ABSTRACT

Machine learning techniques are widely used to develop Intrusion Detection Systems (IDS) to detect and classify cyber attacks at the network level and the host level in a timely and automated manner. However, many challenges arise as malicious attacks are constantly changing and occurring in very large volumes requiring a scalable solution. Therefore, this study conducted a comparison of the K-Nearest Neighbor Knn and Naive Bayes algorithms. The dataset used in this study is the Ddos features-IDS 2017 dataset published in 2019. This research analyzes the comparison of methods generated from the classification process based on metric accuracy, specificity and sensitivity parameters. The classification process using the K-Nearest Neighbor (KNN) and Naive Bayes algorithms, it can be concluded that the results of the three tests with a percentage split of 60%, 70% and 80% show that the K-Nearest Neighbor (KNN) algorithm gets a higher value than Naive Bayes except the error rate because the error rate indicates that the data failed to be classified properly. Testing on a percentage split of 60% KNN parameter accuracy gets a value of 99.53%, specificity 94.05%, sensitivity 75.20%, testing on a percentage split 70% KNN parameter accuracy gets a value of 99.69%, specificity 94.59%, sensitivity 78.40% and testing on percetage split 80%, KNN parameter accuracy parameter got a value of 99.70%, specificity 94.44%, sensitivity 75.85%.

Keywords : IDS, K-nearest Neighbour, Naive Bayes

ABSTRAK

Teknik *machine learning* banyak digunakan untuk mengembangkan *Intrusion Detection System* (IDS) untuk mendeteksi dan mengklasifikasikan serangan dunia maya di tingkat jaringan dan tingkat *host* secara tepat waktu dan cara otomatis. Namun, banyak tantangan muncul karena serangan jahat terus berubah dan terjadi dalam volume yang sangat besar yang membutuhkan solusi yang dapat diskalakan. Oleh karena itu penelitian ini melakukan perbandingan algortima *K-Nearest Neighbour* knn dan naive bayes. *Dataset* yang digunakan dalam penelitian ini adalah dataset *Ddos features*-IDS 2017 yang diterbitkan pada tahun 2019. Peneitian ini menganalisis perbndingan metode yang dihasilkan dari proses klasifikasi berdasarkan parameter metric accuracy, specifity dan sensitivity. Proses klasifikasi dengan menggunakan algoritma *K-Nearest Neighbour* (KNN) dan Naive Bayes, maka dapat disimpulkan hasil dari ketiga pengujian dengan *percentage split* 60%, 70% dan 80% menunjukkan bahwa algoritma *K-Nearest Neighbour* (KNN) mendapatkan nilai yang lebih tinggi dari Naive Bayes kecuali *error rate* karena *error rate* menunjukan bahwa data gagal diklasifikasi dengan baik. Pengujian pada percetage split 60% KNN parameter *accuracy* mendapatkan nilai 99,53%, *specificity* 94,05%, *sensitivity* 75,20%, pengujian pada percentage split 70% KNN parameter *accuracy* mendapatkan nilai 99,69%, *specificity* 94,59%, *sensitivity* 78,40% dan pengujian pada percetage split 80% parameter KNN parameter *accuracy* mendapatkan nilai 99,70%, *specificity* 94,44%, *sensitivity* 75,85%.

Kata kunci : IDS, *K-nearest Neighbour, Naive Bayes*