

Prediksi *Employee Attrition* Menggunakan Seleksi Fitur dengan *Information Gain* dan Klasifikasi *Random Forest*

Sindi Fatika Sari¹, Kemas Muslim Lhaksana²

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

⁴Divisi Digital Service PT Telekomunikasi Indonesia

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

Abstrak

Employee attrition adalah hilangnya karyawan dalam suatu perusahaan yang disebabkan oleh beberapa faktor, yaitu karyawan mengundurkan diri, pensiun, atau faktor lainnya. Employee attrition dapat berdampak negatif pada suatu perusahaan jika tidak ditangani dengan baik, antara lain penurunan produktivitas. Perusahaan juga membutuhkan lebih banyak waktu dan tenaga untuk merekrut dan melatih karyawan baru untuk mengisi posisi yang kosong. Prediksi attrition ini bertujuan untuk membantu bagian sumber daya manusia (SDM) pada perusahaan untuk mengetahui faktor-faktor apa saja yang memengaruhi terjadinya attrition karyawan. Penelitian ini mengimplementasikan Random Forest dengan membandingkan metode seleksi fitur Information Gain, Select K Best, dan Recursive Feature Elimination untuk mencari seleksi fitur mana yang menghasilkan performansi terbaik. Penerapan metode-metode tersebut mengungguli penelitian sebelumnya dalam hal akurasi, presisi, recall, dan skor f1. Dalam perancangan penelitian ini, penulis pertama mengumpulkan dataset, membuat program, dan menyusun jurnal. Penulis kedua membantu penulis pertama dalam memprogram dan menyiapkan jurnal. Dari hasil pengujian yang telah dilakukan, Information Gain menghasilkan nilai akurasi tertinggi yaitu sebesar 89.2%, sedangkan Select K Best menghasilkan nilai akurasi sebesar 87.8% dan Recursive Feature Elimination menghasilkan nilai akurasi sebesar 88.8%.

Kata kunci : klasifikasi, *employee attrition*, seleksi fitur, *information gain*, *random forest*

Abstract

Employee attrition is the loss of employees in a company caused by several factors, namely employees resigning, retiring, or other factors. Employee attrition of employees can have a negative impact on a company if it is not handled properly, including decreased productivity. The company also requires more time and effort to recruit and train new employees to fill vacant positions. This attrition prediction aims to help the human resources (HR) department in the company to find out what factors influence the occurrence of employee attrition. This research implements Random Forest while comparing Information Gain, Select K Best, and Recursive Feature Elimination feature selection methods to find which feature selection produces the best performance. The implementation of the aforementioned methods outperforms previous research in terms of accuracy, precision, recall, and f1 scores. In preparing this research, the first author collects data sets, makes programs, and compiles journals. The second author assists the first author in programming and preparing the journal. From the results of the tests that have been carried out, Information Gain produces the highest accuracy value of 89.2%, while Select K Best produces an accuracy value of 87.8% and Recursive Feature Elimination produces an accuracy value of 88.8%.

Keywords: *classification*, *employee attrition*, *feature selection*, *information gain*, *random forest*