BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Komputer saat ini menjadi penting sebab digunakan diberbagai bidang, seperti bisnis, pendidikan, pemerintahan, organisasi dan lain-lain[1]. Penggunaan internet ini semakin meluas dari tahun ke tahun, tercatat pada 2016 terdapat lebih dari 100 juta pengguna internet sebagai media komunikasi[2] digunakannya internet sebagai bisa media komunikasi tanpa keamanan jaringan yang cukup akan berdampak negatif sehingga banyak serangan yang akan terjadi. Serangan tersebut bisa berupa fisik atau serangan jaringan. Serangan fisik bisa berupa menyentuh perangkat secara langsung, penyerangan logika dapat menyerang logika dari sebuah perangkat lunak sehingga dapat membuka celah yang akan meliputi serangan jaringan untuk disusupi sehingga dapat mengontrol jaringan.

Maka dari itu, meningkatkan keamanan jaringan komputer menjadi suatu kebutuhan yang mendesak. Dalam penggunaan internet yang semakin luas dan terhubung secara global, risiko serangan dari pihak yang tidak bertanggung jawab semakin meningkat. Ancaman-ancaman di dunia maya selalu menjadi perhatian utama karena akses informasi melalui internet telah menjadi bagian integral dari kehidupan[3]. Serangan-serangan ini dapat menyebabkan kerusakan yang serius, baik pada jaringan maupun sistem, serta pencurian data yang berpotensi merugikan. Oleh karena itu, menjaga keamanan jaringan dengan meningkatkan akan membuat effisiensi dalam melakukan pencegahan terhadap serangan yang akan terjadi.

Peningkatan keamanan dapat dicapai dengan menerapkan klasifikasi data untuk mendeteksi keadaan normal atau serangan dalam jaringan. Salah satu pendekatannya adalah dengan menggunakan teknologi IDS *(Intrusion Detection System)* yang dapat membantu dalam mendeteksi serangan secara efektif, sehingga memungkinkan penyelesaian masalah dan pencegahan dilakukan dengan lebih cepat.

Intrusion Detection System(IDS) adalah merupakan salah satu upaya untuk medeteksi serangan dan laporan yang tidak diinginkan atau pola yang tidak normal dalam sebuah sistem[4]. IDS akan menjadi lebih baik apabila memiliki sebuah klasifikasi data untuk mendeteksi anomali di dalam jaringan. Klasifikasi dari sebuah data jaringan juga dapat mengidentifikasi tingkatan berbahaya atau tidak untuk

memperkirakan risiko yang akan terjadi. Klasifikasi juga bertujuan memprediksi kelas tujuan dengan presisi yang tinggi[5].

Berdasarkan penelitian terkait dengan judul ”Deteksi Serangan Pada *Intrusion Detection System* (IDS) untuk klasifikasi serangan dengan algoritma *Naïve Bayes*, C.45, dan K-NN” menjelaskan pentingnya klasifikasi data penyerangan untuk melindungi jaringan dan sistem informasi. Adanya IDS yang dilengkapi klasifikasi serangan yang akurat dapat mengidentifikasi serangan dengan lebih effisien dan meningkatkan keamanan jaringan[6]. Hasil pengujian menunjukkan bahwa algoritma C.45 memiliki tingkat akurasi yang optimal. Sedangkan algoritma KNN memiliki performa yang rendah karena menganggap atribut yang seharusnya tidak ada serangan sebagai adanya serangan.

Penelitian lain yang terkait dengan judul ”Klasifikasi Anomali *Intrusion Detection System* (IDS) Menggunakan Algoritma *Naïve Bayes Classifier* dan *Correlation-Based Feature Selection*” bertujuan untuk mengembangkan sistem deteksi intrusi yang efektif dengan menggunakan algoritma *Naïve Bayes Classifier* dan metode seleksi fitur berbasis korelasi[7]. Menggabungkan algoritma *Naïve Bayes* yang memiliki kemampuan dalam mengklasifikasikan anomali dalam *Intrusion Detection System* (IDS) dan seleksi fitur berbasis korelasi untuk memilih fitur yang memiliki korelasi tinggi dengan target, sistem deteksi intrusi dengan seleksi berdasarkan korelasi dapat memberikan klasifikasi yang lebih akurat dan efisien, meningkatkan keamanan jaringan komputer, dan melindungi data dari serangan yang berpotensi merugikan. Hasil yang diperoleh dari penelitian tersebut *Naïve Bayes* tanpa seleksi sebesar 71,2%, yang lebih rendah dibandingkan dengan akurasi *Naïve Bayes* dengan seleksi sebesar 74,8%.

Dari beberapa penelitian yang sudah disebutkan, algoritma C4.5 dianggap dapat menjadi pilihan untuk melakukan klasifikasi data penyerangan pada *Intrusion Detection System* (IDS). Penelitian sebelumnya menunjukkan bahwa algoritma C4.5 memiliki tingkat akurasi yang optimal, serta tingkat *recall* dan presisi yang tinggi[6]. Oleh karena itu, penerapan algoritma C4.5 dalam IDS dapat memberikan identifikasi serangan yang lebih efisien dan meningkatkan keamanan jaringan secara keseluruhan.

Penelitian ini memfokuskan pada penggunaan algoritma C4.5 dalam klasifikasi IDS. Menggunakan algoritma C4.5, IDS dapat mempelajari pola serangan yang ada dan membuat keputusan yang cerdas dalam mengklasifikasikan serangan baru yang muncul, sehingga memberikan perlindungan yang lebih baik terhadap jaringan dan sistem informasi yang terhubung.

Berbeda dengan penelitian sebelumnya yang menggunakan aplikasi RapidMiner atau WEKA, penelitian ini melibatkan penulisan kode program dari awal*(Scratch*. Pendekatan ini memungkinkan peneliti untuk memiliki kontrol penuh atas implementasi algoritma dan menyesuaikannya dengan kebutuhan penelitian.

Dalam penelitian ini, UNSW-NB15 dipilih sebagai *dataset* yang digunakan. Pemilihan *dataset* ini didasarkan pada bahwa *dataset* tersebut merupakan *dataset* baru yang dikembangkan pada tahun 2015 dibandingkan Cup 1999 dan NSL-KDD yang merupakan *dataset* IDS lama. *dataset* UNSW NB-15 terdiri dari kombinasi data serangan normal daan tidak normal[8]. *dataset* UNSW NB-15 ini merupakan representasi nyata dari lingkungan ancaman modern adalah untuk mengumpulkan data yang mencerminkan situasi serangan aktual. Kumpulan data ini diberi label kategori serangan, yaitu *"attack cat"* dan "label". Untuk rekaman yang merepresentasikan situasi normal, diberi label 0 (nol), sedangkan untuk rekaman yang merepresentasikan serangan, diberi label 1 (satu). Selanjutnya, serangan diklasifikasikan menjadi sembilan kelompok yang berbeda, berdasarkan karakteristik dan jenis serangannya.[9]

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah tertulis, adapun rumusan masalah adalah:

1. Bagaimana penerapan algoritma *Decision Tree Classifier* C.45 dalam melakukan analisa klasifikasi data *Intrusion Detection System*.
2. Bagaimana performa dari algoritma *Decision Tree Classifier* C.45 dalam melakukan klasifikasi serangan yang terjadi didalam jaringan.

1.3 Batasan Permasalahan

Batasan-batasan masalah pada penelitian ini adalah, sebagai berikut:

1. Pengolahan data yang ada hanya berasal dari *dataset* UNSW NB-15 yang berisikan data penyerangan jaringan.
2. Klasifikasi jenis serangan berdasarkan serangan atau normal hanya berada dalam lingkup yang ada di dalam *dataset* UNSW NB-15.

1.4 Tujuan Penelitian

1. Menerapkan metode *Decision Tree Classifier* C.45 sebagai klasifikasi data *Intrusion Detection System*
2. Mengetahui performa dari *Decision Tree Classifier* dalam melakukan klasifikasi serangan yang terjadi

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini dapat menambahkan referensi terhadap peningkatan klasfikasi data menggunakan IDS dengan implementasi *Decision Tree Classifier* sehingga mempermudah penanganan dari serangan luar jaringan

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran.

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
  Pada bab ini berisi latar belakang masalah, masalah dan rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.
- Bab 2 LANDASAN TEORI
  Bagian ini menjabarkan teori yang digunakan penelitian secara menyeluruh. Teori tersebut mencakup implementasi algoritma *Decision Tree Classifier* C.45 untuk klasifikasi data penyerangan jaringan dan *pre-processing*.
- Bab 3 METODOLOGI PENELITIAN
  Bab ini menjelaskan analisis sistem yang akan dijalankan dan perancangan sistem secara detail untuk implementasi menggunakan algoritma *Decision Tree Classifer C4.5*.
- Bab 4 HASIL DAN DISKUSI
  Bab ini menampilkan hasil dari pengujian yang telah dilakukan serta pembahasan dari hasil pengujian.

- Bab 5 SIMPULAN DAN SARAN

  Bab ini memberikan kesimpulan mengenai hasil dari keseluruhan isi bab sebelumnya serta memberikan saran dari hasil yang diperoleh untuk penelitian selanjutnya dengan harapan bermanfaat dalam pengembangan selanjutnya.