A Machine Learning Approach to Detecting Evasive PDF Malware

Authors : Vareesha Masood, Ammara Gul, Nabeeha Areej, Muhammad Asif Masood, Hamna Imran

Abstract : The universal use of PDF files has prompted hackers to use them for malicious intent by hiding malicious codes in their victim's PDF machines. Machine learning has proven to be the most efficient in identifying benign files and detecting files with PDF malware. This paper has proposed an approach using a decision tree classifier with parameters. A modern, inclusive dataset CIC-Evasive-PDFMal2022, produced by Lockheed Martin's Cyber Security wing is used. It is one of the most reliable datasets to use in this field. We designed a PDF malware detection system that achieved 99.2%. Comparing the suggested model to other cutting-edge models in the same study field, it has a great performance in detecting PDF malware. Accordingly, we provide the fastest, most reliable, and most efficient PDF Malware detection approach in this paper.

Keywords : PDF, PDF malware, decision tree classifier, random forest classifier
Conference Title : ICCSIS 2024 : International Conference on Computer Science and Intelligent Systems
Conference Location : Montreal, Canada
Conference Dates : May 23-24, 2024