1. Pendahuluan

1.1. Latar Belakang

Pada era industri 4.0 teknologi sangatlah berkembang dengan pesat, khususnya teknologi informasi. Teknologi ini digunakan untuk memudahkan pekerjaan manusia dan keandalan suatu aplikasi sangatlah dibutuhkan oleh manusia. Salah satu teknologi informasi yang sering digunakan oleh manusia adalah android. Pada tahun 2017, *Gertner* melaporkan penjualan Smartphone di seluruh Dunia naik 9 persen pada kuartal pertama dengan 380 juta perangkat seluler yang terjual dan perangkat seluler yang menjadi primadona di pasar adalah android dengan 84,1% [1]. Laporan Kementerian komunikasi dan informatika memperkirakan jumlah pengguna aktif smarfphone di indonesia pada 2018 melebihin 100 juta orang [2], beserta *statcounter* melaporkan tahun 2018 jumlah pengguna android di indonesia 91% [3], sehingga dapat perkirakan jumlah pengguna android di indonesia kurang lebih 91 juta orang. Meningkatnya pengguna android yang cukup besar membuat sistem operasi android menjadi target serangan malware, sebuah sistem operasi yang telah terserang malware dapat mengalami kerusakan dan bahkan dengan niat yang lebih jahat malware dapat digunakan untuk pencurian data-data penting yang ada di sistem.

Malicious Software atau sering dikenal sebagai malware merupakan sebuah software yang diprogram agar dapat menyusup ke sebuah sistem operasi yang dapat merusak cara kerja sistem dan bahkan digunakan untuk mencuri data-data penting pada perangkat korban. Malware memiliki beberapa jenis dan mempunyai cara kerja yang beragam, salah satu contoh serangan malware melalui aplikasi berbahaya yang melakukan akses *permission* secara ilegal tanpa izin dari pengguna dan sistem operasi. Oleh karena rentannya serangan malware dan merugikan para pengguna android sehingga diperlukan analisis lebih lanjut terhadap malware. Pada tahun 2019, perusahaan keamanan siber Kaspersky melaporkan bahwa malware yang terdeteksi sebanyak 556.486 dan pada tahun 2020 perusahaan Kaspersky juga melaporkan terjadi penurunan serangan malware di indonesia sebanyak 31,89% sehingga menjadi 378,973. Oleh hal ini menjadikan Indonesia negara dengan jumlah anacama malware terbesar yang terdeteksi se-Asia Tenggara dan menduduki posisi ke-4 secara global [4]

Analisis statis dilakukan dengan membongkar source code dari malware tersebut lalu mempelajari dan memahami perilaku jahat melalui source code, data dan file biner, sehingga proses analisis statis tidak memerlukan eksekusi terhadap malware. Analisis statis digunakan untuk mengekstrak karakteristik dan mengidentifikasi setiap aplikasi, menggunakan ApkTool 2.0.3 [5].

Berdasarkan data serangan malware dan pengguna android yang sangat banyak, ini menjadi sebuah masalah maka penelitian ini mengimplementasikan sebuah sistem *machine learning* untuk mengetahui suatu metode klasifikasi yang menghasilkan akurasi yang tinggi dalam melakukan deteksi dini terhadap serangan malware. Beberapa penelitian sebelumnya telah menggunakan beberapa metode *machine learning* seperti *Naïve bayes*, *Decision Tree*, *KNN* dan metode lainnya , seperti yang dilakukan oleh Lopez dkk [5] telah melakukan penelitian menggunakan beberapa metode *machine learning* dengan menghasilkan akurasi yang tidak lebih dari 95%. Dengan dikembangkan metode-metode *machine learning* untuk melakukan klasifikasi, sehingga pada penelitian ini mengimplementasikan sebuah sistem *machine learning* untuk mengetahui suatu metode *machine learning* yang menghasilkan performansi matrik yang lebih unggul dari penelitian [5] dalam melakukan deteksi dini terhadap serangan malware.

Penelitian ini menggunakan metode klasifikasi *Support Vector Machine* (SVM) dan *Random Forest* untuk mengetahui metode mana yang terbaik untuk melakukan klasifikasi dengan baik pada dataset malware serta membandingkan hasil perfomasi matrik dengan penelitian sebelumnya yang dilakukan oleh Lopez A. et al [5]. Hasil penelitian ini diharapkan dapat membantu *developer google play store* dan pengguna dalam mengunduh aplikasi di *google play store*. Aplikasi yang ingin dipublikasikan oleh perorangan maupun perusahaan dapat diketahui apakah aplikasi tersebut berbahaya atau tidak. Sehingga aplikasi yang ada di *google play store* lebih terjamin dari serangan malware dan aman untuk diunduh oleh pengguna.

Topik dan Batasannya

Pada tugas akhir ini membahas masalah bagaimana cara membatasi android terkena malware dengan mengklasifikasi akses izin (*permission*) yang dilakukan secara ilegal. Deteksi pada tugas akhir ini menggunakan metode linear *Support Vector Machine* dan *Random Forest*. Dataset yang digunakan untuk mendeteksi malware adalah "*Dataset malware/begingn permission Android*".

Tujuan

Tujuan dari penelitian ini berdasarkan masalah yang ada adalah untuk mengimplementasikan sebuah algoritma *Support Vector Machine* (SVM) dan *Random Forest* untuk mengklasifikasikan dataset *permission* android (Malware), menganalisis data dengan membuat fitur umum dan fitur ciri dari dataset yang dipilih. Serta membandingkan hasil performansi matrik antara kedua metode tersebut dan membandingkan hasil performansi matrik dengan penelitian sebelumnya