# Using Random Forest and KNN techniques to Detect Anomalies in Various IoT Traffic

Rashi Pachino[1] and Noah Weiss[2]

[1]Department of Computer Science, Ariel University

November 2022

## 1 Introduction

The Internet of Things (IoT) pertains to the budding network of interconnected physical devices, home appliances, vehicles and even people that are provided with unique identifiers and do not require human-to-human or human-to-computer interaction to transfer data. A "thing" can be any natural or man-made object that can be assigned an Internet Protocol (IP) address and is able to transfer data over a network. IoT has facilitated the integration between the physical world and communication networks and its application in the environment we live in. Consequently, the need to address security issues affiliated with IoT, such as spoofing, intrusion and Distributed DoS (DDoS), is more prevalent than ever. IoT is split into four abstraction layers. Physical layer collecting data using IoT sensors; network layer used for transferring data among devices for processing; processing layer responsible for performing some processes and computational tasks; and the application layer delivered by the devices of end users. All these layers are subject to security threats. Whether the intended victim is a human with a heart monitor, a self-driving vehicle or a smart home security device, cyber criminals have more opportunities to target devices, with the potential to cause harm at far greater scale than ever seen before.

Machine learning (ML) and Deep Learning can play a vital role in detecting anomalous activity through training a supervised, unsupervised or semi-supervised model on relevant datasets. ML allows robots to learn from their interactions rather than directly programmed. Human assistance, complex math algorithms, or performing in complex networks are not needed for ML, making it a valuable asset in cybersecurity. ML strategies for IoT protection have advanced considerably in recent years. Techniques include, but are not limited to, SVM, RF, Naïve Bayes, Decision Tree, Random Forest, KNN, and MLP based machines.

## 2 Abstract

As the world of technology advances, and the Internet of Things expands, the need for anomalous detection models increases. Denial of Service (DOS), Distributed DoS (DDoS) attacks, among others, are the most common attack types that face the IoT networks. In response, a proficient detection model should be implemented. We propose a classification model that combines several supervised classifiers together to amplify the accuracy, precision and recall. Algorithms include Random Forest, K-Nearest-Neighbors, XGBoost and Weighted Majority Voting. In addition, an Autoencoder will be added to pre-process the data in order to retrieve the most relevant features. With this, we hope to higher the recall score These techniques are trained and tested on the UNSW-NB15 benchmark dataset containing normal and malicious traffic from numerous devices.

## 3 Related Works

In this section, we discuss related works on anomaly and attack detection using machine learning algorithms. Protecting IoT devices is a critical task, tried by many. Considering sensitive information is shared among the networks and devices are updating constantly, proper action must be taken to prevent third party intervention. We have separated past works by approaches: supervised, unsupervised and semi-supervised. In reference to machine learning-based anomaly detection, supervised models are trained on labeled data, while unsupervised models are trained on unlabeled data. Semi-supervised models use both labeled and unlabeled data in training. In the following paragraphs, we evaluate recent machine learning techniques for anomaly detection.

Al-Akhras et al. proposed joining several supervised algorithms together to build a classification model for identifying attack signatures for IoT networks. Using the Weka platform, an assortment of classification algorithms including Random Forest (RN), K-Nearest-Neighbors (KNN), and Naïve Bayes were applied. Additionally, the voting method was exercised. The model was applied on the UNSW-NB15 dataset which covers a large number of normal and malicious network traffic and is explored in the data exploration section of this article. Based on the experimental results, the percentage of correctly classified instances was 100% for the RF and KNN classifiers, while the Naïve Bayes classifier showed the lowest results with 95%. Even After adding noise, results show RF and KNN classifiers as having a higher performance over the evaluation metrics, accuracy, precision and recall, while the Naïve Bayes classifier still shows the lowest results. In addition, the voting method had the best performance over the evaluation metrics, accuracy, precision and recall. To summarize, training three supervised classifiers, and adopting the voting method to create a classification model, revealed the best performance over all evaluation metrics.

Apostol et al. proposed an IoT botnet anomaly-detection solution based on a deep autoencoder model. This Artificial Nerual Network (ANN), requires

unlabeled normal data to be separated from malicious data to train the model. An ANN is an unsupervised model that consists of layers of nodes, with each node behaving like a neuron. The first layer is the input layer, while the last layer is the output layer. Intermediate layers are called hidden layers. If there are at least three hidden layers, the ANN becomes a Deep Neural Network (DNN). Since the autoencoder was trained only on the normal pattern of a system, it can reconstruct normal patterns with minimum error. In case of anomalies, the reconstruction error should exceed a certain threshold. Meaning, by training the model on normal traffic, then testing the model on anomalous traffic, if the autoencoder cannot replicate the data below a certain score,the data is considered malicious. The dataset used in this experiment was the UNSW Canberra at ADFA BOT-IoT dataset. The dataset's source files are provided in different formats, including the original pcap files, the generated argus files and csv files. The captured pcap files are 69.3 GB in size, with more than 72.000.000 records, and extracted flow traffic, in csv format is 16.7 GB in size. The dataset includes DDoS, DoS, OS and Service Scan, Keylogging and Data exfiltration attacks, with the DDoS and DoS attacks further organized, based on the protocol used. Experimental results showed that their proposed method achieved scores of 99.7 for accuracy, 0.99 for precision, and 0.99 for recall.

Abusitta et al. proposed a deep learning-enabled anomaly detection model for IoT systems. Their goal was to detect malicious data among traffic between various types of IoT devices in a network. The framework consists of stacked denoising autoencoders which are used as building block for the development of Deep Neural Networks based on IoT. The denoising autoencoders are trained on benign data and tested on anomalous data in order to extract robust features and isolate unnecessary features (neural features). The parameters are then sent to a supervised machine learning classifier like binary logistic regression to differentiate between benign and anomalous data. They used the DS2OS traffic traces dataset which contains data from diverse IoT devices including light controller, thermometer, movement sensors, washing machines, batteries, thermostats, smart doors and smart phones. With an accuracy score of 94.6%, their model outscored other popular machine-learning models such as MLP-based (85.5% accuracy rate), SAE-based (84.7% accuracy rate), RBM-based (85.7% accuracy rate) and other Stacked Denoising Autoencoder-based anomaly detection (85.5% accuracy rate). In short, the results show the validity of this semi-supervised method in heightening the accuracy of detecting anomalous data.

# 4 Data Description and Exploration

After researching and exploring numerous datasets exercised in past works, we decided to implement the UNSW-NB15 dataset. Creators of this dataset utilized the tcpdump tool to capture 100 GB of the raw traffic (e.g., Pcap files). It has nine types of attacks, particularly, Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode and Worms. The Argus, Bro-IDS tools are

used and twelve algorithms are developed to generate totally 49 features with the class label. Although we receive the data already processed into features, we implement further feature selection on the data. In the following images, we explore the given dataset in favor of further understanding the power of the data.

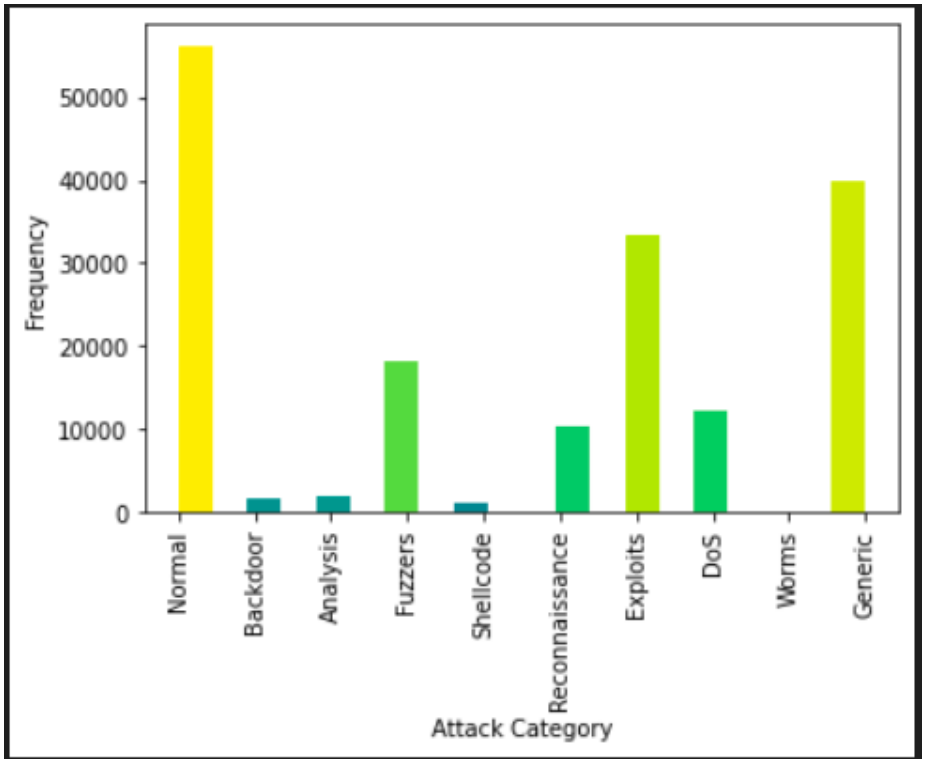


Figure 1: Distribution of Attacks in Dataset

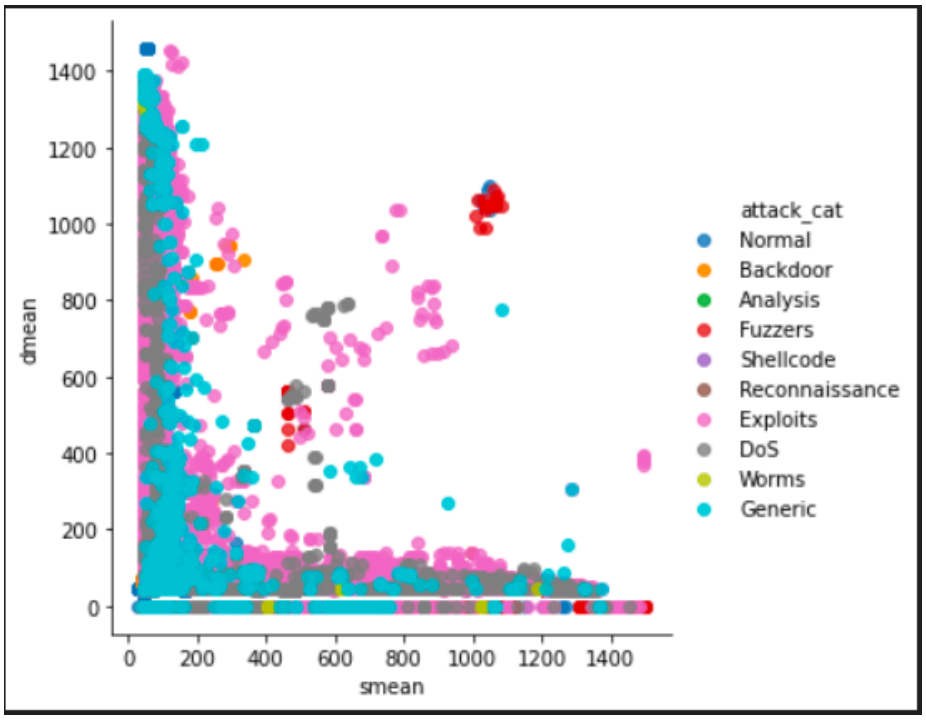


Figure 2: Smean VS Dmean

As shown in Figure 1, the distribution of attacks across the dataset are vast and spread out. However, in Figure 2, we can see that regarding the given features of this dataset, (in this example the mean size of the packet sent from the source vs the mean of the size of the packet transmitted by the dest) it can be observed that there is a huge overlap between various attacks, implying that UNSW- NB15 is a complex dataset to deal with for identifying specific attack categories. Nonetheless, our research focuses on anomaly detection rather than specific attack apprehension.

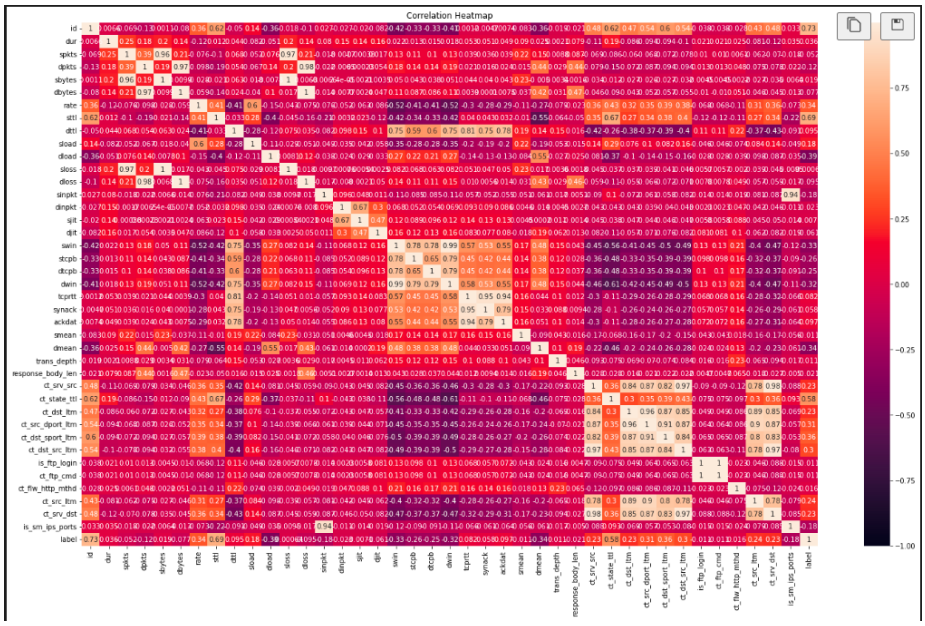


Figure 3: Correspondence Between Features

Figure 3 outlines the correlation between features of the dataset. This is an important factor of data exploration. Pairs of features with high correlation rates can be filtered, as keeping both features is redundant for our learning algorithm.

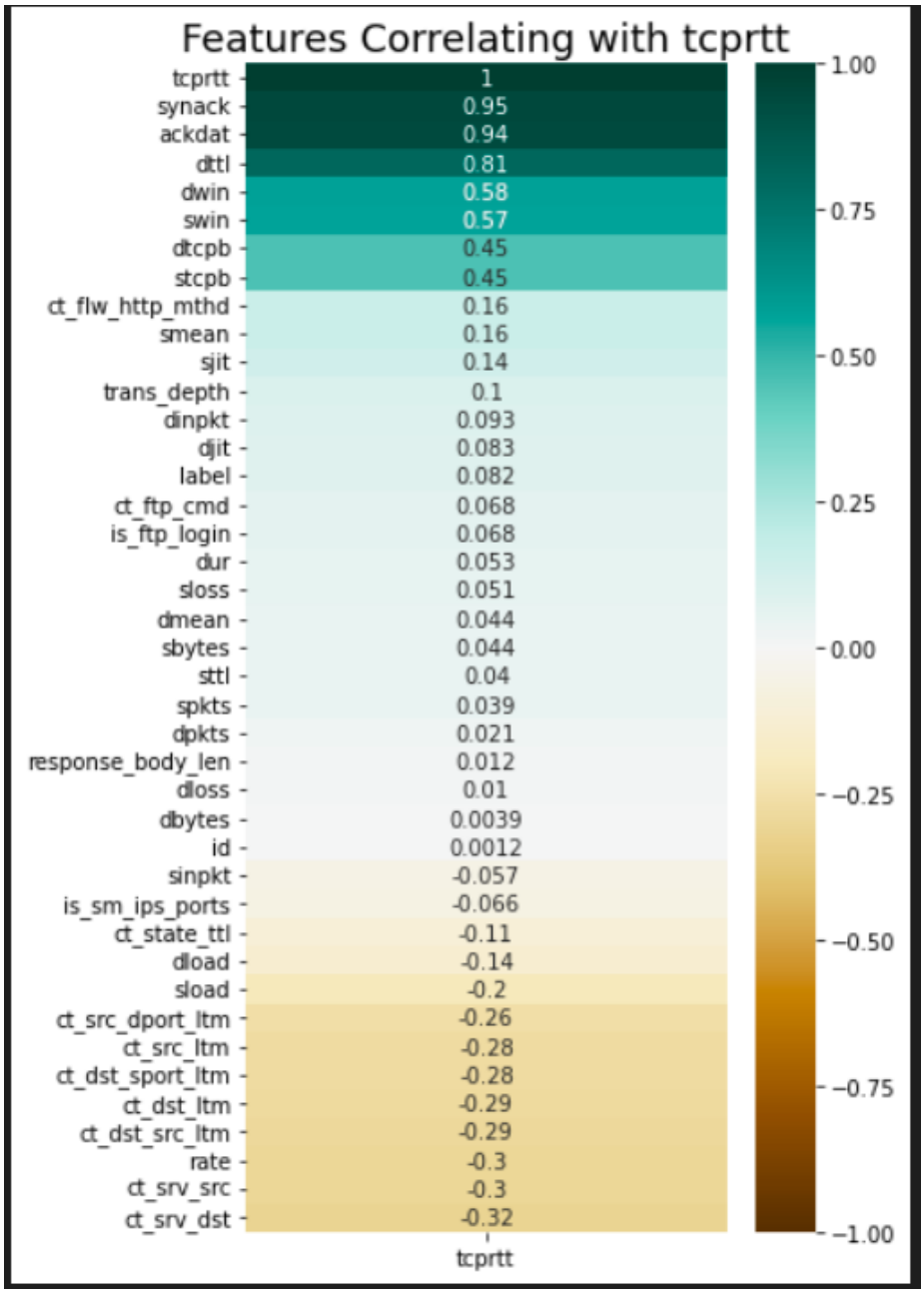


Figure 4: Features Correlating With Tcprtt Feature

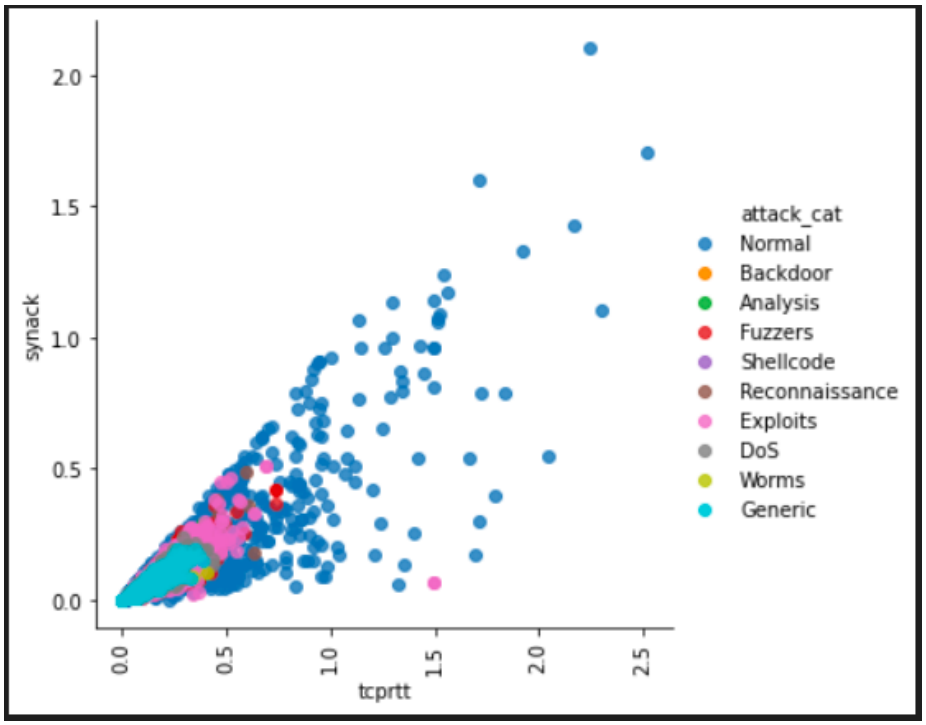


Figure 5: Tcprtt VS Synack

For example, Figure 4 describes the features that correlate the most with a specific feature tcprtt. This feature measures the TCP connection setup round-trip time, the sum of synack and ackdat. The highest correlating feature is synack. In Figure 5, we can see that besides for outliers, most of the data is almost linear. Therefore, using both tcprtt and synack features would prove to be unnecessary.

# 5 Results

As identifying malicious content is a crucial task, allowing for little mistake regarding false negatives, we hoped to higher the recall score and our model did just that. While starting with an already featurized dataset, the addition of an Autoencoder as a pre-processor causes the most relevant features to be selected and sent to the chosen classifiers to obtain the highest results. Our model produces an accuracy of 97%, recall of 99%, precision of 97% and an F1-score of 98%.

# 6 Summary and Open Issues

To conclude, Over the past few years, IoT has become one of the most important technologies of the 21st century. Connecting everyday objects—kitchen appliances, cars, thermostats and baby monitors to the internet via embedded devices, has made seamless communication possible between people, processes, and things. As a result, the need for IoT security has skyrocketed. Concurrently, Machine learning has become widely used technology for cyber security solutions. Thus it is only natural to combine the two to form a reliable and accurate model to detect anomalous data. Our proposed solution will use supervised algorithms such as Random Forest, KNN, XGBoost and Majority Voting to identify malicious data from the UNSW-NB15 benchmark dataset. As well as an unsupervised Autoencoder to pre-process and select features. Based on past studies, these classifiers are accurate and precise, delivering a high accuracy rate. However, with the use of an added Autoencoder, we hope to raise the recall score, an important factor in anomaly detection. Used together, we can create a ML model to detect and identify anomalous traffic, aiding in the solution to this major cyber security threat.