SKRIPSI

ANALISIS STATIS *MALWARE* ANDOROID BERDASARKAN FITUR *OPCODE* DAN *PERMISSION* MENGGUNAKAN *RANDOM FOREST*

STATIC ANALYSIS OF ANDROID MALWARE BASED ON OPCODE AND PERMISSION FEATURES USING RANDOM FOREST

Disusun Oleh:

MUHAMMAD AMRAN AZIS

D0221102

PROGRAM STUDI

INFORMATIKA FAKULTAS

TEKNIK UNIVERSITAS

SULAWESI BARAT MAJENE

2025

USULAN PENELITIAN

ANALISIS STATIS *MALWARE* ANDOROID BERDASARKAN FITUR *OPCODE* DAN *PERMISSION* MENGGUNAKAN *RANDOM FOREST*

Skripsi

Disusun Oleh:

MUHAMMAD AMRAN AZIS

D0221102

PROGRAM STUDI

INFORMATIKA FAKULTAS

TEKNIK UNIVERSITAS

SULAWESI BARAT MAJENE

2025

LEMBAR PERSETUJUAN

SKRIPSI

ANALISIS STATIS *MALWARE* ANDROID BERDASARKAN FITUR *OPCODE* DAN *PERMISSION* MENGGUNAKAN *RANDOM FOREST*

Telah disiapkan oleh:

MUHAMMAD AMRAN AZIS

D0221102

Telah dipertahankan di depan Penguji

Pada tanggal 02 Oktober 2025

Susunan Tim Penguji

Dosen Pembimbing I

Arnita Irianti S.Si., M.SI

NIP. 198708062018032001

Dosen Pembimbing II

Wawan Firgiawan, S.T., M.Kom

NIDK. 8948080023

Dosen Penguji I

Dr. Eng. Sulfayanti, S.Si., M.T

NIP. 198903172020122011

Dosen Penguji II

Chairi Nur Insani, S.Kom., M.T

NIP. 199407272025062011

Dosen Penguji III

Diny Anggriani Adnas, S.ST., M.T

NIP. 198904152024061001

LEMBAR PENGESAHAN

ANALISIS STATIS *MALWARE* ANDROID BERDASARKAN FITUR *OPCODE* DAN *PERMISSION* MENGGUNAKAN *RANDOM FOREST*

SKRIPSI

Untuk memenuhi sebagian persyaratan

memperoleh gelar Sarjana Komputer

MUHAMMAD AMRAN AZIS

D0221102

Telah diperiksa dan disetujui oleh:

Pembimbing I	Pembimbing II
Arnita Irianti S.Si., M.SI	**Wawan Firgiawan, S.T., M.Kom**
NIP. 198708062018032001	**NIDK:89480880023**

Dekan Fakultas Teknik,	Ketua Program Studi Informatika
Prof. Dr. Ir. Hafsah Nirwana,. M.T	**Muh. Rafli Rasyid. S.Kom., M.T**
NIP. 19640405199003220002	**NIP. 198808182022031006**

HALAMAN PERNYATAAN

Dengan ini saya menyatakan bahwa karya tulis ini merupakan hasil saya sendiri, penelitian ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar Strata 1 (S1) Perguruan Tinggi Universitas Sulawesi Barat dan sepanjang pengetahuan saya, tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disebutkan dalam kutipan, dikutip dengan cara yang benar, serta dicantumkan dalam daftar pustaka.

Majene, 02 Oktober 2025

Yang membuat pernyataan

Muhammad Amran Azis

HALAMAN PERSEMBAHAN

Tiada lembar skripsi yang paling indah dalam laporan skripsi ini kecuali lembar persembahan. Dengan tulus, skripsi ini penulis persembahkan untuk:

Allah SWT yang telah memberikan kemudahan dan pertolongan sehingga penulis dapat menyelesaikan skripsi ini dengan baik.

Kedua orang tua penulis, bapak Azis dan ibu Hadijah dua orang yang sangat berjasa dalam hidup penulis. Dua orang yang selalu mengusahakan anak keduanya ini menempuh Pendidikan setingi-tingginya, meskipun mereka berdua hanya bisa menempuh Pendidikan sampai tahap dasar. Kepada bapak, terima kasih atas setiap cucuran keringat dan kerja keras yang engkau tukarkan menjadi nafkah demi anakmu bisa sampai tingkat ini. Dan terima kasih telah menjadi contoh untuk menjadi seorang laki-laki yang bertanggung jawab penuh terhadap keluarga. Untuk ibu, terima kasih atas segala motivasi, pesan, do'a dan harapan yang selalu mendampingi setiap langkah dan mengusahakan anakmu menjadi seseorang yang berpendidikan, atas langkah salah yang pernah penulis ambil dalam menolak melanjutkan Pendidikan dahulu. Terima kasih atas kasih sayang tanpa batas yang tak pernah dimakan oleh waktu, atas kesabaran dan pengorbanan yang selalu mengiringi perjalanan hidup penulis. Terima kasih telah menjadi sumber kekuatan dan inspirasi, serta pelita yang tak pernah padam dalam setiap langkah yang penulis tempuh. Terima kasih atas segala apa yang kalian berikan yang tak terhitung jumlahnya. Terakhir terima kasih kepada diri sendiri yang sanggup bertahan sampai saat ini.

BAB I

PENDAHULUAN

A. Latar Belakang

Sistem operasi Android telah mengalami pertumbuhan yang pesat dan menjadi platform seluler paling populer di dunia. Namun, di balik popularitas tersebut, Android juga menjadi target utama serangan siber, khususnya dalam bentuk *malware* (Hadiprakoso, Aditya dan Pramitha, 2022). Seiring dengan terus berkembangnya teknologi sistem operasi dan perangkat keras, permasalahan keamanan menjadi semakin kompleks dan menimbulkan tantangan yang signifikan dalam dunia digital saat ini. Android menerapkan model berbasis izin, yang memungkinkan aplikasi untuk mengakses berbagai informasi penting, termasuk data sistem, informasi perangkat, data pengguna, dan sumber daya eksternal (Razeed dan Nowfeek, 2022). Namun sistem ini juga menimbulkan kerentanan untuk dieksploitasi jika tidak dikelola dengan baik. *Malware* sebagai perangkat lunak berbahaya, didefinisikan sebagai program yang dirancang untuk mendapatkan akses tanpa izin, menyadap, atau merusak sistem dan jaringan komputer (Alhogail dan Alharbi, 2025). *Opcode* adalah fitur statis yang diekstraksi dengan membaca kode program tanpa mengeksekusi aplikasi, dan hal tersebut mencerminkan perilaku dari suatu program (Kakisim, Gulmez dan Sogukpinar, 2022). Urutan *Opcode* pada aplikasi Android berbahaya sering kali berbeda dari aplikasi yang sah, sehingga urutan *Opcode* menjadi fitur penting untuk membedakan keduanya (Lakshmanarao, J. S Mantena, *dkk.*, 2025).

Data dari Kaspersky Security Bulletin 2024 menunjukkan bahwa selama periode November 2023 hingga Oktober 2024, lebih dari 302 juta serangan *malware* berhasil diblokir oleh sistem keamanan Kaspersky. Selain itu, lebih dari 85 juta URL berbahaya terdeteksi, dan 303 ribu pengguna terlindungi dari serangan ransomware (AMR, 2024). Di Indonesia, situasi serupa juga terjadi. Badan Siber dan Sandi Negara (BSSN) mencatat bahwa jumlah serangan siber tertinggi terjadi pada Desember 2024, dengan total 107,9 juta insiden (Setianingsih *dkk.*, 2024). Pada tingkat lokal, Tim Tanggap Insiden Komputer Sulawesi Barat (CIRT)

melaporkan Serangan *malware* berbasis APK yang beredar menjelang Pemilu 2024. *Malware* tersebut di desain untuk mencuri informasi serta kredensial perangkat yang terinfeksi, dengan metode penyebaran melalui pesan WhatsApp yang menyerupai dokumen atau undangan palsu. Kasus ini menegaskan pentingnya kewaspadaan digital, terutama dengan tidak mengunduh maupun membuka file dari sumber yang tidak terpercaya, demi mencegah potensi peretasan dan pencurian data pribadi. (SulbarProv-CSIRT, 2024).

Deteksi *malware* berbasis signature adalah pendekatan tradisional yang bergantung pada pencocokan pola atau karakteristik file dengan basis data signature yang telah ditentukan sebelumnya. Namun, pendekatan ini memiliki keterbatasan yang signifikan karena tidak mampu mendeteksi *malware* baru atau yang sebelumnya tidak dikenal. Oleh karena itu, pendekatan berbasis *Machine Learning* semakin banyak diadopsi untuk meningkatkan efektivitas deteksi, terutama terhadap serangan zero-day (Chatterjee, 2021).

Penelitian sebelumnya telah mengusulkan berbagai metode untuk mendeteksi *malware* Android. Penelitian yang dilakukan oleh (A. Pathak, Kumar dan Barman, 2024), menggunakan analisis statis berdasarkan fitur izin dan menunjukkan bahwa algoritma *Random Forest* mampu mencapai akurasi hingga 97,5%, bahkan hanya dengan menggunakan 48 fitur izin. Hal ini menunjukkan bahwa fitur izin dapat berfungsi sebagai indikator penting dalam membedakan aplikasi berbahaya dari aplikasi benign. Penelitian lain oleh (Pandey dan Lal, 2021), menggunakan unigram *Opcode* yang diekstraksi dari bytecode aplikasi Android dengan teknik *reverse engineering*. Dengan memilih 20 *Opcode* yang paling sering muncul sebagai fitur klasifikasi dan menerapkan algoritma *Random Forest*, penelitian tersebut berhasil mencapai akurasi hingga 95,39%.

Berdasarkan penelitian sebelumnya, penelitian ini mengusulkan pendekatan baru untuk klasifikasi *malware* Android dengan menggabungkan dua jenis fitur statis, urutan *Opcode* dan izin aplikasi. Sebanyak 1000 data di gunakan dalam penelitian ini yang di ambil dari Androzoo dengan ketentuan 500 data *malware* dan 500 data benign. Proses ekstraksi dilakukan langsung dari file APK, di mana *Opcode* di peroleh dari folder smali, sedangkan fitur izin di ekstraksi dari file AndroidManifest.xml.

Dengan demikian, tujuan utama dari penelitian ini Adalah untuk menganalisis, menerapkan dan mengevaluasi *Random Forest* dalam melakukan klasifikasi *malware* android berdasarkan fitur *Opcode* dan izin aplikasi. Nilai inovatif dari penelitian ini terletak pada eksplorasi penggabungan fitur *Opcode* dan izin aplikasi, yang sebelumnya sebagian besar diteliti secara terpisah. Dengan demikian, kebaruan penelitian ini terletak pada integrasi dua fitur statis yaitu *Opcode* dan *Permission* serta penerapan *Hyperparameter tuning*. Diharapkan pendekatan ini dapat memberikan kontribusi yang bermakna dalam meningkatkan sistem keamanan siber, khususnya dalam deteksi dan pencegahan ancaman *malware*.

B. Rumusan Masalah

Adapun Manfaat dari penelitian ini adalah sebagai berikut :

1. Bagaimana mengklasifikasikan aplikasi Android antara *malware* dan benign dengan mengombinasikan fitur *Opcode* dan *Permission* menggunakan algoritma *Random Forest*?

C. Batasan Masalah

Adapun Manfaat dari penelitian ini adalah sebagai berikut :

1. Penelitian ini hanya menggunakan 1.000 sampel aplikasi (500 *malware*, 500 benign) dari dataset AndroZoo. dan
2. Menganalisis fitur *Opcode* dan *Permission* secara statis tanpa melibatkan analisis dinamis.
3. Hanya berfokus pada *binary classification* tidak berfokus pada *family*.

D. Tujuan Penelitian

Adapun Tujuan dari penelitian ini adalah sebagai berikut :

1. Mengembangkan dan mengevaluasi model klasifikasi Andorid berbasis analisis statis dengan memanfaatkan kombinasi fitur *Opcode* dan *Permission* menggunakan algoritma *Random Forest*.

E. Manfaat Penelitian

Adapun Manfaat dari penelitian ini adalah sebagai berikut :

1. Memberikan kontribusi pada pengembangan sistem deteksi *malware* Android berbasis *Machine Learning*, sehingga dapat membantu meningkatkan keamanan pengguna dari ancaman aplikasi berbahaya.

BAB V

PENUTUP

A. Kesimpulan

1. Berdasarkan hasil dari dua belas pengujian menggunakan total data 212, 147 urutan *Opcode*, dan 65 izin sampel, dapat disimpulkan bahwa *Random Forest* cukup baik dalam melakukan klasifikasi, di peroleh akurasi terbaik pada pengujian ketiga dengan penyetelan *Hyperparameter*, akurasi sebesar 98%, presisi 96%, recall 96%, dan F1-score 98%. Akurasi yang lebih optimal jatuh pada pengujian ketiga menggunakan *Random Forest* dengan penyetelan *Hyperparameter* terbaik ada pada parameter GridSearch max_depth = 10, max_features = "sqrt", min_samples_leaf = 1, dan n_estimators = 200, yang menghasilkan akurasi sebesar 99%, presisi 99%, rekal 99%, F1-skor 99%. Hal ini menunjukkan bahwa pemilihan fitur dan penyetelan parameter yang tepat dapat menghasilkan model klasifikasi *malware* yang akurat.
2. Dengan demikian, kebaruan penelitian ini terletak pada integrasi dua fitur statis yaitu *Opcode* dan *Permission* serta penerapan *Hyperparameter tuning*. Hal ini terbukti meningkatkan performa klasifikasi secara signifikan dan berpotensi diimplementasikan dalam sistem deteksi *malware* Android secara real-time.

B. Daftar Pustaka

Ahmed, A.A. *dkk.* (2024) "Android Ransomware Detection Using *Supervised Machine Learning* Techniques Based on Traffic Analysis," *sensors*, 24(1), hlm. 2–21. Tersedia pada: https://doi.org/10.3390/s24010189.

Alfatah, D. (2025) "Analisis Forensik Digital Pada Perangkat Android: Studi Kasus Serangan *Malware*," *Jurnal Komputer*, 3(2), hlm. 37–42. Tersedia pada: https://doi.org/https://doi.org/10.70963/jk.v3i2.102.

Alhogail, A. dan Alharbi, R.A. (2025) "Effective ML-Based Android *Malware* Detection and Categorization," *Electronics (Switzerland)*, 14, hlm. 2–22. Tersedia pada: https://doi.org/10.3390/electronics14081486.

Alotaibi, O., Pardede, E. dan Tomy, S. (2023) "Cleaning Big Data Streams: A Systematic Literature Review," *Technologies*. Multidisciplinary Digital Publishing Institute (MDPI), hlm. 3. Tersedia pada: https://doi.org/https://doi.org/10.3390/technologies11040101.

AMR (2024) *Kaspersky Security Bulletin 2024. Statistics*, *Kaspersky Lab*. Tersedia pada: https://securelist.com/ksb-2024-statistics/114795/ (Diakses: 8 Mei 2025).

Asrori, N., Prastowo, A.T. dan Putra, A.D. (2021) "MEDIA PEMBELAJARAN OLAHRAGA SENAM LANTAI DENGAN AUGMENTED REALITY BERBASIS ANDROID," *Jurnal Informatika dan Rekayasa Perangkat Lunak (JATIKA)*, 2(4), hlm. 559–569. Tersedia pada: http://jim.teknokrat.ac.id/index.php/informatika.

Chatterjee, S. (2021) "Advanced *Malware* Detection in Operational Technology: Signature-Based Vs. Behaviour-Based Approaches," *ESP Journal of Engineering & Technology Advancements*, 1(2), hlm. 273. Tersedia pada: https://doi.org/10.56472/25832646/JETA-V1I2P128.

Falah, M.D. *dkk.* (2021) "Analisis Kepatuhan Keamanan pada Aplikasi Olahraga," *JURNAL INFORMATIKA UPGRIS*, 7(2).

Hadiprakoso, R.B. *dkk.* (2022) *ANALISIS STATIS DETEKSI MALWARE ANDROID MENGGUNAKAN ALGORITMA SUPERVISED MACHINE LEARNING*.

Hadiprakoso, R.B., Aditya, R.W. dan Pramitha, F.N. (2022) "ANALISIS STATIS DETEKSI *MALWARE* ANDROID MENGGUNAKAN ALGORITMA *SUPERVISED MACHINE LEARNING*," *CyberSecurity dan Forensik Digital*, 5, hlm. 1–5. Tersedia pada: https://doi.org/https://doi.org/10.14421/csecurity.2022.5.1.3116.

Ilhamdi, Y. dan Kunang, Y.N. (2021) "ANALISIS *MALWARE* PADA SISTEM OPERASI WINDOWS MENGGUNAKAN TEKNIK FORENSIK," dalam *Bina Darma Conference on Computer Science*. Palembang, Indonesia, hlm. 256–264.

jamaluddin *dkk.* (2024) "Klasifikasi Kanker Payudara Menggunakan Algoritma Neural Network dan *Random Forest*," *jurnal Manajemen Informatika & Sistem Informasi (MISI)*, 7, hlm. 77. Tersedia pada: https://doi.org/10.36595/misi.v5i2.

Kakisim, A.G., Gulmez, S. dan Sogukpinar, I. (2022) "Sequential *Opcode* embedding-based *malware* detection method," *Computers and Electrical Engineering*, 98, hlm. 2–11. Tersedia pada: https://doi.org/10.1016/j.compeleceng.2022.107703.

Khan, K. (2023) "Detecting android *malware* and Prevention Using *Supervised Learning*," *Integrated Journal for Research in Arts and Humanities*, 3(1), hlm. 139–149. Tersedia pada: https://doi.org/10.55544/ijrah.3.1.25.

Kusnadi, S.A. dan Wijaya, A.U. (2021) "Perlindungan Hukum Data Pribadi Sebagai Hak Provasi," *JA: Jurnal Al-Wasath*, 2, hlm. 19–32.

Lakshmanarao, A., Mantena, J. S, *dkk.* (2025) "Android *malware* detection through *Opcode* sequences using deep *learning* LSTM and GRU networks," *Indonesian Journal of Electrical Engineering and Computer Science*, 38(2), hlm. 1106. Tersedia pada: https://doi.org/10.11591/ijeecs.v38.i2.pp1106-1114.

Lakshmanarao, A., Mantena, J. S., *dkk.* (2025) "Android *malware* detection through *Opcode* sequences using deep *learning* LSTM and GRU networks," *Indonesian Journal of Electrical Engineering and Computer Science*, 38(2), hlm. 1106–1114. Tersedia pada: https://doi.org/10.11591/ijeecs.v38.i2.pp1106-1114.

Larasati, F.A., Ratnawati, D.E. dan Hanggara, B.T. (2022) *Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest*. Tersedia pada: http://j-ptiik.ub.ac.id.

Mahendra, M.H., Murdiansyah, D.T. dan Lhaksmana, K.M. (2023) "Analisis Sentimen Tweet COVID-19 Menggunakan Metode K-Nearest Neighbors dengan Ekstraksi Fitur TF-IDF dan CountVectorizer," *Jurnal Ilmu Multidisiplin*, 1, hlm. 39.

Mahmuda, S. (2024) "Implementasi Metode *Random Forest* pada Kategori Konten Kanal Youtube," *Jurnal Jendela Matematika*, 2, hlm. 23.

Matin, I.M.M. *dkk.* (2023) "DETEKSI *MALWARE* MENGGUNAKAN *MACHINE LEARNING* DENGAN METODE ENSEMBLE," *Prosiding Sains Nasional dan Teknologi*, 13(1), hlm. 265. Tersedia pada: https://doi.org/10.36499/psnst.v13i1.9224.

Nurhalizah, R.S., Ardianto, R. dan Purwono, P. (2024) "Analisis *Supervised* dan Un*supervised Learning* pada *Machine Learning*: Systematic Literature Review,"

Jurnal Ilmu Komputer dan Informatika, 4(1), hlm. 63. Tersedia pada: https://doi.org/10.54082/jiki.168.

Pandey, S. dan Lal, R. (2021) "*Opcode*-Based Android *Malware* Detection Using *Machine Learning* Techniques," *International Research Journal of Innovations in Engineering and Technology (IRJIET)*, 5(7), hlm. 56–61. Tersedia pada: https://doi.org/10.47001/IRJIET/2021.507010.

Pathak, A., Kumar, T.S. dan Barman, U. (2024a) "Static analysis framework for *Permission*-based dataset generation and android *malware* detection using *Machine Learning*," *Eurasip Journal on Information Security*, 2024. Tersedia pada: https://doi.org/10.1186/s13635-024-00182-3.

Pathak, Amarjyoti, Kumar, T.S. dan Barman, U. (2024) "Static analysis framework for *Permission*-based dataset generation and android *malware* detection using *Machine Learning*," *Eurasip Journal on Information Security*, 2024(1). Tersedia pada: https://doi.org/10.1186/s13635-024-00182-3.

Pathak, A., Kumar, T.S. dan Barman, U. (2024b) "Static analysis framework for *Permission*-based dataset generation and android *malware* detection using *Machine Learning*," *Eurasip Journal on Information Security*, 2024(1), hlm. 2–12. Tersedia pada: https://doi.org/10.1186/s13635-024-00182-3.

Ramadan, N.S. dan Darwis, D. (2025) "PERBANDINGAN METODE NAÏVE BAYES DAN SVM UNTUK SENTIMEN ANALISIS MASYARAKAT TERHADAP SERANGAN RANSOMWARE PADA DATA KIP-K," *Jurnal Sistem Informasi dan Informatika (Simika)*, 8, hlm. 15.

Razeed, M. dan Nowfeek, M. (2022) "IJRSI) |Volume IX, Issue I," *International Journal of Research and Scientific Innovation*, 9, hlm. 26–30. Tersedia pada: www.rsisinternational.org.

Rininda, G., Santi, I.H. dan Kirom, S. (2023) "Penerapan SVM Dalam Analisis Sentimen Pada Edlink Menggunakan Pengujian Confusion Mantrix," *JATI Jurnal Mahasiswa Teknik Informatika*, 7, hlm. 3337.

Setianingsih, A.Y. *dkk.* (2024) *Laporan tahunan honeynet bssn*. Tersedia pada: https://www.bssn.go.id/honeynet/ (Diakses: 9 Mei 2025).

Sihombing, D.J.C. (2024) "Application of Feature Engineering Techniques and *Machine Learning* Algorithms for Property Price Prediction," *JITSI : Jurnal Ilmiah Teknologi Sistem Informasi*, 5(2), hlm. 72–76. Tersedia pada: https://doi.org/10.62527/jitsi.5.2.241.

Sitorus, Y.W., Sukarno, P. dan Mandala, S. (2021) "Analisis Deteksi *Malware* Android menggunakan metode Support Vector Machine & *Random Forest*," 8, hlm. 12500.

SulbarProv-CSIRT (2024) *Imbauan Keamanan: Penipuan dengan Modus Berkas Aplikasi Berbasis Android (APK) melalui PPS Pemilu 2024*. Tersedia pada: https://csirt.sulbarprov.go.id/imbauan-keamanan-bahaya-*malware*-apk-atas-nama-pemilu-2024 (Diakses: 8 Mei 2025).

Tjahjadi, E. V., Santoso, B. dan Serwin (2023) "*MALWARE* CLASSIFICATION USING *MACHINE LEARNING* TECHNIQUES," *Jurnal Ilmiah Ilmu Komputer*, 2, hlm. 60–70.

Turnip, T.N. *dkk.* (2023) "Klasifikasi *Malware* Android Aplikasi Menggunakan *Random Forest* Berdasarkan Fitur Statik," *Teknik Informatika dan Sistem Informasi*, 10, hlm. 926–936.

Wicaksono, M.H., Purbolaksono, M.D. dan Faraby, S.A. (2023) "Perbandingan Algoritma *Machine Learning* untuk Analisis Sentimen Berbasis Aspek pada Review Female Daily," 10(3), hlm. 3594.