

Phishing website is a fake web page that mimics legitimate website using social engineering techniques to lure unsuspecting users to web page for the purpose of stealing their personal information such as credit card, username and password. Phishing attack is on the increase on a daily basis due to the inability of the existing systems to fully identify a phishing web page from legitimate. Machine learning technique is a trending and intelligent approach for detecting a phishing web page. but, identifying efficient algorithm with the ability to classify and identify web page as either legitimate or phishing in real-time continues to pose a challenge as the existing systems are characterized by misclassifications resulting into low detection rate, high false positive, high running time. The purpose of this study is to develop an efficient machine learning based model with the ability of detecting whether a web page is phishing or not. A performance analysis of some popular classification algorithms was performed and revealed Random Forest as the best classifier on the phishing dataset. A machine learning based model for the detection of phishing attack was built based on Random Forest with wrapper based on classifier attributes evaluator and ranker (CAER) feature selection method. The performance of the proposed model was evaluated using phishing dataset that comprises of static and dynamic features of websites. The experimental results show that the proposed Random Forest based model with feature selection outperformed some of the existing solutions including the best performance of the Random Forest when the full features were used with high accuracy of 97.3% in addition to better precision, sensitivity and lower false positive rate of 0.03 achieved.