PROJECT REPORT

COMPARING MACHINE LEARNING ALGORITHM TO BE USED IN SIGNATURE-BASED MALWARE DETECTION

VIANDRA GEOVAN

21.K1.0047

Faculty of Computer Science

Soegijapranata Catholic University

2025

ABSTRACT

In today’s digital era, computer security is one of the most significant worries due to technology and the continual threat from malware. However, traditional signature-based malware detection tends to be good against known threats but it fails in cases related to newly created or unknown malwares. The main objective of this study involves comparing machine learning algorithms to see if it can be used to enhance signature-based malware detection to make it more efficient and robust. This research will therefore aim at providing answers to questions on traditional detection methods and the potential for enhancement by ML as well as determining which ML models are best used for malware detection. The study is focused on developing a method that can combine both methods, by considering common malwares and determining the efficiency of various machine learning techniques such as support vector machines, decision trees and convolutional neural networks. This research methodology involves gathering data from credible sources from the internet followed by data preprocessing steps such as cleaning, normalization, and data splitting. Models are trained and evaluated using metrics like accuracy, precision, recall and F1-score. Legal ethical economic aspects were not included in this study thus only technological advancements were focused. The study concludes that using ML with signature-based detection significantly improves the system's robustness against evolving malware threats. The findings suggest that ML can predict unknown malware patterns, enhancing traditional detection capabilities and providing a more adaptive solution to cybersecurity challenges. This research offers a detailed methodology for replication and aims to contribute to the development of more effective malware detection systems.

Keyword: Malware, Signature-based malware detection, cybersecurity, machine learning