BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Kemajuan dalam teknologi informasi dan komunikasi telah memberikan keunggulan yang signifikan bagi masyarakat *modern*, tetapi juga membawa tantangan baru terkait dengan keamanan *cyber*. Menurut Dony Koesmandarin, *Territory Manager Kaspersky for* Indonesia, pada tahun 2023 silam tercatat ada 80 juta serangan *cyber* di Indonesia. Serangan terhadap website di Indonesia terjadi lebih dari 80 ribu kali per harinya dan mereka berhasil melakukan pemblokiran terhadap sekitar 29 juta serangan sepanjang 2023 [1].

Dalam dekade terakhir, terdapat fokus penelitian yang signifikan pada deteksi *malicious network traffic* dan identifikasi jenis serangan [2], [3]. *Malicious traffic* adalah setiap aktivitas jaringan yang dirancang untuk mengganggu, merusak, atau mengakses sistem komputer atau jaringan secara ilegal. Ini mencakup berbagai pertukaran data yang *unauthorized* atau berbahaya, yang sering kali dilakukan oleh *cyber criminal*, untuk mengeksploitasi kerentanan, mencuri data, atau mengkompromikan integritas sistem [4]. Sebagian besar *malicious traffic* dihasilkan oleh berbagai jenis serangan jaringan. Penyerang menggunakan malware untuk merusak sistem komputer dengan mengeksploitasi kerentanan keamanan [5]. Beberapa contoh aktivitas *malicious traffic* adalah distribusi *malware*, *phising*, serangan *Denial of Service* dan masih banyak lagi. Cara-cara setiap jenis *malicious traffic* berbeda-beda, dari *email* dengan *attachment* berisi *malware*, *overload* menggunakan DDoS, iklan *pop-up* dengan kode *malicious* dan sebagainya [4].

Serangan *cyber* semakin kompleks dan merugikan, menyebabkan kerugian finansial, pencurian data, gangguan dalam operasi bisnis dan sebagainya. Oleh karena itu, deteksi dini dan analisis yang akurat terhadap ancaman *cyber* menjadi krusial dalam menjaga keberlangsungan organisasi dan melindungi keamanan informasi [6]. *Machine learning* (ML) mampu memproses data secara cepat dan menemukan pola yang kompleks, yang memungkinkan untuk mendeteksi ancaman *cyber* dengan lebih akurat dan efisien. ML juga telah digunakan untuk berbagai keamanan jaringan seperti analisis *network traffic* [7], deteksi intrusi [8], dan deteksi *botnet* [9].

Algoritme ML yang akan digunakan untuk mendeteksi jaringan berbahaya

dalam penelitian ini ada dua yaitu Naive Bayes dan *Support Vector Machine*. Kedua algoritme ML tersebut merupakan algoritme *supervised* yang dipilih karena sudah ada penelitian sebelumnya yang menggunakan Naive Bayes untuk deteksi url berbahaya [10], begitu juga dengan penelitian yang menggunakan *Support Vector Machine* untuk mendeteksi *malware* [11] dan memberikan hasil yang baik. Penelitian ini akan membahas seberapa akurat kedua algoritme tersebut dalam mendeteksi *malicious traffic*.

1.2 Rumusan Masalah

Seberapa akurat hasil klasifikasi yang didaptkan dari model Naive Bayes dan *Support Vector Machine* dalam mendeteksi *malicious traffic* dalam jaringan?

1.3 Batasan Permasalahan

Dalam penelitian ini, ada beberapa batasan yang diterapkan yaitu:

1. Fokus pada penerapan model Naive Bayes dan *Support Vector Machine* untuk menganalisis jaringan berbahaya dalam data *traffic log*.
2. Pengujian dilakukan pada dataset yang tersedia dan hanya memiliki dua label penentu.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk mengukur tingkat keakuratan model yang dibuat menggunakan algoritme Naive Bayes dan *Support Vector Machine* dalam mendeteksi *traffic* berbahaya dalam jaringan dengan menggunakan *confusion matrix*.

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat berikut:

1. Membantu organisasi dan individu dalam mengimplementasikan solusi pendeteksi *malicious traffic* yang akurat.
2. Menyumbang pada perkembangan teori dan praktik keamanan *cyber* dengan menggunakan pendekatan Machine Learning.

1.6 Sistematika Penulisan

Berikut ini merupakan sistematika penulisan untuk setiap bab dalam laporan berdasarkan penelitian yang sudah dilakukan.

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
  Bab pertama meliputi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.
- Bab 2 LANDASAN TEORI
  Bab kedua merupakan penjelasan singkat tentang teori-teori yang digunakan dalam penelitian ini, yaitu *Support Vector Machine*, dan *Naive Bayes*.
- Bab 3 METODOLOGI PENELITIAN
  Bab ketiga adalah tahap-tahap yang dilakukan dalam proses penelitian. Di antaranya adalah studi literatur, pengumpulan set data, *preprocessing*, *testing* dan penulisan laporan.
- Bab 4 HASIL DAN DISKUSI
  Bab keempat membahas dan menjelaskan kode dan mengevaluasi hasil dari penelitian yang sudah dilakukan.
- Bab 5
  Bab kelima menunjukkan kesimpulan yang diambil dari hasil penelitian dan saran yang diajukan berdasarkan hasil.