

Abstrak

Phishing merupakan salah satu jenis serangan rekayasa sosial yang bertujuan untuk mendapatkan informasi pribadi korban, dengan metode menyamar sebagai entitas yang dapat dipercaya. Salah satu cara untuk mendeteksi situs *phishing* adalah dengan klasifikasi dari fitur-fitur yang menjadi ciri-ciri situs tersebut. Namun, beberapa penelitian terkait justru menunjukkan terdapat beberapa fitur yang tidak penting dan tidak relevan. Terlebih lagi, penelitian-penelitian tersebut menggunakan fitur yang berbeda-beda. Penelitian ini memiliki tujuan berupa: terciptanya sistem yang dapat mengidentifikasi fitur yang paling optimum dalam pengklasifikasian situs *phishing*. Metode yang digunakan berupa *feature selection* dengan teknik perankingan menggunakan formula *Chi-Square*, lalu diikuti dengan proses klasifikasi menggunakan algoritma *Support Vector Maching* (SVM), *Multilayer Perceptron* (MLP), *K-Nearest Neighbors* (KNN), dan *Random Forest*. Tujuan utama dari *feature selection* adalah memilih fitur terbaik dari suatu kumpulan fitur data. Eksperimen dilakukan sebanyak dua kali yaitu sebelum dan sesudah proses *feature selection*. Hasil dari penelitian ini adalah algoritma *Random Forest* menghasilkan nilai akurasi paling tinggi yaitu 96.65% jika menggunakan dataset sebelum *feature selection* dan meningkat menjadi 96.92% pada dataset setelah *feature selection*.

Kata kunci : fitur, *chi-square*, situs *phishing*, *feature selection*, klasifikasi