

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi tidak hanya berpengaruh pada bidang industri saja, namun juga berdampak pada kehidupan sosial masyarakat. Salah satu indikasinya adalah dengan terciptanya media sosial berbasis online. Pada media tradisional, pengguna hanya sebagai konsumen saja. Informasi mengalir dalam satu arah, yakni dari penerbit konten ke pengguna. Pada media sosial berbeda, pengguna dapat menjadi penerbit konten dan konsumen konten dalam waktu yang bersamaan[1].

Media sosial dari waktu ke waktu semakin berkembang dan semakin banyak penggunaannya. Tahun 2010 sampai dengan tahun 2013 adalah masa dimana perkembangan media sosial sangat pesat[2]. Perkembangan media sosial tidak lepas dari bertambahnya kebutuhan untuk saling berinteraksi satu sama lain. Selain itu perkembangan ponsel pintar juga mempengaruhi perkembangan media sosial.

Di media sosial pengguna bebas mengekspresikan diri. Mulai dari membagikan kegiatan yang sedang dilakukan, menyampaikan kritikan, memberi pujian, dan bahkan sentimen negatif terhadap seseorang atau sebuah produk. Dengan perkembangan media sosial yang pesat, maka bertambah juga data yang ada pada media sosial. Data-data tersebut mengandung informasi jika diolah dengan benar. Oleh karena lahir data mining, yang digunakan untuk menggali informasi yang berguna dari sejumlah data yang sangat banyak[3].

Dengan banyaknya pengguna yang menggunakan media sosial, maka banyak juga sentimen yang ditulis oleh pengguna, khususnya yang ditujukan untuk seseorang, suatu produk, dan peristiwa tertentu. Untuk mengetahui sebuah sentimen terhadap suatu hal, ada beberapa yang dapat dilakukan, salah satunya adalah dengan melakukan survei. Dengan survei, maka akan diketahui sentimen-sentimen masyarakat dari berbagai lapisan, dari pedesaan sampai dengan perkotaan. Namun, untuk melakukan hal tersebut butuh biaya yang sangat besar. Selain itu, juga membutuhkan waktu yang lama.

Media sosial dapat dijadikan alternatif untuk menganalisa sentimen masyarakat. Salah satu media yang populer yang digunakan masyarakat adalah Twitter. Karena data di media sosial sangat banyak, maka perlu diolah terlebih dahulu, kemudian setelah itu diproses dengan metode

tertentu untuk analisis sentimen. Metode yang dapat digunakan untuk analisis sentimen antara lain *Support Vector Machine*, *Naïve Bayes*, dan *Maximum Entropy*[4].

Metode yang digunakan dalam penelitian ini adalah *naïve bayes*. Metode ini mempunyai kelebihan pada akurasi yang cukup baik dan kecepatan yang baik ketika digunakan dalam database berukuran besar[3], dan tidak butuh data dalam jumlah banyak untuk memulai pembelajaran[5]. Sedangkan untuk mereduksi dimensi dari fitur-fitur yang diproses, maka perlu dilakukan seleksi fitur. Dengan menggunakan seleksi fitur, maka diharapkan waktu komputasi akan semakin cepat dan akurasi lebih baik. Metode yang digunakan untuk seleksi fitur dalam penelitian ini adalah *information gain*. Metode ini dapat menyeleksi fitur secara efektif [6].

1.2. Rumusan Masalah

Permasalahan yang akan dibahas dalam penelitian ini adalah :

1. Bagaimana performa metode *naïve bayes* dan *information gain* untuk klasifikasi sentimen.
2. Bagaimana mengatasi teks yang tidak baku.

1.3. Tujuan dan Manfaat Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah :

1. Mengetahui seberapa akurat metode *naïve bayes* dan *information gain* untuk dapat mengklasifikasi sentimen.
2. Menerapkan *preprocessing* untuk mengatasi teks yang tidak baku.

Manfaat dari hasil penelitian ini adalah :

1. Menghasilkan prototipe perangkat lunak klasifikasi sentimen.
2. Berguna untuk penelitian selanjutnya yang ingin memperbandingkan metode ini dengan metode lain.

1.4. Batasan Masalah

Batasan masalah dalam penelitian ini adalah :

1. Data yang akan diproses berupa data teks.
2. Data yang digunakan berasal dari tweet media sosial Twitter.
3. Data tweet yang diambil dicari dengan kata kunci 'via vallen'.

4. Data yang digunakan untuk training sejumlah 100 data sentimen positif, 100 data sentimen negatif, dan 100 data netral.

1.5. Metode Penelitian

Secara umum, penelitian dilakukan dengan beberapa tahapan, yaitu :

- a. Studi Literatur

Pada tahap ini pengumpulan informasi yang dibutuhkan. Informasi didapat dari membaca, mempelajari literatur, makalah, jurnal ilmiah dan juga informasi yang disediakan oleh situs-situs web yang berkaitan dengan penelitian tugas akhir.

- b. Analisa Sistem

Analisa dan perancangan pada tahap ini dibangun sesuai dengan pemahaman dan hasil studi pustaka.

- c. Implementasi Sistem

Mengimplementasikan metode kedalam program sesuai dengan perancangan yang telah dilakukan.

- d. Uji Coba Sistem

Sistem yang telah dibuat akan diuji coba untuk dijalankan.

- e. Analisa dan Evaluasi

Melakukan analisa terhadap hasil klasifikasi untuk mengetahui hasil akurasi dari sistem tersebut serta mengetahui kekurangan dan kelebihan dari metode yang telah digunakan.

- f. Dokumentasi

Dilakukan penyusunan buku sebagai laporan tugas akhir yang berisi dokumentasi tahap awal sampai tahap akhir.

1.6. Sistematika Penulisan

Adapun sistematika pembahasan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini berisikan latar belakang dari penelitian ini, perumusan masalah, batasan masalah, tujuan dan manfaat penelitian, metode penelitian dan sistematika penulisan.

BAB II STUDI PUSTAKA

Pada bab ini dijelaskan mengenai materi-materi yang menjadi landasan dalam penyusunan penelitian ini. Materi-materi tersebut mengenai analisis sentimen, media sosial, *text mining*, *preprocessing*, seleksi fitur, klasifikasi, dan *cross validation*.

BAB III PERANCANGAN SISTEM

Pada bab ini akan menjelaskan rancangan dari seluruh sistem yang akan dibuat dan alur dari sistem yang akan dikembangkan.

BAB IV IMPLEMENTASI SISTEM

Pada bab ini akan dijelaskan implementasi sistem dari rancangan yang telah dibuat pada bab sebelumnya, dimana sistem akan diuji coba berdasarkan skenario yang telah ditentukan.

BAB V PENUTUP

Pada bab ini berisi kesimpulan dari hasil uji coba serta saran yang dapat digunakan untuk pengembangan selanjutnya.