BAB I
PENDAHULUAN

1.1 Latar Belakang

Seiring dengan berkembangnya teknologi kecerdasan buatan, *machine learning* yang merupakan cabang dari kecerdasan buatan juga ikut berkembang dan semakin diterapkan dalam berbagai aspek kehidupan manusia, dengan *machine learning* kita dapat melatih mesin untuk belajar tanpa diprogram secara eksplisit, pembelajaran ini dilakukan dengan menerapkan algoritma *machine learning* menggunakan data dan pengalaman masa lalu sembari meningkatkan kinerja seiring data dan informasi yang diterima (V. Rameshbabu et al., 2023).

Salah satu penggunaan *machine learning* adalah pemanfaatannya dalam melakukan klasifikasi, klasifikasi adalah proses penggolongan atau pengelompokan fungsi yang akan menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk menghasilkan perkiraan kelas dari suatu objek dengan label yang belum diketahui atau pembagian sesuatu berdasarkan kelas-kelas nya (Dinata et al., 2020). Salah satu contoh penggunaan algoritma klasifikasi diatas adalah dalam klasifikasi *ransomware*.

Ransomware adalah jenis *malware* yang dirancang untuk menghalangi akses ke sistem komputer, *file* atau data hingga tebusan dibayar oleh korban. *Ransomware* mengenkripsi *file* pada komputer atau jaringan, membuatnya tidak dapat diakses dan menuntut pembayaran seringkali dalam mata uang kripto sebagai tebusan untuk mengembalikan akses ke data tersebut (Noorbehbahani et al., 2019).

Serangan *ransomware* umumnya terjadi ketika korban tanpa sadar mengklik tautan atau mengunduh *file* berbahaya yang kemudian memasang *ransomware* ke dalam sistem komputer mereka, kemudian akan ditampilkan pesan pada sistem komputer korban yang menuntut tebusan agar korban mendapatkan kembali akses terhadap komputer dan *file* mereka.

Melihat besarnya ancaman dan potensi kerugian yang dapat disebabkan oleh *ransomware*, penelitian untuk mengembangkan metode deteksi, identifikasi dan

klasifikasi *ransomware* terus dilakukan, disisi lain para kriminal juga menggunakan teknologi canggih untuk terus mengembangkan jenis *ransomware* baru yang lebih kompleks dan sulit untuk ditangani dibandingkan dengan *ransomware* tradisional (Adamu et al., 2019).

Deteksi dan klasifikasi *ransomware* dapat dilakukan dengan analisis statis dan dinamis. Analisis statis berfokus pada pemeriksaan karakteristik dan fitur sampel *ransomware* tanpa menjalankan *file* sementara analisis dinamis dilakukan dengan memantau perilaku *ransomware* selama *file* berjalan (Bahaa et al., 2022).

Antivirus modern umumnya menggunakan analisis statis berbasis *signature* dalam deteksi *ransomware,* namun cara ini dinilai tidak efektif dalam menghadapi pesatnya perkembangan *ransomware* karena jenis *ransomware* baru terus bermunculan seiring waktu, selain itu para pelaku kriminal juga dapat mengeksploitasi kerentanan dalam sistem operasi untuk menghindari mekanisme perlindungan ini, sedangkan jika analisis dinamis dilakukan akan memakan sumber daya, biaya dan waktu proses yang ekstra (Manabu et al., 2022).

Untuk mengatasi batasan-batasan tersebut dikembangkanlah metode alternatif dalam deteksi dan klasifikasi *ransomware* memanfaatkan *machine learning*. Dalam penelitiannya (Masum et al. 2022) secara sukses menerapkan algoritma *machine learning* untuk klasifikasi *ransomware* berdasarkan fitur yang diekstraksi dari *file ransomware*. Melanjutkan penelitian sebelumnya (Masum et al. 2022), dalam penelitian ini penulis memilih algoritma klasifikasi *random forest* sebagai metode *machine learning* yang akan digunakan tidak hanya untuk klasifikasi namun juga untuk mendeteksi *file* yang dicurigai sebagai *ransomware*.

Penulis memilih *random forest* sebagai algoritma dalam penelitian ini karena kelebihannya yang telah terbukti sesuai untuk diterapkan dalam klasifikasi *ransomware,* kelebihan tersebut mencakup kemampuan *random forest* menangani *dataset* dengan jumlah besar, waktu pelatihan yang relatif singkat dan hasil prediksi yang lebih akurat dibanding algoritma lain, selain itu *random forest* juga dapat mengurangi risiko *overfitting* yang membuat penulis memilih algoritma ini sebagai algoritma yang dianggap paling tepat untuk digunakan dalam penelitian ini.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dirumuskan permasalahan terkait dengan penelitian ini, antara lain:

1. Bagaimana cara menerapkan algoritma *random forest* dalam deteksi dan klasifikasi *file ransomware*.
2. Melihat akurasi algoritma *random forest* dalam deteksi dan klasifikasi *file ransomware*.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang diuraikan diatas, maka tujuan dari perancangan ini, antara lain:

1. Mengetahui cara implementasi algoritma *random forest* dalam deteksi dan klasifikasi *file ransomware*.
2. Mengetahui akurasi algoritma *random forest* dalam deteksi dan klasifikasi *file ransomware*.

1.4 Manfaat Penelitian

Manfaat yang didapatkan dari penelitian ini antara lain:

A. Bagi peneliti

1. Menambah wawasan pengetahuan mengenai penerapan algoritma *random forest* dalam deteksi dan klasifikasi *ransomware*.
2. Memenuhi sebagian syarat untuk menyelesaikan program studi Strata-1 Teknik Informatika.

B. Bagi masyarakat

1. Memberikan informasi dan meningkatkan kesadaran tentang bahaya *ransomware* yang mana diharapkan dapat membantu mengurangi korban dari serangan *ransomware*.
2. Sebagai rujukan untuk penelitian lebih lanjut dan dapat menjadi dasar untuk pengembangan perangkat lunak anti *ransomware*.

1.5 Batasan Penelitian

Penelitian ini ditentukan pada ruang lingkup tertentu antara lain :

1. Penelitian ini terbatas pada analisis statis berdasarkan fitur yang diekstraksi dari *file ransomware* yaitu metadata (seperti nama *file*, ukuran *file* dan lain-lain) yang kemudian diolah menggunakan algoritma klasifikasi *random forest.*
2. Analisis dinamis maupun hibrida dengan menjalankan *file ransomware* tidak dilakukan pada penelitian ini.