BAB I
PENDAHULUAN

1.1 Latar Belakang

Penggunaan internet mengalami perkembangan yang signifikan hal ini ditandai dengan mudahnya mendapatkan akses ke internet dimana kemajuan teknologi informasi menjadi pendukung dari perkembangan penggunaan internet yang secara langsung juga meningkatkan perkembangan teknologi jaringan data baik secara lokal maupun global. Meningkatnya jumlah perangkat seperti komputer dan ponsel yang terhubung pada internet membuat keamanan akan serangan siber menjadi isu kritis dengan banyaknya perangkat ini membuat peluang lebih banyak celah bagi penyerang siber untuk melakukan serangan terhadap perangkat tersebut yang mana salah satunya adalah serangan botnet.

Di Indonesia sendiri, *MyloBot* salah satu bentuk dari *Botnet* menjadi sumber trafik anomali tertinggi pada tahun 2022 dengan 254.260.339 jumlah kasus sebagaimana tertera dalam dokumen "Lanskap Keamanan Siber Indonesia Tahun 2022" oleh BSSN. Alasan mengapa tingginya angka serangan botnet disebabkan karena deteksi serangan *botnet* bisa menjadi sebuah tantangan, sulitnya mendeteksi serangannya dan rendahnya tingkat akurasi pendeteksian serangan menyebabkan tingginya *false detection* sehingga data normal dianggap menjadi serangan. *Botnet* ini kumpulan dari aplikasi *bot* (robot) yang dikonfigurasi untuk dapat berjalan secara otomatis dalam jaringan maka setiap komputer yang terinfeksi dan tergabung kedalam jaringan botnet akan mengeksekusi

perintah atau instruksi yang diberikan oleh *Botmaster* dengan dilakukan dari jarak jauh. *Botnet* mampu menyediakan platform yang dapat didistribusikan pada kegiatan illegal seperti *spam*, *phishing, click fraud*, pencurian kata sandi dan *Distributed Denial of Service (DdoS)* (Xing et al., 2021).

Metode yang akan digunakan dalam penelitian ini menggunakan *machine learning*. Algoritma *machine learning* mampu mempelajari data trafik jaringan dan mengidentifikasi aktivitas yang mencurigakan, hal ini membuat *machine learning* menjadi metode yang efektif. *Machine learning* dengan berbasis fitur seleksi memiliki akurasi diatas 90% dalam mendeteksi serangan DdoS pada dataset(Maslan et al., 2020), algoritma machine learning juga mampu mendeteksi serangan dan data normal(Darryl & Subali, 2021). Dengan demikian, dalam penelitian ini *Machine learning* akan diterapkan kedalam sebuah dataset yang berisi data botnet dan normal dan mengeksplorasi bagaimana algoritma *machine learning* dapat secara efektif mendeteksi yang mana pola serangan *botnet* dan yang mana data normal. Adapun *dataset* yang akan digunakan pada penelitian ini diambil dari *GitHub* yang diolah dari *CTU-13 Dataset*.

Salah satu masalah utama dalam mendeteksi serangan botnet adalah jumlah data dalam jaringan dan kompleksitasnya, banyak teknik deteksi tidak mampu mengatasi jumlah data dengan tingkat akurasi yang tinggi, selain itu fitur yang tidak relevan dalam dataset juga membuat performa model deteksi menjadi menurun. Mengingat akan hal ini, penelitian ini akan memfokuskan pada deteksi serangan botnet dengan

menggunakan fitur seleksi dari dataset yang ada. Penggunaan fitur seleksi ini bertujuan untuk mengurangi kompleksitas analisis dan meningkatkan akurasi dalam deteksi, dengan tetap menjaga keterwakilan dataset. Penelitian ini akan mengeksplorasi bagaimana algoritma *machine learning*, seperti *Naive Bayes*, *Support Machine Vector*, dan *K–Nearest Neighbors*, dapat diterapkan pada data untuk mendeteksi pola serangan *botnet* secara efektif. Dengan menggunakan algoritma *machine learning* berbasis *feature selection*, masalah diatas bisa diatasi dengan lebih efisien. *Feature selection* bisa membantu men- *filter* fitur yang tidak relevan, sehingga model deteksi serangan bekerja lebih akurat.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah dijelaskan, identifikasi masalah pada penelitian ini adalah sebagai berikut :

1. Tingginya trafik anomali, hal ini membuat sulitnya membedakan antara aktifitas normal dan aktifitas mencurigakan.
2. Tingkat akurasi dalam mendeteksi *botnet* rendah, ini dikarenakan kesulitan dalam mendeteksi serangan oleh banyak metode deteksi yang menghasilkan *false positive* dimana data normal diklasifikasikan menjadi serangan.
3. Banyak fitur dalam *dataset* yang digunakan namun tidak relevan, ini membuat kinerja proses menjadi lambat dan akurasi deteksi yang berkurang.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Penelitian ini hanya menggunakan 3 algoritma *machine learning*, yaitu *Naïve Bayes, Support Machine Vecotr* dan *K–Nearest Neighbors* sebagai tolak ukur untuk membandingkan hasil deteksi tanpa menggunakan algoritma lain.
2. Penelitian ini menggunakan *dataset* tertentu, yang mungkin memiliki karakteristik dan struktur yang berbeda, dan tidak dapat digeneralisasi dengan *dataset* lain.
3. Teknik pemilihan *feature selection*, teknik yang digunakan akan dibatasi pada metode tertentu, tanpa mengeksplorasi semua metode fitur yang ada.

1.4 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, penelitian ini akan berfokus pada pendeteksian antara serangan botnet dan data normal menggunakan dataset. Untuk lebih memperjelas arah penelitian, maka rumusan masalah yang dapat diajukan adalah sebagai berikut:

1. Bagaimana melakukan klasifikasi serangan *botnet* ?
2. Bagaimana melakukan seleksi fitur terbaik dalam mendeteksi serangan *botnet* ?
3. Bagaimana kinerja Algoritma *Machine learning* KNN, SVM dan *Naive Bayes* ?

1.5 Tujuan Penelitian

Tujuan penelitian ini bisa dijabarkan sebagai berikut :

1. Untuk melakukan klasifikasi serangan *botnet* menggunakan algoritma *machine learning*.
2. Untuk mengetahui fitur terbaik dalam mengklasifikasikan serangan *botnet.*
3. Untuk mengevaluasi kinerja Algoritma *Machine learning* KNN, SVM dan *Naive Bayes.*

1.6 Manfaat Penelitian

Manfaat penelitian ini adalah adanya peningkatan akurasi deteksi dengan memilih fitur yang relevan dengan kebutuhan deteksi dan pemahaman lebih baik tentang botnet melalui identifikasi karakteristik serangannya.

1.6.1 Secara Teoritis

1. Penelitian ini akan menambah literatur ilmiah dalam bidang keamanan siber, khususnya dalam deteksi serangan botnet menggunakan algoritma machine learning.
2. Penelitian ini membantu dalam mengukur efektivitas berbagai metode deteksi dan memungkinkan perbandingan antara teknik yang berbeda.

1.6.1 Secara Praktis

1. Penelitian ini dapat membantu dalam mengidentifikasi dan memitigasi ancaman botnet lebih cepat, mengurangi risiko kerugian finansial dan operasional yang diakibatkan oleh serangan siber.

2. Hasil dari penelitian ini dapat diterapkan untuk mengembangkan sistem deteksi intrusi yang lebih baik, meningkatkan kemampuan jaringan untuk mendeteksi dan merespons serangan botnet secara *real-time*.