BAB I
PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi dan perkembangan internet yang semakin cepat menyebabkan semakin banyak informasi dan data yang perlu dilindungi karena banyak cara yang bisa dilakukan oleh peretas untuk mendapatkan informasi atau data. Kerentanan sebuah jaringan juga disebabkan karena semakin terbukanya pengetahuan tentang *hacking* dan *cracking*. Sehingga, banyak kelompok yang tidak bertanggung jawab mencoba untuk mencuri informasi. Peretasan biasanya dilakukan untuk mencuri data atau meghabiskan sumber daya jaringan yang mengakibatkan komputer atau server tidak bisa bekerja dengan baik. Salah satu jenis serangannya yaitu *Distributed Denial of Service* (DDOS).

Distributed Denial of Service (DDOS) adalah penyerangan dengan lebih dari satu *Attacker* membanjiri dengan paket-paket menuju server, sehingga server sibuk melayani permintaan paket yang sangat banyak dan membuat kinerja server menurun. Apabila permintaan paket yang datang lebih banyak lagi, maka akan menyebabkan kerusakan pada perangkat keras jaringan (Kolahi, 2015). Ada beberapa jenis serangan dari DDOS yang sering terjadi, seperti UDP *Flooding*, SYN *Flooding*, *Ping Of Death*, dan *Remote Controlled Attack*. Serangan DDOS mengakibatkan sistem yang diserang mengalami gangguan berupa *error request*, *halt*, kegagalan sistem, dan sebagainya.

Inilah yang menjadi alasan mengapa keamanan informasi di dalam jaringan diperlukan. Menurut Charles P. Pfleeger (2009), keamanan komputer adalah tindakan atau pencegahan dari serangan pengguna komputer yang tidak memiliki hak akses oleh seorang peretas.

Berbagai penelitian deteksi serangan DDoS telah berhasil mengembangkan berbagai teknik dan metode. Penelitian (Chen, 2014) mengusulkan deteksi serangan DDoS menggunakan teknik *entropy*, dengan membandingkan *entropy* alamat IP sumber dan *entropy* alamat IP tujuan dan berhasil mendeteksi DDoS secara efisien. Untuk menghasilkan klasifikasi data sebelum diproses oleh algoritma maka digunakan *information gain* karena dapat digunakan dalam

memproses teks untuk mendapatkan kategori/klasifikasi (M. Zekri, 2017), dan pentingkan durasi hasil komputasi adalah untuk mengetahui kecepatan algoritma yang digunakan. Penelitian (Nezhad, 2016) yang mengusulkan model *time series* ARIMA dan *chaotic system*, model mampu mengklasifikasikan serangan hingga 99,5%. Pada penelitian (Saied, 2016) ANN (*Artificial Neural Network*) diaplikasikan untuk mendeteksi DDoS dengan karakterisitik fitur khusus, solusi yang ditawarkan memiliki akurasi hingga 98%. Sedangkan (Zekri, 2018) merancang sistem deteksi DDoS pada lingkungan *cloud computing* dengan *algorithma* C.4.5. hasil eksperimen memperlihatkan hasil yang akurat dibandingan algorithma pembelajaran mesin yang lain.

Seleksi fitur merupakan proses pemilihan fitur-fitur penting yang terdapat pada dataset. Fitur-fitur penting yang dimaksud adalah fitur-fitur yang memiliki hubungan erat dengan sebuah label tertentu. Hal ini penting untuk dilakukan agar model klasifikasi yang dibuat dapat merepresentasikan dataset yang akan diuji, menghindari *overfitting*, dan mempercepat proses komputasi (Yu, 2003).

Dalam penelitian ini digunakan algoritma naive bayes classifier dimana Naïve Bayes Classifier memiliki beberapa kelebihan yaitu, cepat dalam proses perhitungan, algoritma yang sederhanadan akurasi yang tinggi (Muhamad, Prasojo, Sugianto, Surtiningsih, & Cholissodin, 2017). Untuk pembanding algoritma naive bayes maka digunakan algoritma random forest yang merupakan salah satu metode yang digunakan untuk klasifikasi dengan membangun banyak pohon klasifikasi (Sandag, 2020). Metode ini dapatmeningkatkan hasil akurasi, dengan cara membangkitkan simpul anak untuk setiap node (simpul diatasnya) dan dilakukan pemilihan secara acak, Kemudian hasil klasifikasi dari setiap pohon diakumulasikan dan dipilh hasil klasifikasi yang paling banyak muncul.

1.2 Perumusan Masalah

Berdasarkan latar belakang diatas, maka peneliti dapat merumuskan permasalahan yaitu:

1. Bagaimana menerapkan metode seleksi fitur menggunakan *Information Gain*?
2. Bagaimana cara mengklasifikasikan DDoS pada dataset CICIDS 2018

menggunakan metode Naïve Bayes dan *Random Forest*?

3. Bagaimana perbandingan akurasi dari menggunakan metode Naïve Bayes dan *Random Forest*?

1.3 Pernyataan Masalah

Berdasarkan latar belakang diatas dapat di pernyataan masalah yang ada seperti:

1. Pentingnya melakukan seleksi fitur sebelum mengklasifikasikan data.
2. Belum terdapat sistem yang dapat membuktikan perbandingan antar algoritma yang digunakan untuk klasifikasi data DDoS CICDS 2018.
3. Belum diketahui tingkat akurasi klasifikasi Naïve Bayes dan *Random Forest* menggunakan dataset DDoS CICIDS 2018.

1.4 Tujuan

Tujuan dari penelitian ini adalah menjawab berbagai masalah yang telah penulis uraikan pada perumusan masalah, yaitu :

1. Menerapkan metode seleksi fitur menggunakan *Information Gain*.
2. Menerapkan Naïve Bayes dan *Random Forest* untuk mengklasifikasikan serangan menggunakan dataset DDoS CICIDS 2018.
3. Mengetahui perbandingan akurasi dari metode Naïve Bayes dan *Random Forest.*

1.5 Batasan Masalah

Mengingat luasnya kemungkinan pembahasan masalah yang akan dilakukan, maka penelitian ini membatasi ruang lingkup masalah agar pembahasan dapat lebih terfokus dan tujuan penulisan dapat tercapai. Pembatasan ruang lingkup permasalahan dalam penelitian ini meliputi:

1. Metode seleksi fitur yang digunakan adalah *Information Gain*
2. Dataset yang digunakan DDoS *balanced* CICDS 2018.
3. Metode klasifikasi yang digunakan Naïve Bayes dan *Random Forest.*
4. Perhitungan tingkat akurasi menggunakan Confusion Matrix.

1.6 Hipotesis

Hipotesis dari penelitian ini adalah:

4. *Information Gain* merupakan metode yang tepat untuk seleksi fitur.
5. Algoritma Naïve Bayes dan *Random Forest* dapat mengklasifikasi serangan DDoS terhadap server.
6. Tingkat akurasi dari metode *Random Forest* lebih baik dibandingkan metode Naïve Bayes untuk klasifikasi serangan DDoS.

1.7 Sistematika Pembahasan

Berikut adalah sistematika penyusunan dalam penelitian ini:

BAB I : PENDAHULUAN

Bab ini berisikan latar belakang, perumusan masalah, ruang lingkup dan batasan masalah, tujuan, manfaat serta metodologi dan sistematika penulisan yang dipakai.

BAB II : KAJIAN PUSTAKA

Bab ini berisikan berbagai teori yang berhubungan dengan Perancangan *Bot* Untuk *Monitoring Server* Dari Serangan *DDOS*.

BAB III : METODOLOGI DAN DESAIN SISTEM

Bab ini membahas Analisis berisi kajian awal terhadap sistem yang ada berdasarkan data yang diperoleh melalui wawancara, pustaka, dokumen-dokumen, penilaian kelebihan-kekurangan terhadap sistem yang ada, kebutuhan sistem yang ideal, usulan sistem yang baru. Perancangan berisi alur kerja sistem yang lama, alur kerja sistem yang baru/usulan, diagram konteks usulan, diagaram alir data usulan, relasi antar entitas usulan, rancangan tabel usulan.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini berisikan uraian tentang hasil implementasi monitoring server.

BAB V : KESIMPULAN DAN SARAN

Bab ini berisikan kesimpulan dari hasil penelitian yang dilakukan dan saran-saran terhadap kekurangan dari penelitian tersebut