VISUALISASI SERANGAN PADA *MALWARE SPYWARE* MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat Memperoleh Gelar Sarjana Komputer

OLEH:

SARTIKA

09011381924138

JURUSAN SISTEM KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2023

LEMBAR PENGESAHAN

VISUALISASI SERANGAN PADA *MALWARE SPYWARE* MENGGUNAKAN METODE *NAÏVE BAYES CLASSIFIER*

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer

Oleh:

SARTIKA

09011381924138

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

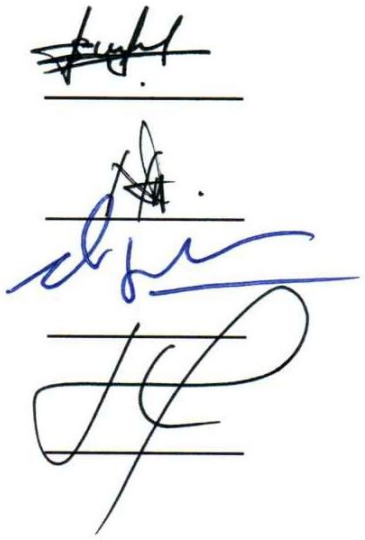
Telah diuji dan lulus pada

Hari : Kamis

Tanggal : 25 Mei 2023

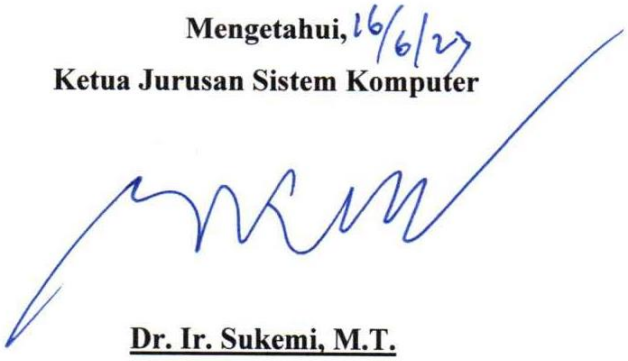
Tim Penguji

1. Ketua : Sarmayanta Sembiring, M.T.
2. Sekretaris : Nurul Afifah, M.Kom
3. Pembimbing : Deris Stiawan, M.T., Ph.D
4. Penguji : Huda Ubaya, M.T



Mengetahui, 16/6/23

Ketua Jurusan Sistem Komputer



Dr. Ir. Sukemi, M.T.

NIP. 196612032006041001

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Sartika

NIM : 09011381924138

Judul : Visualisasi Serangan Pada Malware Spyware Menggunakan Metode Naïve Bayes Classifier

Hasil Pengecekkan *Software iThenticate/Turnitin* : 13%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil 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 2023

Sartika

09011381924138

HALAMAN PERSEMBAHAN DAN MOTTO

بِسْـــــــــــــــمِ اللهِ الرَّحْمَنِ الرَّحِيْمِ

Q.S Al Baqarah:286 “Allah tidak akan membebani seorang melainkan sesuai dengan kemampuannya. Dia mendapat (pahala) dari (kebajikan) yang dikerjakannya dan dia mendapat siksa dari kejahatan yang diperbuatnya”

Kelemahan terbesar kita adalah bersandar pada kepasrahan. Jalan yang paling jelas menuju kesuksesan adalah mencoba, setidaknya satu kali daripada tidak sama sekali.

~Thomas A. Edison~

Skripsi ini saya persembahkan untuk :

Pertama, untuk diri saya sendiri yang telah berjuang dan bertahan hingga saat ini dapat menyelesaikan perkuliahan tepat pada waktunya.
Kedua, untuk kedua orang tuaku yang tanpa lelah dengan penuh kasih sayang memanjatkan doa yang luar biasa untuk anaknya serta memberikan dukungan baik moril ataupun materi. Terimakasih banyak atas segala pengorbanan dan kerja keras dalam mendidik saya.

VISUALISASI SERANGAN PADA *MALWARE SPYWARE* MENGGUNAKAN METODE *NAÏVE BAYES CLASSIFIER*

SARTIKA (09011381924138)

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email : sartikaa2912@gmail.com

ABSTRAK

Jumlah *Malware* terus meningkat. Sebagian besar *Malware* merupakan modifikasi dari data *Malware* sebelumnya. Dataset yang digunakan dari *CIC-MalMem-2022* yaitu *Benign* dan *Spyware-CWS*. Penelitian ini menggunakan algoritma *Naïve Bayes Classifier*. *Naive Bayes* merupakan salah satu algoritma klasifikasi yang memiliki akurasi dalam membuat prediksi dan memiliki reputasi yang baik dalam klasifikasi, terutama dalam kecepatan dalam belajar dibandingkan dengan algoritma klasifikasi *Machine Learning* lainnya. Untuk mendapatkan hasil akurasi terbaik terdapat lima fitur dengan nilai tertinggi untuk di training menggunakan *K-fold* 3 yang dihitung menggunakan *Confusion Matrix*. Hasil penelitian menunjukkan bahwa metode *Naïve Bayes Classifier* dapat menganalisis tingkat akurasi menggunakan *K-Fold* 3 dengan Akurasi 94,11%. Dasi hasil penelitian menyatakan bahwa visualisasi serangan pada *Malware Spyware* menggunakan metode *Naïve Bayes Classifier* mendapatkan hasil yang efisien dan akurat.

Kata Kunci : *Malware, Spyware, Visualisasi, Naïve Bayes Classifier, K-Fold, Machine Learning.*

VISUALIZATION OF ATTACKS ON *SPYWARE MALWARE* USING THE *NAÏVE BAYES CLASSIFIER* METHOD

SARTIKA (09011381924138)

Department of Computer Systems, Faculty of Computer Science, University Sriwijaya

Email : sartikaa2912@gmail.com

ABSTRACT

The amount of *Malware* is constantly increasing. Most *Malware* is a modification of previous *Malware* data. The datasets used from *CIC-MalMem-2022* are *Benign* and *Spyware-CWS*. This study used the *Naïve Bayes Classifier* algorithm. *Naïve Bayes* is one of the classification algorithms that has accuracy in making predictions and has a good reputation in classification, especially in learning speed compared to other *Machine Learning* classification algorithms. To get the best accuracy results, there are five features with the highest value for training using *K-fold 3* which are calculated using the *Confusion Matrix*. The results showed that the *Naïve Bayes Classifier* method can analyze the accuracy level using *K-Fold 3* with an Accuracy of 94.11%. The tie of the results of the study stated that visualization of attacks on Spyware Malware using the *Naïve Bayes Classifier* method obtained efficient and accurate results.

Keywords: *Malware, Spyware, Visualization, Naïve Bayes Classifier,K-Fold, Machine Learning.*

KATA PENGANTAR

السَّلاَمُ عَلَيْكُمْ وَرَحْمَةُ اللهِ وَبَرَكَاتُهُ

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 Pada *Malware Spyware* Menggunakan Metode *Naïve Bayes Classifier*"**

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 Subhanahu Wata'ala 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. selaku Dosen Pembimbing Tugas Akhir yang telah berkenan meluangkan waktunya untuk membimbing, memberikan saran dan motivasi serta bimbingan terbaik untuk penulis dalam menyelesaikan Tugas Akhir ini.
6. Bapak Aditya.P.P.Prasetya,S.KOM.,M.T. selaku Pembimbing Akademik Jurusan Sistem Komputer.
7. Mbak Sari selaku admin Jurusan Sistem Komputer yang telah membantu mengurus seluruh berkas.

8. Mbak Nurul Afifah M.Kom yang selalu memberikan masukan dan saran kepada Tim Riset Malware *Spyware* (*Spyware-CWS)*
9. Amelia Pramudita dan Zarah Fitriani yang memberikan dukungan dan saran selama proses penyusunan laporan skripsi.
10. Grup riset COMNETS.
11. Seluruh teman Sistem Komputer bukit 2019.
12. Untuk pemilik NIM 1930102075 terimakasih telah memberikan bantuan baik motivasi dan semangat untuk penulisan skripsi ini.

Penulis menyadari bahwa laporan 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 laporan 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.

وَالسَّلاَمُ عَلَيْكُمْ وَرَحْمَةُ اللهِ وَبَرَكَاتُهُ

Palembang, Juni 2023

Penulis,

Sartika

NIM. 09011381924138

DAFTAR ISI

LEMBAR PENGESAHAN ..... **ii**
HALAMAN PERNYATAAN ..... **iii**
HALAMAN PERSETUJUAN ..... **iv**
HALAMAN PERSEMBAHAN ..... **v**
ABSTRAK ..... **vi**
ABSTRACT ..... **vii**
KATA PENGANTAR ..... **viii**
DAFTAR ISI ..... **x**
DAFTAR GAMBAR ..... **xiii**
DAFTAR TABEL ..... **xiv**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Rumusan Masalah 3
1.3 Tujuan 3
1.4 Manfaat 4
1.5 Batasan Masalah 4
1.6 Metodologi Penelitian 4
1.7 Sistematika Penulisan 5
BAB II TINJAUAN PUSTAKA **7**
2.1 Penelitian Terkait 7
2.2 *Malicious Software* 11
2.2.1 Jenis – Jenis *Malware* 11
2.3 *Spyware* 13
2.4 CoolWebSearch (CWS) 14
2.5 Visualisasi Data 14
2.6 Machine Learning 18
2.7 Naïve Bayes Classiffier 18

2.8 Confusion Matrix 21
2.9 Google Colaboratory 22
2.10 Cross Validation 23
2.10.1 K-Fold Cross Validation 23
2.11 Dataset CIC-MalMem-2022 24
2.12 *Correlation Based Feature Selection (CFS)* 25
BAB III METODOLOGI PENELITIAN 27
3.1 Pendahuluan 27
3.2 Kerangka Kerja Penelitian 27
3.3 Perancangan Sistem 29
3.4 Spesifikasi Perangkat Keras dan Lunak 30
3.4.1 Perangkat Keras (Hardware) 30
3.4.2 Perangkat Lunak (Software) 30
3.5 Dataset Malware 31
3.6 Pre-Processing 34
3.6.1 Konversi Label dari String ke Number 34
3.6.2 Feature Selection 35
3.6.3 Visualisasi 36
3.7 Skenario Percobaan 37
3.8 Validasi Hasil **37**
3.8.1 Cross Validation 38
BAB IV HASIL DAN ANALISA 39
4.1 Pendahuluan 39
4.2 Analisis Dataset 39
4.2.1 Dataset 39
4.2.2 Pemilihan Data Menggunakan Data Balance 41
4.3 Analisa Visualisasi Hasil 43
4.3.1 *Correlation Based Feature Selection* 43
4.3.2 *Naïve Bayes Classifier* 45

4.4 Validasi Pengujian K-Fold 47
4.5 Validasi Pola dan Perhitungan *Confusion Matrix* *48*
BAB V KESIMPULAN DAN SARAN 50
5.1 Kesimpulan 50
5.1 Saran 51
DAFTAR PUSTAKA 52

DAFTAR GAMBAR

Gambar 2.1 Grafik Parallel Coordinates 15

Gambar 2.2 Grafik Pie Chart 15

Gambar 2.3 Grafik Bar Chart 16

Gambar 2.4 Grafik Line Chart 16

Gambar 2.5 *Naïve Bayes Classifier* 19

Gambar 2.6 K-Fold Cross Validation 23

Gambar 3.1 Kerangka Kerja Penelitian 28

Gambar 3.2 Perancangan Sistem 29

Gambar 3.3 *Flowchart* Konversi Label dari String ke Number 35

Gambar 3.4 *Flowchart* Feature Selection 36

Gambar 4.1 Dataset CIC-MalMem-2022 40

Gambar 4.2 Pemilihan Data Menggunakan Data Balance 42

Gambar 4.3 Tampilan Heatmap Correlation 43

Gambar 4.5 Hasil visualisasi Menggunakan *Parallel Coordinates* *46*

Gambar 4.6 Hasil Visualisasi *Parallel Coordinates* *47*

Gambar 2.7 Nilai *Confusion Matrix* 10% Data *training* 48

DAFTAR TABEL

Tabel 2.1 Matrix Penelitian Terkait 7
Tabel 2.2 *Representasi Confusion Matrix* 21
Tabel 2.3 *Tabel Confusion Matrix* 21
Tabel 2.4 Family Malware *Spyware* 24
Tabel 3.1 Spesifikasi Hardware Komputer yang digunakan 30
Tabel 3.2 Daftar Software yang digunakan 30
Tabel 3.3 Fitur di Dalam Dataset CIC-MalMem-2022 31
Tabel 3.4 Skenario Percobaan 37
Tabel 4.1 Tampilan Lengkap Fitur Variabel Pada Kolom 40
Tabel 4.2 Feature Selection 44
Tabel 4.3 Fitur yang dipilih dari metode NBC 46
Tabel 4.4 Pengujian Menggunakan K-Fold 48

BAB I
PENDAHULUAN

1.1 Latar Belakang

Cyber security adalah salah satu resiko terpenting untuk semua jenis cyber physical system (CPS). Dalam penelitian ini [1] Untuk mengevaluasi resiko cyber security pada jaringan CPS, model penilaian hirarki kuantitatif berdasarkan tinggi serangan yang mengancam, probabilitas keberhasilan serangan suatu malware dan konsekuensi serangan yang diusulkan dapat mengevaluasi resiko yang dibawa oleh malware yang berkelanjutan di tingkat host dan tingkat sistem. Kemudian definisi dan metode perhitungan indikator tersebut dibahas secara rinci.

Diperkirakan bahwa lebih dari 286 juta varian unik malware dirilis pada tahun 2010, perhari malware menciptakan rata – rata 784.000 varian dikutip menurut sebuah laporan oleh Symantec Corp1. Jumlah varian malware unik meningkat 41% pada tahun 2011, dan terus meningkat pada tahun 2012. The *Statistic Sclearly* menunjukkan bahwa malware adalah masalah serius dan dengan demikian, tidak mengherankan lagi bahwa ada sejumlah besar studi yang dilakukan oleh para peneliti untuk menganalisis banyak sampel malware secara manual[2].

Ciri khas utama malware yaitu jika perangkat lunak menunjukkan aktivitas berbahaya seperti mencuri data pengguna, mereplikasi, menonaktifkan fitur keamanan tertentu (aplikasi anti virus), atau menjalankan perintah yang tidak dimaksudkan oleh pengguna, maka itu dapat dianggap sebagai malware. Teknik baru untuk visualisasi malware yang menyoroti aspek perilaku malware akan diungkapkan. Teknik ini menampilkan perilaku malware dalam bentuk gambar (disebut Malware Behaviour Image). Metode yang disajikan dapat digunakan untuk analisis malware, deteksi dan identifikasi varian malware[3].

Secara umum, *Spyware* adalah perangkat lunak berbahaya yang tanpa diketahui diinstal pada mesin korban untuk memonitor perilaku pengguna dan mengumpulkan informasi pengguna, merekam histori web, dan mencuri kata sandi pengguna. *Spyware* ada dalam berbagai bentuk dan melakukan tindakan dari

berbagai tingkat kejahatan. Alat anti-*spyware* saat ini beroperasi dengan cara yang mirip dengan anti virus [4].

Klasifikasi Bayesian[5] didasarkan dari teorema Bayes, klasifikasi bayes merupakan model klasifikasi probabilistik berdasarkan pada asumsi yang cukup sederhana dan seringkali disebut "Naïve" karena tidak tergantung pada nilai-nilai fitur lainnya. Klasifikasi Bayesian dapat memprediksi probabilitas kelas atau label dari data yang tidak dikenal berdasarkan data yang telah dipelajari, seperti probabilitas bahwa sampel yang diberikan termasuk dalam kelas tertentu. Teorema bayes ini dibuat untuk menyederhanakan perhitungan dan juga dapat di mengerti sebagai algoritma pembelajaran sederhana yang menggunakan aturan Bayes bersama dengan asumsi yang kuat bahwa atribut tersebut independen, Naïve Bayes sering memberikan akurasi klasifikasi yang kompetitif, Ditambah dengan efisiensi komputasi dan banyak fitur lain yang diinginkan, yang menyebabkan Naïve Bayes sering diterapkan secara luas dalam berbagai penelitian[6].

Pada penelitian sebelumnya[7] data yang di peroleh melalui analisis memori dapat memberikan wawasan penting tentang perilaku serta pola malware, pada penelitian ini, penggunaan data memori pada deteksi malware disarankan. Deteksi malware dilakukan dengan menggunakan berbagai pendekatan deep learning dan machine learning di lingkungan big data menggunakan data memori. Studi ini dilakukan menggunakan Pyspark pada platform big data Apache Spark di Google Colaboratory. Penelitian dilakukan pada data CIC-MalMem-2022 menggunakan data yang seimbang. Klasifikasi biner dirancang menggunakan Random Forest, Decision Tree, Gradient Boosted Tree, Logistic Regression, Naïve Bayes, Linear Vector Support Machine, Multilayer Perceptron, Deep Feed Forward Neural Network, dan algorimta Long Short-Term Memory.

Hasilnya dievaluasi menggunakan metrik kinerja Accuracy,F1-score, Precision, Recall, dan AUC. Hasilnya, pendeteksian malware yg paling berhasil diperoleh menggunakan algoritma Regresi Logistik, dengan tingkat akurasi 99,97% pada pendeteksian malware menggunakan analisis memori. Gradient Boosted Tree mengikuti algoritma Regresi Logistik menggunakan akurasi 99,94%. algoritma Naive Bayes memberikan performa terendah dalam analisis malware menggunakan data memori, menggunakan akurasi 98,41%.

Pada penelitian sebelumnya[8] klasifikasi adalah proses untuk membedakan kelas data, dengan tujuan untuk dapat memperkirakan kelas suatu objek menggunakan label. Salah satu metode yang sering digunakan untuk mengklasifikasikan data adalah Naïve Bayes Classifier. Kelebihan dari metode ini yaitu algoritma yang sederhana serta akurasi yang tinggi. di penelitian ini akan ditunjukkan kelebihan Naïve Bayes Classifier untuk mengklasifikasikan kualitas suatu jurnal yang biasa dianggap dengan Quartile. Penelitian ini memakai dataset sebanyak 1491 data. Hasinya menunjukkan akurasi sebesar 71,60% serta tingkat kesalahan sebanyak 8,40%.

Pada penelitian sebelumnya[9] Makalah ini mengusulkan metode pengisian celah air permukaan sesuai klasifikasi Naive Bayes. Metode ini mempertimbangkan korelasi antara badan air yang terputus serta tidak bergantung pada data medan. saat citra tertutup awan tebal, metode ini pula dapat merekonstruksi luasan air lengkap secara akurat. lima wilayah studi menggunakan skenario yang berbeda termasuk sungai, danau atau waduk, dipilih untuk mengevaluasi metode tersebut. hasil menunjukkan bahwa rata-homogen akurasi pengisian celah di kelima daerah studi ialah lebih dari 90%.

1.2 Rumusan Masalah

Perumusan masalah pada penelitian ini akan membahas tentang bagaimana cara memvisualisasikan Serangan *Malware Spyware* menggunakan metode *Naïve Bayes Classifier*, yang bertujuan untuk membedakan data *Spyware* dan data normal. Selain itu, bagaimana memvisualisasikan *Spyware* dalam bentuk grafik sehingga dapat lebih mudah di pahami dan mengklasifikasikan *Naïve Bayes* agar mendapatkan hasil akurasi dan model terbaik.

1.3 Tujuan

Berikut tujuan yang akan dicapai pada penelitian ini adalah sebagai berikut :

1. Menerapkan metode *Naïve Bayes Classifier* dalam visualisasi serangan *Malware Spyware*
2. Menganalisis tingkat akurasi dari pengklasifikasian *Spyware* menggunakan algoritma *Naïve Bayes Classifier*

3. Memvisualisasikan dan mengelompokkan data *Spyware* dan data normal dalam bentuk grafik dengan metode *Naïve Bayes Classifier*

1.4 Manfaat

Adapun manfaat yang dapat diambil dari penelitian ini adalah :

1. Mampu membedakan data *Spyware* dan data normal
2. Mampu mempelajari proses dalam visualisasi dan kemudahan dalam membedakan data *Spyware* dan data normal
3. Mampu mengetahui tingkat akurasi algoritma Naïve Bayes Classifier dalam mengklasifikasikan serangan *Spyware*.

1.5 Batasan Masalah

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

1. Algoritma yang digunakan dalam visualisasi *spyware* yaitu algoritma Naïve bayes classifier.
2. Dataset yang digunakan berasal dari *Canadian Institute for Cybersecurity (CIC)* yaitu CICMalMem2022 dengan jenis *spyware* dan benign.
3. Dalam penelitian ini tidak membahas tentang bagaimana pencegahan terhadap malware *spyware*.
4. Malware yang digunakan dalam penelitian ini hanya *Spyware-CWS.*

1.6 Metodologi Penelitian

Pada metodologi yang digunakan pada penelitian ini akan melewati beberapa tahapan penelitian diantaranya :

1. Studi Pustaka/Literatur

   Pada tahapan ini dilakukan setelah masalah yang dibahas sudah sesuai untuk dijadikan penelitian, dengan membaca literatur yang relevan dengan topik penelitian dan mencari dataset yang akan digunakan.
2. Pengolahan Data

Pada tahapan ini akan membahas proses tentang bagaimana mengolah suatu data mentah agar dapat di olah, menerapkan metode serta memvisualisasikan data.

3. Visualisasi

   Pada tahapan ini dilakukan proses visualisasi data botnet dan data normal hingga selesai, dilanjutkan pada proses validasi.

4. Analisa

   Pada tahapan ini melakukan analisis menggunakan hasil yang telah didapatkan sebelumnya agar mendapatkan hasil yang objektif.

5. Kesimpulan dan Saran

   Pada tahapan terakhir membuat kesimpulan dari permasalahan, studi pustaka, metodelogi, dan analisa hasil visualisasi. Selain itu terdapat beberapa saran yang dapat dijadikan dasar pada penelitian selanjutnya.

1.7 Sistematika Penulisan

Dalam penyusunan laporan tugas akhir ini, penulis menerapkan sistematika penulisan sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini berisi penjelasan mengenai topik penelitian yang diambil secara sistematis yang meliputi latar belakang, tujuan, manfaat, rumusan masalah, batasan masalah, metodologi penelitian, dan sistematika penulisan.

BAB II. TINJAUAN PUSTAKA

Bab ini berisikan teori dari penelitian yang berkaitan langsung dengan pokok penelitian ini mengenai serangan *malware spyware, naïve bayes classifier,* dan visualisasi.

BAB III. METODOLOGI PENELITIAN

Pada bab ini akan membahas proses penelitian sistematis untuk memenuhi syarat tugas akhir. Penjelasan pada sub bab ini meliputi tahapan pengaturan sistem dan penerapan metode pada suatu sistem dalam penelitian.

BAB IV. HASIL DAN PEMBAHASAN

Pada bab ini merupakan hasil dari pengumpulan data dari serangan Malware *Spyware* menggunakan metode Naïve Bayes Classifier. Dan Analisa dari hasil deteksi serangan berdasarkan parameter yang ditentukan.

BAB V. KESIMPULAN DAN SARAN

Bab ini akan menyajikan kesimpulan akhir dari data penelitian yang dilakukan, serta saran yang diperlukan untuk penelitian selanjutnya.

DAFTAR PUSTAKA

[1] W. Wu, R. Kang, and Z. Li, "Risk assessment method for cyber security of cyber physical systems," *Proc. 2015 1st Int. Conf. Reliab. Syst. Eng. ICRSE 2015*, pp. 0–4, 2015, doi: 10.1109/ICRSE.2015.7366430.

[2] Y. Ye, T. Li, D. Adjeroh, and S. S. Iyengar, "A survey on malware detection using data mining techniques," *ACM Comput. Surv.*, vol. 50, no. 3, 2017, doi: 10.1145/3073559.

[3] M. F. A. Muhammad Rashidi Wahab, "Jurnal Teknologi," *J. Teknol.*, vol. 5, pp. 31–39, 2013.

[4] M. Egele, C. Kruegel, E. Kirda, and D. Song, "Dynamic Spyware Analysis," *Analysis*, pp. 233–246, 2007, doi: http://portal.acm.org/citation.cfm?id=1364403.

[5] Y. Karaca and C. Cattani, "7. Naive Bayesian classifier," *Comput. Methods Data Anal.*, pp. 229–250, 2018, doi: 10.1515/9783110496369-007.

[6] G. I. Webb, "Encyclopedia of Machine Learning and Data Science," *Encycl. Mach. Learn. Data Sci.*, no. January 2016, 2020, doi: 10.1007/978-1-4899-7502-7.

[7] 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.

[8] A. P. Wibawa *et al.*, "Naïve Bayes Classifier for Journal Quartile Classification," *Int. J. Recent Contrib. from Eng. Sci. IT*, vol. 7, no. 2, p. 91, 2019, doi: 10.3991/ijes.v7i2.10659.

[9] B. Bai *et al.*, "Naive Bayes classification-based surface water gap-filling from partially contaminated optical remote sensing image," *J. Hydrol.*, vol. 616, no. November 2021, 2023, doi: 10.1016/j.jhydrol.2022.128791.

[10] H. Saputra, S. Basuki, and M. Faiqurahman, "Implementasi teknik seleksi fitur pada klasifikasi malware Android menggunakan support vector machine (SVM)," *Repositor*, vol. 1, no. 1, p. 1, 2019, doi: 10.22219/repositor.v1i1.1.

[11] J. Zhou, W. Niu, X. Zhang, Y. Peng, H. Wu, and T. Hu, "Android Malware

Classification Approach Based on Host-Level Encrypted Traffic Shaping," *2020 17th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. ICCWAMTIP 2020*, pp. 246–249, 2020, doi: 10.1109/ICCWAMTIP51612.2020.9317429.

[12] H. M. Kim, H. M. Song, J. W. Seo, and H. K. Kim, "Andro-Simnet: Android Malware Family Classification using Social Network Analysis," *2018 16th Annu. Conf. Privacy, Secur. Trust. PST 2018*, pp. 1–8, 2018, doi: 10.1109/PST.2018.8514216.

[13] 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.

[14] S. A. Roseline, S. Geetha, S. Kadry, and Y. Nam, "Intelligent Vision-Based Malware Detection and Classification Using Deep Random Forest Paradigm," *IEEE Access*, vol. 8, pp. 206303–206324, 2020, doi: 10.1109/ACCESS.2020.3036491.

[15] L. Massarelli, L. Aniello, C. Ciccotelli, L. Querzoni, D. Ucci, and R. Baldoni, "AndroDFA: Android malware classification based on resource consumption," *Inf.*, vol. 11, no. 6, pp. 31–38, 2020, doi: 10.3390/INFO11060326.

[16] R. Chaganti, V. Ravi, and T. D. Pham, "Image-based malware representation approach with EfficientNet convolutional neural networks for effective malware classification," *J. Inf. Secur. Appl.*, vol. 69, no. August, p. 103306, 2022, doi: 10.1016/j.jisa.2022.103306.

[17] S. Talukder and Z. Talukder, "A Survey on Malware Detection and Analysis Tools," *Int. J. Netw. Secur. Its Appl.*, vol. 12, no. 2, pp. 37–57, 2020, doi: 10.5121/ijnsa.2020.12203.

[18] I. A.Saeed, A. Selamat, and A. M. A. Abuagoub, "A Survey on Malware and Malware Detection Systems," *Int. J. Comput. Appl.*, vol. 67, no. 16, pp. 25–31, 2013, doi: 10.5120/11480-7108.

[19] M. Sikorski and A. Honig, *Practical malware analysis: the hands-on guide to dissecting malicious software. no starch press*. 2012.

[20] N. Hoque, D. K. Bhattacharyya, and J. K. Kalita, "Botnet in DDoS Attacks:

Trends and Challenges," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 4, pp. 2242–2270, 2015, doi: 10.1109/COMST.2015.2457491.

[21] B. Stone-Gross, "Malware Analysis of the Lurk Downloader," pp. 1–8, 2014, [Online]. Available: https://www.secureworks.com/research/malware-analysis-of-the-lurk-downloader.

[22] S. Kim, J. Park, K. Lee, I. You, and K. Yim, "A Brief Survey on Rootkit Techniques in Malicious Codes," *J. Internet Serv. Inf. Secur.*, vol. 3, no. 4, pp. 134–147, 2012, [Online]. Available: http://isyou.info/jisis/vol2/no34/jisis-2012-vol2-no34-12.pdf.

[23] R. K. Shahzad and N. Lavesson, "Detecting scareware by mining variable length instruction sequences," *2011 Inf. Secur. South Africa - Proc. ISSA 2011 Conf.*, 2011, doi: 10.1109/ISSA.2011.6027523.

[24] L. Nataraj and B. S. Manjunath, "SPAM: Signal Processing to Analyze Malware [Applications Corner]," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 105–117, 2016, doi: 10.1109/msp.2015.2507185.

[25] R. Tahir, "A Study on Malware and Malware Detection Techniques," *Int. J. Educ. Manag. Eng.*, vol. 8, no. 2, pp. 20–30, 2018, doi: 10.5815/ijeme.2018.02.03.

[26] H. M. Salih and M. S. Mohammed, "Spyware Injection in Android using Fake Application," *Proc. 2020 Int. Conf. Comput. Sci. Softw. Eng. CSASE 2020*, pp. 100–105, 2020, doi: 10.1109/CSASE48920.2020.9142101.

[27] J. Stewart, "This business of malware," *Inf. Secur. Tech. Rep.*, vol. 9, no. 2, pp. 35–41, 2004, doi: 10.1016/S1363-4127(04)00023-8.

[28] J. J. Van Wijk, "The value of visualization," *Proc. IEEE Vis. Conf.*, p. 11, 2005, doi: 10.1109/VIS.2005.102.

[29] M. Khan and S. Shah Khan, "Data and Information Visualization Methods, and Interactive Mechanisms: A Survey," *Int. J. Comput. Appl.*, vol. 34, no. 1, pp. 975–8887, 2011.

[30] M. Batta, "Machine Learning Algorithms - A Review," *Int. J. Sci. Res.*, vol. 18, no. 8, pp. 381–386, 2018, doi: 10.21275/ART20203995.

[31] K. Vembandasamy, R. Sasipriya, and E. Deepa, "Heart Diseases Detection Using Naive Bayes Algorithm," vol. 2, no. 9, pp. 441–444, 2015.

[32] S. Taheri and M. Mammadov, “Learning the naive bayes classifier with optimization models,” *Int. J. Appl. Math. Comput. Sci.*, vol. 23, no. 4, pp. 787–795, 2013, doi: 10.2478/amcs-2013-0059.

[33] J. C. Griffis, J. B. Allendorfer, and J. P. Szaflarski, “Voxel-based Gaussian naïve Bayes classification of ischemic stroke lesions in individual T1-weighted MRI scans,” *J. Neurosci. Methods*, vol. 257, pp. 97–108, 2016, doi: 10.1016/j.jneumeth.2015.09.019.

[34] A. Anagaw and Y. L. Chang, “A new complement naïve Bayesian approach for biomedical data classification,” *J. Ambient Intell. Humaniz. Comput.*, vol. 10, no. 10, pp. 3889–3897, 2019, doi: 10.1007/s12652-018-1160-1.

[35] J. T. Townsend, “Erratum to: Theoretical analysis of an alphabetic confusion matrix.,” *Percept. Psychophys.*, vol. 10, no. 4, p. 256, 1971, doi: 10.3758/BF03212817.

[36] S. Visa, B. Ramsay, A. Ralescu, and E. Van der Knaap, “Edited by Sofia Visa, Atsushi Inoue, and Anca Ralescu,” *Maics*, vol. 710, pp. 120–127, 2011.

[37] P. Kanani and M. Padole, “Deep learning to detect skin cancer using google colab,” *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6, pp. 2176–2183, 2019, doi: 10.35940/ijeat.F8587.088619.

[38] C. Bergmeir and J. M. Benítez, “Forecaster performance evaluation with cross-validation and variants,” *Int. Conf. Intell. Syst. Des. Appl. ISDA*, pp. 849–854, 2011, doi: 10.1109/ISDA.2011.6121763.

[39] S. Yadav and S. Shukla, “Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification,” *Proc. - 6th Int. Adv. Comput. Conf. IACC 2016*, no. Cv, pp. 78–83, 2016, doi: 10.1109/IACC.2016.25.

[40] N. Gopika and A. E. A. Meena Kowshalaya, “Correlation Based Feature Selection Algorithm for Machine Learning,” *Proc. 3rd Int. Conf. Commun. Electron. Syst. ICCES 2018*, no. Icces, pp. 692–695, 2018, doi: 10.1109/CESYS.2018.8723980.