ABSTRACT

This research aims to develop a phishing detection model using the K-Nearest Neighbors (KNN) algorithm and compare its performance with Decision Tree (DT). With phishing threats on the rise, effective detection is critical to protecting user data. In this research, a dataset consisting of 10,000 URLs of phishing and legitimate sites was used. After data pre-processing including cleaning, normalization, and feature extraction, the KNN and DT models were trained and tested. The research results show that the KNN model achieves an accuracy of 95%, while the DT achieves an accuracy of 93%. KNN shows superiority in accuracy and consistency, while DT provides better interpretation through decision tree visualization. This research provides insight into the effectiveness of each method and identifies the most relevant features in the classification process.

Keywords : Cybersecurity, Phishing Detection, K-Nearest Neighbor (KNN), Decision Trees (DT), URLs