

PENGEMBANGAN MODEL *MACHINE LEARNING* UNTUK DETEKSI SERANGAN SIBER

ABSTRAK

Abstract: As cyber attacks become more complex, conventional detection systems such as signature-based Intrusion Detection Systems (IDS) are starting to experience limitations in accurately recognizing new threats. This study aims to develop a deep learning-based machine learning model to detect cyber attacks by combining the Convolutional Neural Network (CNN) architecture as feature extraction and Random Forest (RF) as classification. The dataset used in this study is CICIDS2018, which has several types of network attacks. The study was conducted in two Incremental stages, namely model development and model implementation into a web application that can detect cyber attacks using input from .pcap files and output in .csv format. The evaluation results showed that the combination of the CNN model as feature extraction and RF as classification obtained an accuracy of 91%, making this model combination better than the combination of the CNN-XGBoost model which only obtained an accuracy of 89%. This model also proved to be better than the single CNN model which only obtained an accuracy of 87% and better than the previous research architecture which obtained an accuracy of 90%. The results of all tested models can be concluded that using a combination of CNN-RF models produces better accuracy results.

Keywords: Machine Learning, Convolutional Neural Network, Random Forest, CICIDS2018

Abstrak: Semakin kompleksnya serangan siber, sistem deteksi konvensional seperti Intrusion Detection System (IDS) berbasis signature mulai mengalami keterbatasan dalam mengenali ancaman baru secara akurat. Penelitian ini bertujuan untuk mengembangkan model machine learning berbasis deep learning untuk mendeteksi serangan siber dengan menggabungkan arsitektur Convolutional Neural Network (CNN) sebagai ekstraksi fitur dan Random Forest (RF) sebagai klasifikasi. Dataset yang digunakan pada penelitian ini adalah CICIDS2018, yang mempunyai beberapa jenis serangan jaringan. Penelitian dilakukan dengan dua tahap Inkremental, yaitu pengembangan model dan implementasi model ke dalam aplikasi web yang dapat mendeteksi serangan siber menggunakan input dari file .pcap dan mengeluarkan output dalam format .csv. Hasil evaluasi menunjukkan bahwa penggabungan model CNN sebagai ekstraksi fitur dan RF sebagai klasifikasi mendapatkan akurasi sebesar 91% yang membuat penggabungan model ini semakin baik dibandingkan dengan penggabungan model CNN- XGBoost yang hanya mendapatkan akurasi sebesar 89%. Model ini juga terbukti lebih baik dibandingkan dengan model CNN tunggal yang hanya mendapatkan akurasi sebesar 87% serta lebih baik dari arsitektur penelitian terdahulu yang mendapatkan akurasi sebesar 90%. Hasil dari semua model yang diuji coba dapat disimpulkan bahwa menggunakan penggabungan antara model CNN-RF mendapatkan hasil akurasi yang lebih baik.

Kata kunci: Machine Learning, Convolutional Neural Network, Random Forest, CICIDS2018