Abstract

Malware is currently one of the biggest data security dangers to people or associations and has kept on expanding as far as volume and modernity to assault PC frameworks. PDF document is considered as the most commonly transferred medium for exchange of data. The popularity in the PDF document have made it one of the most attractive misuse vehicles. The structure of a PDF file contains different objects, features and characteristics that make it easier for the attackers to invade into a PDF document. We proposed new features from the PDF file structure that played a very vital role in improving the success of the neural network and machine learning classifiers by increasing the overall accuracy up to 99.9%. The proposed features provided better results when compared with the results of already identified features. The newly proposed features combine with the previously identified features to make the neural network and machine learning classifiers stronger in making the prediction increasing the file safety.