

Machine Learning-Powered Protection Against Phishing Crimes

Vishal Borate¹, Dr. Alpana Adsul², Rohit Dhakane³, Shahuraj Gawade⁴,
Shubhangi Ghodake⁵, Pranit Jadhav⁶

Assistant Professor, Department of Computer Engineering¹

Associate Professor, Department of Computer Engineering²

Students, Department of Computer Engineering^{3,4,5,6}

Dr. D. Y. Patil College of Engineering & Innovation, Talegaon, Pune, India

Abstract: Phishing attacks are a serious cybersecurity risk that uses phony emails, websites, or messages to target users and organizations in an attempt to steal confidential data. By examining data patterns and spotting questionable activity, machine learning (ML) offers creative ways to identify and stop these attacks. This study examines several machine learning (ML) techniques for phishing detection, including Principal Component Analysis (PCA), Random Forest (RF), and Decision Trees (DT). According to studies, RF models are very effective and can attain up to 97% accuracy. But there are still issues with feature extraction, data imbalance, and changing phishing strategies. Phishing detection capabilities can be further improved by incorporating real-time detection systems, hybrid approaches, and sophisticated deep learning models. Additionally, to increase detection accuracy and reduce detection errors, cybersecurity researchers and organizations must collaborate and update datasets continuously.

Keywords: Phishing attack, machine learning, Random Forest, decision tree, Principal Component Analysis, Cyber-security, deep learning

