* Introduction
  + Technology has advanced rapidly and with that advancement comes many benefits such as email or SMS messages. With today’s technology we can instantly send a message to any recipient around the world. Unfortunately, these systems can be abused by marketers or threat actors via the use of spam or malware. Spam is unwanted or unsolicited digital communication that gets sent out in bulk. While malware is any software intended to do harm such as a virus, worm, or other code-based entity that infects a host or intends to cause harm. Threat actors leverage spam methods to send out mass emails or SMS messages as phishing attempts in an effort to obtain PII (personally identifiable information), credit information, and/or unauthorized system access using spam or malware.

In order to help lower the amount of spam we receive to lower the attack surface or amount of spam one receives we must utilize a spam filter which is usually an application or a rules-based system that identifies the difference between a spam message and a legitimate message. This process is typically done by a spam filter application reviewing the email or SMS message for anomalies such as misspelled words, malicious links, and other relevant threats. If it finds any of these issues it will either flag the message as spam or alert the user to verify if the spam is indeed spam. First iterations of these systems were based on rules and usually did not work as well as methods we deploy today.

Initially this rules-based system was created to identify spam messages however as threats advanced the amount of spam messages increased, we run into issues such as accuracy and performance of these spam filter applications. To mitigate these concerns, we begin leveraging the power of machine learning. Machine learning is a field of computer science that dates back as early as the 1940’s which researchers first began exploring ways to teach computers to learn from data. Machine learnings strength within spam and anomaly detection lies within its ability to detect and respond to new and evolving threats in which a rules-based system could not address. Machine learning models can be trained and used to detect and filter out spam emails and phishing attempts with an understanding of basic machine learning principles. Using machine learning algorithms can be trained on large datasets of labeled email or SMS messages to help the model learn to recognize patterns and features that indicate spam or phishing. By leveraging the dataset, we can train and fit out model to ensure we have the most effective means for detecting spam messages.

To understand and compare the best methods for filtering spam messages we have selected SMS datasets in which we will leverage spam vs ham algorithms to classify our SMS messages as spam or legitimate (ham). Spam vs ham methods are used for detecting and preventing threats such as spam and phishing attacks. Using these algorithms allows the system to adapt to new unknown types of spam since it is trivial for an attacker to modify the content of the spam to circumvent the rules-based system. Although effective they are not perfect as attackers can deliberately craft message to evade detection by the algorithm.

We start in section 1 in which we will introduce our data set and configuration. In this section we will import and review the current state of the data set for any potential issues. Once we have verified the data set is adequate the dataset will then be analyzed for various properties such as count, top, and frequency values. This will allow us to gauge the performance and accuracy of our tests carried out by the various spam detection algorithms. Section 2 will take a dive into modifying the dataset to include metadata such as discrete values for classifying the data as spam or ham. In addition, the baselines of the data will be examined prior to running the data through the models. Section 3 will explain the methods of each model and why the model was chosen for spam detection. We will analyze the data to see if our chosen model was effective as classifying spam accurately and efficiently. Finally, section 4 will compare and contrast the results of the various models used in the dataset to determine which model provide the best accuracy and performance for spam detection.