BAB I
PENDAHULUAN

1.1 Latar Belakang Penelitian

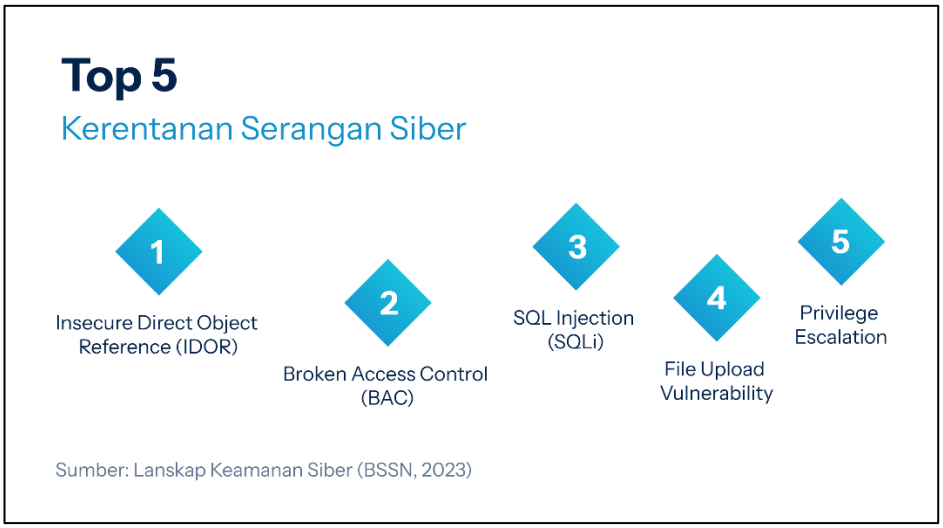
Dalam era revolusi digital saat ini, sistem keamanan jaringan menjadi salah satu aspek yang krusial, terutama dalam menghadapi berbagai ancaman siber yang terus berkembang. Dikutip dari data yang dirilis oleh Badan Siber dan Sandi Negara (BSSN) dalam Lanskap Keamanan Siber Indonesia pada tahun 2023, Indonesia mengalami total trafik anomali sebesar 403.990.813 anomali. Angka ini mencerminkan adanya gangguan atau aktivitas yang tidak biasa pada lalu lintas data yang dapat mengindikasikan potensi masalah di sistem jaringan atau perangkat yang digunakan.

Gambar 1.1 Trafik Anomali Serangan Siber

(Badan Siber dan Sandi Negara, 2023)

Trafik anomali tertinggi terjadi pada bulan Agustus, dengan total 78.464.385 anomali, yang mungkin menunjukkan adanya lonjakan atau serangan siber yang lebih intensif pada bulan tersebut. Sebaliknya, trafik anomali terendah tercatat pada bulan November, dengan 19.296.439 anomali, yang menunjukkan adanya periode dengan gangguan yang relatif lebih sedikit. Aktivitas anomali trafik ini berpotensi

menimbulkan berbagai dampak negatif yang cukup serius bagi organisasi atau perusahaan yang terlibat. Salah satu dampak utama yang dapat terjadi adalah penurunan performa perangkat dan jaringan, di mana kinerja sistem menjadi terganggu akibat lonjakan trafik yang tidak normal. Selain itu, terdapat risiko pencurian data sensitif yang dapat membahayakan informasi penting seperti data pribadi, finansial, atau informasi bisnis. (Badan Siber dan Sandi Negara, 2023).



Gambar 2.2 Top 5 Kerentanan Serangan Siber

(Badan Siber dan Sandi Negara, 2023)

Berdasarkan hasil ITSA (*Information Technology Security Assessment*) yang dilakukan pada tahun 2023, ditemukan sebanyak 2.860 celah keamanan yang berpotensi mengancam keamanan sistem dan data, terdapat lima kerentanan yang memiliki tingkat risiko *critical* dan perlu mendapat perhatian lebih besar karena dampaknya yang sangat merugikan jika tidak segera ditangani. Pertama, *Insecure Direct Object Reference* (IDOR), yaitu kerentanan IDOR terjadi ketika aplikasi website gagal melakukan validasi atau otorisasi yang memadai pada permintaan akses objek atau data. Kedua, *Broken Access Control* (BAC), adalah celah keamanan yang terjadi ketika kontrol akses pada sistem atau aplikasi tidak diterapkan secara tepat. Kerentanan ini memungkinkan *threat actor* untuk memperoleh akses yang tidak sah ke bagian-bagian sistem atau data yang seharusnya terlindungi. Ketiga, SQL *Injection* (SQLi), adalah serangan yang mengeksploitasi celah pada aplikasi website yang tidak memadai dalam memproses

input dari pengguna. *Threat actor* dapat menyisipkan kode SQL berbahaya ke dalam input aplikasi, yang kemudian dapat dieksekusi pada basis data. Keempat, *File Upload Vulnerability*, *case* kerentanan ini terjadi ketika aplikasi *website* memungkinkan pengguna untuk mengunggah file tanpa melakukan validasi yang memadai terhadap file yang diunggah. Dan terakhir, *Privilege Escalation*, *case* ini terjadi ketika threat actor berhasil meningkatkan hak akses atau hak istimewa mereka untuk memperoleh kontrol lebih besar atas sistem atau aplikasi. Dengan memperoleh hak akses yang lebih tinggi, penyerang dapat mengeksekusi perintah yang tidak sah, merusak data, atau mengambil alih kontrol penuh terhadap sistem. (Badan Siber dan Sandi Negara, 2023)

Dari top 5 serangan ancaman anomali yang terjadi berdasarkan data di atas salah satu serangan yang sering terjadi adalah SQL *Injection*. SQL *Injection* merupakan teknik serangan yang memanfaatkan celah keamanan dalam aplikasi website untuk menyisipkan perintah SQL berbahaya. Serangan ini dapat dimanfaatkan untuk mencuri informasi sensitif pengguna, merusak basis data, atau bahkan mengambil alih kendali situs website. Penyerang dapat memanfaatkan celah keamanan ini dengan menyisipkan perintah SQL berbahaya ke dalam input pengguna, seperti pada formulir autentikasi atau URL situs website, guna memperoleh akses tidak sah terhadap sistem (Natanael dkk., 2024).

Karena dampak serangan siber dalam metode SQL *injection* menjadi top 5 dengan kerentanan tinggi dan cukup krusial yang dapat mengancam keamanan pada sebuah website, untuk itu perlu diupayakan langkah-langkah metode mitigasi atau meminimalisir terjadi nya serangan siber berupa SQL *injection*, seperti yang dilakukan (Paul dkk., 2024) dalam upaya mencegah serangan SQL *injection* menurutnya metode untuk mitigasi SQL *injection* terbagi menjadi tiga, yaitu: *Traiditional framework, machine learning based framework, dan deep learning based framework.*

Seiring berkembangnya teknologi penanganan mitigasi SQL *injection* lebih difokuskan dengan berbasis *machine learning*, dan *deep learning*, selain akurasi yang baik kedua metode tersebut juga performa dan efisiensi nya dapat diandalkan. Untuk SQL *injection prevention* berdasarkan *machine learning based* terdiri dari

beberapa teknis algoritma seperti *Naïve bayes*, *SVM*, *Logistic Regression,* dan *Random Forest.* (Abdullah & Mohsin Abdulazeez, 2024). Dari riset yang dilakukan oleh (Triloka dkk., 2022), dalam menguji lima algoritma untuk pendeteksian SQL *injection* menerangkan bahwa metode *Logistic Regression* menjadi salah satu hasil yang paling baik dengan nilai *accuracy* 0,996.

Selain metode metode preventif di atas, sebagai langkah awal ketika terjadi insiden serangan siber yang perlu dilakukan adalah menyimpan *log server* sebagai dasar untuk tindak lanjut analisis mendeteksi ada tidak nya serangan SQL *injection* dalam proses ini dilakukan *text mining* untuk melakukan pembobotan data pada *log server* dengan metode TF-IDF (Asnawi dkk., 2023). Penelitian yang dilakukan oleh (Li & Zhang, 2019) mencoba membuktikan beberapa algoritma *machine learning based* yang dikombinasikan dengan TF-IDF sebagai pendeteksian SQL *injection*, beberapa di antara nya adalah dengan menggabungkan metode TF-IDF dengan *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), dan *Decision Tree* (DT).

Dari referensi penelitian yang dilakukan oleh (Li & Zhang, 2019), belum ada uji coba kombinasi TF-IDF dan *Logistic Regression* sebagai pendeteksian SQL *injection*, untuk itu penulis mencoba melakukan penelitian tersebut dengan judul Implementasi Algoritma *Term Frequency Inverse Document Frequency* dan *Logistic Regression* Untuk Deteksi SQL *Injection.*

1.2 Rumusan Masalah Penelitian

Berdasarkan permasalahan yang telah dijelaskan sebelumnya, peneliti merumuskan beberapa rumusan masalah dalam penelitian ini diantaranya:

1. Bagaimana menerapkan metode TF – IDF dan algoritma *Machine Learning Logistic regression* untuk mendeteksi SQL *Injection* pada aplikasi berbasis website?
2. Bagaimana kinerja sistem dalam melakukan deteksi SQL *Injection*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan sebelumnya, maka tujuan dari penelitian ini untuk:

1. Menerapkan metode *pre-processing* TF-IDF dan algoritma *Machine Learning Logistic regression* untuk melakukan deteksi terhadap input menggunakan dataset SQL *Injection Attack* pada website.
2. Mengevaluasi kinerja sistem yang di bangun dalam mendeteksi SQL *Injection*.

1.4 Batasan Masalah

Penelitian ini memiliki beberapa batasan masalah sebagai berikut:

1. Penelitian ini tidak menggunakan database sebagai objek pengujian.
2. Penelitian akan difokuskan pada perancangan model deteksi SQL *Injection* Attack menggunakan metode *pre-processing* TF – IDF dan algoritma *Logistic regression*.
3. Aplikasi yang dirancang mencakup *login* dan *monitoring*.

1.5 Manfaat Penelitian

Berdasarkan tujuan penelitian yang telah dipaparkan sebelumnya, diharapkan penelitian ini dapat bermanfaat bagi perkembangan teknologi terutama di bidang *machine learning* dan *network security*. Berikut beberapa manfaat dari penelitian ini di antaranya:

1.5.1 Manfaat Teoritis

Diharapkan manfaat teoritis dalam penelitian ini yaitu:

1. Memberikan referensi yang baru bagi penelitian dengan tema dan topik yang sama dengan harapan dapat mengembangkan penelitian ini menjadi lebih baik dan lebih bermanfaat lagi, serta dapat dikaji lebih baik lagi untuk penelitian – penelitian ke depannya
2. Memberikan inspirasi dalam mengintegrasikan *network security* dan *machine learning* dalam menciptakan aplikasi pertahanan serangan siber sehingga menjadi ruang gerak baru bagi para peneliti selanjutnya untuk mengembangkan.

1.5.1 Manfaat Praktis

Adapun manfaat praktis dari penelitian ini yaitu:

1. Bagi pengembang aplikasi website, Aplikasi ini bisa menjadi dasar bagi para pengembang untuk mengintegrasikan secara langsung fitur deteksi yang didasarkan pada metode TF-IDF dan algoritma *machine learning* ke dalam aplikasinya.
2. Bagi peneliti, peneliti dapat mengimplementasikan pembelajaran dan pengalaman yang telah didapatkan selama masa perkuliahan berlangsung. Selain itu, Peneliti dapat mengembangkan jiwa kreatif dan inovatif peneliti dengan memadukan beberapa keilmuan secara sekaligus, yaitu *machine learning* dan network *security* yang tentunya akan bermanfaat bagi banyak pihak.

1.6 Struktur Organisasi Skripsi

Penelitian ini ditulis dengan pedoman yang sebagian besar mengacu pada Pedoman Penulisan Karya Ilmiah Universitas Pendidikan Indonesia Tahun 2024. Sistematika penulisan penelitian ini adalah sebagai berikut:

1. PENDAHULUAN

   Bagian Bab I ini memberikan penjelasan tentang latar belakang penelitian yang juga mencakup gap penelitian, rumusan masalah penelitian, tujuan penelitian, manfaat teoritis dan praktis, batasan masalah penelitian, hipotesis penelitian, dan struktur organisasi skripsi.

2. TINJAUAN PUSTAKA

   Tinjauan Pustaka membahas mengenai studi literatur terkait penelitian. Beberapa bagian dalam bab ini menjelaskan tentang konsep, metode, teori, dan teknologi yang digunakan.

3. METODE PENELITIAN

   Metode Penelitian membahas tentang prosedur penelitian. Ini mencakup jenis dan metode penelitian yang digunakan, perancangan desain website, perancangan sistem, perancangan algoritma, teknik pengujian dan rencana analisis.

4. HASIL DAN PEMBAHASAN

Pada Bab IV Hasil dan Pembahasan akan memaparkan hasil yang diperoleh dari penelitian yang telah dilakukan, berupa hasil perancangan aplikasi, hasil perancangan sistem deteksi.

5. SIMPULAN DAN SARAN

Bab terakhir, yaitu Simpulan dan Saran, yang merangkum temuan penelitian. Simpulan dari hasil penelitian disajikan, diikuti dengan pembahasan sistem pengamanan. Terakhir, penulis menyampaikan rekomendasi untuk pengembangan penelitian selanjutnya, memberikan arah bagi peneliti masa depan untuk mengembangkan konsep dan aplikasi lebih lanjut dalam bidang ini.