UNIVERSITI TEKNOLOGI MARA

COMPARISON OF MALWARE DETECTION MODEL USING SUPERVISED MACHINE LEARNING ALGORITHMS

SYAMIR BIN MOHD SHAHIRUDIN

BACHELOR OF SCIENCE (Hons.) DATA COMMUNICATION AND NETWORKING

JULY 2022

Table Contents

SUPERVISOR’S APPROVAL

COMPARISON OF MALWARE DETECTION MODEL USING SUPERVISED MACHINE LEARNING ALGORITHMS

By

SYAMIR BIN MOHD SHAHIRUDIN
2019290626

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 Programme’s Name.

Approved by

Sir 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.

………………………

SYAMIR BIN MOHD SHAHIRUDIN

2019290626

JULY 14, 2022

ABSRACT

Because of various security concerns and cyberattacks, cybersecurity is crucial in today's environment. In addition, malware has evolved quickly in recent years. Machine learning is utilised for malware detection with the advancement of malware analysis. The comparison of malware detection model utilising supervised machine learning techniques is the main goal of this project. The objective of this project is to develop the Windows malware detection model using supervised machine learning in Decision Tree, K-NN and Naïve Bayes, to evaluate the performance of malware detection in term of testing and training of the features selection and to compare the accuracy detection model in all three machine learning algorithms. The Windows malware dataset has been trained and tested by these three machine learning algorithms to get the percentage detection accuracy. Then, the outcomes demonstrated that the best classifier for categorizing our data with 0.96% accuracy is the Decision Tree machine learning algorithm. When comparing the accuracy of a malware detection model, it is excellent if there are numerous machine learning algorithms and more malware datasets included.