



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Social media merupakan salah satu tren yang berkembang di masyarakat sebagai tempat untuk berinteraksi. Selain untuk berinteraksi dengan lingkungan sosialnya pengguna *social media* seringkali menceritakan mengenai kehidupan sehari-harinya serta berkomentar tentang kejadian yang dialaminya.

Menurut penelitian yang dilakukan oleh Pak (2010), *microblogging* telah menjadi alat komunikasi yang populer diantara pengguna internet. Jutaan pesan disampaikan setiap harinya melalui *web-site* yang menyediakan fitur *microblogging* seperti Twitter, Tumblr, dan Facebook. Pesan tersebut dapat berisi mengenai kehidupan mereka sampai opini mengenai berbagai topik dan diskusi akan isu saat ini. Karena makin banyaknya pesan yang disampaikan mengenai produk dan servis yang mereka gunakan ataupun ekspresi akan pandangan politik dan agama, *microblogging* menjadi sumber berharga untuk opini atau sentimen penggunannya.

Selain itu Pak (2010) juga mengatakan bahwa Twitter sangat cocok untuk menjadi sumber dari analisa sentimen dan *opinion mining* karena penggunaan Twitter untuk mengekspresikan opini mereka terhadap berbagai topik. Kemudian pengguna Twitter juga bervariasi dari pengguna biasa hingga selebritis, perwakilan perusahaan, politikus bahkan presiden.

Berdasarkan penggunaan Twitter sebagai wadah komentar dapat disimpulkan bahwa reputasi baik atau buruk seorang tokoh di masyarakat dapat dilihat dari komentar yang diberikan pengguna Twitter kepada tokoh tersebut. Komentar pada *social media* Twitter dilakukan dengan bahasa sehari-hari atau *natural language* sehingga sulit bagi mesin untuk mengklasifikasikan komentar tersebut bersifat positif atau negatif dan dibutuhkan algoritma untuk mengklasifikasikan komentar.

Willi (2011) membuat penelitian mengenai klasifikasi kepribadian seseorang melalui *social media* Twitter menggunakan algoritma *Naïve Bayes*. Setelah melakukan penelitian tersebut ia berpendapat bahwa akurasi *Naïve Bayes* dalam mengklasifikasi memberikan hasil yang kurang akurat.

Dalam bidang lain, Adiswara (2011) melakukan penelitian klasifikasi data untuk pemantauan distribusi penjualan sepeda motor menggunakan *decision tree*. Berbeda dengan *Naïve Bayes*, penelitian tersebut dianggap berhasil karena klasifikasi dengan *decision tree* sesuai dengan kebutuhan dan mudah dipahami.

Berdasarkan kedua penelitian tersebut maka dibuatlah sebuah sistem yang dapat mengolah *natural language* dengan mengklasifikasikan komentar menggunakan algoritma *classification tree*. *Classification tree* merupakan bagian dari *decision tree* yang biasa digunakan untuk pengklasifikasian. Dengan adanya sistem ini komentar yang tersebar di *social media* mengenai tokoh-tokoh dapat diklasifikasikan dan dapat diketahui tokoh tersebut dinilai negatif atau positif di mata pengguna Twitter.

1.2 Perumusan Masalah

Berdasarkan kebutuhan di atas maka permasalahan yang akan dirumuskan adalah bagaimana membuat suatu sistem yang dapat mengumpulkan data komentar dari Twitter mengenai tokoh terkait dalam bahasa sehari-hari atau natural language dan kemudian mengklasifikasikan komentar tersebut sebagai komentar positif ataupun negatif menggunakan algoritma *classification tree*.

1.3 Batasan Masalah

Luasnya penggunaan social media dan lingkup bahasa yang digunakan maka batasan masalah dapat dijabarkan sebagai berikut.

1. Data komentar Twitter yang digunakan diambil menggunakan Twitter API sebanyak 30 komentar. Hal ini dikarenakan Twitter tidak memberikan seluruh data komentar namun hanya dalam jangka waktu tertentu saja.
2. Tokoh yang akan dijadikan pengujian memiliki akun Twitter. Nama dari akun Twitter tergolong unik dan komentar didapatkan hanya yang ditujukan kepada akun Twitter tersebut.
3. Kata-kata yang mengindikasikan positif atau negatif menggunakan bahasa Indonesia yang baik dan benar, tidak termasuk *slang language* seperti 'woles', 'masbuloh', 'ciyus', 'miapah' dan lain-lain. Selain itu kalimat yang berbentuk sarkasme tidak dapat diklasifikasikan.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah menguji dan mengimplementasikan algoritma *classification tree* untuk mengklasifikasikan kalimat dalam *natural language* dan persentase reputasi suatu tokoh dari komentar di Twitter.

1.5 Manfaat Penelitian

Dengan adanya sistem yang dapat mengklasifikasikan komentar di *social media* maka dapat dihitung persentase komentar positif dan negatif dari akun tokoh tersebut. Masyarakat dan tokoh tersebut dapat mengetahui bagaimana pendapat publik mengenai tokoh tersebut.

1.6 Sistematika Penulisan

Penelitian ini menggunakan sistematika penulisan yang dapat dijabarkan sebagai berikut.

1. Bab I Pendahuluan

Pada bab ini terdapat latar belakang, rumusan masalah, batasan masalah, tujuan penulisan, manfaat penulisan, dan sistematika penulisan.

2. Bab II Tinjauan Pustaka

Bab ini berisi landasan teori yang digunakan untuk mendukung skripsi ini meliputi analisa sentimen, *natural language*, *stemming*, *decision tree*, *classification tree*, opini, dan *social media*.

3. Bab III Metodologi dan Perancangan

Bab ini berisi penjelasan mengenai metode penelitian serta analisa dalam membangun penelitian dengan diagram alir untuk mendukung analisa.

4. Bab IV Implementasi dan Uji Coba

Bab ini berisi tentang implementasi algoritma beserta hasil uji coba sistem yang dibangun.

5. Bab V Simpulan dan Saran

Dalam bab ini dijabarkan kesimpulan dari penelitian dan saran untuk penelitian selanjutnya.