BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Internet merupakan salah satu bagian terutama dari kehidupan modern. Dengan adanya internet, mendapatkan pengetahuan dan melakukan komunikasi pada sesama menjadi lebih mudah. Akan tetapi, tidak semua mengetahui bahwa pengguna internet rentan dari berbagai macam ancaman siber. Salah satu ancaman siber yang terjadi adalah mendapatkan *malware* dalam berbagai bentuk seperti *adware, spyware dan ransomware*. *Ransomware* sendiri telah menjadi sebuah masalah yang cukup serius. Terdapat jutaan kasus yang dilaporkan secara global setiap tahunnya. Pada tahun 2021 terdapat peningkatan secara pesat pada serangan *ransomware* yaitu lebih dari 150% yang memberi banyak dampak terhadap berbagai sektor seperti kesehatan, keuangan dan pemerintahan [5]. Oleh sebab itu, dibutuhkan sebuah *system* yang bisa mendeteksi *malware-malware* dengan akurat dan cepat sebelum *malware* tersebut dapat memasuki jaringan pengguna internet seperti *Intrusion Detection System* (IDS).

IDS merupakan sebuah teknologi keamanan jaringan yang digunakan untuk mendeteksi eksploitasi kerentanan terhadap aplikasi atau komputer target [6]. Berdasarkan penelitian-penelitian terdahulu, IDS dapat digunakan untuk mendeteksi serangan siber seperti *ransomware* akan tetapi IDS sendiri memiliki tantangan-tantangan yang harus dihadapi seperti *false alarm rate, low detection rate, unbalanced dataset* dan *response time* sehingga dibutuhkan sebuah solusi yang dapat mengatasi tantangan-tantangan tersebut [7]. Terdapat penelitian IDS terdahulu yang dilakukan menggunakan *Principal Component Analysis* (PCA) dengan *Random Forest* (RF) dapat menghasilkan tingkat akurasi sebesar 96.78 %, tingkat *error* 0.21% dan waktu yang digunakan sebanyak 3.42 menit [8]. Akan tetapi, penelitian IDS tersebut tidak digunakan untuk mendeteksi *ransomware* secara khusus melainkan hanya mendeteksi kerentanan secara umum.

Maka dari itu, peneliti melakukan penelitian mengenai IDS yang dapat digunakan untuk mendeteksi *ransomware*. Setelah itu, peneliti melakukan peningkatan akurasi pada IDS tersebut memanfaatkan algoritma RF. Algoritma RF dipilih karena dapat menyelesaikan data yang rumit dengan kuantitas yang besar dan juga memberikan penjelasan yang dapat dipahami terhadap hasil prediksi [9].

Algoritma ini dimungkinkan cocok untuk mendeteksi *ransomware* dikarenakan algoritma ini memiliki beberapa keuntungan seperti tahan terhadap *overfitting* yang dapat membuat model menjadi lebih efektif dalam mendeteksi *ransomware* dalam berbagai situasi dengan jenis data yang tidak pernah dilihat sebelumnya dan variasi berkurang seiring dengan bertambahnya jumlah pohon tanpa menyebabkan bias yang dapat membuat model menjadi lebih stabil dan akurat karena tidak terlalu terpengaruh oleh data tertentu atau noise [10]. Terdapat juga penelitian terdahulu yang meneliti dengan pendekatan RF dalam mendeteksi *ransomware* dan menunjukkan hasil akurasi sebesar 97.74% pada *byte level ransomware* [10]. Akan tetapi, Algoritma RF memiliki kekurangan yaitu daya komputasi yang dibutuhkan besar sehingga dibutuhkan sebuah optimasi pada algoritma tersebut. Optimasi perlu dilakukan agar dapat mengurangi daya komputasi yang dibutuhkan oleh algoritma dan memastikan model mendapatkan hasil dengan kinerja yang maksimal. Terdapat penelitian terdahulu yang melakukan optimasi algoritma RF menggunakan *Particle Swarm Optimization* (PSO) pada klasifikasi diabetes dan menghasilkan tingkat akurasi, *recall* dan *precision* yang lebih tinggi dibandingkan dengan menggunakan algoritma RF tanpa optimasi sehingga terdapat peran penting pada penggunaan PSO dalam peningkatan kinerja algoritma RF [11].

Oleh karena itu, peneliti melakukan optimasi pada algoritma RF menggunakan PSO. PSO dapat membantu algoritma RF dengan cara memilih fitur paling relevan yang akan digunakan oleh algoritma RF sehingga terjadi peningkatan pada hasil kinerja algoritma RF.

1.2 Rumusan Masalah

1. Bagaimana cara mengoptimalkan akurasi dan waktu komputasi IDS dalam mendeteksi *ransomware* menggunakan algoritma RF dengan PSO?
2. Apakah hasil dari optimasi algoritma RF dengan PSO lebih baik dibandingkan algoritma RF dengan PCA dalam mendeteksi *ransomware*?

1.3 Batasan Permasalahan

1. Algoritma yang diteliti adalah RF dengan PSO dan RF dengan PCA, tanpa membandingkan algoritma lainnya dan kedua algoritma menggunakan dataset yang sama

2. Penelitian dilakukan pada *dataset CIC-MalMem* [12] dan mungkin tidak mencakup semua jenis *ransomware*.

3. Penelitian ini tidak mempertimbangkan aspek non-teknis seperti kebijakan dan regulasi terkait *ransomware*.

1.4 Tujuan Penelitian

1. Mengoptimalkan akurasi dan waktu komputasi IDS dalam mendeteksi *ransomware* menggunakan algoritma RF dengan PSO.

2. Mengetahui jika hasil dari optimasi algoritma RF dengan PSO lebih baik dibandingkan algoritma RF dengan PCA dalam mendeteksi *ransomware*.

1.5 Manfaat Penelitian

1. Menghasilkan algoritma RF dengan PSO yang dapat mengoptimalkan akurasi dan waktu komputasi IDS dalam mendeteksi *ransomware*.

2. Menambah wawasan mengenai optimasi RF dengan PSO dan RF dengan PCA dalam mendeteksi *ransomware*.

3. Mendapatkan hasil perbandingan dari optimasi algoritma RF dengan PSO dan algoritma RF dengan PCA dalam mendeteksi *ransomware*.

1.6 Sistematika Penulisan

Sistematika penulisan yang digunakan peneliti adalah sebagai berikut:

- BAB 1 PENDAHULUAN
  Bab ini melakukan pembahasan mengenai bagian-bagian dari pendahuluan seperti latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian serta sistematika penulisan

- BAB 2 LANDASAN TEORI
  Bab ini membahas penjelasan mengenai teori-teori yang akan dipakai oleh laporan skripsi ini. Teori-teori berikut merupakan *ransomware*, *intrusion detection system*, *ensemble learning*, *decision tree learning*, *random forest*, *paritcle swarm optimization*, dan *principal component analysis*

- BAB 3 METODOLOGI PENELITIAN

  Bab ini menjelaskan alur dari penelitian yang akan dilaksanakan, yaitu dimulai dengan identifikasi masalah, telaah literatur, pengumpulan data, pembuatan model *machine learning*, serta evaluasi

- BAB 4 HASIL DAN DISKUSI

  Bab ini membahas mengenai hasil penelitian yang didapatkan. Bab ini mencakup langkah-langkah dan hasil optimasi yang dilakukan pada algoritma RF menggunakan PSO dan perbandingannya dengan optimasi algoritma RF menggunakan PCA

- Bab 5 KESIMPULAN DAN SARAN

  Bab ini membahas kesimpulan dan saran untuk dilaksanakan pada penelitian selanjutnya.