Keylogger Detection Analysis Using Machine Learning Algorithm

Muhammad Faiz Hazim Abdul Rahman[1], Abidah Hj Mat Taib[2], Nor Alifah Rosaidi[3]
[1]*faizhazim.rahman@gmail.com*
[2]*abidah@uitm.edu.my*
[3]*alifah.rosaidi@uitm.edu.my*

ABSTRACT

Malware is one of the most harmful forms of attack on computers because of its passive approach and hidden execution. The most widespread type of malicious software that discreetly monitors user activities and logs keystrokes is called keylogging malware. Accordingly, the goal of this study are to create a detection model based on both supervised machine learning on keylogger dataset. Plus, to analyse the efficiency of a detection model on keylogger dataset by evaluating a selection of attributes. Besides, to test the accuracy of detection models on keylogger dataset comparing two machine learning algorithms. This study is carried out through the utilisation of two machine learning techniques, namely Decision Tree and Naive Bayes, on Jupyter Notebook in order to conduct an analysis of the Keylogger Detection dataset obtained from a trustworthy website known as Kaggle. There are a few outcomes that have been achieved to decide between those two machine learning methods that have better accuracy to carry out analysis on the dataset which of the two, but rather Decision Tree, have the greater accuracy. Early identification of a keylogger malware attack could prevent hackers from accessing personal user data and reduce the likelihood of infiltration, which could reveal account information, credit cards, usernames, passwords, and other data. In this way, we can decrease the likelihood of being the victim of a spyware attack and losing our information. It is intendedthat this initiative would deliver benefits to all of the users and be useful to them.

Keywords: Decision Tree, Naïve Bayes, Jupyter notebook