# Abstract

## NLP-Based Phishing Email Detection

This code implements a robust email classification system that leverages machine learning to categorize emails as "Spam" or "Not Spam" (Ham). The process begins by loading and preprocessing the dataset, where missing values in the subject and body columns are filled, and the labels are preprocessed to standardize the email classification. Text cleaning is performed on both the subject and body of each email, where punctuation, stopwords, and case differences are removed, ensuring that the model focuses on meaningful content. The system extracts two types of features: textual features using a TF-IDF (Term Frequency-Inverse Document Frequency) vectorizer that converts the cleaned text into numerical representations, and engineered features that capture additional patterns, such as the number of hyperlinks, uppercase words, special characters, email length, and word count. These features are combined into a single feature set using sparse matrix stacking and are used to train a Logistic Regression model. The model is then evaluated using accuracy, precision, recall, and F1-score, providing a comprehensive classification report. A prediction function is built to classify new emails by extracting features from the subject and body, transforming them, and passing them through the trained model for classification. Sample predictions demonstrate the model's ability to accurately classify emails with professional content as "Not Spam" and those with fraudulent or promotional content as "Spam." This system can be applied in real-world scenarios, such as email clients or corporate email servers, to automatically filter spam messages, significantly reducing the manual effort required by users and improving email system efficiency. Furthermore, the model's performance can be further enhanced by incorporating additional machine learning techniques or refining the feature engineering process.