

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG MASALAH

Internet of Things (IoT) telah menjadi salah satu teknologi yang mengalami kemajuan pesat dalam beberapa tahun terakhir. Berdasarkan penelitian terbaru, pada tahun 2022 tercatat sebanyak 14,4 miliar perangkat IoT yang aktif di seluruh dunia, jumlah ini diperkirakan akan meningkat hingga mencapai 30 miliar pada tahun 2027 [1]. Penelitian lain memprediksi bahwa jumlah perangkat IoT dapat melonjak hingga 75 miliar dalam beberapa tahun mendatang [2].

Namun, seiring dengan peningkatan penggunaan teknologi IoT, muncul berbagai masalah, seperti masalah privasi, keamanan, konfigurasi sistem, kontrol akses, verifikasi [3]. Selain itu, jaringan IoT juga sangat rentan terhadap serangan, menurut laporan [4] pada tahun 2023, serangan *cyber* yang menargetkan perangkat IoT mengalami peningkatan 400% dibandingkan dengan tahun 2022, dan menurut [5] 1,5 miliar serangan IoT terdeteksi dalam 6 bulan pertama pada tahun 2023.

Serangan *brute force* masih menjadi salah satu teknik yang sering digunakan oleh peretas untuk membobol sistem[6]. Variasi dari serangan ini, seperti *dictionary attack*, sering memanfaatkan kecenderungan pengguna dalam menggunakan kata sandi umum atau mudah ditebak [7]. Untuk itu, dibutuhkan sistem yang mampu mendeteksi upaya-upaya intrusi secara cepat dan tepat. Salah satu sistem yang sering digunakan adalah *Intrusion Detection System* (IDS), yang dirancang untuk mengawasi dan menganalisis lalu lintas jaringan guna mengenali adanya aktivitas mencurigakan [8].

Namun, kompleksitas data dalam lingkungan IoT menuntut metode deteksi yang tidak hanya akurat, tetapi juga efisien dalam memproses data berskala besar. Dalam hal ini, *random forest* menjadi salah satu metode yang menjanjikan. Metode ini telah banyak digunakan dalam berbagai penelitian karena mampu memberikan hasil yang stabil dalam kondisi data yang bervariasi [9].

Untuk mendukung kinerja *random forest*, pemilihan fitur yang relevan juga menjadi langkah penting. Teknik seperti *Recursive Feature Elimination* (RFE) membantu menyederhanakan jumlah fitur yang digunakan, sehingga proses pelatihan model menjadi lebih cepat dan hasilnya lebih mudah dipahami, tanpa harus mengorbankan akurasi [10].

Gabungan metode *random forest* dan *Recursive Feature Elimination* (RFE) dipilih dalam penelitian ini karena keduanya telah terbukti efektif dalam studi sebelumnya. Penelitian [11] menunjukkan bahwa *random forest* berhasil mendeteksi serangan *brute force* pada protokol SSH, FTP, dan XSS dengan akurasi 97.77%. Penelitian lain [12] mencatat akurasi tertinggi 99.9% dalam mendeteksi serangan SSH dan FTP menggunakan metode yang sama. Namun, kedua studi tersebut menggunakan *dataset* CSE CIC IDS2018 yang tidak mewakili karakteristik khas jaringan IoT, seperti ukuran paket yang kecil dan frekuensi komunikasi yang tinggi.

Sementara itu, penelitian [10] menggabungkan *random forest* dengan RFE untuk seleksi fitur pada data berdimensi tinggi dan mencatat akurasi 91.9%. Sayangnya, penerapannya masih terbatas pada konteks umum, tanpa fokus pada serangan *cyber* di lingkungan IoT.

Berdasarkan studi-studi sebelumnya dapat disimpulkan bahwa metode *random forest* dan RFE memiliki kinerja yang konsisten kuat dalam deteksi serangan *cyber*. Namun, penerapannya masih terkendala pada keterbatasan *dataset* yang digunakan, baik karena bersifat umum maupun hanya berfokus pada satu protokol tertentu, sehingga belum mencerminkan karakteristik khas jaringan IoT. Kondisi ini menimbulkan celah riset yang penting untuk diisi. Oleh karena itu, penelitian ini menggunakan *dataset* CIC IoT 2023 yang lebih representatif terhadap kondisi nyata IoT, sehingga dapat memberikan gambaran lebih akurat mengenai efektivitas *random forest* dan RFE dalam mendeteksi serangan di lingkungan IoT.

Melalui pendekatan *random forest* yang dioptimalkan dengan RFE, penelitian ini diharapkan mampu membangun model deteksi *dictionary brute force* yang lebih akurat dan sesuai dengan kebutuhan sistem IoT saat ini. Fokus utama dari penelitian ini adalah mengatasi tantangan keamanan pada jaringan perangkat IoT, khususnya dalam mendeteksi serangan *dictionary brute force* secara efektif dan tepat sasaran. Atas dasar itu, penelitian ini diberi judul: “Klasifikasi Serangan *Dictionary Brute Force* pada Jaringan IoT Menggunakan *Random Forest* dan *Recursive Feature Elimination*.”

1.2 RUMUSAN MASALAH

Jaringan *Internet of Things* (IoT) rentan terhadap serangan *dictionary brute force* yang mengancam keamanan data dan operasional sistem, diperparah oleh kompleksitas dan pertumbuhan perangkat IoT [13]. Oleh karena itu penelitian ini bertujuan untuk menjawab:

1. Bagaimana penerapan metode *random forest* dalam mendeteksi serangan *dictionary brute force*?
2. Bagaimana performa model *random forest* dalam mendeteksi serangan *dictionary brute force*?
3. Bagaimana pengaruh *recursive feature elimination* terhadap hasil evaluasi model?

1.3 BATASAN MASALAH

Untuk Menghindari pembahasan diluar judul penelitian, maka penulis melakukan pembahasan yang lebih spesifik, berdasarkan judul tersebut. Adapun Batasan masalahnya mencakup:

1. *Dataset* yang digunakan pada penelitian ini adalah CIC IoT 2023 (<https://www.unb.ca/cic/datasets/iotdataset-2023.html>) yang memiliki 47 fitur, yang dikelompokkan menjadi 7 kategori utama:
 - a. Waktu dan durasi: *flow duration, Duration, IAT (Inter-Arrival Time)*.
 - b. Karakteristik dan statistik paket: *Header Length, Tot size, Number, Tot sum, Min, Max, AVG, Std*.
 - c. Protokol dan layanan: *Protocol Type, TCP, UDP, DHCP, ARP, ICMP, IPv, LLC, HTTP, HTTPS, DNS, Telnet, SMTP, SSH, IRC*.
 - d. Tanda dan frekuensi TCP: *fin flag number, syn flag number, rst flag number, psh flag number, ack flag number, ece flag number, cwr flag number, ack count, syn count, fin count, urg count, rst count*.
 - e. Kecepatan transfer: *Rate, Srate, Drate*.

- f. Analisis numerik: *Magnitude, Radius, Covariance, Variance, Weight*.
 - g. *Label* klasifikasi: *label*. *Label* 0 menandakan aktivitas normal, sedangkan *label* 1 menunjukkan serangan *dictionary brute force*.
2. *Tool* yang akan digunakan adalah *Google Colab*.
 3. Metode yang akan digunakan adalah *random forest* dan *recursive feature elimination*.
 4. Hanya fokus pada serangan *dictionary brute force*.
 5. Pengujian tidak dilakukan pada lalu lintas *real-time*.
 6. Tidak membahas langkah pencegahan.
 7. Tidak mencakup perbandingan dengan metode lain.
 8. Tidak membahas *hyperparameter tuning* secara mendalam.

1.4 TUJUAN DAN MANFAAT PENELITIAN

Penelitian ini bertujuan untuk menghasilkan model deteksi serangan *dictionary brute force* yang efektif pada jaringan IoT. Untuk mendukung pencapaian tujuan tersebut, berikut dijelaskan secara terpisah mengenai tujuan yang ingin dicapai serta manfaat yang diharapkan dari penelitian ini.

1.4.1 Tujuan Penelitian

1. Mengimplementasikan metode *random forest* untuk mendeteksi serangan *dictionary brute force* menggunakan *dataset* CIC IoT 2023.
2. Mengevaluasi performa model deteksi yang telah dilatih untuk mengidentifikasi serangan *dictionary brute force*.

3. Menganalisis pengaruh *recursive feature elimination* terhadap evaluasi model deteksi serangan.

1.4.2 Manfaat Penelitian

1. Membuat deteksi yang lebih baik dengan mengembangkan model yang lebih sensitif dan cepat dalam mengenali pola serangan.
2. Mengurangi dampak serangan dengan mendeteksi serangan lebih awal, sehingga dapat diambil tindakan pencegahan untuk mencegah kerusakan lebih lanjut pada sistem.
3. Berkontribusi pada pengembangan metode deteksi serangan *brute force* yang lebih baik secara keseluruhan, membantu meningkatkan keamanan jaringan secara menyeluruh.

1.5 SISTEMATIKA PENELITIAN

Secara garis besar penulisan ini akan terdiri dari lima bab. Gambaran umum dari penelitian dapat dilihat dalam sistematika penelitian sebagai berikut:

BAB I : PENDAHULUAN

Pada bab ini, akan dijelaskan tentang latar belakang penelitian, termasuk urgensi dan relevansi topik penelitian, serta perumusan masalah yang akan diselesaikan. Tujuan dari bab ini adalah memberikan gambaran umum tentang topik penelitian dan mengapa penelitian ini penting untuk dilakukan.

BAB II : LANDASAN TEORI

Bab ini akan membahas landasan teori yang merangkum penelitian terdahulu sehubungan dengan deteksi serangan *brute force*, penerapan *random forest* dan *recursive feature elimination*. Selain itu, akan dipaparkan juga landasan teori yang mencakup konsep-konsep penting dalam deteksi serangan dan konsep-konsep dasar yang relevan dengan penelitian ini.

BAB III : METODOLOGI PENELITIAN

Pada bab ini, akan dijelaskan tentang rancangan penelitian, langkah-langkah pengumpulan data, pemilihan *dataset*, dan implementasi algoritma *machine learning* dan *feature selection*.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini akan berisi tentang analisis data dan hasil eksperimen yang telah dilakukan. Selain itu, akan dibahas juga pembahasan terkait hasil penelitian, interpretasi temuan, serta implikasi dari hasil tersebut.

BAB V : KESIMPULAN DAN SARAN

Pada bab terakhir ini, akan disajikan rangkuman dari seluruh penelitian, termasuk kesimpulan dari temuan utama dan saran untuk penelitian selanjutnya.

