

PENERAPAN ALGORITMA *COSINE SIMILARITY* DAN TF-IDF DALAM MENENTUKAN RUMPUN JABATAN

Rangga Saputra

ABSTRAK

Pada tahun 2019, sebuah instansi pemerintah memperkenalkan sistem rumpun jabatan untuk meningkatkan efisiensi dalam penugasan jabatan pegawai. Tantangan muncul ketika data pegawai tahun sebelumnya tidak memiliki klasifikasi rumpun jabatan, dan data yang tersedia berupa teks dalam jumlah besar. Rumpun jabatan hanya dikelompokkan berdasarkan kesamaan karakteristik, termasuk keluarga jabatan, fungsi jabatan, dan peran jabatan. Rumpun Jabatan adalah kelompok dari jabatan administrasi yang saling terkait dalam tugas, fungsi, dan kompetensi untuk melaksanakan pekerjaan tertentu. Karena jumlah pegawai yang banyak, informasi yang melimpah, dan data yang dikelola merupakan data teks dalam jumlah yang besar, maka pengklasifikasian manual menjadi tidak efisien. Kondisi ini menyulitkan Biro SDM yang bertanggung jawab atas data pegawai. Untuk mengatasi permasalahan ini, diperlukan metode bantu yang dapat memproses data dengan cepat dan akurat. Salah satu pendekatan yang digunakan adalah *Cosine Similarity* menggunakan metode TF-IDF. Evaluasi hasil menunjukkan bahwa metode ini memberikan nilai *precision* untuk keluarga jabatan dan fungsi secara berurutan mencapai 89% dan 81%. Sementara itu, dalam mengklasifikasikan peran, tingkat *precision* yang dihasilkan rendah sebesar 52%. Meskipun terdapat *error rate* sebesar 26%, model ini dianggap cukup akurat dalam mengidentifikasi kesamaan antara data pegawai dan rumpun jabatan. Dengan toleransi *error rate* sebesar 30%, model ini dapat dipertimbangkan untuk implementasi oleh Biro SDM guna otomatisasi penentuan rumpun jabatan. Meskipun terdapat perbedaan kinerja antara kategori rumpun jabatan, model ini dianggap dapat diandalkan terutama untuk keluarga jabatan dan fungsi. Penelitian ini menyoroti potensi positif dalam penerapan algoritma tersebut sebagai solusi dalam konteks pengelolaan data pegawai dan klasifikasi rumpun jabatan.

Kata kunci: Instansi Pemerintahan; Rumpun Jabatan; *Cosine Similarity*

PENERAPAN ALGORITMA *COSINE SIMILARITY* DAN TF-IDF DALAM MENENTUKAN RUMPUN JABATAN

Rangga Saputra

ABSTRACT

In 2019, a government agency introduced a job cluster system to enhance efficiency in assigning employee positions. Challenges arose as the previous year's employee data lacked job cluster classifications, and the available data consisted of extensive text. Job clusters were only grouped based on shared characteristics, including job family, job function, and job role. Job clusters represent a grouping of administrative positions interconnected by tasks, functions, and competencies for specific job execution. Due to a large number of employees and copious textual data managed, manual classification became inefficient. This posed difficulties for the Human Resources Department responsible for employee data. To address this issue, a supporting method capable of processing data quickly and accurately was necessary. One approach employed was Cosine Similarity using the TF-IDF method. Evaluation results revealed that the method provided precision values for job family and function sequentially reaching 89% and 81%. However, in classifying roles, the generated precision level was low at 52%. Despite a 26% error rate, the model was considered sufficiently accurate in identifying similarities between employee data and job clusters. With a 30% error rate tolerance, this model could be considered for implementation by the Human Resources Department to automate job cluster determinations. Despite performance differences between job cluster categories, the model was considered reliable, especially for job family and function. This research highlights the positive potential of implementing this algorithm as a solution in the context of employee data management and job cluster classification.

Keywords: *Government Agencies; Position Family; Cosine Similarity*