KLASIFIKASI *SCAREWARE MALWARE* PADA ANDROID MENGGUNAKAN METODE *RANDOM FOREST*

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer

OLEH :

SANDI NOPRIANSYAH

09011381722111

JURUSAN SISTEM KOMPUTER

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2021

LEMBAR PENGESAHAN

KLASIFIKASI *SCAREWARE MALWARE* PADA ANDROID MENGGUNAKAN METODE *RANDOM FOREST*

TUGAS AKHIR

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
Oleh :

Sandi Nopriansyah
09011381722111

Palembang, November 2021
Mengetahui,

Pembimbing I Tugas Akhir	Pembimbing II Tugas Akhir
Ahmad Heryanto, M.T.	**Ahmad Fali Oklilas, M.T.**
NIP. 198701222015041002	**NIP. 197210151999031001**

Ketua Jurusan Sistem Komputer

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

ii

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Senin

Tanggal : 18 Oktober 2021

Tim Penguji :

1. **Ketua** : **Sarmayanta Sembiring, M.T**
2. **Sekretaris** : **Rendyansyah, S.Kom., M.T**
3. **Penguji** : **Huda Ubaya, M. T**
4. **Pembimbing 1** : **Ahmad Heryanto, M.T**
5. **Pembimbing 2** : **Ahmad Fali Oklilas, M.T**

Mengetahui,
Ketua Jurusan Sistem Komputer

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

iii

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Sandi Nopriansyah

NIM : 09011381722111

Judul : Klasifikasi *Scareware Malware* pada Android Menggunakan Metode *Random Forest*

Hasil pengecekan Plagiat/Turnitin : 15%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil plagiat atau penjiplakan. 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 tidak ada paksaan dari pihak manapun.

Palembang, November 2021

Yang menyatakan,

Sandi Nopriansyah

NIM : 09011381722111

Halaman Persembahan

MOTTO :

Jika kau dalam kesusahan dan keterpurukan dan segala
Usaha telah kau kerahkan dan itu masih belum cukup
Ingat kau tak sendirian
Minta ke tuhan agar semua dapat baik- baik saja.

Jangan karena mengejar gelar sarjana kau
Lupa gelar sajadah.

Ku persembahkan untuk:

- **Terkhusus untuk Ayah dan Ibu terimakasih banyak atas segala harapan yang telah kalian berikan, sampai kapan pun saya sangat berterima kasih kepada kalian, terima kasih keringat,doa,serta senyum, yang selalu terpikir di dalam pikiran ku agar dapat menyelesaikan ini.**
- **Teman-teman seperjuangan Sistem Komputer Universitas Sriwijaya angkatan 2017**
- **Almamaterku Universitas Sriwijaya**

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

Puji syukur Alhamdulillah penulis panjatkan atas kehadirat Allah SWT yang telah memberikan karunia dan rahmat-Nya, sehingga penulis dapat menyelesaikan penulisan Laporan Tugas Akhir ini yang berjudul "KLASIFIKASI *SCAREWARE MALWARE* PADA ANDROID MENGGUNAKAN METODE *RANDOM FOREST*" .

Dalam Laporan tugas akhir ini penulis menjelaskan mengenai Klasifikasi *malware* jenis *Scareware* pada *platform android* dengan menggunakan metode *random forest* berserta dengan data-data hasil penelitian yang saya lakukan. Harapan saya agar tulisan ini dapat bermanfaat serta menjadi penambah wawasan bagi pembaca.

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

1. Allah Subhanahu Wata'ala yang telah memberikan berkah serta nikmat kesehatan dan kesempatan kepada penulis dalam melaksanakan tugas akhir.
2. Orang tua dan saudaraku tercinta yang telah memberikan do'a dan dukungan baik secara oril maupun materil dan tidak lupa keluarga besar penulis.
3. Bapak Ahmad Heryanto, S.Kom., M.T., selaku Dosen Pembimbing Tugas Akhir 1 dan Bapak Ahmad Fali Oklilas, M.T., selaku Dosen Pembimbing Tugas Akhir 2 yang telah berkenan meluangkan waktunya guna membimbing, memberikan saran dan motivasi serta bimbingan terbaik untuk penulis dalam menyelesaikan Tugas Akhir ini.
4. Bapak Dr. Ir. Sukemi, M.T. selaku Ketua Jurusan Sistem Komputer Universitas Sriwijaya.
5. Bapak Dr. Ir. Bambang Tutuko. M.T. selaku Pembimbing Akademik di Jurusan Sistem Komputer Universitas Sriwijaya.

6. Teman-teman dan kakak-kakak Selaku yang menjadi teman berdiskusi dalam tugas akhir penulis.
7. Teman- teman seperjuangan dari semester awal hingga semester akhir yang telah menyemagati dan menghibur penulis.
8. Teman-teman Seperjuangan Sistem Komputer Angkatan 2017 serta pihak-pihak yang terlibat dalam membuat Tugas Akhir ini yang tidak dapat disebutkan satu-persatu.
9. Mbak Sari selaku Admin Jurusan Sistem Komputer yang telah membantu penulis dalam hal-hal administrasi.
10. Seluruh pihak yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan semangat serta do'a.
11. Almamater.

Penulis menyadari dalam penyusunan laporan tugas akhir ini masih terdapat banyak kekurangan, karenanya penulis mengharapkan kritik dan saran untuk perbaikan. Semoga laporan tugas akhir ini dapat bermanfaat bagi siapa saja yang membacanya

Wassalamu'alaikum Wr. Wb.

Palembang, November 2020

Penulis,

Sandi Nopriansyah

NIM. 09011381722111

KLASIFIKASI *SCAREWARE MALWARE* PADA ANDROID MENGGUNAKAN METODE *RANDOM FOREST*

Sandi Nopriansyah (09011381722111)
Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya
E-mail : sandinopriansyah00@gmail.com

Abstrak

Malware ialah semua perangkat lunak bersifat jahat, Virus atau Malware pada komputer dapat dengan mudah bekerja melalui cara menempel pada suatu file komputer. Pada dasarnya diciptakanya malware digunakan sebagai tindakan illegal untuk merugikan user ataupun pengguna seperti mencuri,merusak data di dalam media penyimpanan. Contoh *Malware* yang berbahaya salah satunya adalah Scareware, ialah salah satu jenis perangkat lunak yang cukup berbahaya dan terbaru yang bisa menimbulkan ancaman keuangan dan privasi atau data pribadi bagi pemula. Pada penelitian ini mencoba untuk mengkasifikasikan Malware jenis Scareware keluarga *VirusShield* Dataset yang digunakan bersumber dari internet tepatnya di CICMaldroid 2017, hasil akurasi dengan menerapkan *Algoritma Random Forest* ini adalah 90.36%.

Kata kunci **:** Klasifikasi**,** Scareware, *Malware, Virus, Random Forest.*

CLASSIFICATION OF SCAREWARE MALWARE ON ANDROID USING RANDOM FOREST METHOD

Sandi Nopriansyah (09011381722111)
Departement of Computer Engineering, Faculty of Computer Science, Sriwijaya University
E-mail : sandinopriansyah00@gmail.com

Abstract

Malware is all malicious software, Viruses or Malware on a computer can easily work by attaching to a computer file. Basically, the creation of malware is used as an illegal act to harm users or users such as stealing, destroying data on storage media. Examples of dangerous malware, one of which is Scareware, is a type of software that is quite dangerous and the latest that can pose financial and privacy threats or personal data for beginners. In this study, we try to classify the type of Scareware Malware of the VirusShield Dataset family used from the internet, precisely at CICMaldroid 2017, the accuracy result by applying the Random Forest Algorithm is 90.36%.

Keywords: Classification, *Scareware*, *Malware*, Virus, *Random Forest.*

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar belakang

Android adalah sebuah sistem yang berbasiskan linux. Salah satu kelebihan android antara lain adalah pengembangnya yang bersifat terbuka dan tidak harus hanya satu individu, sehingga dapat mempermudahkan untuk para individu pengembang lainya untuk minciptakan fitur terbaru dan lebih baik dari pada sebelumnya sehinga sistem operasi android sesuai dengan keingin dan kehendak mereka tersendiri[1]. semakin berkembangnya technology mulai terciptalah yang namanya situs-situs dan file file jahat yang biasanya dikenal atau disebut malware.

Malware ialah semua perangkat lunak bersifat jahat, Virus/Malware pada komputer dapat dengan mudah bekerja melalui cara menempel pada suatu file computer[2]. Pada dasarnya diciptakanya malware digunakan sebagai tindakan illegal untuk merugikan user ataupun pengguna seperti mencuri,merusak data di dalam media penyimpanan. Contoh *Malware* yang berbahaya salah satunya adalah Scareware, ialah salah satu jenis perangkat lunak yang cukup berbahaya dan terbaru yang bisa menimbulkan ancaman keuangan dan privasi atau data pribadi bagi pemula pengguna. Penanggulanganya secara tradisional, seperti perangkat lunak anti-virus, memerlukan pembaruan rutin dan seringkali tidak memiliki kemampuan untuk mendeteksi contoh baru (tak terlihat).

Pada penelitian yang dilakukan sebelumnya [2].Ada beberapa metode populer yang digunakan oleh peneliti untuk mengklasifikasi *malware*, seperti *Decision Tree(DT)* dan *K-Nears (KNN)*. akan tetapi pada penelitian sebelumnya dengan metode *K-Nears Nighbor*[3][4] dan *Decision Tree*(DT)[2] menyajikan hasil yang kurang dari 50% untuk test evaluasi dan tidak dapat mengkalsifikasi dengan baik . Untuk mengatasi kekurangan tersebut perlu dilakukan penelitian yang lebih baik dan menyajikan hasil lebih dari 50% untuk hasil test evaluasi dan mengklasifikasi data dengan baik dari penelitian sebelumnya. Dan dari penelitian sebelumunya mendapatkan akurasi pada kategori malware untuk metode Random forest dengan nilai recall adalah 48.50% dan untuk nilai presisi-nya adalah 49.90%[4]

Berdasarkan uraian diatas maka dibutuhkan penelitiaan untuk Klasifikasi yang berfokus pada *malware scareware* dan *Benign* ini ada banyak teknik yang bisa digunakan, dimana salah satu teknik yang bisa digunakan adalah *Random Forest*, merupakan algoritma pembelajaran ensemble yang dikembangkan oleh Breiman[5].

Untuk itu penelitan ini mengacu pada penelitian sebelumnya untuk mendapatkan hasil yang lebih baik dan nantinya penulis mengimplementasikan metode *Random Forest* untuk mengkalsifikasikan *Malware Scareware* dengan mengusulkan klasifikasi dan karakteristik dengan baik dari penelitian sebelumnya[4]. Algoritma *Random Forest* salah satu metode pelajar bahasa *assembly* yang menghasilkan banyak pelajar individu lalu menggabungkan hasil akhirnya. *Random Forest* adalah sebuah algoritma atau metode yang dapat digunakan dalam membantu pengkalsifikasian data yang besar. Random Forest mempunyai parameter yang dinamai *decision tree*, *Random Forest* tidak hanya berskala baik ketika ada banyak fitur per vektor fitur, tetapi juga membantunya dalam mengurangi saling ketergantungan (korelasi) antara atribut fitur.

1.2 Rumusan Masalah

Untuk Rumusan masalah pada tugas akhir dan penelitian ini adalah bagaimana penerapan algoritma *Random Forest* dan klasifikasi *malware* dan didapatlah akurasinya.

a. Bagaimana pengklasifikasian jenis *Malware* yaitu *malware Scareware* dan file normal atau data Normal dengan menggunakan metode *Random Forest Classifiers*.
b. Bagaimana membangun sistem analisis dengan mnggunakan program *Pyhton Jyupiter* untuk menganalisa *Malware* secara otomatis.

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

a. Mengklasifikasi *Malware* dengan menggunakan algoritma *Random forest Classifiers* pada *dataset CICAndMal2017*
b. Menerapkan metode *Random Forest* dengan *Feature Selection* untuk klasifikasi *Scarware Malware*.
c. Membandingkan nilai evaluasi terhadap penelitian sebelumnya.

1.3.2 Manfaat

Dari latar belakang yang telah di uraikan sebelumnya, maka manfaat dalam penelitian ini adalah:

a. Mendapatkan hasil dari penerapan metode random forest dalam klasifikasi malware *scareware* dan *Benign*
b. Didapatlah hasil tingkat akurasi klasifikasi dengan menggunakan *metode Random forest* .
c. Agar mendapatkan nilai akurasi yang lebih baik dari *recall* dan *Precision* dari peneliti sebelumnya.

1.4 Batasan Masalah

Dari latar belakang yang telah di uraikan sebelumnya, maka di dapatlah batasan masalah untuk penelitian anatara lain:

a. Pada penelitian ini Data yang digunakan hanya menggunakan dataset
b. Mengklasifikasi *Malware* dan data normal dengan pemerograman Python dan metode *Random forest* kalasifikasi.
c. Untuk penelitian ini tidak ada bahasan tentang bagaiman *Malware* masuk ke dalam komputer.

1.5 Metodelogi Penelitian

Metodologi yang digunakan dalam tugas akhir ini sebagai berikut:

1. Tahapan Studi Pustaka

   Tahapan pertama penelitian ini mencari refrensi ataupun informasi yang berhubungan dengan penelitian ini melalaui paper,jurnal,dan lain-lain.

2. Tahap Perancangan

   Selanjutnya adalah metode perancangan dimana penelitian ini merancang tahapan selajutnya seperti pengambilan data.

3. Tahap Pengambilan Data

   Pada penelitian ini adalah melakukan pengambilan beberapa data sesuai dengan sistem perancanagan sebelumnya.

4. Tahap Observasi

   Tahapan terakhir dilakukanya pengumoulan data dan disimpan lalu diolah.

1.6 Sistematika Penulisan

BAB 1 PENDAHULUAN

Pada bab pertama dan awal ini memiliki isi berdasarkan penjabaran dan topic yanag telah diambil antara lain berisikan latar belakang, batasan masalah, dan rumusan masalah,lalu tujuan dan manfaat.

BAB II TINJAUAN PUSTAKA

Pada bab kedua ini berisikan dasar dari teori penelitian yang menjelaskan Pembahasan tentang apa itu *malware,scareware,random* forest dan lain-lain.

BAB III METODOLOGI PENELITIAN

Pada bab ketiga ini menjelaskan dan membahas tentang tahapan atau rancangan untuk penelitian yang berisikan pengembangan,flowchart dan lainnya.

BAB IV HASIL DAN ANALISIS

Untuk penelitian ini berisikan penjelasan hasil ataupun analisa dari penelitian yang telah dilakukan sebelumnya.

BAB V KESIMPULAN

Dan terakhir adalah bab kelima yaitu menarik dan memberikan kesimpulan dan penambahan saran berdasarkan hasil dan analisa yang telah di lakukan di dalam penelitian ini.

DAFTAR PUSTAKA

[1] 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.

[2] A. H. Lashkari, A. F. A. Kadir, L. Taheri, and A. A. Ghorbani, “Toward Developing a Systematic Approach to Generate Benchmark Android Malware Datasets and Classification,” *Proc. - Int. Carnahan Conf. Secur. Technol.*, vol. 2018-Octob, no. Cic, pp. 1–7, 2018, doi: 10.1109/CCST.2018.8585560.

[3] V. Rahmayanti *et al.*, “Klasifikasi Malware Family Menggunakan Metode K-Nearest Neighbor,” pp. 319–323, 2020.

[4] N. Udayakumar, V. J. Saglani, A. V. Cupta, and T. Subbulakshmi, “Malware Classification Using Machine Learning Algorithms,” *Proc. 2nd Int. Conf. Trends Electron. Informatics, ICOEI 2018*, pp. 1007–1012, 2018, doi: 10.1109/ICOEI.2018.8553780.

[5] M. S. Alam and S. T. Vuong, “Random forest classification for detecting android malware,” *Proc. - 2013 IEEE Int. Conf. Green Comput. Commun. IEEE Internet Things IEEE Cyber, Phys. Soc. Comput. GreenCom-iThings-CPSCom 2013*, pp. 663–669, 2013, doi: 10.1109/GreenCom-iThings-CPSCom.2013.122.

[6] 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.

[7] R. Kesuma Dinata and N. Hasdyna, *Machine Learning*. 2020.

[8] E. Retnoningsih and R. Pramudita, “Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python,” *BINA Insa. ICT J.*, vol. 7, no. 2, pp. 156–165, 2020.

[9] T. Wahyono, *Fundamental of Python for Machine Learning: Dasar-Dasar Pemrograman Python untuk Machine Learning dan Kecerdasan Buatan*. 2018.

[10] O. W. Treatment and C. Proceedings, “E Valuation of the a Pplication U Niformity of S Ubsurface,” vol. 2004, no. 701, pp. 73–83, 2004.

[11] C. Nguyen, Y. Wang, and H. N. Nguyen, “Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic,” *J. Biomed. Sci. Eng.*, vol. 06, no. 05, pp. 551–560, 2013, doi: 10.4236/jbise.2013.65070.

[12] Y. L. Pavlov, “Random forests,” *Random For.*, pp. 1–122, 2019, doi: 10.1201/9780429469275-8.

[13] E. S. Lamdompak Sistem Komputer and F. Ilmu Komputer, “Klasifikasi Malware Trojan Ransomware Dengan Algoritma Support Vector Machine (SVM),” vol. 2, no. 1, pp. 122–127, 2016.

[14] P. A. Zariyah, A. Widayanti, and M. Adrian, “Aplikasi Pengadaan Persediaan Bahan Baku Cepat Basi Dengan Pendekatan Material Requirements Planning (MRP) : Studi Kasus Usaha Ayam Taliwang Khas Eyang Padalarang,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 3, pp. 2660–2667, 2019, doi: 10.25126/jtiik.