Introduction

As cell phones adopt new capabilities, they become privy to more sensitive information. Credit cards, medical history, browsing data, passwords, and other protected data is accessible via mobile phone. Therefore, it makes sense that there would be an influx of security attacks in the form of phishing texts, commonly known as SPAM. Identity thieves rely on the curious or unknowing user to press an embedded link or respond to their SPAM messages to gain access to private information or begin a dialogue. Unsavory marketers also use SPAM to broadcast unsolicited ads to users on a global scale.

It is imperative that provider software be able to filter SPAM messages from those that are not SPAM, known as HAM. The goal of this project is to use a dataset to build various machine learning models that will solve the binary classification problem: Is this message HAM or SPAM? First the data will be studied using exploratory data analysis (EDA) and cleaned/preprocessed before use in the machine learning classification models. The machine learning models that will be tested include Naïve Bayes, K-Nearest Neighbors, Random Forest Classifier, and Gradient Boosting Machine.

Confusion matrices will be used to evaluate each model's performance. When designing a filter, it is important to prioritize the reduction of false positives over false negatives. It would be better to receive a SPAM message than to not receive a HAM message.