

Yopi Rohendi Saputra. 2014. **Aplikasi Search Engine Pencarian Judul Skripsi Tugas Akhir dan Kerja Praktek dengan Algoritma VSM (Vector Space Model)** Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Gunung Djati Bandung. Pembimbing: (1) Edi Mulyana, ST., MT. (2) Wisnu Uriawan, ST., M.Kom.

Kata kunci: Judul Skripsi, Judul Kerja Praktek, Pembimbing, Pencarian, *Text Mining*, *Vector Space Model*.

Banyak skripsi dan kerja praktek yang tersebar dalam berbagai media sehingga menyulitkan dalam pengkategorian, oleh karena itu dibutuhkan media untuk pencarian judul skripsi dan kerja praktek. Media tersebut salah satunya dapat berupa *search engine* pencarian skripsi dan tugas akhir. Banyak referensi pencarian judul skripsi dan kerja praktek yang dihimpun berupa bentuk buku yang didalamnya terdapat kumpulan-kumpulan skripsi dan kerja praktek untuk bahan referensi. Perkembangan teknologi Internet sekarang ini, maka dibangun *search engine* pencarian judul skripsi dan kerja praktek berbasis web yang didalamnya diterapkan aplikasi pencarian. Dengan berkembangnya teknologi web sekarang ini, dapat memberikan kemudahan dalam pengaksesan dan pencarian konten yang berkaitan tentang isi dari *search engine* tersebut serta diharapkan dapat membantu pengguna dalam menemukan informasi sesuai referensi yang dibutuhkan pengguna.

Penelitian ini menggunakan teknik *text mining* dengan algoritma *vector space model* untuk pengukuran kemiripan hasil pencarian terhadap dokumen. Teknik *text mining* terdapat tiga tahapan penting yaitu: *text preprocessing*, *text transformation*, dan *pattern discovery*. Tahapan *text preprocessing* terdiri tahapan pembersihan teks dan pemecahan kalimat menjadi kata-kata (*tokenizing*) sehingga menjadi *stem*. Kemudian, hasil *stem* tersebut akan dianalisis dengan menghitung bobot kata dan kesesuaian dengan *keyword*. Tahapan *text transformation* terdiri dari tahapan *filtering*. Tahapan *pattern discovery* adalah tahapan pembobotan, pembobotan yang digunakan yaitu pembobotan *TF-IDF*. Hasil perhitungan bobot dengan *TF-IDF* akan dilakukan perhitungan bobot kemiripan *document* dengan kata kunci. Dalam mengukur tingkat kemiripan (*similarity*) antara *keyword* terhadap *document* menggunakan algoritma *vector space model* dengan metode *cosine similarity*.

Data yang digunakan untuk melakukan uji coba diketahui bahwa dalam melakukan pencarian agar dapat ditemukan hasil pencarian yang relevan maka masukan kata kunci harus sesuai dengan aturan pencarian judul skripsi dan kerja praktek.