UNIVERSITI TEKNOLOGI MARA

COMPARISON OF DETECTION MODEL USING MACHINE LEARNING ON ANDROID MALWARE

MUHAMMAD IKMAL BIN IHSAN

BACHELOR OF COMPUTER SCIENCE (HONS.)
DATA COMMUNICATION AND NETWORKING

JULY 2022

Universiti Teknologi MARA

Comparison of Detection Model using Machine Learning on Android Malware

Muhammad Ikmal bin Ihsan

Thesis submitted in fulfilment of the requirement for Bachelor of Computer Science (Hons.) Data Communication and Networking
Faculty of Computer and Mathematical Sciences

July 2022

SUPERVISOR'S APPROVAL

COMPARISON OF DETECTION MODEL USING MACHINE LEARNING ON ANDROID MALWARE

By

MUHAMMAD IKMAL BIN IHSAN
2019627244

This thesis was prepared under the supervision of the project supervisor, Sir Mohd Faris bin Mohd Fuzi. It was submitted to the Faculty of Computer and Mathematical Sciences and was accepted in partial fulfilment of the requirements for the degree of Bachelor of Computer Science (Hons.) Data Communication and Networking.

Approved by

……………………………
Mohd Faris bin Mohd Fuzi
Project Supervisor

JULY 14, 2022

STUDENT DECLARATION

I certify that this thesis and the project to which it refers is the product of my own work and that any idea or quotation from the work of other people, published or otherwise are fully acknowledged in accordance with the standard referring practices of the discipline.

………………………………………
MUHAMMAD IKMAL BIN IHSAN
2019627244

JULY 14, 2022

ABSTRACT

Mobile devices have experienced tremendous growth during the past ten years. As gadgets become more pervasive and people save more sensitive data on their mobile devices, the prevalence of mobile malware has increased. Malicious software, commonly known as malware, poses a greater risk to these mobile devices nowadays. Recently, several articles have been published regarding the proliferation of Android malware. Many modern technologies, such as smartphones, have been used into Android malware development, enabling it to advance. It has been used for a long time but is now worthless due of the evolution of Android malware and the inability to detect it. This project utilized supervised machine learning techniques such as SVMs, Naive Bayes and Random Forest to build an android malware detection model. It also evaluated and train the selection of Android characteristics to evaluate the malware detection model's performance. Then, the project examined the effectiveness of several machine learning detection models in identifying Android malware. This project has been divided into five distinct parts, each with a distinct purpose. Initialization, planning, development, evaluation, and documentation are all part of the process. In the end of the project, the result has been discussed and been compared for each machine learning used to get the highest accuracy to achieve the project objectives. The result of the comparison using the machine learning techniques for Android malware dataset discovered that SVM machine learning get the highest percentage with accuracy of 0.93. It also recorded that SVM machine learning got the lowest FPR and highest TPR among other machine learning used in the project. For future references to do this project, the project can be improved by using the project’s own Android malware dataset and use more than three machine learning when training and evaluating the dataset to discover the true potential for each machine learning