1. PENDAHULUAN

1.1. Latar Belakang

Dalam era digital yang semakin berkembang dan kompleks, keamanan data menjadi salah satu prioritas utama dalam pengembangan teknologi yang terus berevolusi. Terutama dalam konteks pengamanan terhadap serangan SQL Injection yang semakin canggih dan dapat menyusup kedalam sistem dengan berbagai cara[1]. Serangan SQL Injection merupakan serangan yang memanfaatkan celah yang ada dengan menginjeksi kode SQL berbahaya. Dampak yang terjadi jika seorang penyerang berhasil melakukan serangan SQL Injection adalah penyerang dapat dengan mudah mengetahui data-data sensitif yang terdapat dalam *database*, tidak hanya itu penyerang dapat mengubah serta menghapus data yang tersimpan di dalam *database* [2]. Ini berarti bahwa serangan SQL Injection tidak hanya dapat mengungkap informasi rahasia, tetapi juga dapat menganggu *integritas* dan ketersediaan data yang ada di dalam *database.* Serangan SQL Injection merupakan serangan yang memiliki peringkat tertinggi menurut Open Web Application Security Project[3].

Analisis deteksi serangan SQL Injection ini memerlukan perhatian yang serius karena insiden-insiden SQL Injection semakin sering terjadi. Konsep yang diterapkan dalam melakukan analisis deteksi serangan ini melibatkan penggunaan algoritma pembelajaran mesin sebagai pendekatan utama untuk mengidentifikasi serta mengantisipasi serangan SQL Injection yang mungkin terjadi. Oleh karena itu, melakukan analisis yang mendalam terhadap model pembelajaran mesin ini menjadi esensial guna men-deteksi serangan SQL Injection yang berpotensi merusak[4]. Dalam kasus ini memerlukan adanya sebuah algoritma pembelajaran mesin yang dapat digunakan untuk melakukan deteksi serangan SQL Injection.

Beberapa penelitian mengenai performa model machine learning telah dilakukan. Pada tahun 2022, Triloka dan rekan-rekannya melakukan penelitian terhadap dataset SQL Injection sebanyak 30.904 baris dengan menggunakan algoritma Support Vector Machine (SVM). Penelitian tersebut menghasilkan akurasi hingga 99.77% dibandingkan algoritma Logistic Regression (LR)

memiliki akurasi hingga 99.60%, Dan algoritma K-Nearest Neighbor (KNN) memiliki akurasi hingga 99.70%[4]. Melalui perbandingan ini dapat disimpulkan bahwa SVM memiliki keunggulan terkait tingkat akurasi dalam mendeteksi serangan SQL Injection[4].

Support Vector Machine (SVM) merupakan algoritma yang digunakan untuk menentukan batas keputusan. Batas keputusan menentukan klasifikasi dari algoritma ini. SVM menggunakan model linear sebagai batas keputusan. Pada penelitian ini algoritma SVM merupakan algoritma yang memiliki tingkat akurasi tertinggi terhadap deteksi serangan SQL Injection yaitu mencapai 99.77%[4]. K-Nearest Neighbor (KNN) merupakan algoritma yang menggunakan metode Euclidean Distance untuk melihat jarak prediksi terdekat pada label yang telah ditentukan. KNN bekerja dengan cara mencari k tetangga terdekat dari suatu data uji, kemudian melakukan voting untuk menentukan kelas dari data uji tersebut. Pada penelitian ini algoritma KNN memiliki tingkat akurasi terhadap deteksi serangan SQL Injection yaitu mencapai 99.70% [4]. Logistic Regression adalah algoritma yang digunakan dalam berbagai bidang, termasuk deteksi serangan SQL Injection. Algoritma ini bekerja dengan mengklasifikasikan variabel target diskrit sebagai fungsi dari beberapa variabel fitur. Dalam konteks deteksi serangan SQL Injection, Logistic Regression telah menunjukkan hasil yang sangat mengesankan, dengan tingkat akurasi mencapai 99.60%[4]. Ini berarti bahwa dalam hampir semua kasus, algoritma ini dapat dengan akurat mendeteksi apakah suatu operasi merupakan serangan SQL Injection atau bukan.

Harapannya, dimasa yang akan datang, model deteksi serangan *SQL Injection* ini akan secara luas diterapkan terhadap berbagai kebutuhan aplikasi, khususnya dalam aplikasi berbagai *website*. Ini akan membantu meningkatkan keamanan dan ketahanan terhadap serangan *SQL Injection* yang terus berkembang. Dengan menerapkan pendekatan berbasis pembelajaran mesin, sistem deteksi serangan SQL Injection ini dapat secara proaktif mengidentifikasi dan mengatasi upaya serangan sebelum mereka dapat menyebabkan kerusakan atau merusak integritas data. Hal ini akan memberikan perlindungan data yang lebih baik serta mencegah potensi kebocoran data yang merugikan. Dengan demikian, penerapan deteksi SQL Injection berbasis *machine learning* menjadi

langkah penting dalam mengamankan lingkungan *online* yang semakin kompleks dan rentan terhadap serangan *cyber*.

1.2. Perumusan Masalah

Penelitian ini akan membahas performansi model *machine learning* terkait deteksi serangan SQL Injection. Penelitian ini akan menjawab model *machine learning* mana yang memiliki tingkat akurasi tertinggi terhadap deteksi serangan SQL Injection, serta bagaimana langkah pencegahan terhadap serangan SQL Injection yang dapat dilakukan dari hasil deteksi serangan SQL Injection tersebut.

Adapun batasan masalah pada penelitian ini yaitu model *machine learning* yang digunaka adalah SVM, KNN, dan LR. Penelitian ini juga difokuskan untuk mencari tingkat akurasi tertinggi, *precision,* dan juga *recall/sensitivity.* dari ketiga model tersebut dengan dataset yang diambil dari *platform* Kaggle.

1.3. Tujuan

Tujuan utama dari penelitian ini adalah untuk memperoleh tingkat akurasi tertinggi dari ketiga model *machine learning*, yaitu SVM, KNN, dan LR, dalam melakukan deteksi serangan SQL Injection. Selain Tingkat akurasi, penelitian ini pun membandingkan metrik lain seperti *precision* dan juga *recall/sensitivity.* Penelitian ini juga bertujuan untuk mengimplementasikan strategi pencegahan yang dapat digunakan terhadap serangan SQL Injection.

1.4. Rencana Kegiatan

Pada penelitian ini kajian Pustaka diambil dari beberapa sumber yang dapat di akses melalui GoogleScholar, sumber yang didapatkan berasal dari IEEE Explore dan juga yang lainnya. Kajian Pustaka yang dipakai merupakan jurnal internasional yang memiliki tahun terbit mulai dari 2019 sampai dengan 2023. Pengumpulan *dataset* yang digunakan pada penelitian ini berasal dari *platform* Kagle. Dataset ini berisi kumpulan *script* yang dapat menjadi potensi terjadinya serangan SQL Injection.

Rancangan penelitian yang akan dilakukan pada beberapa tahap. Pertama akan mencari *dataset* SQL Injection yang berada pada *platform* kagle. Melakukan pembuatan model pembelajaran SVM, KNN, dan LR. Melakukan analisis terkait tingkat akurasi tertinggi dari ketiga model tersebut. Melakukan uji coba terhadap ketiga model tersebut untuk mencari tingkat akurasi tertinggi serta mengimplementasikan langkah pencegahan yang dapat dilakukan terhadap serangan SQL Injection menggunakan model *machine Learning*. Terakhir adalah tahap pembuatan kesimpulan beserta laporan terkait hasil Analisa yang didapatkan.

Teknik pengujian dalam penelitian ini melibatkan alokasi 70% dari *dataset* untuk proses pelatihan model dan 30% untuk evaluasi model pada data uji. Proses ini dilakukan dengan memisahkan *dataset* menjadi dua bagian, dimana 70% data digunakan untuk melatih model dan 30% sisanya digunakan untuk menguji performa model pada data yang tidak pernah dilihat sebelumnya.

Dalam proses penarikan kesimpulan dari ketiga model, yaitu *SVM*, *KNN*, dan *LR*, akan diambil kesimpulan terkait tingkat akurasi tertinggi, *precision*, dan *recall/sensitivity* yang dapat dicapai oleh masing-masing model. Evaluasi ini menjadi kunci dalam menilai kinerja relatif dari ketiga model tersebut. Dengan menganalisis hasil akurasi dari SVM, KNN, dan LR, dapat ditentukan model yang paling efektif dalam melakukan prediksi pada dataset yang telah diujikan.

1.5. Jadwal Kegiatan

Jadwal kegiatan pada penelitian ini adalah sebagai berikut :

Table 1 Rencana kegiatan

Kegiatan	**Bulan**					
Mencari *dataset* SQL Injection pada platform kagle untuk pembuatan model	1	2	3	4	5	6
Membuat model SVM, LR, dan KNN dari dataset yang didapatkan.						
Melakukan analisis terkait tingkat akurasi tertinggi dari ketiga model tersebut(SVM,KNN,LR).						
Melakukan uji coba terhadap ketiga model(SVM,KNN,LR) untuk mencari tingkat akurasi tertinggi, *precision,* dan *recall/sensitivity* serta mengimplementasikan langkah pencegahan yang dapat dilakukan terhadap serangan SQL Injection menggunakan model machine Learning						
Pembuatan kesimpulan serta laporan akhir						

*Keterangan: shading warna *grayscale*