

BAB 1 PENDAHULUAN

Bab ini menjelaskan tentang latar belakang, permasalahan, tujuan, ruang lingkup, metodologi dan sistematika penulisan Tugas Akhir. Bagian latar belakang memberikan gambaran awal mengenai klasifikasi dokumen dan perannya dalam perkembangan teknologi informasi. Bagian permasalahan menjelaskan tentang permasalahan yang melatarbelakangi penelitian ini. Bagian tujuan memberikan penjelasan mengenai hasil yang ingin diketahui dari penelitian dalam tugas akhir ini. Bagian ruang lingkup memberikan penjelasan mengenai batasan dalam melakukan penelitian. Bagian metodologi penelitian menjelaskan tentang cara pengembangan penelitian secara sistematis. Bagian sistematika penulisan berisi penjelasan mengenai masing-masing bab yang ada di laporan penelitian Tugas Akhir.

1.1 Latar Belakang

Teknologi informasi telah berkembang sangat pesat sampai saat ini. Salah satu teknologi informasi dan komunikasi yang berkembang pesat adalah internet. Perkembangan internet sudah merambah ke berbagai lapisan masyarakat, mulai dari anak kecil hingga orang dewasa bahkan yang tua sekalipun. Internet sudah menjadi kebutuhan bagi banyak orang karena dengan internet kita dapat menemukan dan mengakses informasi dari seluruh dunia dengan cepat dan mudah.

Informasi yang diperoleh dari internet dapat berupa dokumen teks seperti dokumen berita, suara, video, maupun objek multimedia lainnya. Informasi tersebut dapat diakses melalui halaman *web*. *Web* memuat banyak informasi yang dihasilkan dari waktu ke waktu secara kontinu dari berbagai sumber. Jumlah informasi yang terus bertambah dari waktu ke waktu dapat menyulitkan para pencari informasi dalam menemukan informasi yang relevan. Salah satu cara yang paling berhasil untuk mengorganisasikan informasi dalam jumlah banyak dan dapat dipahami oleh para

pencari informasi adalah dengan melakukan klasifikasi dokumen berdasarkan topiknya. Ada banyak metode yang digunakan untuk melakukan klasifikasi dokumen. Salah satu metode yang digunakan untuk mengklasifikasikan dokumen adalah dengan menggunakan pendekatan *machine learning* Naïve Bayes [WIB05].

1.2 Permasalahan

Kebutuhan akan dokumen pembelajaran untuk melakukan klasifikasi dokumen merupakan salah satu permasalahan yang sering muncul dalam topik klasifikasi dokumen. Permasalahan lain yang muncul adalah seberapa banyak dokumen pembelajaran yang dibutuhkan agar klasifikasi dokumen memberikan akurasi yang maksimal. Apabila jumlah dokumen pembelajaran yang digunakan terlalu sedikit, maka tidak akan menghasilkan tingkat akurasi yang maksimal. Permasalahan dokumen pembelajaran untuk melakukan klasifikasi dokumen ini dapat diatasi dengan pendekatan baru yang tidak memerlukan dokumen pembelajaran. Pendekatan ini dikenal dengan nama pendekatan ontologi.

Pendekatan ontologi melakukan klasifikasi dokumen berdasarkan pada kemiripan fitur yang ada di dokumen dengan fitur yang ada di ontologi [LSL06]. Dalam penelitian ini ontologi dibuat dengan menggunakan konsep hierarki. Ontologi terdiri dari konsep, relasi antar konsep, fitur konsep, dan batasan untuk klasifikasi dokumen.

1.3 Tujuan

Tujuan utama dari tugas akhir ini adalah untuk membuktikan bahwa klasifikasi dokumen berbahasa Indonesia dapat menggunakan metode ontologi. Selain itu, tugas akhir ini juga bertujuan untuk mengetahui tingkat akurasi klasifikasi dokumen menggunakan metode ontologi dan membandingkan tingkat akurasinya dengan metode *machine learning*.

1.4 Ruang Lingkup

Ruang lingkup pengerjaan tugas akhir ini adalah sebagai berikut.

1. Klasifikasi dokumen menggunakan artikel media massa. Artikel media massa didapat dengan melakukan pengambilan data melalui halaman *web* <http://www.kompas.com>.
2. Jenis fitur yang digunakan adalah pembobotan *term frequency – inverse document frequency* (TF-IDF).
3. Klasifikasi dokumen dilakukan dengan menggunakan metode ontologi dan *machine learning* Naïve Bayes. Kedua metode ini mengklasifikasikan dokumen ke tepat satu kategori dari lima kategori yang ada. Pekerjaan yang dilakukan mencakup pengumpulan data, pemilihan fitur, pemodelan ontologi, persiapan data masukan untuk masing-masing metode klasifikasi dokumen, dan pemanfaatan *tools* atau *library* yang tersedia untuk klasifikasi dokumen.

1.5 Metodologi Penelitian

Metodologi yang digunakan dalam penelitian ini adalah metode eksperimental dan dilaksanakan dalam tahapan-tahapan berikut ini.

1. Studi literatur – Pencarian informasi mengenai klasifikasi dokumen, metode-metode yang dapat digunakan, penelitian yang telah dikerjakan dan pembelajaran mengenai metode-metode yang digunakan untuk melakukan klasifikasi dokumen.
2. Perancangan – Melakukan perancangan eksperimen dengan mempersiapkan data, penentuan variabel percobaan, dan perancangan klasifikasi dokumen menggunakan metode ontologi dan *machine learning*.
3. Implementasi – Melakukan implementasi sesuai dengan perancangan yang telah disusun dengan membuat program untuk pengolahan dokumen dan penggunaan dua metode untuk melakukan klasifikasi dokumen.

4. Analisis hasil – Melakukan perbandingan nilai akurasi dari kedua metode yang dilihat dari penggunaan *stopwords*, penggunaan pemotongan imbuhan (*stemming*), dan jumlah kategori.

1.6 Sistematika Penulisan

Sistematika penulisan Tugas Akhir ini adalah sebagai berikut.

- Bab 1 Pendahuluan - Bab ini berisi penjelasan mengenai latar belakang, permasalahan, tujuan dari penelitian. Bagian pendahuluan juga berisi ruang lingkup, metodologi dan sistematika penulisan Tugas Akhir ini.
- Bab 2 Landasan Teori - Bab ini berisi penjelasan mengenai perolehan informasi, klasifikasi dokumen, dan metode yang digunakan untuk melakukan klasifikasi dokumen. Bagian landasan teori juga berisi persiapan dokumen, pembobotan kata, evaluasi, dan penelitian dalam bidang klasifikasi dokumen.
- Bab 3 Metodologi Penelitian - Bab ini berisi penjelasan mengenai data, persiapan dokumen, pembobotan kata, dan pemodelan ontologi. Bagian perancangan juga berisi metode klasifikasi dokumen meliputi Naïve Bayes dan ontologi.
- Bab 4 Implementasi - Bab ini menjelaskan mengenai implementasi dari perancangan klasifikasi dokumen. Bab ini berisi penjelasan mengenai implementasi persiapan dokumen, pemodelan ontologi, dan metode klasifikasi.
- Bab 5 Hasil dan Pembahasan - Bab ini berisi hasil eksperimen dan pembahasan terhadap hasil eksperimen klasifikasi dokumen dengan menggunakan Naïve Bayes dan ontologi.
- Bab 6 Kesimpulan dan Saran - Bab ini merupakan bab penutup dari laporan tugas akhir yang berisi kesimpulan dan kendala dari percobaan yang dilakukan dengan menggunakan metode ontologi dan Naïve Bayes. Selain itu, pada subbab ini juga diberikan saran-saran untuk pengembangan lebih lanjut dalam penelitian klasifikasi dokumen.