

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Belakangan ini masyarakat banyak disuguhkan berbagai macam situs berita mulai dari *thin-content* dengan kategori berita yang sedikit sampai dengan yang *rich-content* seperti detik.com dengan berbagai macam kategori. Tapi biasanya situs berita tersebut dibebani oleh banyaknya *banner* yang ditampilkan sehingga memperlambat akses untuk membaca berita, terlebih lagi jika *bandwidth* untuk koneksi internet yang digunakan terbatas. Selain itu, tidak setiap orang memiliki banyak waktu untuk membaca semua berita yang disuguhkan pada situs berita tersebut.

Menemukan berita yang relevan sesuai dengan yang diinginkan secara cepat saat ini menjadi masalah yang serius karena ketersediaan berita di internet semakin meningkat jumlahnya. Pencarian dengan *search engine* dengan menggunakan *query* biasa menjadi tidak tepat karena biasanya akan menimbulkan masalah baru. Tidak jarang suatu *query* dapat menghasilkan puluhan bahkan ratusan cuplikan berita. Bahkan dengan tingkat presisi yang rendah, kemungkinan besar berita-berita yang tidak relevan juga akan ikut ditampilkan.

Selain dengan itu diperlukan suatu aplikasi yang dapat mengambil berita dari suatu situs berita secara periodik dan otomatis. Ini akan memberikan kemudahan bagi pengguna dalam memperoleh berita tanpa harus mengakses situs berita tersebut secara langsung. Selain itu pengklasifikasian terhadap berita juga dibutuhkan untuk mempermudah menemukan rangkaian berita yang berkaitan.

Karena pengambilan dan pengklasifikasian berita ini melibatkan data text yang tidak terstruktur, teknik dalam *text mining* dapat dijadikan sebagai solusi untuk menemukan kata atau pola yang diinginkan untuk dijadikan kunci dalam mengambil berita serta mengklasifikasikan berita tersebut. *Text Mining* adalah suatu teknik atau proses yang saat ini banyak digunakan untuk mengambil informasi yang bermanfaat dari dokumen yang berbasis text. Dengan menggunakan *text mining*, menemukan text atau informasi yang diinginkan dengan pola yang *non-trivial* sekalipun dapat dilakukan.

Melihat kenyataan yang ada diperlukan suatu aplikasi yang dapat membantu user untuk mendapatkan berita yang diinginkan dengan mudah, cepat dan tepat. *Internet Agent* adalah suatu aplikasi yang saat ini sedang populer dibicarakan sebagai suatu aplikasi yang digunakan untuk mencari informasi yang diinginkan di internet. *Internet Agent* memiliki kemampuan untuk menjelajah internet untuk melakukan pencarian, penfilteran, dan penyajian informasi bagi user, secara mandiri. Oleh karena itu penerapan *Internet Agent* untuk mengklasifikasikan berita secara otomatis dapat dijadikan suatu pilihan atau solusi untuk mendapatkan berita yang diinginkan dengan mudah, cepat dan tepat.

1.2 Perumusan Masalah

Perumusan masalah dalam penyusunan tugas akhir ini adalah :

1. Bagaimana membuat suatu aplikasi internet agent yang mampu mendownload berita utama dari suatu situs berita secara otomatis.

2. Bagaimana cara mengklasifikasikan dan mendeteksi bahwa suatu berita memiliki keterkaitan dengan berita yang lain.

1.3 Pembatasan Masalah

Dalam perancangan sistem informasi ini, agar tidak menyimpang dari tujuan, maka pembahasan masalah dibatasi pada hal-hal sebagai berikut :

1. Pengambilan berita hanya dilakukan pada situs berita Indonesia dan memiliki *Web Content Management*.
2. Pengambilan berita dilakukan pada provider situs berita yang sudah ditentukan sebelumnya antara lain : www.detik.com, www.kompas.com dan www.jawapos.co.id.
3. Apabila terjadi perubahan pada struktur situs, maka skenario *regex*-nya juga perlu diubah.
4. Pengambilan berita tidak termasuk gambar.
5. Bahasa pemrograman yang dipakai adalah PHP.
6. Database yang digunakan dalam aplikasi ini adalah MYSQL.

1.4 Tujuan

Alasan yang menjadi tujuan dari pembuatan Tugas Akhir ini adalah :
Membuat aplikasi internet agent yang mampu mengambil berita pada suatu situs berita dan mengklasifikasikan berita tersebut secara otomatis.

1.5 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam penyusunan tugas akhir ini adalah sebagai berikut :

1. BAB I : PENDAHULUAN

Dalam bab ini dibahas mengenai latar belakang permasalahan, perumusan masalah, pembatasan masalah, tujuan dan sistematika penulisan.

2. BAB II : LANDASAN TEORI

Dalam bab ini dibahas mengenai landasan teori yang digunakan dalam pembuatan tugas akhir ini yaitu teori mengenai Internet Agent, Fuzzy Search, Association Rule dan Regular Expreition.

3. BAB III : METODE PENELITIAN DAN PERANCANGAN SISTEM

Dalam bab ini dibahas mengenai tahapan-tahapan yang dilalui dalam pembuatan tugas akhir ini, mulai dari menganalisa dan mendesain sistem yang akan dibuat, perancangan DFD dan ERD serta desain uji coba.

4. BAB IV : IMPLEMENTASI DAN EVALUASI

Dalam bab ini dibahas secara mendetail mengenai implementasi penggunaan program dalam mengambil berita pada situs berita yang ditentukan dan mengklasifikasikan berita tersebut serta hasil uji coba dan evaluasi.

5. BAB V : PENUTUP

Dalam bab ini dibahas uraian kesimpulan yang telah di buat beserta saran-saran yang dapat berguna untuk penyempurnaan sistem.

STIKOMMP SURABAYA