

ABSTRACT

Azzam Fawwaz Mahmud. 91121170.

Machine Learning Implementation to Detect Phishing Website.

Phishing websites are criminal mechanisms that use social engineering and technical pretexts to extract personal identification data and financial account credentials from customers. In Indonesia alone, according to a report by the Indonesian Internet Domain Name Manager (Pandi), the number of phishing recorded in the last 5 years has reached 34,622. The number of unique phishing attacks reported in Q3 2022 was 7,988. This study aims to find a classification machine learning algorithm with the best performance for detecting phishing websites using URL features. The classification algorithms to be compared are the decision tree, random forest, and KNN. The results of this study are that the first model that uses a decision tree obtains an accuracy of 0.833, a precision of 0.86, a recall of 0.83, and an F1-score of 0.83. The second model that uses the random forest algorithm gets an accuracy of 0.834, a precision of 0.86, a recall of 0.83, and an F1-score of 0.83. The last model that uses the K-Nearest Neighbors algorithm gets an accuracy of 0.482, a precision of 0.24, a recall of 0.50, and an F1-score of 0.48. Thus, of the three algorithms random forest is the best algorithm for detecting phishing websites.

Keywords: Phishing Website, Machine Learning, Classification, Decision Tree, Random Forest, KNN.

xv + 56 + Attachments + 32 Figures + 5 Tables

Bibliography (2015 - 2023)