ABSTRACT

The rapid development of digital communication technology increased the risk of phishing attacks through email, highlighting the need for accurate and reliable email phishing detection methods. Previously, phishing email classification systems primarily relied on rule-based or keyword-based approaches; however, their performance was limited due to difficulties in handling complex text variations and continuously evolving phishing patterns. This study aimed to build a phishing email classification system by comparing three text vectorization methods, namely Term Frequency-Inverse Document Frequency, Count Vectorizer, and Word2Vec CBOW, using two classification algorithms: Support Vector Machine and Naive Bayes. The objective was to identify the most efficient combination of vectorization method and algorithm for detecting phishing in four types of datasets, namely CEAS_08, SpamAssasin, Enron, and Phishing_Email. The results indicated that the combination of TF-IDF and Support Vector Machine consistently achieved the best performance, with the highest accuracy obtained on the CEAS_08 dataset at 0.996, followed by SpamAssassin at 0.993, Enron at 0.989, and Phishing Email at 0.977. These findings confirmed that TF-IDF efficiently represented the textual characteristics of email data, while Support Vector Machine provided stable and highly accurate classification performance. Therefore, the combination of TF-IDF and Support Vector Machine was selected as the optimal method for a phishing email detection system.

Keywords: Email Phishing, Text Classification, Support Vector Machine, Naive Bayes, TF-IDF, Count Vectorizer, Word2Vec