Pengembangan Model-Based Feature Selection untuk Peningkatan Akurasi Sistem Deteksi dan Peringatan Serangan berbasis Machine Learning pada Jaringan Internet = The Development of Model-Based Feature Selection to Improve System's Detection Accuracy and Machine Learning-based Attack Warning System

Yuri Prihantono, author

Abstrak

Pemanfaatan Intrusion Detection System (IDS) untuk mengamankan infrastruktur jaringan internet masih memiliki masalah yang belum terselesaikan, yaitu kurangnya akurasi deteksi serangan sehingga mengakibatkan terjadinya permasalahan false positif dan banyaknya alarm palsu. Salah satu pendekatan yang banyak digunakan untuk mengatasi permasalahan yang terjadi dalam implementasi IDS adalah dengan menggunakan pendekatan machine learning. Pada penelitian ini, penulis mengusulkan sistem yang menggunakan pendekatan machine learning untuk mendeteksi serangan jaringan dan mengirim peringatan serangan. Dataset CSE-CICIDS2018 dan Model-Based Feature Selection digunakan untuk mengevaluasi kinerja delapan algoritma klasifikasi dalam mengidentifikasi serangan jaringan guna menentukan algoritma terbaik. Hasilnya, Model XGBoost dipilih sebagai model yang memberikan hasil kinerja algoritma terbaik dalam perbandingan model machine learning ini, dengan tingkat akurasi untuk klasifikasi two-class sebesar 99%, dan multi-class sebesar 98,4%. <hr /> Utilization of the Intrusion Detection System (IDS) to secure internet network infrastructure still has unresolved problems, namely the lack of attack detection accuracy, resulting in false positives and many false alarms. One approach that is widely used to overcome the problems that occur in the implementation of IDS is to use a machine learning approach. In this study, the authors propose a system that uses a machine learning approach to detect network attacks and send attack warnings. The CSE-CICIDS2018 dataset and Model-Based Feature Selection were used to evaluate the performance of eight classifier algorithms in identifying network attacks to determine the best algorithm. As a result, the XGBoost model was chosen as the model that gives the best algorithm performance results in this machine learning model comparison, with an accuracy rate of 99% for two-class classification and 98.4% for multi-class.