A novel framework for identifying twitter spam data using machine learning algorithms

ABSTRACT

Nowadays, Twitter has become one of the most popular social media in the world. However, its popularity makes it an attractive platform for spammers to spread spam. Twitter spam becomes a severe issue. It is referred to as unsolicited tweets containing malicious links that direct victims to external sites containing malware downloads, terrorists, phishing, drug sales, scams, etc. Previous studies have approached spam detection as a classification problem, high dimension, time-consuming problem, which requires new methods to address the problems. This study introduces a novel framework for identifying Twitter spam data based on machine learning algorithms. By initializing data pre-processing for clean-up, noise removal, and unpredictable unfinished data, reducing the number of features in the tweet dataset using mutual information is the study's methods. The feature selection is introduced to select the most important from the extracted high-dimensional best features and feed the selected features into the minimum Redundancy and Maximal Relevance algorithm and apply random forest for classification. This study allows us to achieve higher classification accuracy and speed. The effectiveness evaluation being confirmed by experiment results show that accuracy is improved by 90% in 0hr 0m 20s time, compared with the existing system, the completion time is 2.022 seconds, and the accuracy is 80%. The research results contribute significantly to the field of cyber-security by forming a real-time system using machine learning algorithms.

Keyword: Natural language processing; Machine learning; Twitter spam; Feature selection; Minimum redundancy and maximal relevance algorithm