Real-Time Detection and Explainable Analysis of Unsafe URLs Using Catboost and Random Forest Ensembles

Authors Uddipan Ghosh, Pratik Saha, Siddhartha Roy, Sangita Dutta, Arijeet Ghosh, Ankan Sarkar

Publication date 2026/3/18

Conference 9th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)

Description While web-based services have expanded at break-neck speed, attacks on the cyber space, most importantly malicious URLs, have also evolved. Traditional detection mechanisms that rely on blacklists or signatures have a tendency to miss newly created or highly obfuscated phishing URLs. To overcome these vulnerabilities, this work presents an explainable hybrid ensemble model that combines the strengths of CatBoost and Random Forest classifiers to accurately and explainably detect malicious URLs. A varied collection of manually crafted features-e.g., suspicious words, character patterns, domain pattern, and uncommon TLDs-were derived from a synthetically balanced dataset meant to simulate malicious and benign URL patterns in the real world. In pursuit of more transparency and user trust-worthiness, we blended two AI explanation approaches: SHAP (SHapley Additive exPlanations) is used for global ...