**Individual In-depth Report**

**Member name: Justin Young**

**Evaluated by:** [Yeshwanth Reddy Chennur](mailto:ychennur@asu.edu)

**Date:** Sep, 11 2023

**Tasks Assigned:**

* Literature review of Comparative Analysis of Machine Learning Models in Computer Network Intrusion Detection.

**Summary:**

* This research paper primarily focuses on evaluating and comparing the performances of some machine learning algorithms used for intrusion detection over the CICIDS 2017 dataset
* The paper presents a supervised machine learning approach for a binary classification of network traffic as either malicious or benign.
* A procedure for the comparative analysis of the models is given
  + Data cleansing (reduce the size of the dataset by removing records containing null and incorrect values)
  + Create training and testing data (split the data)
  + Feature selection (Random Forest Regressor technique was used to narrow down relevant features)
  + Machine learning algorithm implementation (a python training package LazyPredict trains on a subset of the data to determine best algorithms for comparative analysis)
* For this research, the XGB, Decision Tree, Random Forest, K-Nearest Neighbor (KNN) classifiers and Multi-Layer Perceptron (MLP) and Quadratic Discriminant Analysis (QDA) algorithms were chosen for analysis.
* The results of the analysis:
  + XG-Boost had the highest accuracy on 10 and 20 feature datasets, however it had a slow training time.
  + Random Forest and MLP similarly also had high accuracy just below XGB, however both have relatively low training time.
  + Decision tree is the most balanced between accuracy and training time with the following scores:
* Ten features -
  + Accuracy = 94.96%
  + Precision = 94.91%
  + Recall = 95.41%
  + F1 = 98.1%
* Twenty features -
  + Accuracy = 98.1%
  + Precision = 98.0%
  + Recall = 98.2%
  + F1 = 98.1%

**Outcome:**

This research paper concluded that the Decision Tree model had the best overall performance, considering accuracy and training time. When considering an intrusion detection system for a security problem, this paper provides a tested model we can use.

**References** *(with citation)*

[27] E. Osa and O. E. Oghenevbaire, “Comparative Analysis of Machine Learning Models in Computer Network Intrusion Detection”, In *2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development*, Lagos, Nigeria, 2022, pp. 1-5, doi: 10.1109/NIGERCON54645.2022.9803175

**Evaluation of Report**

**Evaluation summary with justification.**

* The research paper successfully evaluates and compares various machine learning algorithms for intrusion detection on the CICIDS 2017 dataset. XG-Boost demonstrated highest accuracy but slower training, while Random Forest and MLP offered a good balance between accuracy and training time. The Decision Tree showed commendable performance across metrics.

**The quality of the major result(s) with justification.**

* The major results demonstrate strong performance across selected algorithms. XG-Boost excels in accuracy but with slower training, while Random Forest and MLP strike a good balance. The Decision Tree shows notable accuracy with efficient training.

**The usefulness of the paper to the overall project.**

* The paper provides crucial insights for selecting effective intrusion detection algorithms, optimizing accuracy, and considering training efficiency.

**Other comments**

**Evaluation Approval  
  
Evaluation by:** Yeshwanth Reddy Chennur **Date:** Sep, 11 2023

**Is the written report of the in-depth study complete with all the major result(s) of the paper(s)? If not, provide as many examples of the major result(s) missing in the written report as possible. (in bullet form). [Normally within 100 words]**

* Yes, In-depth study completes with all the major results.

**Is each section of the guidelines sufficiently completed? If not, point out what is missing. [Normally within 40 words].**

* Yes

**Is the quality of this version of the written report satisfactory? If not, then why not? [Normally within 40 words]**

* Yes, the quality report is satisfactory.

**Approval.  
  
Approved by:** [Krupaben Kothadia](mailto:kkothadi@asu.edu) **Date:** 9/11/23 **Is the quality of this written in-depth study report and Evaluation report satisfactory? If not, then why not? (limit: 40 words)**

Yes, the in-depth report and evaluation report is satisfactory.