Pengembangan dan Evaluasi Sistem Pendeteksian Serangan Siber Berbasis Autoencoder pada Intrusion Detection System (IDS) = Development and Evaluation of an Autoencoder-based Cyber Attack Detection System on an Intrusion Detection System (IDS)

Nabil Mafaza, author

Abstrak

Penggunaan internet telah mengubah hidup dan perilaku manusia. Internet yang awalnya hanya dimanfaatkan segilintir orang, berubah menjadi sebuah hal yang banyak orang manfaatkan. Perubahan perilaku manusia terlihat dalam cara manusia berkomunikasi, belajar, sampai menikmati konten hiburan. Namun, di balik manfaatnya, internet membawa bahaya yang merugikan banyak pihak. Bahaya tersebut timbul dalam bentuk serangan siber. Untuk mengatasi serangan siber, banyak perangkat keras dan lunak yang digunakan, salah satunya adalah intrusion detection system (IDS). Akan tetapi, IDS tidak dapat mendeteksi serangan baru akibat sifat pendeteksiannya yang rule-based. Penelitian ini bertujuan untuk menambah kemampuan IDS dalam mendeteksi serangan siber dengan menggunakan model machine learning (ML), khususnya autoencoder, untuk mendeteksi serangan siber dalam lalu lintas jaringan. Autoencoder digunakan untuk meng-encode lalu lintas jaringan, kemudian men-decode/merekonstruksi hasil encode. Lalu lintas jaringan akan dideteksi sebagai serangan siber apabila perbedaan hasil rekonstruksi dengan lalu lintas jaringan asli melebihi ambang tertentu. Berdasarkan testing yang dilakukan, model autoencoder paling optimal adalah model yang di-train dengan dataset yang dipisah menjadi dense dan sparse berdasarkan nilai quantile 70% fitur tot_l_fwd_pkt dan tot_l_bwd-_pkt, dilakukan feature selection menggunakan random forest dengan nilai importance 0,2, menggunakan activation function ReLU, dan menggunakan empat layer encoder dan decoder serta jumlah neuron 16, 8, 4, 2, 1, 2, 4, dan 16. Model autoencoder untuk dataset dense terbaik memiliki F1-score 84% (lalu lintas benign) dan 83% (lalu lintas malicious), trainable parameter berjumlah 830, dan ukuran model sebesar 71 KB. Sementara, model autoencoder untuk dataset sparse terbaik memiliki F1-score 71% untuk lalu lintas benign dan malicious, trainable parameter berjumlah 890, dan ukuran model sebesar 72 KB.

.....The use of the internet has transformed human lives and behavior. Initially utilized by a few, the internet has become an essential tool for many. This transformation is evident in how people communicate, learn, and enjoy entertainment content. However, alongside its benefits, the internet also poses significant risks in the form of cyber attacks. To combat these threats, various hardware and software solutions, including intrusion detection systems (IDS), are employed. Traditional IDS, however, struggle to detect new attacks due to their rule-based nature. This research aims to enhance IDS capabilities in detecting cyber attacks by using machine learning (ML) models, specifically autoencoders, to detect cyber attacks in network traffic. Autoencoders encode network traffic and then decode/reconstruct the encoded data. Network traffic is identified as a cyber attack if the reconstruction error exceeds a certain threshold. Based on the testing conducted, the most optimal autoencoder model was trained on a dataset split into dense and sparse categories based on the 70% quantile values of the tot_l_fwd_pkt and tot_l_bwd_pkt features. Feature selection was performed using random forest with an importance threshold of 0.2, employing the ReLU activation function, and using four encoder and decoder layers with neuron counts of 16, 8, 4, 2, 1, 2, 4, and

16. The best autoencoder model for dense dataset achieved an F1-score of 84% for benign traffic and 83% for malicious traffic, with 830 trainable parameters and a model size of 71 KB. Meanwhile, the best autoencoder model for sparse dataset achieved an F1-score of 71% for both benign and malicious traffic, with 890 trainable parameters and a model size of 72 KB.