

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini dilakukan untuk membandingkan kinerja klasifikasi tunggal dengan klasifikasi menggunakan fitur seleksi dan mengetahui fitur paling dominan yang bisa dijadikan referensi untuk mendeteksi adanya web-phising. Pearson Correlation, Information Gain, Gain Ratio, Chi-Square dan Recursive Feature Elimination adalah fitur seleksi yang digunakan untuk mengetahui fitur paling dominan. Algoritma klasifikasi yang digunakan yakni Decision Tree karena Decision Tree populer untuk memecahkan permasalahan klasifikasi. Berdasarkan eksperimen yang dilakukan menggunakan Web Phising Dataset dari UCI Machine Learning Repository, kinerja tertinggi yakni menggunakan recursive feature elimination yang menghasilkan kinerja sebesar 97,88% dibandingkan Pearson Correlation, Information Gain, Gain Ratio dan Chi-Square. Berdasarkan hasil seleksi dari 5 fitur seleksi yang digunakan, terdapat 5 fitur dominan yang mana fitur tersebut selalu ada pada hasil seleksi menggunakan 5 fitur seleksi. Fitur dominan dari 5 fitur seleksi yakni: SSL Final State, URL of Anchor, Prefix Suffix, Web Traffic dan Having Sub Domain.

5.2 Saran

Hasil penelitian menunjukkan fitur seleksi memiliki pengaruh terhadap kinerja dari model dikarenakan fitur yang akan diklasifikasi merupakan fitur dominan terhadap kelas data. Berdasarkan hasil dari penelitian menggunakan fitur seleksi, terlihat adanya potensi peningkatan kinerja klasifikasi untuk mendeteksi web-phising. Oleh karena itu, pada penelitian berikutnya disarankan melakukan percobaan berikut:

1. Mengganti algoritma Decision Tree dengan algoritma klasifikasi populer lainnya seperti Naïve Bayes, Support Vector Machine atau Random Forest
2. Menggunakan metode fitur seleksi lainnya, seperti wrapper subset selection, hybrid feature selection ataupun metode fitur seleksi populer lainnya.
3. Menggunakan proses dimension reduction menggunakan algoritma Principal Component Analysis (PCA) sebagai langkah untuk proses ekstraksi fitur.