BAB I
PENDAHULUAN

1.1 Latar Belakang

Di era digital ini keamanan jaringan komputer semakin krusial. Seiring dengan pertumbuhan pesat teknologi informasi, jumlah perangkat yang terhubung ke jaringan terus bertambah, baik untuk keperluan bisnis, pemerintahan, maupun personal. Seiring meningkatnya kompleksitas integrasi infrastruktur jaringan, risiko terhadap keamanan server pun turut meningkat secara substansial [1]. Berbagai bentuk serangan server, mulai dari serangan *Denial of Service* (DoS), *Remote to Local* (R2L), *User to Root* (U2R), dan *Probe*, menjadi ancaman serius, ditambah munculnya serangan baru yang lebih beragam seperti *Fuzzers, Generic, Analysis, Reconnaissance, Backdoor, Worms, Exploits, dan Shellcode* yang makin menantang untuk dideteksi [2]. Dengan demikian, kemampuan mengidentifikasi anomali pada jaringan komputer secara dini menjadi krusial dalam upaya mitigasi potensi kerugian akibat serangan terhadap server [3].

Pendekatan tradisional dalam mendeteksi serangan siber umumnya bergantung pada metode *signature-based* [4]. *Signature-based* adalah metode deteksi ancaman yang bekerja dengan cara mencocokkan pola atau *signature* tertentu dari serangan yang terdokumentasi sebelumnya untuk mengenali aktivitas berbahaya[5]. Metode ini hanya efektif dalam menghadapi serangan dengan pola atau signature yang telah terdokumentasi [6]. Namun, dengan semakin canggihnya ancaman, mekanisme keamanan perlu dikembangkan agar dapat secara proaktif mendeteksi anomali dalam arus data dan perilaku mencurigakan [7]. Dengan semakin seringnya serangan yang belum pernah teridentifikasi, metode deteksi lama menjadi kurang efektif, sehingga dibutuhkan strategi yang lebih efisien dalam menghadapi ancaman baru ini [8].

Prediksi dini serangan terhadap server menjadi fokus utama dalam meningkatnya keamanan siber. Pemanfaatan teknologi *Machine Learning* (ML) menjadi semakin krusial dalam upaya identifikasi serangan yang tidak biasa [3].

ML memungkinkan sistem untuk mempelajari pola dari data yang tersedia dan mengidentifikasi anomali tanpa bergantung pada signature spesifik [9]. Beberapa tahun terakhir ini, berbagai model algoritma *ML* telah diterapkan untuk mendeteksi anomali jaringan. Salah satu contohnya adalah *Support Vector Machine* (SVM), yang telah terbukti efektif dalam klasifikasi dan regresi. SVM sering digunakan dalam deteksi anomali jaringan karena kemampuannya yang unggul dalam memisahkan data normal dari data yang mencurigakan [10].

Berdasarkan permasalahan yang telah diidentifikasi, penelitian ini menggunakan lima algoritma *Machine Learning*, yaitu *Logistic Regression, Random Forest, Support Vector Machine (SVM), Naive Bayes*, dan *K-Nearest Neighbors (KNN)*, karena masing-masing mewakili pendekatan klasifikasi yang berbeda: *linear, ensemble, margin-based, probabilistik*, dan *instance-based*. Pemilihan algoritma ini bertujuan untuk mengevaluasi efektivitas dan efisiensi deteksi anomali dari berbagai perspektif pemodelan, sehingga dapat ditentukan model yang paling adaptif terhadap variasi serangan. Penelitian ini juga mempertimbangkan konteks implementasi sistem deteksi pada infrastruktur keamanan seperti *Security Information and Event Management (SIEM)*. Dalam sistem semacam ini, data dapat bersumber dari dua arah utama: trafik jaringan dan log aktivitas system. Dataset *NSL-KDD* dan *UNSW-NB15* dipilih karena menyediakan representasi dari keduanya baik dalam bentuk fitur lalu lintas jaringan maupun pencatatan aktivitas yang menyerupai log. Oleh karena itu, kedua dataset ini relevan untuk menguji efektivitas model deteksi dalam konteks dunia nyata, di mana sistem keamanan memerlukan kemampuan analisis terhadap data dari berbagai sumber. Harapannya, model yang dikembangkan dapat meningkatkan deteksi dini terhadap serangan siber dan memperkuat mitigasi risiko pada sistem server.

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini sebagai berikut :

1. Bagaimana kinerja model *Machine Learning* dalam mengidentifikasi anomali pada sistem jaringan komputer?

1.3 Batasan Masalah

Beberapa Batasan masalah yang diambil dalam penelitian ini adalah :

1. Evaluasi model hanya dilakukan terhadap performa dari lima algoritma machine learning yang telah dipilih: *Logistic Regression, Random Forest, Support Vector Machine (SVM), Naive Bayes,* dan *K-Nearest Neighbors (KNN).* Penelitian tidak membandingkan dengan metode lain seperti *deep learning*.
2. Proses klasifikasi dan evaluasi model dilakukan menggunakan data yang telah dipra-proses, tanpa mempertimbangkan proses pengumpulan data dari jaringan secara real-time.

1.4 Tujuan Penelitian

Tujuan dari penelitian sebagai berikut :

1. Mengembangkan dan mengevaluasi model deteksi anomali menggunakan algoritma Machine Learning untuk mengukur efektivitas identifikasi pada jaringan komputer.
2. Menganalisis performa dari masing-lima algoritma Machine Learning yang dipilih dalam mengidentifikasi anomali pada data jaringan komputer.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut :

1. Menjadi dasar penelitian lanjutan dalam pengembangan algoritma baru atau peningkatan algoritma yang sudah ada untuk deteksi serangan siber.
2. Memberikan wawasan dan pengetahuan yang diperlukan untuk pengembangan teknologi keamanan canggih, yang dapat bermanfaat bagi pengembang dan peneliti di bidang keamanan siber.
3. Memberikan Gambaran performa dari algoritma *Logistic Regression,*

Random Forest, Support Vector Machine (SVM), Naive Bayes, dan *K-Nearest Neighbors (KNN).*

1.6 Sistematika Penulisan

Untuk memudahkan pembaca dalam memahami dan mengikuti penelitian ini, penulis telah menyusun struktur penulisan sebagai berikut :

BAB I PENDAHULUAN

Bab ini membahas latar belakang yang menjelaskan pentingnya deteksi anomali pada jaringan komputer, perumusan masalah yang menjadi fokus penelitian, Batasan masalah untuk membatasi ruang lingkup penelitian, tujuan yang ingin dicapai, serta manfaat penelitian. Di bagian akhir, dijelaskan sistematika penulisan untuk memandu pembaca memahami alur penelitian ini.

BAB II TINJAUAN PUSTAKA

Memaparkan konsep konsep kunci terkait penelitian, meliputi Jenis jenis serangan pada server, serta penjelasan mendalam tentang algoritma *Support Vector Machine* (SVM), *Random Forest, Naive Bayes, K-Nearest Neighbors* (KNN), dan *Logistic Regression*.

BAB III METODE PENELITIAN

Menjabarkan flowchart alur penelitian yang dimulai dari pengumpulan data, pengolahan data, implementasi *Support Vector Machine* (SVM), *Random Forest, Naive Bayes, K-Nearest Neighbors* (KNN), dan *Logistic Regression*

BAB IV HASIL DAN PEMBAHASAN

Menyajikan temuan penelitian beserta interpretasinya. Perbandingan efektivitas *Support Vector Machine* (SVM), *Random Forest, Naive Bayes, K-Nearest Neighbors* (KNN), dan *Logistic Regression* dalam mendeteksi anomaly dan serangan, termasuk analisis kelebihan dan keterbatasan masing masing metode.

BAB V PENUTUP

Merangkum hasil utama penelitian terkait deteksi anomali jaringan, disertai

rekomendasi untuk pengembangan dan studi lanjutan di masa depan