



Engineering
& Computing

School of Computing
& Information Sciences

School of Computing & Information Sciences

Summer 2023 Capstone II Project



Enhancing Intrusion Detection Systems

Student: Christopher Puig

Mentor: Yanzhao Wu

Instructor: Dr. Masoud Sadjadi, Florida International University

Problem

Our decision to undertake the project was driven by an interest in creating a more robust and accurate network security solution. This is achieved by merging two individual IDS models using ensemble techniques, we aimed to leverage the unique strengths of each model and create an ensemble IDS that surpasses the performance of the individual models. The project sought to explore the potential benefits of ensemble approaches in intrusion detection, providing valuable insights into how combining different models can enhance accuracy and precision in identifying network intrusions. Ultimately, the project aimed to contribute to the field of cybersecurity by demonstrating the effectiveness of ensemble methods in improving network security and enhancing the reliability of intrusion detection systems.

Project Overview

Our primary objective was to develop a merged IDS solution by integrating multiple open-source models. The specific goals include:

1. Investigate the architecture and compatibility of the open-source models to identify areas for integration and determine the viability of merging these systems.
2. Evaluate the performance of each IDS model individually using standard metrics such as accuracy, recall, precision, and F1 score. This evaluation will provide insights into the strengths and weaknesses of each model in detecting network intrusions.
3. Combine the two pre-trained IDS models using ensemble techniques such as Voting, Stacking, or Weighted Averaging. The aim is to leverage the diverse predictions of both models and create an ensemble IDS that potentially outperforms the individual models in accuracy and precision

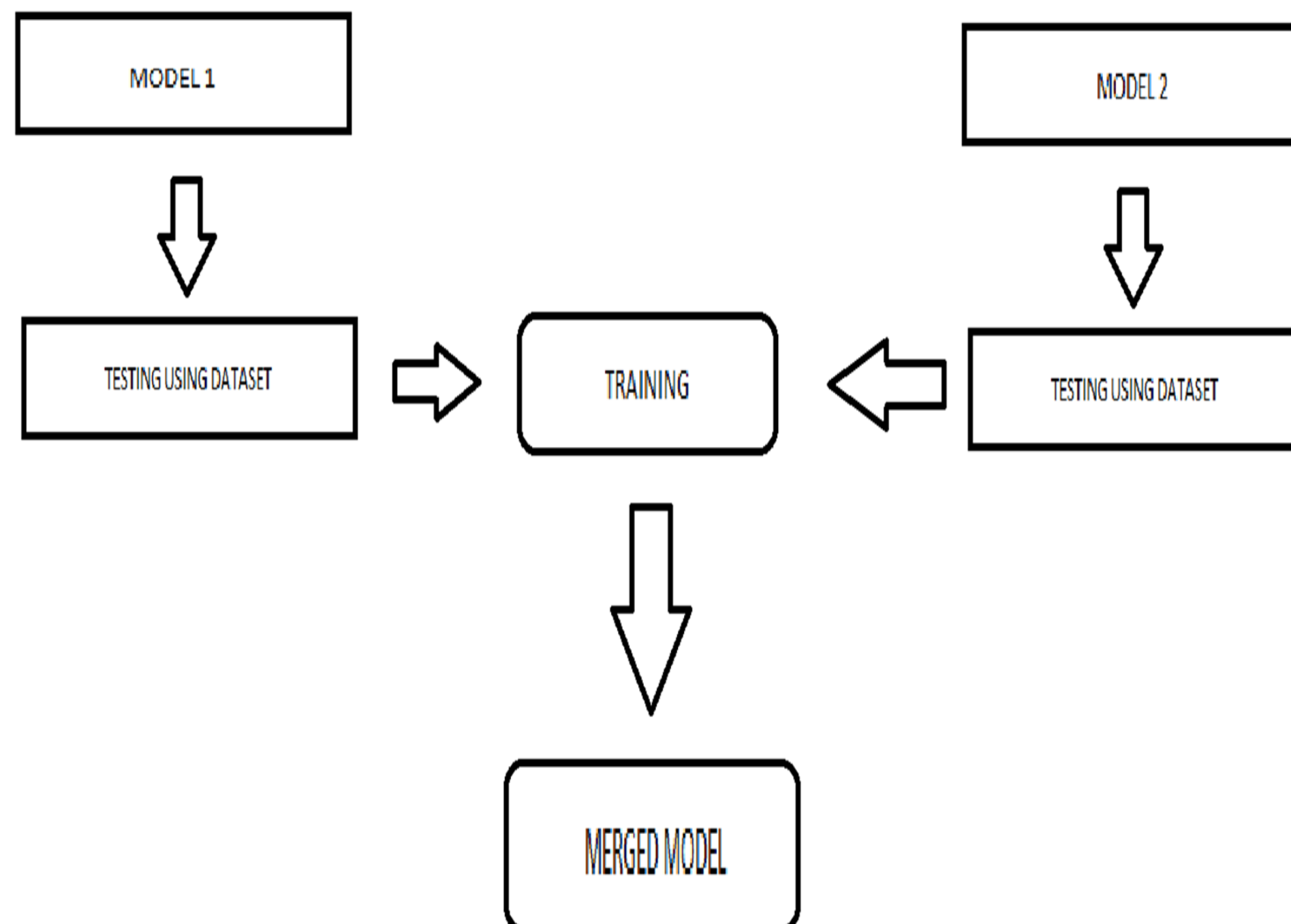
Tools Used

Programming languages: Python.

Applications: Oracle, Virtual Box, Kali Linux VM, Windows 10 VM.

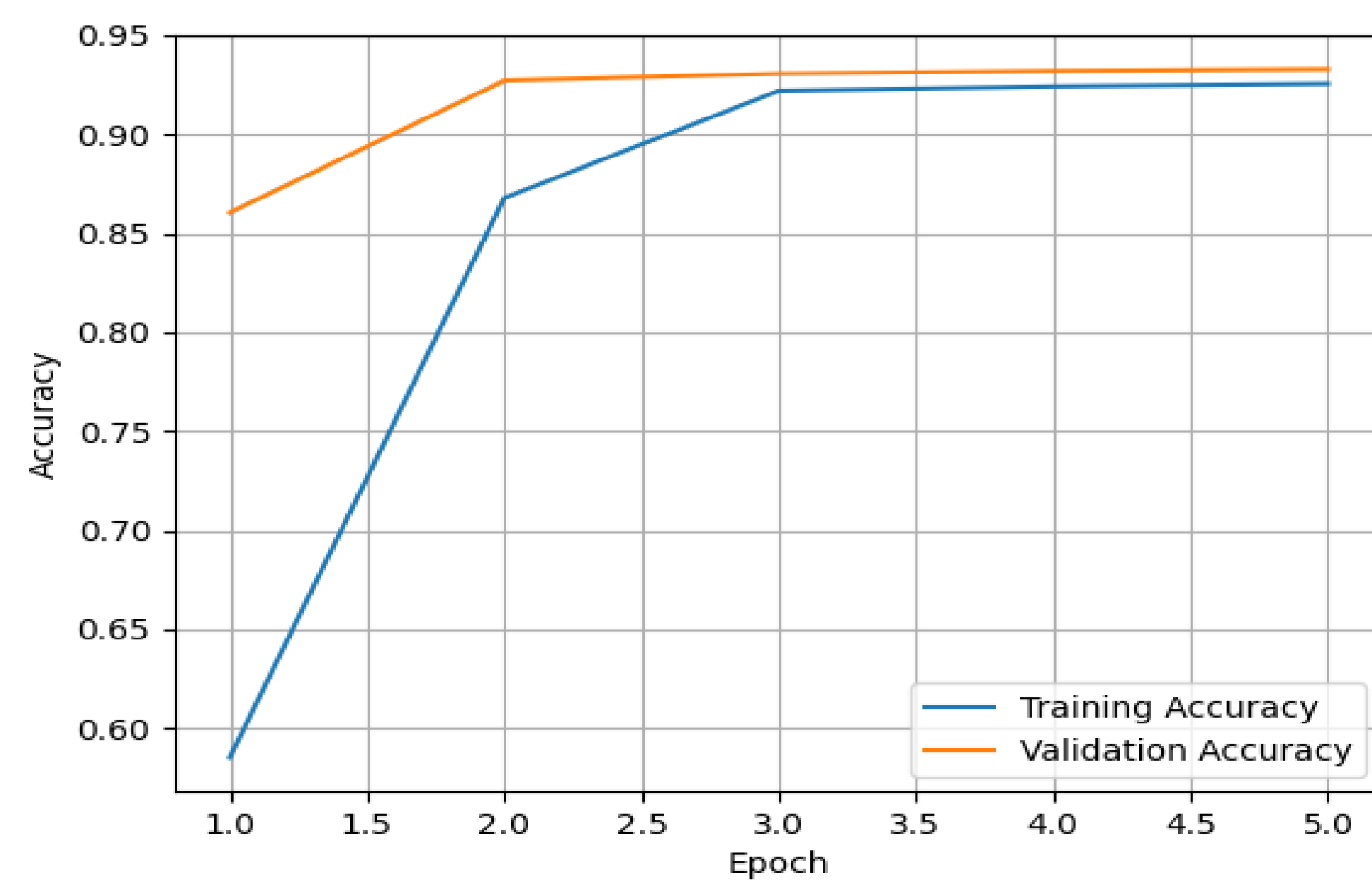
Datasets: KDD – 99

System Design



System Results

The validation accuracy line is accuracy of the IDS system on new data, while the training accuracy line is the accuracy of the model on the data it was originally trained on. The results are showing that as the IDS is trained with new data it improves its accuracy, at first its going to be less accurate as depicted below but eventually become on par with the validation accuracy.



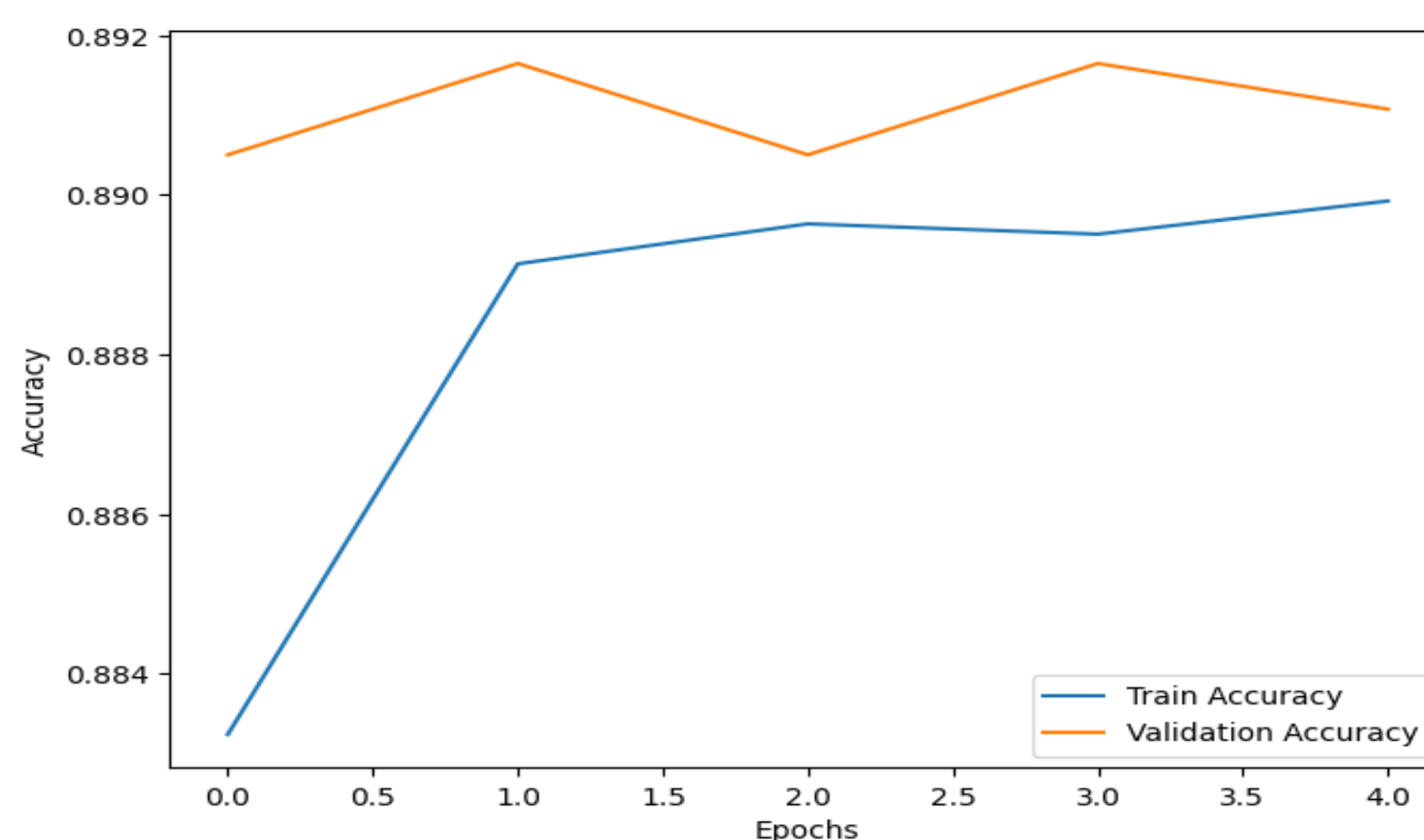
Implementation



Individual Contributions

Throughout, the sprints my focus was on team management and acting as a second product owner. My tracks were keeping track of all individual work, group meetings, accomplishments, and the overall project completion process. These tasks simulated a product owner or scrum master in an enterprise setting. My goal was to keep the team on track without sacrificing efficiency and making sure that anyone team member who refused to complete their work was reported to the professor.

Screenshots



```
Run merge
1944/1944 [=====] - 1s 468us/step
1944/1944 [=====] - 1s 468us/step
1944/1944 [=====] - 1s 463us/step
1944/1944 [=====] - 1s 390us/step
1944/1944 [=====] - 1s 446us/step
1944/1944 [=====] - 1s 391us/step
1944/1944 [=====] - 1s 468us/step
1944/1944 [=====] - 1s 415us/step
1944/1944 [=====] - 1s 436us/step
1944/1944 [=====] - 1s 399us/step
1944/1944 [=====] - 1s 434us/step
Best Weight for Model1: 0.2
Best Weight for Model2: 0.9
9720/9720 [=====] - 4s 417us/step
9720/9720 [=====] - 4s 448us/step
-----
accuracy
0.845
recall
0.845
precision
0.998
f1score
0.845
```

Summary

The goal of this project was to improve the accuracy of an IDS by merging two well known IDS systems. Our project was successful in improving the precision and accuracy statistics against trained data and unknown introduced data. This does not mean the end of the project since more changes and testing can be made to improve upon our existing results and possibly merge another known IDS system and create even better statistics against network intrusion signatures.

Acknowledgement

The material presented in this poster is based upon the work supported by XXXXX (name of the sponsor: federal, state or local, or industry, if applicable). We/I thank XXXX for assistance, cooperation and mentorship that I/we received throughout this process