Logo

Description automatically generated

Capstone Project – Intrusion Detection using AI

Craig Perkins – [perkins.cr@northeastern.edu](mailto:perkins.cr@northeastern.edu)

Northeastern University

January 2023 – EAI6980

Capstone Project – Intrusion Detection using AI

For the capstone project, I will be using artificial intelligence to develop an intrusion detection system. Over the last few decades, more and more companies have been relying on technology to host mission critical applications that keep industry running. In the last decade, cloud computing has emerged as an enormous field where small companies can leverage the same infrastructure as major technology companies like Microsoft, Google and Amazon. In addition to cloud services, many large companies also host applications on-premise in a data center that is managed by the same company or co-located in another data center. When securing a physical location like a data center, companies can install cameras to monitor for unauthorized entry to restricted locations to detect intrusions. In addition to protecting physical assets, it’s also important to monitor the digital infrastructure inside the data center which is connected to the internet and accessible by the outside world.

There are many methods for securing servers like protecting the server with username and password, restricting traffic to certain ports and adding a firewall to block traffic. Even with all the security measures in place, hackers can still find methods to exploit to remotely takeover machines. There have been prominent examples of zero-day vulnerabilities that allow attackers to remotely execute code on a vulnerable application server. Examples include heartbleed and log4shell.

In this project, I seek to find a solution for automated monitoring for intrusion detection to determine anomalous behavior within a cluster that should trigger a security alert. I will be examining network traffic to determine if and when a machine was compromised.

**Technologies Used**

In this project, I used python with scikit-learn, keras, AutoGluon and jupyter notebooks to demonstrate the process of model building for an intrusion detection system. The jupter-notebook is in the same repository and is viewable on Github and running locally by following the directions in the README to install the dependencies and run the Jupyter notebook. All data for this project is stored in the `data/` directory of the repository. The jupyter notebook runs through steps like the Exploratory Data Analysis, Data Preparation, Model Building and finally with Model Evaluation and Selection.

This is a one-pager for the midpoint checkin and at this point in the project, I have built a few models to get a sense of the base level for accuracy for which to beat with upcoming models. This project builds both a binary classifier to determine if a record in the dataset is benign or malicious and also builds a multi-class classifier to predict the category of attack of each record in the test dataset.

**Challenges**

When building an Intrusion Detection System, it is important to build a model that is sensitive enough to pick up on anomalous behavior while also minimizing inadvertently flagging normal behavior that would result in security alert fatigue for a security analyst monitoring the system.