

BAB 5

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil uji coba dan analisis yang telah dilakukan, terdapat hasil yang dapat disimpulkan, yaitu penggunaan *feature selection Recursive Feature Elimination* (RFE) pada algoritma *Support Vector Machine* (SVM) dapat disimpulkan bahwa :

1. Algoritma RFE dan SVM telah berhasil diimplementasikan pada dataset *website phishing*. Penggunaan *feature selection* RFE ditambah dengan evaluasi menggunakan *cross validation* 10, membuat eliminasi fitur semakin akurat dan semakin baik. Penggunaan standarisasi juga penting untuk dataset yang digunakan sehingga tidak terdapat data *outlier* sehingga akurasi yang dihasilkan pada saat melakukan prediksi pada training data dengan algoritma SVM menjadi lebih akurat dan lebih tinggi
2. Pada penelitian ini, hasil prediksi memiliki akurasi yaitu 96.09%. Penggunaan *GridSearchCV* juga memberikan optimasi yang lebih baik dalam menentukan *hyperparameter* SVM. Pada percobaan dengan dataset yang berbeda dengan algoritma SVM dan *feature selection RFE*, terjadi peningkatan yang signifikan. Tanpa RFE, dataset yang berbeda hanya mendapatkan akurasi 56.05%, sedangkan menggunakan RFE meningkatkan akurasi hingga 95.23%. Sehingga dari beberapa uji coba yang dilakukan dapat diberikan kesimpulan bahwa penggunaan *feature selection Recursive Feature Elimination* memberikan dampak yang cukup baik dalam menangani fitur-fitur yang banyak dan membantu algoritma *Support Vector Machine* dalam optimasi akurasi untuk klasifikasi *website phishing*.

5.2 Saran

Berdasarkan penelitian yang sudah dilakukan, terdapat beberapa saran untuk penelitian selanjutnya, yaitu sebagai berikut.

1. Penelitian yang dilakukan memiliki keterbatasan dan kelemahan, yaitu komputasi pada *feature selection* RFE sangat besar karena pengulangan pengecekan fitur yang dilakukan sehingga perlu dilakukan penelitian dengan metode untuk membuat komputasi lebih baik seperti *hybrid feature IGRF-RFE (Iterative Group Recursive Feature Elimination)*
2. Penggunaan metode pencarian *hyperparameter GridSearchCV* memakan komputasi yang tinggi sehingga perlu dilakukan penelitian dengan metode pencarian *hyperparameter* yang lain seperti *Bayesian Optimization* atau *RandomizedSearchCV* yang memiliki pendekatan yang lebih efisien.