UNIVERSITI TEKNOLOGI MARA

PEARSON CORRELATION AND HYPERPARAMETER OPTIMIZATION FOR INTRUSION DETECTION SYSTEM

FUAD BIN MAT ISA

Thesis submitted in fulfillment
of the requirements for the degree of
Master of Science
(Computer Science)

College of Computing, Informatic and Media

June 2023

ABSTRACT

Internet has become vulnerable place due to the rapid advancement of technology. The need of security approaches that will protect this sensitive information has never been more crucial. Intrusion detection system (IDS) is playing a pivotal role to provide an extra layer of security in the network. In order to effectively detecting network intrusion, one requires to gathering a large volume of data. This high dimension data needs to be synthesized and filter from redundant and irrelevant features to improve the performance of IDS. However, the existing technique unable to performed well in terms of accuracy, detection rate and false positive rate. Therefore, this research proposed a Pearson Correlation and Hyperparameter Optimization for improving the performance of IDS. Pearson Correlation is used for selecting the optimal number of features which are most useful for identifying the attacks. To further improve the classifier algorithms, an optimization module named Tune Model Hyperparameter is applied. There are 3 datasets used in this research, namely NSL-KDD, KDD 99 and CICIDS2017. After the relevant features are selected, these synthesized data was tested on three classifier algorithms, Support Vector Machine (SVM), Decision Forest and Neural Network The results indicate that the proposed method performs better with Decision Forest algorithm in terms of accuracy, detection rate and false positive rate at the value of 99.9%. Moreover, compared with most of the existing state-of-the-art and legacy techniques, this proposed method exhibits better performance under classification metrics in the context of classification accuracy.

ACKNOWLEDGEMENT

Praise and thank to Almighty God and His blessings for giving me the opportunity to embark on my Master and for completing this long and challenging journey successfully. I would like to take this golden opportunity to express my sincere appreciation to my supervisor, Dr Alya Geogiana binti Buja for her encouragement, guidance, advice and moral support throughout the completion of this thesis.

I would like to dedicate my deep appreciation to Assoc. Professor Dr Mohamad Yusof Darus and Mr Shahadan Saad for their encouragement and kindness in providing me the information needed in order for me to complete this thesis. Also, thanks to my beloved parents for their faith, continuous encouragement and support.

Finally, a very sincere appreciation for those who had involved in contributing their help and support either directly or indirectly in making this thesis successful.

TABLE OF CONTENTS