

ABSTRAK

Smart contract pada platform Ethereum rentan terhadap berbagai jenis kerentanan yang dapat menimbulkan kerugian finansial signifikan. Tantangan utamanya adalah satu kontrak dapat mengandung lebih dari satu kerentanan secara bersamaan, hasil deteksi antar alat analisis bersifat heterogen, serta pendekatan yang memanfaatkan metadata hasil konsensus *multi-tool* sebagai fitur untuk klasifikasi *multi-label* masih relatif terbatas. Berdasarkan permasalahan tersebut, penelitian ini mengusulkan klasifikasi *multi-label* kerentanan *smart contract* berbasis metadata hasil konsensus alat analisis menggunakan Regresi Logistik dan Linear-SVM.

Penelitian ini menggunakan dataset ScrawlD yang memuat *smart contract* Ethereum dunia nyata dengan label kerentanan yang dihasilkan melalui majority voting dari lima alat analisis, yaitu Slither, Mythril, SmartCheck, Oyente, dan Osiris. Tahap *preprocessing* mencakup *data cleaning*, *label mapping*, *metadata parsing*, serta pembagian data dengan mekanisme anti-leakage, menghasilkan 22 fitur biner dari pasangan alat-kerentanan yang didukung. Model dibangun dengan skema *One-vs-rest* pada tujuh label efektif (ARTHM, LE, RENT, TimeM, TimeO, TX-Origin, dan UE), dengan interpretasi model menggunakan SHAP LinearExplainer. Antarmuka pengguna berbasis web juga dibangun menggunakan framework Streamlit.

Hasil evaluasi pada 1.133 data uji menunjukkan Regresi Logistik memperoleh *Micro-f1* sebesar 0,8689, *Macro-f1* sebesar 0,6828, dan *Hamming loss* sebesar 0,0523, sedangkan Linear-SVM memperoleh *Micro-f1* sebesar 0,8676, *Macro-f1* sebesar 0,6763, dan *Hamming loss* sebesar 0,0528. Kedua model menunjukkan performa tinggi pada label LE dan TimeO (*F1-score* 1,00) serta ARTHM dan RENT, namun menurun pada label minoritas akibat ketidakseimbangan distribusi label. Analisis SHAP mengungkap bahwa fitur dominan pada sebagian besar label berkorespondensi langsung dengan kerentanan targetnya, seperti *osiris_ARTHM*, *slither_LE*, dan *oyente_TimeO*, sementara pada label minoritas ekstrem model memanfaatkan sinyal lintas label sebagai *proxy*. Penelitian ini berkontribusi sebagai baseline klasifikasi *multi-label* yang interpretatif dan mudah direplikasi, sekaligus menunjukkan bahwa metadata konsensus *multi-tool* mengandung sinyal yang cukup untuk mendukung prediksi dan pemahaman pola kerentanan *smart contract*.

Kata Kunci: *smart contract*, klasifikasi *multi-label*, metadata konsensus, Regresi Logistik, Linear-SVM

ABSTRACT

Smart contracts on the Ethereum platform are susceptible to security vulnerabilities that can cause significant financial losses. The main challenges include a single contract potentially containing multiple simultaneous vulnerabilities, heterogeneous detection results across analysis tools, and limited research that explicitly uses multi-tool consensus metadata as structured features for multi-label classification. This study proposes multi-label classification of smart contract vulnerabilities based on analysis tool consensus metadata using Logistic Regression and Linear-SVM.

This study uses the ScrawlD dataset containing real-world Ethereum smart contracts labeled through majority voting from five tools: Slither, Mythril, SmartCheck, Oyente, and Osiris. Preprocessing includes data cleaning, label mapping, metadata parsing, and anti-leakage data splitting, producing 22 binary features. Models were built using a One-vs-rest scheme across seven effective labels (ARTHM, LE, RENT, TimeM, TimeO, TX-Origin, and UE), with SHAP LinearExplainer for model interpretation and a Streamlit-based web interface for practical use.

Evaluation on 1,133 test samples shows Logistic Regression achieved Micro-f1 of 0.8689, Macro-f1 of 0.6828, and Hamming loss of 0.0523, while Linear-SVM achieved 0.8676, 0.6763, and 0.0528 respectively. Both models performed well on LE and TimeO (F1-score 1.00) and on ARTHM and RENT, but declined on minority labels due to class imbalance. SHAP analysis reveals that dominant features for most labels correspond directly to their target vulnerability, such as osiris_ARTHM, slither_LE, and oyente_TimeO, while extreme minority labels rely on cross-label signals. This study contributes an interpretable and replicable multi-label classification baseline, demonstrating that multi-tool consensus metadata contains sufficient signals for smart contract vulnerability prediction and pattern understanding.

Keywords: *smart contract, multi-label classification, consensus metadata, Logistic Regression, Linear-SVM*