BAB I
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang begitu pesat telah menjadikan aplikasi *web* sebagai bagian penting dalam berbagai aktivitas *digital* di banyak sektor. Dalam dunia *e-commerce*, aplikasi *web* memudahkan proses jual beli secara cepat dan praktis, sekaligus membuka peluang pasar tanpa batasan *geografis* [1]. Selain itu, di bidang pendidikan, aplikasi *web* memungkinkan proses belajar-mengajar berlangsung secara *daring* dengan fleksibilitas tinggi, sehingga siswa dan guru tetap bisa berinteraksi meski berada di tempat yang berbeda [2]. Sementara itu, dalam sektor layanan publik, masyarakat kini semakin mudah dalam mengurus berbagai surat keterangan dan dokumen *admin*istrasi secara *online* berkat kehadiran aplikasi *web* [3] .

Namun di balik semua kemudahan tersebut, keamanan aplikasi *web* menjadi tantangan besar [4]. Gangguan seperti sistem yang tiba-tiba tidak bisa diakses *(downtime)* atau adanya celah keamanan bisa berdampak serius, mulai dari kerugian finansial, kebocoran data sensitif, hingga menurunnya kepercayaan pengguna terhadap layanan *digital* tersebut [5][6]. Ancaman ini semakin besar dengan adanya berbagai jenis serangan siber yang terus berkembang, seperti *Distributed Denial of Service* (DDoS), *SYN Flood*, *SQL Injection, Port Scanning*, dan *Brute Force* yang dapat mengganggu stabilitas sistem, mengeksploitasi kerentanan, hingga mencuri informasi penting [7]. Itulah sebabnya, keandalan dan keamanan sistem menjadi aspek yang sangat penting untuk dijaga [6].

Seiring meningkatnya penggunaan teknologi *digital*, serangan siber tidak hanya semakin sering terjadi, tetapi juga semakin kompleks dan sulit diprediksi. Dampak yang ditimbulkan pun semakin serius, mulai dari pencurian data, perusakan sistem, hingga melumpuhkan layanan dalam skala besar. Dalam menghadapi hal ini, metode keamanan tradisional seperti sistem berbasis aturan

(*rule-based system*) banyak digunakan. Sistem ini bekerja dengan cara mencocokkan pola aktivitas jaringan terhadap aturan yang telah ditentukan sebelumnya, biasanya berupa tanda tangan (*signature*) serangan yang sudah pernah terjadi. Meskipun efektif dalam mendeteksi serangan yang sudah dikenal, sistem *rule-based* memiliki keterbatasan karena tidak mampu mengenali serangan baru (*zero-day attack*), serta kinerjanya dapat menurun ketika jumlah aturan semakin banyak dan komplek [8].

Keterbatasan inilah yang menunjukkan perlunya pendekatan yang lebih cerdas dan adaptif, salah satunya melalui pemanfaatan teknologi *Machine learning* (ML) [9]. Berbeda dengan metode *rule-based* yang hanya mampu mengenali serangan yang sudah didefinisikan sebelumnya [10], *Machine learning* dapat mempelajari pola dari data historis dan menyesuaikan diri terhadap ancaman baru (*zero-day attack*). Selain itu, ML mampu mengolah lalu lintas jaringan yang besar dan kompleks dengan lebih efisien, sehingga deteksi serangan dapat dilakukan dengan lebih cepat dan akurat [9]. Dalam penelitian ini digunakan *algoritma Random Forest*, yaitu metode *ensemble learning* yang menggabungkan banyak pohon keputusan untuk menghasilkan prediksi yang lebih akurat dan stabil. Keunggulan utama *Random Forest* terletak pada kemampuannya menangani data besar dan kompleks, mengurangi risiko *overfitting*, serta memberikan akurasi tinggi dalam mendeteksi serangan [11].

Dalam penerapannya, sistem ini dikembangkan dengan cara merekam lalu lintas jaringan menggunakan *Scapy* untuk menangkap paket data secara *real-time*, kemudian dilakukan tahap *preprocessing* seperti pembersihan, konversi ke format numerik, penanganan data kosong, dan normalisasi. Selanjutnya, dilakukan seleksi fitur dengan metode *Mutual Information* untuk memperoleh fitur yang paling relevan, yang kemudian digunakan dalam pelatihan model menggunakan algoritma *Random Forest.* Model yang dihasilkan diimplementasikan pada *Server* untuk memantau lalu lintas jaringan, dan ketika terdeteksi serangan seperti *SQL Injection* atau *Brute Force*, sistem akan langsung mengklasifikasikannya serta mengirimkan notifikasi otomatis kepada *admin* melalui Bot Telegram agar dapat segera dilakukan penanganan.

Berdasarkan latar belakang diatas, penulis akan mengusulkan sebuah sistem yang bisa mengklasifikasi serangan siber berbasis *machine learning*. Sistem ini dirancang untuk mengenali pola serangan secara otomatis dari lalu lintas jaringan, sehingga dapat memberikan deteksi yang lebih cepat, akurat, dan efisien. Dengan adanya sistem ini, diharapkan dapat membantu meningkatkan keamanan jaringan dan meminimalkan dampak yang ditimbulkan oleh berbagai ancaman siber.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dapat dinyatakan rumusan masalah pada penelitian ini sebagai berikut:

1. Bagaimana merancang sistem klasifikasi serangan siber berbasis *machine learning* dengan menggunakan algoritma *Random Forest*?

1.3 Batasan Masalah

Penelitian ini dibatasi pengembangan sistem untuk klasifikasi serangan siber berbasis *machine learning* pada lalu lintas jaringan aplikasi *web.* Jenis serangan yang dipilih adalah *Benign, SQL Injection*, dan *Brute Force*. Algoritma yang digunakan adalah *Random Forest*. Pengolahan data meliputi pembersihan, normalisasi, dan seleksi fitur menggunakan *Mutual Information*, serta implementasi sistem dilakukan di *Server Ubuntu*, dengan pengiriman notifikasi jika terdeteksi serangan melalui Bot Telegram. Penelitian ini tidak membahas serangan lain atau pencegahan serangan, tetapi hanya fokus pada deteksi dan klasifikasi serangan secara *real-time*. Selain itu, penelitian ini dibatasi hanya pada aktivitas *login form website* sebagai pengujian serangan, sehingga sistem tidak mencakup halaman lain atau fitur aplikasi *web* di luar proses *login* .

1.4 Tujuan

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Mengolah dataset CIC-IDS-2018 melalui tahap pembersihan data, konversi ke bentuk numerik, penanganan data kosong, normalisasi, serta melakukan seleksi fitur dengan metode *Mutual Information*.

2. Mengembangkan model klasifikasi serangan siber menggunakan algoritma *Random Forest* berdasarkan data yang telah diproses.
3. Menerapkan model yang telah dibangun ke dalam sistem klasifikasi untuk menganalisis lalu lintas jaringan dan mengenali jenis serangan.
4. Mengintegrasikan sistem dengan notifikasi *real-time* melalui Bot Telegram agar *admin* dapat segera memperoleh informasi ketika terjadi serangan.

1.5 Manfaat

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Mengembangkan sistem klasifikasi yang mampu mendeteksi serangan jaringan seperti *SQL Injection* dan *Brute Force*.
2. Mengintegrasikan sistem dengan notifikasi *real-time* melalui Bot Telegram agar *admin* dapat segera mengetahui serangan yang terjadi.
3. Mendukung proses respon keamanan yang lebih efektif dengan memberikan informasi penting kepada *admin* secara cepat dan tepat.