

ABSTRAK

Yohana Setiawati, NPM: 535170032. EKSTRAKSI INFORMASI PADA CURRICULUM VITAE UNTUK DATA PENDAFTARAN KERJA DENGAN METODE SUPPORT VECTOR MACHINE. Skripsi, Jakarta: Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara, Januari 2021.

Pada proses pendaftaran kerja, dibutuhkan beberapa dokumen dari pelamar kerja, salah satunya adalah Curriculum Vitae. Dalam sebuah periode pendaftaran, biasanya perusahaan menerima dokumen CV yang sangat banyak dan memiliki struktur yang berbeda. Oleh karena itu, diperlukan sebuah sistem yang dapat membantu bagian HRD dari perusahaan dalam mengumpulkan informasi-informasi dari CV yang diterima. Sistem ekstraksi informasi CV merupakan sebuah sistem yang dapat mengekstraksi informasi dari CV secara otomatis dan melakukan klasifikasi untuk memisahkan data-data yang ada di dalam CV. Sistem ekstraksi informasi ini dibangun dengan metode Support Vector Machine. Pertama, dikumpulkan CV dalam bentuk pdf yang akan menjadi data latih, lalu CV diubah kedalam bentuk teks. Lalu CV akan melalui tahap tokenization dan tagging, pembobotan fitur dengan menggunakan Boolean Retrieval Model dan RegEx serta pembentukan model dengan menggunakan kernel RBF yang akan digunakan untuk tahap klasifikasi. Hasil dari klasifikasi akan dievaluasi tingkat akurasinya dengan menggunakan Confusion Matrix dengan parameter berupa precision, recall dan F1 score. Hasil yang didapatkan dari pengujian yang dilakukan dengan balanced data set lebih baik dibanding dengan imbalanced dataset, yaitu didapatkan nilai akurasi rata-rata sebesar 69,0628%, precision sebesar 47,6752%, recall sebesar 52,8304%, dan F1 score sebesar 49,7623%.

Kata Kunci: Curriculum Vitae, Support Vector Machine, Kernel, Seleksi Fitur Boolean Retrieval Model, Sistem Ekstraksi Informasi Curriculum Vitae.