

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Di era yang terus berkembang ini, perubahan yang paling terasa adalah perkembangan teknologi. Dengan adanya internet, semua orang bisa mendapatkan informasi-informasi yang diinginkan asalkan mereka terhubung dengan jaringan internet. Informasi yang diperoleh bisa terdapat dalam berbagai macam bentuk yaitu dokumen teks, gambar, suara, maupun objek-objek multimedia lainnya (Herni Februariyanti dan Eri Zuliarso, 2012).

Salah satu informasi dalam bentuk teks yang paling sering dicari adalah berita. Namun, informasi dalam bentuk teks biasanya sulit dimengerti karena terlalu panjang. Ringkasan merupakan pokok permasalahan dari suatu dokumen atau informasi. Dengan melihat suatu ringkasan, pembaca dapat mengetahui garis besar dari informasi yang ada (Budhi Kurniawan, 2014).

*Text summarization* merupakan bagian dari *natural language processing* yang dapat menghasilkan teks yang lebih pendek dibandingkan dengan teks sumber dan tetap memiliki inti dari informasi dalam teks sumber (Ha Nguyen Thi Thu, 2014).

*Natural language processing* adalah proses pengolahan bahasa natural. Bahasa natural adalah bahasa yang biasanya digunakan oleh manusia untuk berkomunikasi dalam kehidupan sehari-hari (Winfred Phillips, 2006).

Salah satu metode yang dapat diterapkan untuk melakukan *text summarization* adalah dengan menggunakan *Naive Bayes Classifier*. Salah satu contoh penerapan metode *Naive Bayes Classifier* terdapat dalam paper berjudul *An Optimization Text Summarization Method Based on Naive Bayes and Topic Word for Single Syllable Language* yang dilakukan oleh Ha Nguyen Thi Thu (2014). *Naive Bayes Classifier* merupakan algoritma probabilitas yang digunakan untuk klasifikasi data dengan efisien. *Naive Bayes Classifier* juga merupakan salah satu algoritma yang paling banyak digunakan untuk klasifikasi data karena tingkat akurasi (Berry & Kogan, 2012). Dalam penelitian ini *text summarization* akan digunakan untuk mendapatkan ringkasan dari teks sumber. Metode yang akan digunakan pada penelitian ini adalah *Naive Bayes Classifier*.

## **1.2 Rumusan Masalah**

Dari permasalahan yang telah dijelaskan sebelumnya, maka rumusan masalah dari penelitian ini sebagai berikut :

1. Bagaimana mengimplementasikan metode *Naive Bayes Classifier* pada sistem peringkas berita otomatis
2. Seberapa besar tingkat kesuksesan hasil ringkasan dari sistem peringkas berita otomatis

## **1.3 Batasan Masalah**

Dalam penelitian ini, berita yang digunakan adalah berita *online* atau berita yang berasal dari *website-website* yang sudah tersedia. Selain itu, topik berita

yang digunakan juga tidak semuanya, melainkan hanya beberapa topik saja, yaitu ekonomi, olahraga, dan nasional.

#### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah untuk merancang suatu sistem yang dapat meringkas berita secara otomatis dengan menggunakan metode *Naive Bayes Classifier*.

#### **1.5 Manfaat Penelitian**

Manfaat bagi peneliti adalah menambah pengetahuan mengenai *text mining*, *text summarization*, dan *Naive Bayes Classifier*.

Manfaat bagi pengguna adalah mempermudah pengguna dalam mencari atau mendapatkan inti dari suatu berita dan juga mempersingkat waktu yang digunakan untuk mendapatkannya.

#### **1.6 Sistematika Penulisan**

Sistematika penulisan yang digunakan dalam penelitian ini adalah sebagai berikut.

##### **1. BAB I PENDAHULUAN**

Menguraikan latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

##### **2. BAB II LANDASAN TEORI**

Bab ini berisi penjelasan mengenai teori-teori yang terkait dengan penelitian yang dilakukan, yaitu *natural language processing*, *data mining*, *Naive Bayes*,

*text summarization*, dan *text mining*.

### 3. BAB III METODE DAN PERANCANGAN SISTEM

Bab ini berisi tentang metode penelitian dan perancangan sistem yang terdiri dari data *flow* diagram, *flowchart*, dan rancangan *user interface* yang dibangun.

### 4. BAB IV IMPLEMENTASI DAN UJI COBA SISTEM

Bab ini berisi hasil implementasi sistem yang telah dibangun dan hasil pengujian sistem dalam meringkas berita.

### 5. BAB V KESIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan dan saran dari sistem yang telah dibangun.



UMN