INTISARI

Penelitian ini mengembangkan model *hybrid* yang menggabungkan *random forest* dan CNN dengan teknik *feature selection* untuk meningkatkan akurasi dan efisiensi deteksi serangan jaringan pada *dataset* UNSW-NB15. Ancaman keamanan siber terus berkembang dengan berbagai jenis serangan seperti DoS, DDoS, *exploits*, *backdoors*, dan *worms* yang menimbulkan kerugian global mencapai USD 10,5 triliun per tahun. Penelitian ini membatasi fokus pada klasifikasi biner antara *traffic* normal dan serangan menggunakan metrik evaluasi akurasi, presisi, *recall*, dan *f1-score*. Metode yang diterapkan meliputi tahap *preprocessing data*, *feature selection* berbasis *random forest* untuk mereduksi 47 fitur menjadi 24 fitur terpenting, pembagian data dengan rasio 80:20, pembangunan model *hybrid* CNN dan RF, dan evaluasi menggunakan *confusion matrix*.

Hasil penelitian menunjukkan bahwa *feature selection* berhasil mengidentifikasi fitur penting seperti ct_state_ttl, sttl, dan srcip sebagai indikator utama serangan. Model *hybrid* mencapai performa terbaik dengan akurasi 99,53%, presisi 98,36%, *recall* 97,92%, dan *f1-score* 98,14%. Performa ini setara dengan *random forest* tunggal dengan akurasi 99,53%, namun jauh melampaui CNN tunggal yang hanya mencapai akurasi 98,47% dengan *recall* 89,20%. Integrasi CNN sebagai *feature extractor hierarkis* dan *random forest* sebagai *classifier* akhir terbukti efektif memanfaatkan kekuatan kedua algoritma.

Kesimpulan penelitian ini membuktikan bahwa pendekatan *hybrid* CNN dan *random forest* dengan *feature selection* merupakan solusi yang efisien untuk sistem deteksi intrusi jaringan. Model berhasil mengintegrasikan kemampuan *deep learning* dalam menangkap pola kompleks dengan stabilitas *machine learning* dalam klasifikasi, menghasilkan sistem deteksi yang *robust* dan dapat diandalkan untuk melindungi infrastruktur digital dari ancaman siber.

Kata kunci: deteksi serangan jaringan, CNN, *random forest*, *feature selection*, UNSW-NB15

ABSTRACT

This study developed a hybrid model combining random forest and CNN with feature selection techniques to improve the accuracy and efficiency of network attack detection on the UNSW-NB15 dataset. Cybersecurity threats continue to evolve, with various types of attacks such as DoS, DDoS, exploits, backdoors, and worms causing global losses reaching USD 10.5 trillion annually. This study focused on the binary classification of normal traffic and attacks using evaluation metrics such as accuracy, precision, recall, and f1-score. The applied methods included data preprocessing, random forest-based feature selection to reduce 47 features to 24 most important ones, data splitting with an 80:20 ratio, building a hybrid CNN and RF model, and evaluation using a confusion matrix.

The results showed that feature selection successfully identified important features such as ct_state_ttl, sttl, and srcip as key indicators of attacks. The hybrid model achieved the best performance with 99.53% accuracy, 98.36% precision, 97.92% recall, and 98.14% f1-score. This performance is comparable to a single random forest with 99.53% accuracy, but far surpasses a single CNN, which only achieved 98.47% accuracy with 89.20% recall. The integration of a CNN as a hierarchical feature extractor and a random forest as the final classifier proved effective, leveraging the strengths of both algorithms.

The conclusion of this study demonstrates that the hybrid CNN and random forest approach with feature selection is an efficient solution for network intrusion detection systems. The model successfully integrates the capabilities of deep learning in capturing complex patterns with the stability of machine learning in classification, resulting in a robust and reliable detection system for protecting digital infrastructure from cyber threats.

Keywords: network attack detection, CNN, random forest, feature selection, UNSW-NB15