan istilah umum yang digunakan terhadap program atau perangkat lunak yang dirancang untuk merusak atau mengganggu sistem komputer dengan menginfeksi komputer pengguna yang sah dan melakukan pencurian informasi [1]. Perangkat lunak ini terus berkembang dan bervariasi hingga menjadi salah satu ancaman paling berbahaya bagi sistem komputer [2]. Adapun jenis - jenis malware yaitu *Worms, Viruses, Botnets, Ransomware, Spyware, Scareware, Rootkits*, dan sebagainya [3].

Spyware merupakan jenis serangan malware berupa kode perangkat lunak yang digunakan untuk mencuri informasi atau data sensitif tanpa sepengetahuan pengguna [4]. *Spyware* secara diam – diam dapat memantau aktivitas dan mengumpulkan rincian tentang suatu individu atau organisasi, kemudian dikirimkan kepada pembuatnya [5]. Malware jenis ini menggunakan fungsi pelacakan untuk mengirim berbagai informasi pribadi, seperti daftar situs web yang dikunjungi, alamat *email* pengguna, dan ketikan pada *keyboard* serta dapat juga merekam tangkapan layar dan aktivitas online pada komputer atau perangkat seluler korban. Sementara itu, informasi atau data yang diperoleh berupa kode PIN, kode keamanan, nomor kartu kredit, dan sebagainya. Lebih lanjut, *spyware* memiliki kemampuan untuk mengaktifkan kamera dan mikrofon yang memungkinkan pemantauan dan mendengarkan pengguna tanpa terdeteksi [6].

KNN (*K-Nearest Neighbors*) merupakan salah satu algoritma *machine learning* yang umum digunakan untuk melakukan tugas klasifikasi dan regresi. Prinsip dasar KNN adalah melakukan prediksi berdasarkan mayoritas kelas dari *k* tetangga terdekat untuk melakukan klasifikasi atau dengan mengambil nilai rata-rata dari tetangga terdekat untuk melakukan regresi (*k* adalah bilangan bulat positif dan biasanya kecil) [7]. Algoritma KNN adalah algoritma pembelajaran berbasis jarak yang umumnya diimplementasikan untuk tugas klasifikasi. KNN

beroperasi dengan melakukan perhitungan jarak antara suatu titik data dengan kelas – kelas data yang ada. Proses klasifikasi pada titik data tertentu dilakukan dengan menentukan titik tersebut termasuk dalam kelas malware atau bukan [8].

Pada penelitian [9] dataset yang digunakan serupa dengan penelitian yang dilakukan, namun peneltian ini menggunakkan semua kategori utama malware pada dataset yaitu *Spyware, Ransomware,* dan *Trojan Horse* dengan berfokus pada model deteksi malware yang tersamarkan dan mengekstraksi fitur – fitur dari *dump memory* menggunakan VolMemLyzer. Hasil penelitian ini mendapatkan algortima *Random Forest* dengan akurasi 97%, *Decision Tree* akurasi 97% KNN dengan akurasi 95%, *Naïve Bayes* akurasi 92%, dan *Support Vector Machine* (SVM) akurasi 90%.

Dalam Penelitian [10] ini berfokus pada pengembangan kerangka deteksi malware untuk membantu mengenali apakah lebih dari 10.000 aplikasi android pada perangkat *Internet of Things* (IoT) adalah malware atau *benign* dengan menggunakan metode *Naive Bayes*, *Decision Tree*, *Support Vector Machine*, dan *K-Nearest Neighbors*. Hasilnya *K-Nearest Neighbors* mendapatkan hasil terbaik dengan akurasi, presisi, *recall*, dan *f1 - score* masing-masing sebesar 93%, 95%, 90%, dan 92%. Selain itu, pada penelitian [8] yang menggunakan dataset dari Kaggle terdiri dari 42.797 *API Call* berbahaya dan 1.079 *API Call* tidak berbahaya mendapatkan hasil klasifikasi menggunakan KNN dengan akurasi mencapai 98,17% pada nilai *k*=3.

Berdasarkan paparan diatas dan ulasan dari penelitian sebelumnya yang menjadi latar belakang penulis untuk melakukan penelitian tugas akhir ini dengan judul **Klasifikasi Serangan *Spyware* Dengan Menggunakan Metode *K* – *Nearest Neighbors* (KNN)**. Diharapkan metode yang digunakan dapat mencapai akurasi yang tinggi mengingat KNN dianggap memiliki tingkat keakuratan yang tinggi dalam melakukan klasifikasi suatu data.

1.2 Rumusan Masalah

Berikut adalah rumusan masalah yang akan dibahas dalam penelitian tugas akhir ini:

1. Teknik seleksi fitur seperti apa yang digunakan untuk memilih fitur yang relevan dalam klasifikasi serangan *spyware*?
2. Bagaimana cara melakukan klasifikasi serangan *spyware* pada dataset CICMalMem2022?
3. Bagaimana cara melakukan evaluasi performa algoritma *K–Nearest Neighbors* dalam klasifikasi serangan *spyware*?

1.3 Tujuan

Dari rumusan masalah diatas, berikut merupakan tujuan yang akan dicapai dalam penelitian tugas akhir ini:

1. Menggunakan *correlation based feature selection* untuk memilih fitur yang relevan dalam klasifikasi serangan *spyware*.
2. Menerapkan algoritma *K–Nearest Neighbors* untuk melakukan klasfiikasi serangan *spyware* pada dataset CICMalMem2022.
3. Melakukan evaluasi performa algoritma *K–Nearest Neighbors* dengan menggunakan *confusion matrix* dan *stratified k-fold cross validation*.

1.4 Manfaat

Adapun manfaat yang dapat diperoleh dalam penelitian tugas akhir ini adalah sebagai berikut:

1. Dapat mengoptimalkan waktu proses komputasi dengan menggunakan fitur yang relevan.
2. Dapat melakukan klasifikasi serangan *spyware* pada dataset CICMalMem2022 menggunakan algoritma *K–Nearest Neighbors*.
3. Dapat mengevaluasi performa algoritma *K–Nearest Neighbors* dengan menggunakan *confusion matrix* dan *stratified k-fold cross validation*.

1.5 Batasan Masalah

Berikut adalah batasan masalah yang tidak akan dibahas dalam penelitian tugas akhir ini:

1. Penelitian ini hanya berfokus pada penggunaan *correlation based feature selection* untuk pemilihan fitur.

2. Algoritma yang diterapkan untuk melakukan klasifikasi serangan *spyware* berfokus pada algoritma *K–Nearest Neighbors*.
3. Hanya menggunakan dataset CICMalMem2022 yang berasal dari *Canadian Institute for Cybersecurity* (CIC), data yang digunakan pun khusus pada data *benign* dan data malware jenis *spyware*.
4. Tidak membahas terkait cara mencegah serangan *spyware*.

1.6 Metodologi Penelitian

Metodologi yang digunakan dalam penelitian tugas akhir ini akan melewati beberapa tahapan, sebagai berikut:

1. Studi literatur/studi pustaka

   Tahapan ini dilakukan dengan cara mencari artikel atau jurnal penelitian sebelumnya untuk menentukan topik yang akan diteliti, kemudian mengumpulkan referensi berupa buku dan *paper* jurnal terkait penelitian yang akan dilakukan.
2. Persiapan data

   Pada tahapan ini melakukan pencarian dan menentukan dataset yang akan digunakan dan dilakukan pembelajaran serta pemahaman terkait data yang akan diolah sehingga keperluan untuk penelitian tercukupi.
3. Konsultasi dan tukar pikiran

   Tahapan ini merupakan kegiatan konsultasi dengan pembimbing tugas akhir serta menerima masukan dan saran dari rekan peneliti lain terkait kemajuan dan hambatan dalam penelitian ini.
4. Pengolahan data

   Tahapan ini mencakup kegiatan mengolah data seperti pembersihan data, pemberian label, seleksi fitur, *split* data latih dan uji, serta menerapkan metode KNN yang digunakan untuk mengklasifikasikan serangan *spyware*.
5. Hasil dan analisa

   Pada tahapan ini dilakukan analisis terhadap hasil yang didapatkan dari penelitian yang telah dilakukan dengan tujuan mengetahui performa yang didapat berdasarkan nilai akurasi, presisi, *recall* dan *f-1 score*.

6. Kesimpulan dan saran

Tahapan ini berisi kesimpulan yang didapat dari penelitian yang telah dilakukan dan selanjutnya memberikan saran yang diharapkan dapat berguna untuk penelitian pada topik yang berkaitan.

1.7 Sistematika Penulisan

Sistematika penulisan dalam penelitian tugas akhir ini akan melewati beberapa tahapan, sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini berisikan tentang latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, metodologi penelitian, dan sistematika penulisan terkait topik dari penelitian ini.

BAB 2 TINJAUAN PUSTAKA

Bab ini membahas tentang teori dasar dan informasi dari penelitian terdahulu yang berkaitan dengan klasifikasi serangan malware, Algoritma *K–Nearest Neightbors, Confusion Matrix* dan *Stratified K-Fold Cross Validation*.

BAB 3 METODOLOGI PENELITIAN

Bab ini berisikan penjelasan secara sistematis bagaimana proses penelitian dilakukan dengan perancangan sistem yang meliputi persiapan dataset hingga penerapan algoritma yang digunakan.

BAB 4 HASIL DAN ANALISIS

Bab ini memuat penjelasan terkait hasil pengujian yang diperoleh dan melakukan analisis terhadap hasil penelitian yang menggunakan metode tersebut.

BAB 5 KESIMPULAN DAN SARAN

Bab ini merupakan penutup yang berisikan kesimpulan yang didapat dari penelitian yang telah dilakukan dan terdapat pula saran untuk penelitian mendatang.

DAFTAR PUSTAKA

[1] I. Shhadat, B. Bataineh, A. Hayajneh, and Z. A. Al-Sharif, "The Use of Machine Learning Techniques to Advance the Detection and Classification of Unknown Malware," *Procedia Comput. Sci.*, vol. 170, no. 2019, pp. 917–922, 2020, doi: 10.1016/j.procs.2020.03.110.

[2] N. S. Selamat and F. H. M. Ali, "Comparison of malware detection techniques using machine learning algorithm," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 16, no. 1, pp. 435–440, 2019, doi: 10.11591/ijeecs.v16.i1.pp435-440.

[3] S. Sharma, C. Rama Krishna, and S. K. Sahay, *Detection of Advanced Malware by Machine Learning Techniques*, vol. 742. Springer Singapore, 2019. doi: 10.1007/978-981-13-0589-4_31.

[4] V. Mahesh and K. A. Sumithra Devi, "Spyware detection and prevention using deep learning AI for user applications," *Int. J. Recent Technol. Eng.*, vol. 7, no. 5, pp. 345–349, 2019.

[5] S. Assitant and C. Science, "Detection of Spyware in Software Using Virtual Enviroment," *IEEE*, pp. 1138–1142, 2019, doi: https://doi.org/10.1109/ICOEI.2019.8862547.

[6] M. Rights, "Spyware Detection Technique Based on Reinforcement Learning," *City Res. Online City*, pp. 307–316, 2020.

[7] K. Atefi, H. Hashim, and M. Kassim, "Anomaly Analysis for the Classification Purpose of Intrusion Detection System with K-Nearest Neighbors and Deep Neural Network," *IEEE 7th Conf. Syst. Process Control*, pp. 269–274, 2019, doi: 10.1109/ICSPC47137.2019.9068081.

[8] T. A. Assegie, "An Optimized KNN Model for Signature-Based Malware Detection," *Int. J. Comput. Eng. Res. Trends*, vol. 8, no. 2, pp. 46–49, 2021, doi: https://doi.org/10.22362/ijcert/2021/v8/i2/v8i206.

[9] T. Carrier, P. Victor, A. Tekeoglu, and A. Habibi Lashkari, "Detecting Obfuscated Malware using Memory Feature Engineering," *Int. Conf. Inf. Syst. Secur. Priv.*, no. Icissp, pp. 177–188, 2022, doi:

10.5220/0010908200003120.

[10] H. Babbar, S. Rani, D. K. Sah, S. A. AlQahtani, and A. Kashif Bashir, "Detection of Android Malware in the Internet of Things through the K-Nearest Neighbor Algorithm," *Sensors*, vol. 23, no. 16, pp. 1–17, 2023, doi: 10.3390/s23167256.

[11] D. T. Dehkordy, "DroidTKM: Detection of Trojan Families using the KNN Classifier Based on Manhattan Distance Metric," *IEEE*, no. Iccke, pp. 136–141, 2020, doi: 10.1109/ICCKE50421.2020.9303720.

[12] R. S. ARSLAN and A. H. Yurttakal, "K-Nearest Neighbour Classifier Usage for Permission Based Malware Detection in Android," *Icontech Int. J.*, vol. 4, no. 2, pp. 15–27, 2020, doi: 10.46291/icontechvol4iss2pp15-27.

[13] D. O. Sahin, S. Akleylek, and E. Kilic, "LinRegDroid: Detection of Android Malware Using Multiple Linear Regression Models-Based Classifiers," *IEEE Access*, vol. 10, pp. 14246–14259, 2022, doi: 10.1109/ACCESS.2022.3146363.

[14] S. Choi, "Combined KNN classification and hierarchical similarity hash for fast malware detection," *Appl. Sci.*, vol. 10, no. 15, pp. 1–16, 2020, doi: 10.3390/app10155173.

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

[16] M. S. Akhtar and T. Feng, "Malware Analysis and Detection Using Machine Learning Algorithms," *Symmetry (Basel).*, vol. 14, no. 11, 2022, doi: 10.3390/sym14112304.

[17] I. Ben, A. Ouahab, and M. Bouhorma, "Classification of Grayscale Malware Images Using the K-Nearest Neighbor Algorithm," *Springer Nat. Switz.*, pp. 1038–1050, 2020, doi: https://doi.org/10.1007/978-3-030-37629-1_75.

[18] M. Al-kasassbeh, M. A. Abbadi, and A. M. Al-Bustanji, "LightGBM Algorithm for Malware Detection," in *Advances in Intelligent Systems and Computing*, 2020, vol. 1230 AISC, pp. 391–403. doi: 10.1007/978-3-030-

52243-8_28.

[19] L. Chen, C. Xia, S. Lei, and T. Wang, “Detection, Traceability, and Propagation of Mobile Malware Threats,” *IEEE Access*, vol. 9, pp. 14576–14598, 2021, doi: 10.1109/ACCESS.2021.3049819.

[20] X. Liu, J. Zhang, Y. Lin, and H. Li, “ATMPA: Attacking Machine Learning-based Malware Visualization Detection Methods via Adversarial Examples,” *Proc. Int. Symp. Qual. Serv. IWQoS 2019*, 2019, doi: 10.1145/3326285.3329073.

[21] Mahesh V and S. Devi K A, “Detection and Prediction of Spyware for user Applications by interdisciplinary approach,” *Int. Conf. Comput. Intell. Smart Power Syst. Sustain. Energy, CISPSSE 2020*, pp. 1–6, 2020, doi: 10.1109/CISPSSE49931.2020.9212222.

[22] V. S. Anatoly Belous, “Computer Viruses, Malicious Logic, and Spyware,” in *Viruses, Hardware and Software Trojans*, 1st ed., Springer Cham, 2020, pp. 101–207. doi: https://doi.org/10.1007/978-3-030-47218-4.

[23] 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.

[24] D. Gibert, C. Mateu, and J. Planes, “The rise of machine learning for detection and classification of malware: Research developments, trends and challenges,” *J. Netw. Comput. Appl.*, vol. 153, p. 102526, 2020, doi: 10.1016/j.jnca.2019.102526.

[25] G. Baldini and D. Geneiatakis, “A Performance Evaluation on Distance Measures in KNN for Mobile Malware Detection,” *Int. Conf. Control. Decis. Inf. Technol. (CoDIT 19)*, pp. 193–198, 2019, doi: 10.1109/CoDIT.2019.8820510.

[26] M. Jurecek and R. Lorencz, “Application of Distance Metric Learning to Automated Malware Detection,” *IEEE Access*, vol. 9, no. Dml, pp. 96151–96165, 2021, doi: 10.1109/ACCESS.2021.3094064.

[27] A. Yudhana, S. Sunardi, and A. J. S. Hartanta, “Algoritma K-Nn Dengan Euclidean Distance Untuk Prediksi Hasil Penggergajian Kayu Sengon,”

Transmisi, vol. 22, no. 4, pp. 123–129, 2020, doi: 10.14710/transmisi.22.4.123-129.

[28] B. Pandey, "An Euclidean Distance based KNN Computational Method for Assessing Degree of Liver Damage," *IEEE*, 2016, doi: 10.1109/INVENTIVE.2016.7823222.

[29] N. S. Sani, M. I. Esa, and B. A. Musawi, "Malware Detection Using Deep Learning and Correlation-Based Feature Selection," *Symmetry (Basel).*, vol. 15, no. 123, pp. 1–21, 2023, doi: https://doi.org/10.3390/sym15010123.

[30] J. O. Sinayobye, S. Kaawaase Kyanda, N. F. Kiwanuka, and R. Musabe, "Hybrid Model of Correlation Based Filter Feature Selection and Machine Learning Classifiers Applied on Smart Meter Data Set," *Proc. - 2019 IEEE/ACM Symp. Softw. Eng. Africa, SEiA 2019*, pp. 1–10, 2019, doi: 10.1109/SEiA.2019.00009.

[31] 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.

[32] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, "Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking," *IEEE Access*, vol. 8, pp. 90847–90861, 2020, doi: 10.1109/ACCESS.2020.2994222.

[33] H. Henderi, "Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer," *IJIIS Int. J. Informatics Inf. Syst.*, vol. 4, no. 1, pp. 13–20, 2021, doi: 10.47738/ijiis.v4i1.73.