E-6 Original

Binary classification of malware by analyzing its behavior in the network using machine learning

Jean Soto
Estudiante de pregrado, Facultad de Ingeniería,
Universidad Tecnológica Centroamericana UNITEC, Tegucigalpa

Introduction

- The International Telecommunications Union (ITU) develops the global cybersecurity index (GCI). In the last GCI published in 2020, Honduras performs last place in the America region and is placed in the 178th position out of 182 countries with a punctuation of 2.2 out of 100 (Index, 2020).
- This research work is based on the premise that any new variant of malware behaves similarly to that of its predecessor [1], together with the fact that most malware communicates with external hosts [2].
- The proposed model bases its operation on the behavior-based approach at the data network level [2]. This model parses files containing frames and packets captured from the network, known as PCAP files.

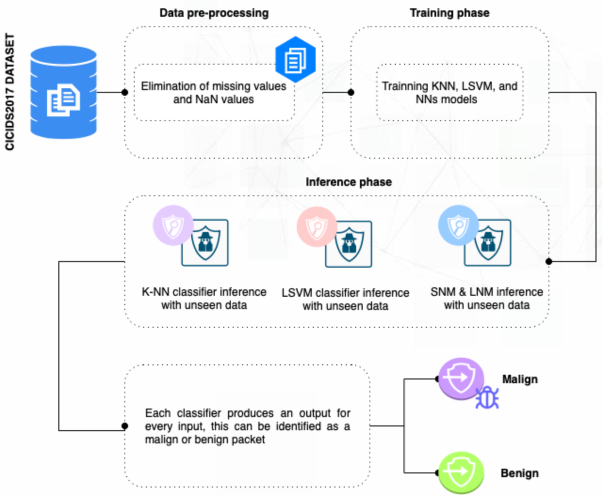
Data pre-processing
Training phase
CICIDS2017 DATASET
Elimination of missing values and NaN values
Trainning KNN, LSVM, and NNs models
Inference phase
K-NN classifier inference with unseen data
LSVC classifier inference with unseen data
SNM & LNM inference with unseen data
Each classifier produces an output for every input, this can be identified as a malign or benign packet
Malign
Benign

Figure 1 - Methodology diagram

Objective

Develop a model to analyze the information in PCAP files and classify network connections as either benign or malicious (malware generated).

Methodology

- The process uses two methods: (a) traditional machine learning algorithms (K-NN and SVM) and (b) various deep neural networks models. Both methods are trained and tested using the CICIDS2017 dataset (Figure 1).
- Neuronal network (NN) model architectures (Figure 2) used the cross-entropy loss function BCEWithLogitsLoss and SGD as the optimization algorithm. Both models use sigmoid as the activation function.

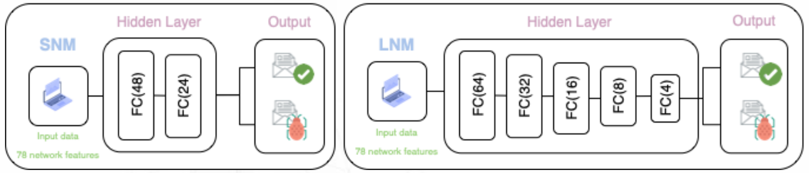

Figure 2 – Neuronal network models architecture

Conclusions

- A process capable of classifying network connections as either benign or malicious (malware-generated) was successfully developed achieving above 95% in the recall, and f1 metric, attained using only traffic information from the data network.
- The method with the best performance without taking execution time into consideration was K-NN of method (a) with non-standardized and standardized data.
- The choice of the best performance was made based on the recall and f1 due to the classifier's context.
- The impact of data standardization had a positive impact on the evaluation metrics. In consequence, the use of data standardization is strongly recommended.

Conflict of interests

The author makes known that there are not conflict of interest.

Contact: jeancasoto@unitec.edu

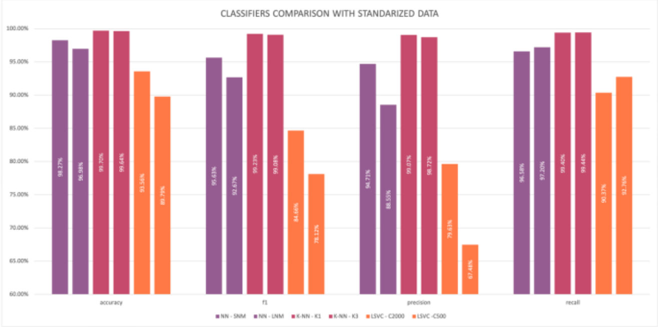

Figure 3 - Best classifiers comparison using standardized data.

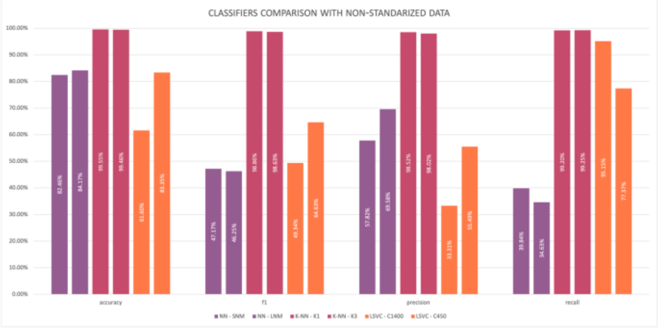

Figure 4 - Best classifiers comparison using non- standardized data.

Acknowledgments

Acknowledgement for Ph.D. Kenny Davila for supporting and mentoring me through development of this research work which was originally developed as a graduation project.

Bibliography

[1] R. Tian, L. Batten, R. Islam, and S. Versteeg, "An automated classification system based on the strings of trojan and virus families," in 2009 4th International conference on malicious and unwanted software (MALWARE). IEEE, 2009, pp. 23–30.

[2] S. Nari and A. A. Ghorbani, "Automated malware classification based on network behavior," in 2013 International Conference on Computing, Networking and Communications (ICNC). IEEE, 2013, pp. 642–647.