Klasifikasi Malware Android Dengan Menggunakan Metode LightGBM Algoritma

TUGAS AKHIR

Diajukan Untuk Memenuhi
Persyaratan Guna Meraih Gelar Sarjana
Informatika Universitas Muhammadiyah Malang

Hadid Ray Aldy
(201910370311329)

Bidang Minat
(Sistem Keamanan Jaringan)

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MALANG
2023

LEMBAR PERSETUJUAN

Klasifikasi Malware Android Dengan Menggunakan Metode LightGBM Algoritma

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1

Informatika Universitas Muhammadiyah Malang

Menyetujui,

Malang, *30 Desember 2023*

Dosen Pembimbing 1	Dosen Pembimbing 2
Ir Denar Regata Akbi S.Kom., M.Kom.	**Didih Rizki Chandranegara S.kom., M.Kom**
NIP. 10816120591PNS.	**NIP. 180302101992PNS.**

LEMBAR PENGESAHAN

Klasifikasi Malware Android Dengan Menggunakan Metode LightGBM Algoritma

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1

InformatikaUniversitas Muhammadiyah Malang

Disusun Oleh :

HADID RAY ALDY

201910370311329

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji pada tanggal 30 Desember 2023

Menyetujui,

Dosen Penguji 1	Dosen Penguji 2
Christian Sri Kusuma Aditya S.Kom., M.Kom	**Luqman Hakim S.Kom., M.Kom.**
NIP. 180327021991PNS.	**NIP. 10819030658PNS.**

Mengetahui,
Ketua Jurusan Informatika

Ir. Galih Wasis Wicaksono S.kom. M.Cs.
NIP. 10814100541PNS.

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

NAMA : HADID RAY ALDY

NIM 201910370311329

Jurusan : Teknik Informatika

Saya, HADID RAY ALDY, dengan ini menyatakan dengan tulus dan sungguh-sungguh bahwa Tugas Akhir yang saya ajukan dengan judul **"Klasifikasi Malware Android Dengan Menggunakan Metode LightGBM Algoritma"** adalah hasil karya saya sendiri. Saya telah mengerjakannya tanpa plagiarisme dan tanpa melakukan penyalinan dari karya orang lain. Sumber informasi yang digunakan dalam Tugas Akhir ini telah saya akui dengan benar. Saya juga menyatakan bahwa Tugas Akhir ini belum pernah diajukan sebagai tugas atau karya yang sama di tempat lain dan belum pernah digunakan untuk memenuhi persyaratan akademik lainnya. Saya memahami bahwa jika Tugas Akhir ini melanggar kode etik akademik dan norma-norma yang berlaku, saya dapat dikenakan sanksi sesuai peraturan lembaga pendidikan saya. Saya bersedia menjalani proses verifikasi, bila diperlukan, untuk menegaskan keaslian dan orisinalitas Tugas Akhir ini.

Mengetahui,
Dosen Pembimbing

(Ir. Denar Regata Akbi, S.Kom., M.Kom)

Malang, 14 Desember 2023
Yang Membuat Pernyataan

(Hadid Ray Aldy)

ABSTRAK

Android pada tahun 2023 sekitar 3,6 miliar orang di seluruh dunia menggunakan HP Android, dan diketahui bahwa dengan pangsa pasar global sebesar 71,75% pada akhir tahun 2022, Android dianggap sebagai sistem operasi yang paling banyak digunakan. Namun dengan kesuksesan Android saat ini menjadikannya target incaran bagi cyber-crime dalam melancarkan aksi malicious yang salah satunya melalui serangan malware. Malware juga kian waktu terus mengalami peningkatan yang menjadikannya semakin sulit untuk dideteksi. Maka dari itu diperlukan metode deteksi yang andal. Pada bidang IT saat ini, machine learning telah menunjukkan hasil yang cukup efisien dalam mendeteksi malware. Penulis ingin mengusulkan metode Algoritma LightGBM dalam Machine Learning sebagai pendekatan klasifikasi malware Android. Banyak boosting tools menggunakan algoritma berbasis pra-sortir (misalnya, algoritma bawaan pada XGBoost) dalam decision tree learning, yang mana ini adalah solusi yang sederhana, namun tetapi tidak mudah untuk dioptimalkan. Sedangkan LightGBM menggunakan algoritma berbasis Histogram, yang mana mengelompokkan nilai fitur (atribut) kontinu ke dalam bin diskrit, yang hal ini dapat mempercepat training dan mengurangi penggunaan memori. Maka dari itu penulis ingin mengusulkan metode algoritma LightGBM pada penelitian kali ini. Setelah dilakukan penelitian, didapatkan bahwa Model LightGBM berhasil memperoleh Detection Accuracy sebesar 96,37% dan Validation Accuracy sebesar 96,34%, untuk F1-Score-nya sendiri diperoleh 0,9638.

KATA KUNCI: Malware, Android, Machine Learning, Light Gradient-boosting Machine, Klasifikasi

KATA PENGANTAR

Alhamdulillah, Segala puji bagi Allah, Tuhan semesta alam, yang dengan rahmat dan pertolongan-Nya penulis dapat menyusun Tugas Akhir ini. Shalawat serta salam semoga tercurahkan kepada Nabi Muhammad Shalallahu ‘Alaihi Wa Sallam. Dengan izin Allah, penulis dengan rendah hati mempersembahkan Tugas Akhir dengan judul **“Klasifikasi Malware Android Dengan Metode LightGBM Algoritma”**. Tugas Akhir ini adalah hasil kerja keras dan bimbingan berharga dari berbagai pihak.

Penulis ingin mengucapkan terima kasih kepada keluarga, teman-teman, dan dosen yang telah memberikan dukungan dan bimbingan dalam perjalanan ini. Semoga hasil karya ini dapat memberikan kontribusi dalam pemahaman dan penanganan masalah keamanan perangkat Android.

Penulis sadar bahwa Tugas Akhir ini jauh dari kalimat sempurna, dan penulis terbuka untuk menerima kritik dan saran yang membangun untuk perbaikan di masa depan, Semoga Allah Subhanahu Wa Ta’ala memberkahi hasil karya penulis.

Malang, 18 - 12 - 2023

Hadid Ray Aldy

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR PUSTAKA

[1] A, Turner. “The Rise of Android: Why is Android Successful?”. Bankmycell. https://www.bankmycell.com/blog/how-many-android-users-are-there. Diakses pada 5 Juni 2023.

[2] Thomas, D. R., Beresford, A. R., & Rice, A. (2015). Security metrics for the android ecosystem. SPSM 2015 - Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices, Co-Located with: CCS 2015, 87–98. https://doi.org/10.1145/2808117.2808118

[3] J, Drake. “Experts Found a Unicorn in the Heart of Android”. Zimperium. https://www.zimperium.com/blog/the-biggest-splash-at-blackhat-and-defcon-2015/. Diakses pada 5 Juni 2023.

[4] Kapersky. Mobile. Kapersky. https://www.kaspersky.com/resource-center/threats/mobile. Diakses pada 5 Juni 2023.

[5] G, Kaur and A, H, Lashkari. “Understanding Android Malware Families (UAMF) – The Foundations (Article 1)”. IT World Canada. https://www.itworldcanada.com/blog/understanding-android-malware-families-uamf-the-foundations-article-1/441562#:~:text=The%20prominent%20Android%20malware%20categories,banker%2C%20and%20trojan%2Ddropper. Diakses pada 6 Juni 2023.

[6] Bandi, A., & Sherpa, L. (2023). Android Malware Detection Using Machine Learning Classifiers. Lecture Notes on Data Engineering and Communications Technologies, 141, 191–200. https://doi.org/10.1007/978-981-19-3035-5_15

[7] X. -w. Chen and J. C. Jeong, "Enhanced recursive feature elimination," Sixth International Conference on Machine Learning and Applications (ICMLA 2007), Cincinnati, OH, USA, 2007, pp. 429-435, doi: 10.1109/ICMLA.2007.35.

[8] Wikipedia. “Android (Operating System)”. Wikipedia. https://en.wikipedia.org/wiki/Android_(operating_system). Diakses pada 12 Juni 2023.

[9] Aslan, O., & Samet, R. (2020). A Comprehensive Review on Malware Detection Approaches. IEEE Access, 8, 6249–6271. https://doi.org/10.1109/ACCESS.2019.2963724

[10] IBM. “What is machine learning?”. IBM. https://www.ibm.com/topics/machine-learning. Diakses pada 12 Juni 2023.
[11] Katrina, W. “A guide to the types of machine learning alorithms and their applications”. SAS. https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html#:~:text=There%20are%20four%20types%20of,%2Dsupervised%2C%20unsupervised%20and%20reinforcement. Diakses pada 12 Juni 2023.
[12] ArcGIS. “How LightGBM algorithm works”. ArcGIS Pro. https://pro.arcgis.com/en/pro-app/latest/tool-reference/geoai/how-lightgbm-works.htm#:~:text=LightGBM%20is%20a%20gradient%20boosting,high%20performance%20with%20distributed%20systems. Diakses pada 13 Juni 2023.
[13] Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. "Lightgbm: A highly efficient gradient boosting decision tree." Advances in neural information processing systems 30 (2017).
[14] Wikipedia. “LightGBM”. Wikipedia. https://en.wikipedia.org/wiki/LightGBM. Diakses pada 14 Juni 2023.
[15] Mathur, A., Podila, L. M., Kulkarni, K., Niyaz, Q., & Javaid, A. Y. (2021). NATICUSdroid: A malware detection framework for Android using native and custom permissions. Journal of Information Security and Applications, 58, 102696.
[16] Lightgbm Docs. “lightgbm.train”. LightGBM Docs. https://lightgbm.readthedocs.io/en/latest/pythonapi/lightgbm.train.html. Diakses pada 27 Juni 2023.
[17] Android Developers. “Manifest.permission”. Developer Android. https://developer.android.com/reference/android/Manifest.permission . Diakses pada 25 Agustus 2023.
[18] Rihad Variawa, “Data Pre-processing & Data Wrangling”. Medium. https://medium.com/swlh/data-pre-processing-data-wrangling-4a6a8624e747 . Diakses pada 29 Agustus 2023.
[19] Mehta, Manish, Rakesh Agrawal, and Jorma Rissanen. “SLIQ: A fast scalable classifier for data mining.” International Conference on Extending Database Technology. Springer Berlin Heidelberg, 1996.
[20] Shafer, John, Rakesh Agrawal, and Manish Mehta. “SPRINT: A scalable

parallel classifier for data mining." Proc. 1996 Int. Conf. Very Large Data Bases. 1996.

[21] Ranka, Sanjay, and V. Singh. "CLOUDS: A decision tree classifier for large datasets." Proceedings of the 4th Knowledge Discovery and Data Mining Conference. 1998.

[22] Machado, F. P. "Communication and memory efficient parallel decision tree construction." (2003).

[23] Li, Ping, Qiang Wu, and Christopher J. Burges. "Mcrank: Learning to rank using multiple classification and gradient boosting." Advances in Neural Information Processing Systems 20 (NIPS 2007).

[24] Taha, A. A., & Malebary, S. J. (2021). Hybrid classification of Android malware based on fuzzy clustering and the gradient boosting machine. *Neural Computing and Applications*, 33(12), 6721–6732. https://doi.org/10.1007/s00521-020-05450-0

[25] Wang, G., & Liu, Z. (2020). Android malware detection model based on lightGBM. *Advances in Intelligent Systems and Computing*, 1031 AISC, 237–243. https://doi.org/10.1007/978-981-13-9406-5_29

[26] Sarah, N. Al, Rifat, F. Y., Hossain, M. S., & Narman, H. S. (2021). An Efficient Android Malware Prediction Using Ensemble machine learning algorithms. *Procedia Computer Science*, 191(2019), 184–191. https://doi.org/10.1016/j.procs.2021.07.023

[27] Onoja, M., Jegede, A., Mazadu, J., Aimufua, G., Oyedele, A., & Olibodum, K. (2022). Exploring the Effectiveness and Efficiency of LightGBM Algorithm for Windows Malware Detection. *Proceedings of the 5th International Conference on Information Technology for Education and Development: Changing the Narratives Through Building a Secure Society with Disruptive Technologies, ITED 2022, November*. https://doi.org/10.1109/ITED56637.2022.10051488

[28] Nguyen, D. T., & Lee, S. (2021). LightGBM-based Ransomware Detection using API Call Sequences. *International Journal of Advanced Computer Science and Applications*, *12*(10), 138–146. https://doi.org/10.14569/IJACSA.2021.0121016

[29] Shi, H. (2007). Best-first Decision Tree Learning (Thesis, Master of Science

(MSc)). The University of Waikato, Hamilton, New Zealand. Retrieved from https://hdl.handle.net/10289/2317

FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : **Hadid Ray Aldy**

NIM : **201910370311329**

Judul TA : **Klasifikasi Malware Android Dengan Menggunakan Metode LightGBM Algoritma**

Hasil Cek Plagiarisme dengan Turnitin

No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%)
1.	Bab 1 – Pendahuluan	10 %	0%
2.	Bab 2 – Tinjauan Pustaka	25 %	3%
3.	Bab 3 – Metodologi Penelitian	25 %	8%
4.	Bab 4 – Hasil dan Pembahasan	15 %	2%
5.	Bab 5 – Kesimpulan	5 %	0%
6.	Makalah Tugas Akhir	20%	13%

Mengetahui,

Pemeriksa (Staff TU)

(.....)