BAB I
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi *mobile* yang pesat telah menjadikan Android sebagai sistem operasi yang dominan di pasar global (Fadhilla et al., 2021). Dominasi ini telah menjadikan Android sebagai target utama penyebaran *malware* (Winarnie, 2024), karena sifat kecanggihannya *malware* semakin hari semakin banyak jenisnya (Hadiprakoso et al., 2022). *Malware* hadir dalam berbagai bentuk seperti *trojan, spyware, ransomware, adware* dan lain-lain (Belous & Saladukha, 2020), yang mengancam keamanan data pribadi, kredensial perbankan, dan stabilitas perangkat pengguna (Chirzah & Ramadhan, 2023).

Metode deteksi *malware* yang ada saat ini masih memiliki berbagai keterbatasan (Alhidamkara et al., 2024). Metode *signature-based detection* tidak efektif dalam mendeteksi varian *malware* baru (Raflie Akbar & Sutabri, 2024) dan *malware* yang menggunakan teknik obfuskasi (Puji Rahayu & Nanang Trianto, 2021), sementara metode *behavior-based* membutuhkan *resources* komputasi yang besar (Aslan et al., 2021). Beberapa penelitian terdahulu telah mencoba mengatasi masalah ini menggunakan pendekatan machine learning, seperti penelitian yang menggunakan algoritma *Support Vector Machine* (SVM) didapati akurasi 97% (Hadiprakoso et al., 2022), pada pengujian dengan algoritma *Random Forest* (RF) didapati akurasi 95% (Wanli Sitorus et al., 2021), dan pada penggunaan algoritma *Convolutional Neural Network* (CNN) didapati akurasi 92% (R. Hidayat et al., 2024).

Dalam penelitian deteksi kendala yang sering dihadapi adalah ketidak seimbangan dataset (Yogi Aptana et al., 2025), minimnya penelitian yang mengkombinasikan teknik boosting dengan penanganan *imbalanced data*, dan beratnya beban komputasi. Untuk mengatasi keterbatasan tersebut, penelitian ini mengusulkan penggunaan kombinasi *Gradient Boosting Decision Tree* (GBDT) dan SMOTE (*Synthetic Minority Over-sampling Technique*). GBDT dipilih karena kemampuannya dalam meningkatkan akurasi prediksi melalui pembelajaran

bertahap dari model-model yang lemah (Mahendra & Putra, 2024), sementara *SMOTE* digunakan untuk mengatasi masalah ketidakseimbangan dataset dengan membangkitkan sampel sintetis dari kelas minoritas (W. Hidayat et al., 2021). Kombinasi kedua metode ini diharapkan dapat memberikan solusi yang lebih efektif dalam mendeteksi *malware* Android, dengan mempertimbangkan aspek akurasi, efisiensi komputasi, dan kemampuan generalisasi terhadap varian *malware* baru.

1.2 Identifikasi Masalah

Banyaknya aktivitas berbahaya atau perangkat lunak perusak (*malware*) yang tersebar mengancam keamanan perangkat. Maka diperlukan metode deteksi *malware* yang efektif serta penanganan terhadap ketidakseimbangan dataset yang digunakan dalam proses deteksi tersebut.

1.3 Perumusan Masalah

Berdasarkan latar belakang diatas, maka peneliti merumuskan permasalahan dalam penelitian ini yaitu "Bagaimana mengimplementasikan model deteksi *malware* yang efektif pada perangkat Android dengan menggunakan algoritma *Gradient Boosting Decision Tree* (GBDT) untuk meningkatkan akurasi deteksi, serta metode *Synthetic Minority Oversampling Technique* (SMOTE) untuk mengatasi ketidakseimbangan data?".

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk menerapkan model deteksi *malware* pada perangkat Android yang efektif dengan memanfaatkan algoritma *Gradient Boosting Decision Tree* (GBDT) untuk meningkatkan akurasi deteksi, serta metode *Synthetic Minority Oversampling Technique* (SMOTE) untuk mengatasi permasalahan ketidakseimbangan data pada dataset *malware*.

1.5 Kebaruan

Kebaruan yang ada pada penelitian ini adalah kombinasi antara Teknik *Gradient Boosting Decision Tree* (GBDT) dan SMOTE dalam deteksi *malware*.

GBDT dikenal sebagai salah satu algoritma *machine learning* yang efektif dalam meningkatkan akurasi prediksi, tetapi penggunaannya untuk mendeteksi *malware* Android masih jarang dieksplorasi. SMOTE merupakan teknik yang dirancang untuk menangani masalah ketidakseimbangan dataset.

1.6 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi dalam meningkatkan keamanan perangkat Android melalui pengembangan model deteksi *malware* yang lebih akurat dan andal, serta menjadi referensi bagi penelitian lanjutan di bidang keamanan siber, khususnya dalam penerapan *machine learning* pada deteksi ancaman keamanan.

1.7 Pembatasan Masalah

Agar penelitian lebih terarah dan tidak menyimpang dari permasalahan yang ada, maka perlu adanya batasan masalah. Batasan masalah dalam penelitian ini yaitu:

a. Dataset yang digunakan dalam penelitian ini adalah "*Android Malware Dataset for Machine Learning*" pada penelitian "*DroidFusion: A Novel Multilevel Classifier Fusion Approach for Android Malware Detection*" (Yerima & Sezer, 2019) dan dataset "TUANDROMD" pada penelitian "*Malware* Dataset Generation and Evaluation" (Borah et al., 2020).

1.8 Sistematika Penulisan

Sistematika Penulisan proposal tesis ini dimaksudkan agar dapat memberikan garis besar secara jelas sehingga terlihat hubungan antara bab yang satu dengan bab yang lainnya. Susunan dan struktur proposal tesis dijabarkan dibawah ini sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini membahas tentang latar belakang, identifikasi masalah, batasan masalah, rumusan masalah, tujuan dan manfaat penelitian, ruang lingkup penelitian, serta susunan dan struktur proposal tesis.

BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI

Pada bab ini membahas tentang tinjauan pustaka, landasan teori yang relevan dari penelitian yang akan dilakukan.

BAB III METODOLOGI PENELITIAN

Pada bab ini pembahasannya yang terdiri dari desain dan jadwal penelitian, data penelitian meliputi jenis data, populasi dan sampel penelitian, kemudian konsep dan metode penelitian yang digunakan, metode pengumpulan data serta teknik analisis data.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil analisis menggunakan model *Gradient Boosting Decision Tree* (GBDT), termasuk interpretasi hasil Akurasi Prediksi.

BAB V SIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari penelitian yang dilakukan, serta saran-saran yang dapat dikemukakan untuk penelitian lebih lanjut atau untuk penerapan model deteksi *Malware Android*.

DAFTAR PUSTAKA

Bagian ini mencantumkan semua sumber referensi yang digunakan dalam penulisan penelitian, termasuk buku, artikel jurnal, dan sumber lainnya yang relevan dengan topik penelitian.

LAMPIRAN

Bagian ini menyertakan informasi tambahan yang mendukung penelitian, seperti data yang digunakan, kode program, atau tabel-tabel pendukung lainnya.