Malware Analysis and Detection Using Machine Learning Algorithm

Abstract

One of the most significant issues facing internet users nowadays is malware. Polymorphic malware is a new type of malicious software that is more adaptable than previous generations of viruses. Polymorphic malware constantly modifies its signature traits to avoid being identified by traditional signature-based malware detection models. To identify malicious threats or malware, we used a number of machine learning techniques. A high detection ratio indicated that the algorithm with the best accuracy was selected for usage in the system. As an advantage, the confusion matrix measured the number of false positives and false negatives, which provided additional information regarding how well the system worked. In particular, it was demonstrated that detecting harmful traffic on computer systems, and thereby improving the security of computer networks, was possible using the findings of malware analysis and detection with machine learning algorithms to compute the difference in correlation symmetry (Naive Byes, SVM, J48, RF, and with the proposed approach) integrals. The results showed that when compared with other classifiers, DT (99%), CNN (98.76%), and SVM (96.41%) performed well in terms of detection accuracy. DT, CNN, and SVM algorithms' performances detecting malware on a small FPR (DT = 2.01%, CNN = 3.97%, and SVM = 4.63%,) in a given dataset were compared. These results are significant, as malicious software is becoming increasingly common and complex.