

INTISARI

ANALISIS PERFORMA RECURSIVE FEATURE ELIMINATION PADA KLASIFIKASI JABATAN SOFTWARE ENGINEER BERBASIS MACHINE LEARNING

Oleh

Puji Resmiati
22/495539/PPA/06310

Tantangan utama dalam rekrutmen IT adalah mengelola data keahlian berdimensi tinggi. Hingga saat ini, belum banyak studi yang secara spesifik menguji dan memvalidasi *subset* fitur keahlian yang paling esensial untuk mempersempit dan memperjelas kriteria penyaringan di tengah tingginya volume pelamar untuk posisi *Software Engineer*.

Penelitian ini bertujuan mengembangkan model prediksi posisi pekerjaan yang efisien berdasarkan pola keahlian yang dimiliki oleh karyawan. Untuk mengatasi kompleksitas dimensi data keahlian, penelitian mengimplementasikan metode seleksi fitur *Recursive Feature Elimination* (RFE) untuk mengidentifikasi fitur-fitur yang paling diskriminatif di setiap posisi pekerjaan. Kinerja model klasifikasi *Logistic Regression* (LR), *Support Vector Machine* (SVM), dan *Random Forest* (RF) kemudian diuji dan dibandingkan pada *subset* fitur terseleksi.

Hasil penelitian menunjukkan bahwa model SVM dan RF mencapai Akurasi Global puncak yang sama sebesar 0.86 setelah optimasi fitur. Strategi RFE terbukti sangat efektif dengan mereduksi 740 atribut data menjadi 50 fitur spesifik (efisiensi 93%), di mana model RF mencatat peningkatan performa paling signifikan sebesar 15% (dari 0.71 menjadi 0.86). Meskipun demikian, SVM diidentifikasi sebagai model yang paling andal karena memiliki stabilitas performa yang lebih tinggi pada berbagai skenario pengujian dan mampu mempertahankan akurasi optimal bahkan pada kondisi data yang sparse. Kesimpulannya, integrasi SVM dengan fitur hasil RFE Global merupakan solusi yang paling efisien, stabil, dan akurat dalam mendukung sistem klasifikasi posisi pekerjaan *Software Engineer*.

Kata Kunci: *Recursive Feature Elimination (RFE), Logistic Regression, Support Vector Machine, Job Classification*

ABSTRACT

PERFORMANCE ANALYSIS OF RECURSIVE FEATURE ELIMINATION IN SOFTWARE ENGINEER JOB CLASSIFICATION BASED ON MACHINE LEARNING

By

Puji Resmiati
22/495539/PPA/06310

The major challenge in IT recruitment is managing high-dimensional skill set data. Currently, only limited studies have specifically tested and validated the most essential subset of features required to narrow and clarify screening criteria amidst a high volume of applicants for Software Engineer positions.

This research aims to develop an efficient job position prediction model based on skill patterns of employed professionals. To address the massive dimensionality of skill features, the study implements the Recursive Feature Elimination (RFE) method to obtain the most discriminative features for each job role. The performance of Logistic Regression (LR), Support Vector Machine (SVM) and Random Forest (RF) classifiers is then tested and compared on the selected feature subsets.

The results of the study indicate that SVM and RF achieved an identical peak Global Accuracy of 0.86 following feature optimization. The RFE strategy proved highly effective, successfully reducing 740 data attributes to 50 specific features (93% elimination efficiency), where the RF model recorded the most significant performance surge of 15% (from 0.71 to 0.86). Nevertheless, SVM was identified as the most reliable model due to its superior performance stability across various testing scenarios and its ability to maintain optimal accuracy even under sparse data conditions. In conclusion, the integration of SVM with the Global RFE feature set provides the most efficient, stable, and accurate solution for the classification system of Software Engineer job positions.

Keywords: Recursive Feature Elimination (RFE), Logistic Regression, Support Vector Machine, Job Classification