

Perbandingan Metode Seleksi Fitur untuk Mengoptimasi Model Support Vector Machine dalam Memprediksi Turnover Pegawai

Ahmad Syafiq Abiyyu¹, Kemas Muslim Lhaksana²

^{1,2}Fakultas Informatika, Universitas Telkom, Bandung

¹ahmadsyafiq@students.telkomuniversity.ac.id, ²kemasmuslim@telkomuniversity.ac.id

Abstrak

Seleksi fitur merupakan salah satu proses yang dilakukan untuk mengurangi dimensi data. Pengurangan dimensi bertujuan untuk meningkatkan performa model algoritma pembelajaran mesin. Turnover pegawai adalah suatu fenomena yang merujuk pada tingkat pegawai yang keluar dari suatu perusahaan. Penelitian mengenai implementasi algoritma pembelajaran mesin dalam memprediksi turnover pegawai sudah banyak dilakukan. Namun, performa model algoritma support vector machine (SVM) secara umum tidak menghasilkan performa yang baik. Dengan menggunakan metode seleksi fitur, hasil performa algoritma SVM diharapkan dapat menjadi lebih baik dalam memprediksi pegawai yang hendak melakukan turnover. Seleksi fitur digunakan pada dataset turnover pegawai sebelum dipelajari oleh model SVM yang dibangun. Metode seleksi fitur yang digunakan adalah filter methods, wrapper methods, dan embedded method. Penelitian ini menampilkan metode seleksi fitur mana yang paling baik dalam meningkatkan performa dari algoritma SVM. Matriks evaluasi seperti akurasi, recall, presisi, dan f1-score digunakan untuk menilai hasil akhir performan dari model SVM setelah dilakukan seleksi fitur. Hasil yang didapatkan adalah metode wrapper method meningkatkan performa dengan lebih baik dibandingkan metode lain. Nilai performa secara keseluruhan naik sebesar 4% dari performa sebelum dilakukan seleksi fitur.

Kata kunci : turnover pegawai, pembelajaran mesin, support vector machine, seleksi fitur

Abstract

Feature selection is one of the processes carried out to reduce the dimensions of data. Dimension reduction aims to improve the performance of machine learning algorithm models. Employee turnover is a phenomenon that refers to the level of employees leaving a company. Research on the implementation of machine learning algorithms in predicting employee turnover has been widely carried out. However, the performance of the support vector machine (SVM) algorithm model generally does not produce a good performance. By using the feature selection method, the results of the SVM algorithm's performance are expected to be better at predicting employees who want to make a turnover. Feature selection is used in the employee turnover dataset before being studied by the SVM model. The feature selection methods used are filter methods, wrapper methods, and embedded methods. This study shows which feature selection method is the best for improving the performance of the SVM algorithm. Evaluation matrices such as accuracy, recall, precision, and f1-score are used to assess the final performance of the SVM model after feature selection. The result obtained is that the wrapper method improves performance better than other methods. Overall performance value increased by 4% from performance before feature selection.

Keywords: employee turnover, machine learning, support vector machine, feature selection

1. Pendahuluan

Latar Belakang

Permasalahan yang dihadapi oleh perusahaan ketika berhadapan dengan karyawannya sendiri salah satunya adalah masalah tingkat perputaran (*turnover*) pegawai dalam perusahaan tersebut. Penting bagi perusahaan untuk dapat mengelola dan mengontrol tingkat perputaran pegawai dengan baik. Tingkat perputaran pegawai sangat berkaitan dengan tingkat rekrutasi suatu perusahaan. Semakin sering pegawai keluar dari perusahaan maka perusahaan juga harus semakin sering untuk melakukan rekrutasi pegawai baru [1]. Hal tersebut, selain dapat menyebabkan pengeluaran biaya yang tinggi, dapat menyebabkan suasana pekerjaan dalam perusahaan menjadi kurang menyenangkan.

Untuk mengatasi permasalahan tersebut, banyak penelitian yang telah membahas tentang bagaimana prediksi *turnover* pegawai menggunakan algoritma pembelajaran mesin. Penggunaan pembelajaran mesin ini dimaksudkan agar hasil analisis mengenai pegawai yang keluar dapat lebih akurat. Algoritma yang digunakan beraneka ragam contoh beberapa di antaranya adalah *decision tree*, *logistic regression*, *support vector machine*, *naïve bayes*, dan algoritma pembelajaran mesin modern lainnya. Dari beberapa hasil penelitian tersebut, beberapa algoritma memiliki hasil yang baik dan beberapa algoritma lain memiliki hasil yang kurang baik. Salah satu algoritma pembelajaran mesin yang memiliki hasil performa yang kurang baik adalah algoritma *support vector machine*.