# KLASIFIKASI UNIFORM RESOURCE LOCATOR BERBAHAYA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)

Ahmad Faisal Fahmi[1], Ginanjar Abdurrahman[2], Ari Eko Wardoyo[3]

Program Studi Teknik Informatika, Fakultas Teknik

Universitas Muhammadiyah Jember

Email : ahmadfaisal00002@gmail.com

## ABSTRAK

Deteksi dini URL berbahaya merupakan langkah penting dalam meningkatkan keamanan siber, terutama dalam menghadapi ancaman seperti *phishing*, *malware*, dan *defacement*. Penelitian ini mengusulkan penerapan algoritma *Support Vector Machine* (SVM) untuk mengklasifikasikan URL berdasarkan tingkat keamanannya. Dataset yang digunakan terdiri dari 1.000 URL dari berbagai kategori, yaitu *benign*, *phishing*, *defacement*, dan *malware*. Serangkaian preprocessing dilakukan, termasuk cleaning, case folding, tokenizing, *filtering*, *stemming*, dan pembobotan TF-IDF. Untuk menangani ketidakseimbangan kelas, digunakan teknik *oversampling SMOTE*. Model divalidasi menggunakan *K-Fold Cross Validation* dengan nilai K=7. Hasil penelitian menunjukkan bahwa SVM menghasilkan rata-rata akurasi 92%, presisi 91%, dan *recall* 92% dari 200 data uji. Nilai akurasi menunjukkan performa model yang baik dalam membedakan URL aman dan berbahaya, meskipun presisi dan *recall* relatif rendah pada beberapa kelas minoritas. Penelitian ini memberikan gambaran bahwa SVM dapat digunakan sebagai pendekatan awal dalam membangun sistem deteksi URL berbahaya.

**Kata Kunci : URL berbahaya, SVM, TF-IDF, SMOTE, Klasifikasi, *Machine Learning***

# *CLASSIFICATION OF DANGEROUS UNIFORM RESOURCE LOCATORS USING THE SUPPORT VECTOR MACHINE (SVM) ALGORITHM*

Ahmad Faisal Fahmi[1], Ginanjar Abdurrahman[2], Ari Eko Wardoyo[3]

Program Studi Teknik Informatika, Fakultas Teknik

Universitas Muhammadiyah Jember

Email : ahmadfaisal00002@gmail.com

## *ABSTRACT*

*Early detection of malicious URLs is an important step in improving cybersecurity, especially in the face of threats such as phishing, malware and defacement. This study proposes the application of the Support Vector Machine (SVM) algorithm to classify URLs based on their security level. The dataset used consists of 1,000 URLs from various categories, namely benign, phishing, defacement, and malware. A series of preprocessing steps were performed, including cleaning, case folding, tokenizing, filtering, stemming, and TF–IDF weighting. To address class imbalance, the SMOTE oversampling technique was used. The model was validated using K-Fold Cross Validation, with a value of K=7. The results showed that SVM produced an average accuracy of 92%, precision of 91%, and recall of 92% from 200 train data. The accuracy value indicates that the model performs well in distinguishing between safe and dangerous URLs, although the precision and recall are relatively low in some minority classes. This study shows that SVM can be used as an initial approach in building a dangerous URL detection system.*

***Keywords: Malicious URLs, SVM, TF-IDF, SMOTE Classification Machine Learning***