

BAB I

Pendahuluan

1.1 Latar Belakang

Sentiment analysis didefinisikan sebagai tugas meneliti pendapat tentang suatu entitas tertentu. Proses pengambilan keputusan dari seseorang dipengaruhi oleh pendapat yang dibentuk oleh para user. Ketika seseorang ingin membeli produk secara online dia biasanya akan mulai dengan mencari ulasan dan opini yang ditulis oleh orang lain pada berbagai macam tawaran. Sistem yang secara otomatis menentukan sudut pandang yang akan memungkinkan pengguna untuk memahami sebagian besar pendapat yang diungkapkan di internet, mulai dari ulasan produk sampai posisi pada kasus politik[15]. Media jejaring sosial banyak dimanfaatkan para politisi untuk berkampanye dan meningkatkan popularitas mereka menjelang pemilihan umum.

Salah satu media jejaring sosial yang telah dimanfaatkan dalam pemilihan umum adalah Twitter. Hal tersebut dikarenakan Twitter merupakan salah satu media jejaring sosial dengan pengguna terbanyak diantara beberapa situs jejaring sosial yang ada. Analisis sentimen terhadap tokoh publik dari Twitter menjelang adanya pemilihan umum sangat bermanfaat dalam memberikan tambahan wawasan serta gambaran bagi masyarakat tentang tokoh publik yang menjadi kandidat dalam pemilihan umum.

Oleh karena itu, digunakanlah Twitter sebagai dasar untuk percobaan klasifikasi sentimen, dengan cara memasukkan informasi antar pengguna dan relasinya, sebagai berikut. Pertama memanfaatkan model yang didasarkan pada jaringan pengikut / yang diikuti yang memiliki ketergantungan tidak hanya antara pendapat pengguna dan pendapat yang diutarakan dalam *tweet* nya, tetapi juga antara pendapat pengguna dan orang-orang yang diikutinya. Varian “@” berbasis jaringan dipertimbangkan untuk digunakan, dimana pengguna memiliki ketergantungan antara pendapat pengguna lain dan pendapat dari orang-orang yang dapat di *mention* melalui “@”. Alasan mengapa perlu diadopsi pendekatan semi-supervised adalah bahwa akuisisi untuk sejumlah data yang besar dan relevan dengan sentimen-label dapat menjadi rawan akan kesalahan dan dapat memakan waktu yang lama. Penelitian ini berfokus pada level opini karena tujuan akhir bagi banyak pengguna teknologi *opinion-mining* adalah untuk mengetahui apa yang orang pikirkan serta menentukan sentimen yang dinyatakan dalam pesan suatu individu. Selain itu, banyak kasus di mana beberapa *tweet* pengguna yang benar-benar ambigu (mungkin karena *tweet*nya yang sangat pendek dan dibatasi maksimal 140 karakter) tetapi pendapat pengguna secara keseluruhan dapat ditentukan dengan melihat koleksi *tweets*nya dan kepada siapa seseorang dapat terhubung. Pengklasifikasian pada kata sifat juga dilakukan karena banyaknya kata ambigu yang memiliki lebih dari satu klasifikasi kata atau makna, dengan dilakukannya pengklasifikasian kata berdasarkan kata sifat, dapat membantu dalam pengertian maksud dari *tweet* seseorang. Alasan lain menggunakan twitter adalah karena twitter memiliki keterbatasan dalam hal penyampaian kata-kata tidak seperti media social lainnya yang bebas mengekspresikan kata-kata tanpa batasan dan twitter

sendiri dapat menimbulkan sentimen yang berbeda tergantung dari sudut pandang dan pendapat user lain terhadap user/obyek tertentu yang akan diamati.

1.2 Perumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah yang akan dibahas adalah sebagai berikut:

- a. Bagaimana mengetahui tokoh publik yang direkomendasikan untuk difollow oleh user lain?
- b. Bagaimana menganalisis *user opinion* pada *twitter* dengan *sentiment analysis* berdasarkan kata sifat?

1.3 Batasan Masalah

Batasan-batasan masalah dalam pengerjaan penelitian ini yaitu:

- a. Penelitian ini hanya melakukan analisis terhadap *tweet* Bahasa Indonesia
- b. Tokoh yang dinilai dalam penelitian ini hanya dibatasi empat tokoh yang telah mencalonkan diri dalam pemilihan presiden Indonesia 2014
- c. Menggunakan nodexl sebagai tools untuk *crawling*.
- d. Suatu *tweet* dinyatakan kalimat opini jika terdapat minimal satu *opinion words*
- e. Tidak menangani masalah *misspelling*.
- f. Penentuan orientasi opini menggunakan kamus *SentiWordNet* yang sudah diterjemahkan ke dalam Bahasa Indonesia
- g. Dataset yang diambil adalah sebelum Pemilihan umum Presiden Indonesia 2014 dilakukan, yakni pada bulan mei minggu pertama dan ketiga.

1.4 Tujuan

Tujuan dari penelitian ini antara lain:

- a. Melakukan klasifikasi *tweet* yang dimiliki dan ditujukan kepada tokoh publik yang dapat direkomendasikan untuk difollow oleh user lain
- b. Melakukan analisis *user opinion* pada *twitter* dengan *sentiment analysis* berdasarkan kata sifat.

1.5 Metodologi Penyelesaian Masalah

Tahapan yang dilakukan di dalam penyelesaian tugas akhir ini yaitu:

- a. Tahap studi literatur
Mencari referensi yang berhubungan dengan topik tugas akhir ini, yaitu *social media*, *sentiment analysis*, *graph based semi-supervised*, *twitter* dalam bentuk buku, jurnal, paper, dan lain-lain. Selain itu, mempelajari dan memahami materi yang berhubungan dengan topik tugas akhir.
- b. Tahap pencarian dan pengumpulan data

Pengumpulan data merupakan hal yang paling krusial dan paling penting dalam penyelesaian tugas akhir ini. Pada tahap ini akan dilakukan pengumpulan data dari situs *twitter* yang berupa opini berbahasa Indonesia.

- c. Tahap perancangan sistem
Dalam tahapan ini, mulai melakukan perancangan sistem karakteristik user berdasarkan referensi yang sudah didapatkan sebelumnya dari studi literatur dan berdasarkan data-data yang sudah berhasil didapatkan dan digambarkan dalam bentuk diagram alir (*flowchart*).
- d. Tahap implementasi
Implementasi sistem dilakukan setelah sistem selesai dilakukan yaitu dengan menuliskan kode program berdasarkan proses bisnis yang telah dirancang sebelumnya
- e. Analisis hasil
Analisis sistem dilakukan untuk mengetahui tingkat akurasi dengan model *heterogenous graph model-learning* yang dihasilkan berdasarkan hasil implementasi dan pengujian sistem sesuai dengan rumusan masalah yang telah dipaparkan.

1.6 Sistematika Penulisan

Adapun sistematika penulisan Tugas Akhir ini adalah sebagai berikut:

1. BAB 1 Pendahuluan
Pada Bab 1 diuraikan isi dan rencana pengerjaan Tugas Akhir secara keseluruhan yang meliputi latar belakang, perumusan masalah, batasan masalah, tujuan, hipotesa dan metode penyelesaian masalah yang diterapkan.
2. BAB 2 Landasan Teori
Bab 2 memaparkan dasar-dasar teori yang berkaitan dengan *social network*, *opinion mining*, , teori *graph*, klasifikasi, *heterogenous graph model*, *wordnet*, *NodeXL* dan *performance evaluation*.
3. BAB 3 Perancangan dan Implementasi
Perancangan sistem dan contoh implementasi dari sistem yang dibangun dipaparkan pada bab ini. Selanjutnya dilakukan proses implementasi.
4. BAB 4 Pengujian dan Analisis
Pada bab ini dibahas skenario dan hasil pengujian yang dilakukan pada hasil implementasi sistem.
5. BAB 5 Kesimpulan dan Saran
Bab ini berisi kesimpulan dan saran yang didapatkan dari hasil implementasi sistem secara keseluruhan.