**IDENTIFIKASI DAN KLASIFIKASI SERANGAN SIBER PADA ELDIRU UNSOED DENGAN ALGORITMA RANDOM FOREST**

**JUSTICIO CAESARIO**
**H1D021055**

## ABSTRAK

Sistem informasi akademik seperti Eldiru Unsoed merupakan aset digital vital yang rentan terhadap serangan siber, sementara *Web Application Firewall* konvensional berbasis aturan menunjukkan kelemahan deteksi. Sistem standar ModSecurity dengan *Core Rule Set* (CRS) terbukti hanya memiliki *recall* 5.34%, yang berarti gagal mengidentifikasi mayoritas serangan aktual dan menciptakan celah keamanan. Penelitian ini merancang sistem deteksi berbasis algoritma *Random Forest* untuk mengatasi masalah tersebut. Data pelatihan utama berasal dari log server Nginx milik sistem Eldiru Unsoed periode Desember 2024 hingga Januari 2025, yang kemudian divalidasi performanya menggunakan dataset publik CSIC 2010. Berdasarkan analisis log tersebut, model dikembangkan dengan merekayasa fitur hibrida yang mencakup analisis leksikal, konteks aturan CRS, dan N-gram untuk mengklasifikasikan lalu lintas web. Hasil evaluasi menunjukkan model *Machine Learning-Random Forest* (ML-RF) yang diusulkan meningkatkan *recall* dari 5.34% menjadi 72.00% dan F1-Score dari 10.10% menjadi 80.00%. Peningkatan metrik ini, dengan tetap menjaga presisi pada 91.00%, menunjukkan bahwa integrasi *machine learning* menghasilkan mekanisme pertahanan siber yang lebih seimbang dan andal untuk menjawab tantangan deteksi ancaman modern dalam melindungi aset digital.

**Kata Kunci:** *Aplikasi Web, Keamanan Siber, Machine Learning, Random Forest, Web Application Firewall*

***IDENTIFICATION AND CLASSIFICATION OF CYBER ATTACKS ON ELDIRU UNSOED USING THE RANDOM FOREST ALGORITHM***

**JUSTICIO CAESARIO**
**H1D021055**

## *ABSTRACT*

*Academic information systems such as Eldiru Unsoed are vital digital assets vulnerable to cyberattacks, while conventional rule-based Web Application Firewalls exhibit detection weaknesses. The standard ModSecurity with Core Rule Set (CRS) system is proven to have a recall of only 5.34%, meaning it fails to identify the majority of actual attacks and creates a security gap. This research designs a detection system based on the Random Forest algorithm to address this problem. The primary training data originates from the Nginx server logs of the Eldiru Unsoed system from the period of December 2024 to January 2025, with its performance subsequently validated using the public CSIC 2010 dataset. Through analysis of these logs, the model was developed by engineering hybrid features that include lexical analysis, CRS rule context, and N-grams to classify web traffic. Evaluation results show the proposed Machine Learning-Random Forest (ML-RF) model increases recall from 5.34% to 72.00% and the F1-Score from 10.10% to 80.00%. This improvement in metrics, while maintaining a precision of 91.00%, indicates that machine learning integration yields a more balanced and reliable cybersecurity defense mechanism to meet modern threat detection challenges in protecting digital assets.*

***Keywords:*** *Cyber Security, Random Forest, Web Application, Web Application Firewall*