PERBANDINGAN MODEL DECISION TREE DAN RANDOM FOREST UNTUK KLASIFIKASI SERANGAN MALWARE

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer

Disusun oleh:

HAIDAR AHMAD MA'ARIF
18.83.0284

Kepada

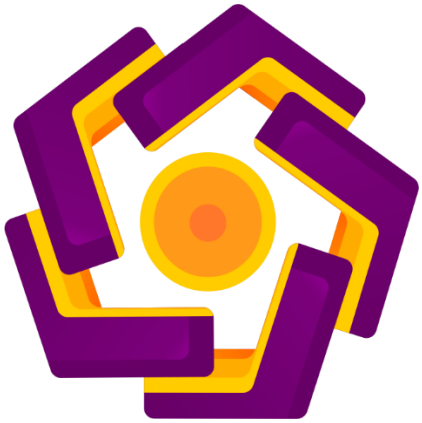
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2023

PERBANDINGAN MODEL DECISION TREE DAN RANDOM FOREST UNTUK KLASIFIKASI SERANGAN MALWARE

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer



Disusun oleh:

HAIDAR AHMAD MA'ARIF
18.83.0284

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2023

HALAMAN PERSETUJUAN

SKRIPSI

PERBANDINGAN MODEL DECISION TREE DAN RANDOM FOREST UNTUK KLASIFIKASI SERANGAN MALWARE

yang disusun dan diajukan oleh

Haidar Ahmad Ma'arif
18.83.0284

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 25 Januari 2023

Dosen Pembimbing

Dony Ariyus M.Kom
NIK. 190302128

HALAMAN PENGESAHAN

SKRIPSI

PERBANDINGAN MODEL DECISION TREE DAN RANDOM FOREST UNTUK KLASIFIKASI SERANGAN MALWARE

yang dipersiapkan dan disusun oleh

Haidar Ahmad Ma'arif

18.83.0284

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 Januari 2023

Susunan Dewan Penguji

Nama Penguji	Tanda Tangan
Wahid Miftahul Ashari, S.Kom., M.T **NIK. 190302452**	
Jeki Kuswanto, M.Kom **NIK. 190302456**	
Dony Ariyus M.Kom **NIK. 190302128**	

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 09 Februari 2023

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al F[illegible], M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Haidar Ahmad Ma'arif
NIM : 18.83.0284

Menyatakan bahwa Skripsi dengan judul berikut:

Perbandingan Model Decision Tree Dan Random Forest Untuk Klasifikasi Serangan Malware

Dosen Pembimbing : Dony Ariyus, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 25 Januari 2023

Yang Menyatakan,

Haidar Ahmad Ma'arif

HALAMAN MOTTO

"Seseorang yang putus asa melihat kesulitan dalam setiap kesempatan, tetapi orang yang optimis melihat peluang dalam setiap kesulitan." – Ali Bin Abi Thalib

HALAMAN PERSEMBAHAN

Segala puji bagi Allah SWT atas limpahan rahmat dan hidayah serta karunia-Nya sehingga skripsi ini selesai dengan sebaik-baiknya. Skripsi ini saya persembahkan untuk :

1. Kedua orang tua saya , Bapak Legiyo dan Ibu Siti Jamzaroh yang selalu mendoakan memberi dukungan, fasilitas serta memberikan jerih payah beliau untuk masa depan anaknya.
2. Bapak Dony Arius , M.Kom Selaku dosen pembimbing yang telah membantu dalam penyusunan skripsi ini.
3. Kepada teman-teman dan sahabat saya , yang ada disaat senang maupun duka selama masa perkuliahan saya hingga skripsi ini diselesaikan.

KATA PENGANTAR

Puji dan syukur ke hadirat Allah Subhanahu Wa Ta'ala, karena berkat rahmat dan hidayahNya maka penulis dapat menyelesaikan skripsi yang berjudul

"**Perbandingan Model Decision Tree Dan Random Forest Untuk Klasifikasi Serangan Malware**"

Shalawat dan salam disampaikan kepada junjungan Rasulullah SAW beserta keluarganya, yang menjadi sumber ilmu yang membawa keselamatan dunia dan akhirat.

Penyusunan skripsi ini merupakan salah satu syarat untuk memperoleh gelar Strata 1 dalam program studi teknik komputer Universitas Amikom Yogyakarta.

Dalam penyusunan skripsi ini berbagai pihak telah memberikan dorongan, bantuan , serta masukan sehingga dalam kesempatan ini penulis menyampaikan terima kasih yang sebesar- besarnya kepada :

1. Ayah dan Ibu tercinta yang senantiasa mendidik , menasehati , dan mendoakan penulis hingga dapat menyelesaikan skripsi ini.
2. Bapak Prof. Dr. M. Suyanto, M.M, selaku Rektor Universitas Amikom Yogyakarta.
3. Bapak Hanif Al-Fatta, M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
4. Bapak Dony Ariyus, M.Kom selaku Kepala Prodi Teknik Komputer dan Dosen Pembimbing yang dengan sabar dan ikhlas telah meluangkan waktu dan memberikan ilmunya kepada penulis dalam memberikan petunjuk dan bimbingan sehingga penulis dapat menyelesaikan skripsi ini.
5. Semua pihak yang selama ini telah memberikan bantuan dalam bentuk moril ataupun materil sehingga penulis dapat menyelesaikan skripsi ini.

Penulis menyadari bahwa penulisan penelitian ini masih banyak terdapat kekurangan ,bila ada benarnya itu atas kehendak Allah, dipersilahkan untuk mengambil manfaatnya, bila ada salahnya itu karena kesalahan dari penulisan sendiri, mohon untuk ditinggalkan.

Terimakasih kepada seluruh pihak yang memberikan dukungan dalam penyelesaian penelitian ini, semoga apa yang telah diberikan dapat bernilai sebagai amalan baik, Akhir kata, mari jadikan ilmu pengetahuan sebagai kekuatan yang dapat mengembalikan system kehidupan menuju arah kebenaran.

Waalamualaikum warahmatullahi wabarakatuh.

Yogyakarta, 06 Januari 2023

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

INTISARI

Serangan malware menjadi semakin meningkat seiring dengan meningkatnya jumlah penggunaan internet dan aplikasi yang terjadi. Oleh karena itu, diperlukan metode deteksi yang efektif untuk mengidentifikasi serangan malware . Salah satu metode yang dapat digunakan adalah machine learning.. Pembelajaran mesin atau *Machine learning* (ML) adalah suatu bidang studi atau bidang keilmuan yang mencakup perencanaan dan pengembangan algoritma yang memampukan suatu komputer dapat mengembangkan perilaku berdasakan data empiris. Di dalam pembelajaran mesin atau *machine learning* (ML) terdapat dua algoritma yaitu pohon keputusan atau *Decision Tree* (DT) dan hutan acak atau *Random Forest* (RF). Pohon Keputusan atau *Decision Tree* (DT) adalah teknik representasi sederhana dari klasifikasi merupakan proses pembelajarn suatu fungsi yang bertujuan memetakan tiap-tiap atribut dan menghasilkan sebuah keputusan. Hutan Acak atau *Random Forest (*RF*)* adalah suatu algoritma yang digunakan untuk klasifikasi data dalam jumlah yang sangat besar dan merupakan gabungan dari pohon (*tree)* dari model pohon keputusan atau *Decision Tree* Dalam penelitian ini, kami menganalisis kinerja model Decision Tree dan model Random Forest dalam mendeteksi serangan malware. Dengan bantuan feature selection yaitu Select K best untuk memilih fitur-fitur yang memiliki nilai korelasi tinggi. Dari hasil kinerja diketahui terdapat 10 fitur yang paling berkorelasi yaitu *Image Base, Check Sum, SizeOf Stack Commit, Size Of Stack Reserve, Size Of Uninitialized Data, Load Configuration Size,Size Of Initialized Data,Resources Max Size,Size Of Heap Reserve dan Loader Flags* Kedua model ini diuji dengan menggunakan data malware yang terdiri dari fitur-fitur yang telah terseleksi.Dan hasil analisis menunjukkan bahwa Random Forest memiliki akurasi 98.2% unggul dari Decision Tree 97.4%.

Kata kunci: *Malware, Decision Tree, Random Forests, Select K best, Machine Learning*

ABSTRACT

Malware attacks are increasing along with the increasing number of internet usage and applications that occur. Therefore, an effective detection method is needed to identify malware attacks. One method that can be used is machine learning.. Machine learning orMachine learning (ML) is a field of study or scientific field that includes the design and development of algorithms that enable a computer to develop behavior based on empirical data. In machine learning or machine learning (ML) there are two algorithms namely Decision Tree orDecision Tree (DT) and Random Forest orRandom Forest (RF). Decision Tree orDecision Tree (DT) is a simple representation technique of classification which is a process of learning a function that aims to map each attribute and produce a decision. Random Forest orRandom Forest (RF) is an algorithm that is used to classify very large amounts of data and is a combination of trees (tree) from a Decision Tree model orDecision Tree In this study, we analyzed the performance of the Decision Tree model and the Random Forest model in detecting malware attacks. With the help of feature selection, namely Select K best to select features that have a high correlation value. From the performance results, it is known that there are 10 features that are most correlated, namely: ImageBase, CheckSum, SizeOfStackCommit,SizeOfStackReserve,SizeOfUninitializedData,LoadConfigurationSize, SizeOfInitializedData, ResourcesMaxSize, SizeOfHeapReserve, and LoaderFlags. Both of these models were tested using malware data which consists of certain features. The results of the analysis show that the Random Forest has an accuracy of 98.2% superior to the Decision Tree of 97.4%.

Keyword *Malware, Decision Tree, Random Forests, Select K best, Machine Learning*