

**DETEKSI *MALWARE* PADA PERANGKAT ANDROID  
MENGUNAKAN *GRADIENT BOOSTING*  
*DECISION TREE* DAN *SMOTE***



**TESIS**

**Qois Al Qorni**

**ENTERPRISE IT INFRASTRUCTURE**

**232420020**

**PROGRAM STUDI TEKNIK INFORMATIKA – S2**

**PROGRAM PASCASARJANA**

**UNIVERSITAS BINA DARMA**

**PALEMBANG**

**2025**

**DETEKSI MALWARE PADA PERANGKAT ANDROID  
MENGUNAKAN GRADIENT BOOSTING DECISION TREE  
DAN SMOTE**



**Tesis ini diajukan sebagai salah satu syarat  
untuk memperoleh gelar**

**MAGISTER KOMPUTER**

**QOIS AL QORNI  
ENTERPRISE IT INFRASTRUCTURE  
232420020**

**PROGRAM STUDI TEKNIK INFORMATIKA – S2  
PROGRAM PASCASARJANA  
UNIVERSITAS BINA DARMA  
PALEMBANG  
2025**

## HALAMAN PENGESAHAN PEMBIMBING TESIS

Judul Tesis: DETEKSI MALWARE PADA PERANGKAT ANDROID  
MENGUNAKAN *GRADIENT BOOSTING DECISION TREE*  
DAN *SMOTE*

Oleh QOIS AL QORNI, NIM 232420020, Tesis ini telah disetujui dan disahkan oleh Pembimbing Program Studi Teknik Informatika – S2 konsentrasi ENTERPRISE IT INFRASTRUCTURE, Program Pascasarjana Universitas Bina Darma pada 27 Agustus 2025 dan telah dinyatakan LULUS.

Palembang, 27 Agustus 2025

Mengetahui,

Program Studi Teknik Informatika – S2

Universitas Bina Darma

Ketua,



.....  
Dr. Usman Ependi, S.Kom., M.Kom.

Pembimbing,


A handwritten signature in black ink is written over a large, faint watermark of the Universitas Bina Darma logo. Below the signature is a dotted line.


.....  
Dr. Usman Ependi, S.Kom., M.Kom.


## HALAMAN PENGESAHAN PENGUJI TESIS


Judul Tesis: DETEKSI MALWARE PADA PERANGKAT ANDROID  
MENGUNAKAN *GRADIENT BOOSTING DECISION TREE*  
DAN *SMOTE*

Oleh QOIS AL QORNI, NIM 232420020, Tesis ini telah disetujui dan disahkan oleh Tim Penguji Program Studi Teknik Informatika – S2 konsentrasi ENTERPRISE IT INFRASTRUCTURE, Program Pascasarjana Universitas Bina Darma pada 27 Agustus 2025 dan telah dinyatakan LULUS.

Palembang, 27 Agustus 2025  
Mengetahui,  
Program Pascasarjana  
Universitas Bina Darma  
Direktur,  
  
PROGRAM PASCASARJANA

Penguji I,  
  
.....  
Dr. Usman Ependi, S.Kom., M.Kom.

Penguji II,  
  
.....  
Dr. Ari Muzakir, M.Cs.

Penguji III,  
  
.....  
Dr. A. Haidar Mirza, S.T., M.Kom.

## SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : QOIS AL QORNI

NIM : 232420020

Dengan ini menyatakan bahwa:

1. Karya tulis Saya (Tesis, Skripsi, Tugas Akhir) ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik baik (Magister, Sarjana, dan Ahli Madya) di Universitas Bina Darma;
2. Karya tulis ini murni gagasan, rumusan dan penelitian Saya sendiri dengan arahan tim pembimbing;
3. Dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam daftar pustaka;
4. Karena yakin dengan keaslian karya tulis ini, Saya menyatakan bersedia Tesis/Skripsi/Tugas Akhir, yang Saya hasilkan di unggah ke internet;
5. Surat Pernyataan ini Saya tulis dengan sungguh-sungguh dan apabila terdapat penyimpangan atau ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima sanksi dengan aturan yang berlaku di perguruan tinggi ini.

Demikian Surat Pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang, 27 Agustus 2025  
Yang Membuat Pernyataan,



QOIS AL QORNI  
NIM: 232420020

## ABSTRAK

Malware merupakan ancaman signifikan terhadap keamanan perangkat Android karena dapat menyamar sebagai aplikasi sah untuk kemudian mencuri data, merusak sistem, maupun memata-matai pengguna. Penelitian ini bertujuan untuk mendeteksi malware pada perangkat Android menggunakan metode *Gradient Boosting Decision Tree* (GBDT) dengan teknik *Synthetic Minority Oversampling Technique* (SMOTE) untuk mengatasi ketidakseimbangan data. Dataset yang digunakan adalah Drebin dan Tuandromd. Proses penelitian meliputi pra-pemrosesan data, pembagian data dengan proporsi 70:30, 80:20, dan 90:10, serta pembangunan model menggunakan algoritma *Decision Tree* dan GBDT baik dengan maupun tanpa SMOTE. Hasil pengujian pada dataset Drebin menunjukkan bahwa model GBDT dengan SMOTE memberikan performa terbaik dengan akurasi hingga 96,42%, presisi 94,60%, recall 90,98%, dan f1-score 92,76% pada proporsi data 80:20. Sementara itu, pada dataset Tuandromd, GBDT dengan SMOTE juga menghasilkan kinerja tertinggi dengan akurasi 93,98%, presisi 90,00%, recall 84,38%, dan f1-score 87,10% pada proporsi 80:20. Temuan ini membuktikan bahwa kombinasi GBDT dan SMOTE efektif meningkatkan kemampuan deteksi malware Android dengan kinerja seimbang pada kelas mayoritas dan minoritas, meskipun memerlukan sumber daya komputasi yang lebih besar.

**Kata Kunci** : *Android, Malware, Gradient Boosting Decision Tree, SMOTE, Deteksi.*

## **ABSTRACT**

*Malware is a significant threat to the security of Android devices because it can disguise itself as a legitimate application to steal data, damage the system, or spy on users. This study aims to detect malware on Android devices using the Gradient Boosting Decision Tree (GBDT) method with the Synthetic Minority Oversampling Technique (SMOTE) technique to address data imbalance. The datasets used are Drebin and Tuandromd. The research process includes data pre-processing, data division with a proportion of 70:30, 80:20, and 90:10, and model building using the Decision Tree and GBDT algorithms both with and without SMOTE. The test results on the Drebin dataset show that the GBDT model with SMOTE provides the best performance with an accuracy of up to 96.42%, a precision of 94.60%, a recall of 90.98%, and an f1-score of 92.76% on a data proportion of 80:20. Meanwhile, on the Tuandromd dataset, GBDT with SMOTE also produced the highest performance with 93.98% accuracy, 90.00% precision, 84.38% recall, and 87.10% f1-score at 80:20 ratio. These findings prove that the combination of GBDT and SMOTE effectively improves Android malware detection capabilities with balanced performance on the majority and minority classes, although it requires greater computational resources.*

**Keywords** : *Android, Malware, Gradient Boosting Decision Tree, SMOTE, Detection.*

## **MOTTO DAN HALAMAN PERSEMBAHAN**

### **MOTTO**

“Seperti model yang tak bisa generalisasi tanpa data yang cukup, hidup pun tak bisa membaik tanpa usaha yang nyata”

### **PERSEMBAHAN**

Karya ini kupersembahkan kepada:

1. Allah SWT, sumber segala ilmu dan kekuatan, yang senantiasa membimbing langkah dalam setiap proses pencarian makna.
2. Keluarga, atas doa yang tak pernah putus, kasih sayang yang tulus, dan dukungan yang menjadi fondasi dalam setiap pencapaian.
3. Dosen dan pembimbing, yang telah menyalakan lentera pengetahuan dan membentuk cara berpikir yang kritis dan bijak.
4. Teman seperjuangan, yang hadir sebagai penguat di tengah tantangan, dan sebagai pengingat bahwa perjalanan ini bukanlah milik satu orang saja.
5. Diriku sendiri, atas keberanian untuk terus belajar, bertumbuh, dan tidak menyerah meski jalan kadang tak mudah.
6. Beliau, yang entitasnya tidak bisa dideskripsikan pada karya ini.

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT, atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan tesis yang berjudul “ **DETEKSI MALWARE PADA PERANGKAT ANDROID MENGGUNAKAN *GRADIENT BOOSTING DECISION TREE* DAN *SMOTE***”. Tesis ini disusun sebagai salah satu syarat untuk menyelesaikan studi pada Program Magister Teknik Informatika di Universitas Bina Darma.

Dalam proses penyusunan tesis, penulis menyadari bahwa keberhasilan penyelesaian tidak lepas dari bantuan, dukungan, dan doa dari berbagai pihak. Dengan segala kerendahan hati penulis mengucapkan terima kasih dan penghargaan yang setulusnya kepada:

1. Bapak Dr. Usman Ependi, S.Kom., M.Kom selaku pembimbing yang telah memberikan bimbingan, arahan, dan motivasi dalam menyusun dan menyempurnakan tesis ini.
2. Seluruh dosen dan staf akademik Program Magister Teknik Informatika atas ilmu, arahan, dan bantuannya selama masa perkuliahan.
3. Orang tua dan keluarga yang selalu memberikan doa, semangat, dan cinta yang tak ternilai dalam setiap langkah penulis.
4. Rekan-rekan seperjuangan yang selalu memberikan semangat dan kerja sama yang baik.

Penulis menyadari bahwa Tesis ini masih jauh dari sempurna. Maka penulis mengharapkan kritik dan saran yang membangun untuk penyempurnaan di masa yang akan datang. Semoga Tesis ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan teknologi, khususnya di bidang analisis malware machine learning.

Palembang, Juli 2025

Penulis

## DAFTAR ISI

HALAMAN DEPAN .....	i
HALAMAN PENGESAHAN PEMBIMBING TESIS .....	ii
HALAMAN PENGESAHAN PENGUJI TESIS.....	iii
SURAT PERNYATAAN .....	iv
ABSTRAK .....	v
<i>ABSTRACT</i> .....	vi
MOTTO DAN HALAMAN PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI .....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR .....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah .....	2
1.3 Perumusan Masalah .....	2
1.4 Tujuan Penelitian.....	2
1.5 Kebaruan .....	2
1.6 Manfaat Penelitian .....	3
1.7 Pembatasan Masalah .....	3
1.8 Sistematika Penulisan .....	3
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	5
2.1 Tinjauan Pustaka .....	5
2.2 Landasan Teori .....	9
2.2.1 Android.....	9

2.2.2 <i>Malware</i> .....	10
2.2.3 Machine Learning .....	15
2.2.4 <i>Gradient Boosting Decision Tree (GBDT)</i> .....	16
2.2.5 Ketidakseimbangan Data .....	18
BAB III METODOLOGI PENELITIAN.....	20
3.1 Waktu dan Tempat Penelitian.....	20
3.2 Prosedur Penelitian.....	20
3.2.1 Dataset.....	21
3.2.2 Pra-Pemrosesan Data .....	22
3.2.3 Pembangunan Model.....	26
3.2.4 Pengujian dan Evaluasi Model.....	28
BAB IV HASIL DAN PEMBAHASAN.....	31
4.1 Distribusi Dataset .....	31
4.2 Pra-Pemrosesan Data .....	34
4.3 Desain Eksperimen.....	40
4.4 Hasil Evaluasi Model pada Dataset Drebin .....	41
4.5 Hasil Evaluasi Model pada Dataset Tuandromd .....	44
4.6 Pembahasan.....	46
4.7 Perbandingan dengan Penelitian Terdahulu .....	48
BAB V PENUTUP.....	50
5.1 Kesimpulan .....	50
5.2 Saran.....	51
DAFTAR PUSTAKA .....	52
DAFTAR RIWAYAT HIDUP .....	59
LAMPIRAN-LAMPIRAN.....	60

## DAFTAR TABEL

Tabel 3. 1 Dataset <i>Malware</i> Android <i>DroidFusion</i> .....	21
Tabel 3. 2 Dataset <i>Malware</i> Android TUANDROMD.....	22
Tabel 3. 3 <i>Feature Dataset Malware Drebin</i> .....	23
Tabel 3. 4 <i>Feature Dataset Malware Tuandromd</i> .....	23
Tabel 3. 5 <i>Struktur Confusion Matrix</i> .....	28
Tabel 4.2. 1 Ekstraksi Fitur Dataset Drebin .....	34
Tabel 4.2. 2 Ekstraksi Fitur Dataset Tuandromd.....	35
Tabel 4.2. 3 Missing Value Dataset Drebin.....	35
Tabel 4.2. 4 Data Duplikat Dataset Drebin .....	36
Tabel 4.2. 5 Missing Value Dataset Tuandromd .....	36
Tabel 4.2. 6 Data Duplikat Dataset Tuandromd.....	37
Tabel 4.4.1 Evaluasi Model pada Dataset Drebin.....	41
Tabel 4.5.1 Evaluasi Model pada Dataset Tuandromd.....	44

## DAFTAR GAMBAR

Gambar 2. 1 Arsitektur Android.....	10
Gambar 2. 2 Ciri utama malware .....	11
Gambar 2. 3 Klasifikasi Jenis Malware .....	13
Gambar 2. 4 Flowchart <i>Gadiant Boosting Decision Tree</i> .....	16
Gambar 2. 5 Arsitektur <i>Gadiant Boosting Decision Tree</i> (GBDT).....	17
Gambar 2. 6 Ilustrasi SMOTE .....	19
Gambar 3. 1 Diagram Alur Penelitian.....	20
Gambar 3. 2 Flowchart SMOTE .....	25
Gambar 3. 3 Model <i>Gradient Boosting Decision Tree</i> .....	27
Gambar 4.1. 1 Distribusi Dataset Drebin .....	31
Gambar 4.1. 2 Diagram Kategori Fitur Drebin .....	32
Gambar 4.1. 3 Distribusi Dataset Tuandromd.....	33
Gambar 4.1. 4 Diagram Kategori Fitur Tuandromd.....	33
Gambar 4.2. 1 Sebelum dan Sesudah SMOTE Proporsi 70:30 Dataset Drebin ...	38
Gambar 4.2. 2 Sebelum dan Sesudah SMOTE Proporsi 80:20 Dataset Drebin ...	38
Gambar 4.2. 3 Sebelum dan Sesudah SMOTE Proporsi 90:10 Dataset Drebin ...	39
Gambar 4.2. 4 Sebelum dan Sesudah SMOTE Proporsi 70:30 Dataset Tuandromd .....	39
Gambar 4.2. 5 Sebelum dan Sesudah SMOTE Proporsi 80:20 Dataset Tuandromd .....	40
Gambar 4.2. 6 Sebelum dan Sesudah SMOTE Proporsi 90:10 Dataset Tuandromd .....	40
Gambar 4.4.1 Evaluasi Model pada Dataset Drebin .....	42
Gambar 4.4. 2 Hasil Pengujian Proporsi 70:30 Dataset Drebin.....	42
Gambar 4.4. 3 Hasil Pengujian Proporsi 80:20 Dataset Drebin.....	43
Gambar 4.4. 4 Hasil Pengujian Proporsi 90:10 Dataset Drebin.....	43
Gambar 4.5.1 Evaluasi Model pada Dataset Tuandromd.....	44
Gambar 4.5. 2 Hasil Pengujian Proporsi 70:30 Dataset Tuandromd .....	45
Gambar 4.5. 3 Hasil Pengujian Proporsi 80:20 Dataset Tuandromd .....	45
Gambar 4.5. 4 Hasil Pengujian Proporsi 90:10 Dataset Tuandromd .....	46