IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, November 2025
Impact Factor: 7.67

PhishAI-Sim: A Hybrid Feature-Free and Deep Learning Framework for Adaptive Phishing Website Detection

Dnyaneshwar Jagannath Tribhuvan[1], Prof. V. S. Chaudhari[2], Mr. Gokul Pawade[3]
[1,2,3]Department of Computer Engineering
[1,2]Vishwabharati Academy's College of Engineering, Sarola Baddi, Ahmednagar
[3]Pravara Rural College of Engineering, Loni, Ahilyanagar (MS) India

Abstract: *Phishing websites continue to evolve rapidly, employing advanced obfuscation and design replication techniques that challenge conventional feature-based detection systems. This paper presents PhishAI-Sim, a hybrid feature-free and deep learning framework designed to enhance phishing website detection accuracy and adaptability. The proposed system integrates the Normalized Compression Distance (NCD)—a universal, parameter-free similarity metric—with deep neural embeddings to measure structural and contextual similarities between webpages without manual feature extraction. A dynamic incremental learning mechanism enables continuous model adaptation to newly emerging phishing patterns, reducing concept drift over time. Experimental results on a large-scale dataset demonstrate that PhishAI-Sim achieves superior detection performance with a true positive rate exceeding 92% and a false positive rate below 1%, outperforming traditional feature-engineered and standalone deep learning models. This study highlights the potential of combining compression-based similarity and intelligent learning for a robust, scalable, and future-ready phishing detection system.*

Keywords: Phishing Detection, Feature-Free Approach, Deep Learning, Normalized Compression Distance (NCD), Incremental Learning, Cybersecurity, Webpage Similarity, Artificial Intelligence

DOI: 10.48175/568