

BAB I

PENDAHULUAN

I.1 Latar Belakang

Semakin berkembangnya teknologi menuntut agar pekerjaan dapat diselesaikan dengan cepat dan meminimalkan campur tangan dari manusia. Semakin banyak pekerjaan yang dikerjakan oleh suatu teknologi maka semakin mempermudah manusia dalam menyelesaikan pekerjaannya. Kategorisasi teks yang sering dijumpai masih banyak yang manual sehingga sistem hanya dapat menentukan kategori teks tersebut berdasarkan yang telah pengguna tetapkan sebelumnya. Kategorisasi teks berdasarkan topik merupakan salah satu metode untuk dapat mengetahui kategori berdasarkan pada suatu topik dari teks yang ada.

Oleh karena itu, dibutuhkanlah teknik khusus untuk mengidentifikasi kata yang relevan terhadap topik dalam sebuah dokumen. Salah satu metode klasifikasi yang terkenal adalah Naive Bayes Classifier (yang kemudian akan ditulis NBC). Kelebihan dari metode ini adalah tingkat akurasi yang tinggi, juga waktu komputasi yang lebih cepat.

I.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dari penelitian ini adalah bagaimana pengimplementasian metode naive bayes dalam proses klasifikasi dokumen teks berbahasa Indonesia.

I.3 Tujuan Penelitian

Adapun tujuan yang dilakukan dalam penelitian ini sebagai berikut :

- a. Dapat mengetahui penerapan metode Naive Bayes dalam melakukan klasifikasi dokumen teks berbahasa Indonesia.
- b. Melakukan pengklasifikasian kata berdasarkan dokumen berbahasa Indonesia bertema pendidikan.

I.4 Batasan Masalah

Adapun batasan-batasan masalah dalam penelitian ini sebagai berikut :

- a. Data yang digunakan adalah data artikel Bahasa Indonesia.
- b. Koleksi dokumen yang digunakan untuk penelitian ini merupakan berkas teks dengan query yang telah ditentukan sebelumnya.
- c. Data yang digunakan sebagai penelitian ini diambil dari e-kompas edisi Desember 2014 sampai dengan Februari 2015.

I.5 Luaran dan Manfaat

Luaran yang akan dihasilkan dari penelitian ini adalah berupa pengklasifikasian kata menggunakan metode naïve bayes untuk dokumen berbahasa Indonesia bertema pendidikan.

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

DAFTAR RIWAYAT HIDUP

LAMPIRAN

