

BAB I

PENDAHULUAN

I.1 Latar Belakang

Perkembangan informasi global menuntut penyediaan informasi tersebut dapat dinikmati/dirasakan secara cepat dan tepat. Informasi berupa teks yang diinginkan dapat diakomodasi oleh teknologi komputer khususnya internet, karena internetlah yang menjadi acuan utama beberapa penelitian mengenai penambangan data berbasis teks dilakukan atau yang sering disebut dengan teks *mining*. Oleh karena itu, diperlukan teknik khusus penggalian teks untuk mendeteksi informasi yang relevan dengan cara penyaringan informasi secara berkala supaya menghasilkan informasi yang baik. Informasi yang diolah berupa klasifikasi dokumen teks Bahasa Indonesia yang diambil dari media elektronik. Salah satunya adalah penelitian terhadap dokumen Ekonomi dan Politik, bertujuan untuk mengetahui sejauh mana tingkat akurasi dan perkembangan konten-konten tersebut disediakan.

Salah satu metode klasifikasi dokumen yang dapat digunakan adalah metode *Naive Bayes Classifier (NBC)* dan *support vector machine (svm)*. Penelitian ini berusaha untuk mengklasifikasikan dokumen dengan metode tersebut. Klasifikasi ini ditekankan untuk dokumen berbahasa Indonesia, sementara keterkaitan antar dokumen diukur berdasarkan probabilitas kemudian dibuktikan dengan algoritma klasifikasi lainnya.

I.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan maka permasalahan dari penelitian ini adalah :

- a. Bagaimana mencari nilai probabilitas kata didalam dua dokumen Bahasa Indonesia dengan menggunakan metode *naïve bayes classifier (NBC)*.
- b. Bagaimana membuktikan hasil perhitungan metode sebelumnya dengan metode *Support Vector Machine (SVM)* dalam proses klasifikasi dokumen teks Bahasa Indonesia

I.3 Batasan Masalah

Ruang lingkup dari penelitian ini adalah

- a. Algoritma yang digunakan dalam mencari nilai probabilitas adalah *Naive Bayes Classifier (NBC)*.
- b. Algoritma *Support Vector Machine (SVM)* digunakan untuk membuktikan hasil kata yang dihitung oleh *Naive Bayes Classifier (NBC)*.
- c. Kategori yang digunakan adalah berita Ekonomi dan politik yang diambil dari media *online*.
- d. Berita yang digunakan adalah berita Bahasa Indonesia.

I.4 Tujuan Penelitian

Adapun tujuan yang dilakukan dalam penelitian ini sebagai berikut :

- a. Menerapkan metode *Naive Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)* dalam melakukan klasifikasi dokumen teks berbahasa Indonesia.
- b. Melakukan pengklasifikasian kata berdasarkan dokumen berbahasa Indonesia bertema Ekonomi dan Politik.

I.5 Luaran dan Manfaat

Luaran yang akan dihasilkan dari penelitian ini adalah berupa pengklasifikasian kata menggunakan metode *Naive Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)* untuk dokumen berbahasa Indonesia bertema Ekonomi dan Politik. Adapun manfaat dari penulisan ini adalah sebagai berikut :

- a. Bagi Iptek Sebagai bahan analisis untuk membantu dalam pengambilan keputusan proses pengolahan data dalam jumlah yang besar dan penerapan teknik data mining
- b. Bagi User Dapat membantu mempermudah user dalam memilih dan mengkategorikan dokumen, sehingga meminimalkan waktu dan sumber daya manusia dalam pengklasifikasian dan pencarian dokumen teks itu sendiri.

I.6 Sistematika Penulisan

Skripsi ini ditulis dan diuraikan dalam lima bagian besar, yaitu BAB I sampai dengan BAB 5. Adapun isi dari bab-bab tersebut adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini memberikan gambaran secara garis besar mengenai isi skripsi sehingga pembaca dapat memahaminya dengan mudah. Isi dari bab ini adalah latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, output dan manfaat, serta sistematika penulisan yang digunakan untuk menyusun laporan ini.

BAB II LANDASAN TEORI

Bab ini berisi tentang penjelasan mengenai Klasifikasi Teks, Metode Pengklasifikasian Teks, Teks Mining, Naïve bayes untuk Klasifikasi Teks, Database(MySQL), HTML, PHP, dan XAMPP.

BAB III METODOLOGI PENELITIAN

Bab ini berisi tentang penjelasan mengenai bahan-bahan seperti perangkat keras dan perangkat lunak serta metodologi yang digunakan dalam penelitian ini.

BAB IV PEMBAHASAN

Bab ini berisi tentang penjelasan mengenai hasil uji coba yang telah dilakukan dalam penelitian ini, serta pembahasan atau analisa dari hasil uji coba tersebut.

BAB 5 PENUTUP

Bab ini berisi tentang penjelasan mengenai kesimpulan dan saran yang dapat membangun serta meningkatkan maksud dan tujuan dari penelitian ini menuju arah yang lebih baik

DAFTAR PUSTAKA

RIWAYAT HIDUP

LAMPIRAN