

# 1. Pendahuluan

## Latar Belakang

Android merupakan sistem operasi yang sifatnya *open source*. Sistem operasi Android dapat ditemukan di berbagai perangkat *smartphone* dan harganya relatif murah sehingga banyak peminatnya. Karena sifatnya yang *open source*, pengguna khususnya *developer* dapat membuat dan mengembangkan aplikasi. Dalam pengunduhan aplikasi, pengguna dapat mengunduh dari berbagai sumber baik dari Google Play Store maupun aplikasi *third-party*

Google Play Store merupakan wadah pengunduhan dan menggugah aplikasi tanpa peninjauan keamanan yang detail sehingga *malware developer* dapat menggugah aplikasinya. Google telah mengeluarkan aplikasi deteksi *malware* yang bernama Google Bouncer yang dapat melakukan *scan* terhadap aplikasi yang digugah pada Google Play Store. Namun, pengembang menemukan cara agar tidak dapat terdeteksi oleh Google Bouncer. Terlebih, Android mengizinkan pengguna untuk mengunduh aplikasi dengan menggunakan *third-party* [1].

Teknologi *artificial intelligent* (AI) dan *machine learning* diperlukan dalam penanganan masalah keamanan siber, khususnya pada perangkat Android yang mampu mengidentifikasi dan merespon secara *real time*. Selain itu, diperlukan akurasi yang lebih tinggi agar dapat terdeteksi dengan baik. Pada penelitian [2], digunakan fitur-fitur yang berasal *file manifest* yaitu *permission* dan *intent* dan digunakan tiga algoritma yaitu Naive Bayes, *Support Machine Vector* (SVM), dan *Random Forest*. Pada tugas akhir ini, penelitian [2] dilanjutkan dengan menggunakan *Random Forest* karena menghasilkan nilai akurasi yang lebih baik jika dibandingkan dengan Naive Bayes dan SVM.

## Topik dan Batasannya

Topik penelitian ini adalah analisis algoritma *Random Forest* dalam melakukan pendeteksian *malware* pada Android, dengan menguji nilai *permission* dan *intent* terbaik. Batasan masalah pada topik penelitian ini adalah:

1. *Dataset* utama berasal dari Koodous [3] merupakan *dataset* yang berisikan aplikasi *malware* sebanyak enam aplikasi dan aplikasi normal sebanyak lima, kemudian sebanyak empat aplikasi normal didapat dari [4], [5], [6], dan [7] serta Drebin dan Malgenome didapat dari [8].
2. Fokus utama pada penelitian ini adalah dilakukan pencarian fitur terbaik dari fitur *permission* dan *intent*. Pada penelitian ini, fitur API tidak digunakan karena tidak semua aplikasi yang ter-*unzip* dapat mengeluarkan *file API*.

Tujuan dari penelitian ini adalah sebagai berikut:

1. Melakukan implementasi algoritma *Random Forest* dengan deteksi statis memanfaatkan *information gain* untuk mencari *permission* dan *intent* terbaik agar mendapatkan akurasi terbaik ketika melakukan pendeteksian.
2. Menganalisis performansi algoritma *Random Forest* dengan memanfaatkan fitur *permission* dan *intent* yang telah diurutkan berdasarkan *information gain*.

## Organisasi Tulisan

Penelitian ini disusun sebagai berikut:

1. Pendahuluan  
Bagian ini memiliki sub-bagian latar belakang penelitian, topik dan batasannya, tujuan, serta organisasi penulisan
2. Studi terkait  
Bagian ini berisikan teori yang terkait dengan topik di mana sumber pustakanya sebagai acuan
3. Sistem yang dibangun  
Bagian ini berisikan penjelasan rancangan sistem yang dibangun
4. Evaluasi  
Bagian ini berisikan penjelasan hasil dan analisis pengujian
5. Kesimpulan  
Bagian ini berisikan kesimpulan dan saran untuk pembaca