BAB 1
PENDAHULUAN

1.1 Latar Belakang

Internet telah menjadi sarana berbagi informasi dan mengekspresikan opini dengan respon yang cepat oleh pengguna di jejaring sosial [1]. Layanan jejaring sosial memungkinkan setiap orang saling berinteraksi dan berbagi informasi tanpa batasan ruang dan waktu. Salah satu jejaring sosial yang populer dan banyak digunakan adalah Twitter.

Twitter adalah *microblogging service* yang telah memperoleh popularitas untuk berbagi informasi dan opini terhadap topik yang bermacam-macam dalam bentuk 140 Karakter pesan yang dinamakan *tweet*, twitter juga dikenal sebagai sarana penyebaran informasi dan mengekspresikan opini [2]. Contohnya selama peristiwa besar terjadi seperti bencana gempa bumi, gunung meletus, badai ataupun peristiwa seperti pemilihan umum, ataupun pendapat terhadap tokoh. Hal-hal tersebut selalu ditanggapi oleh aktivitas di internet [2].

Pengguna twitter menyebarkan *tweet* dengan konten yang berbeda-beda sesuai dengan pola kebiasaan dan gaya menulis pengguna. Konten seperti ini dapat mengandung informasi yang bernilai. Konten yang disebarkan pengguna twitter dapat dikategorikan, secara umum kategorinya yaitu informasi/berita, opini/keluhan, promosi, pernyataan, *random tweet*, pertanyaan, membuktikan eksistensi, obrolan dan *tweet* mengenai pengguna tersebut [3]. Pada penelitian ini akan dilakukan analisis *tweet* yang bersifat opini terhadap partai politik peserta PEMILU 2014.

Analisis sentimen terhadap partai politik dilakukan untuk menganalisa opini negatif dan positif dari pengguna twitter terhadap partai peserta PEMILU 2014. Dalam menentukan polaritas positif atau negatif suatu opini dapat dilakukan secara manual, tetapi mempertimbangkan bertambahnya opini menjadi semakin banyak, tentunya banyak waktu yang akan semakin banyak terpakai. Maka diajukan sebuah metode *machine learning* untuk mengklasifikasikan konten opini dari sumber data yang sangat banyak. Metode *machine learning* untuk melakukan

analisis terhadap konten *tweet* yaitu menggunakan pohon keputusan atau *decision tree learning* yang dibangun dengan algoritma C 4.5. metode ini dipilih karena proses analisis bersifat klasifikasi. Algoritma C4.5 adalah salah satu sistem untuk membangun model klasifikasi yang umum digunakan [4].

Berdasarkan masalah yang diuraikan diatas dibutuhkan sebuah sistem untuk menganalisis sentimen pengguna twitter terhadap partai politik peserta PEMILU 2014 menggunakan algoritma C4.5 *decision tree learning* untuk mengklasifikasikan *tweet*. Hasil dari penelitian tugas akhir ini adalah untuk mengetahui apakah model pohon keputusan yang dihasilkan algoritma C4.5 dalam aplikasi dapat digunakan untuk menganalisa sentimen positif dan negatif pengguna twitter terhadap partai politik. Penelitian sebagai topik untuk tugas akhir ini mengangkat judul **"Analisis Sentimen Pengguna Twitter Terhadap Partai Politik Peserta PEMILU 2014 Menggunakan Algoritma C4.5 *Decision Tree Learning*"**.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan di atas, yang menjadi permasalahan dalam penelitian ini adalah bagaimana menerapkan Algoritma C4.5 untuk membangun pohon keputusan dan membangun aplikasi analisis sentimen pengguna twittter terhadap partai politik peserta PEMILU 2014 untuk menjadi informasi tambahan mengenai partai politik bagi masyarakat.

1.3 Maksud dan Tujuan

Maksud dari penelitian ini adalah untuk menerapkan algoritma C4.5 untuk membangun pohon keputusan untuk analisis sentimen pengguna twitter terhadap partai politik peserta PEMILU 2014.

Sedangkan tujuan yang dicapai untuk penelitian ini adalah:

1. Dapat mengetahui tingkat akurasi pohon keputusan yang dibentuk algoritma C4.5 dalam mengklasifikasi *tweet* sentimen.
2. Dapat menjadi referensi tambahan masyarakat terhadap partai politik.

1.4 Batasan Masalah

Mengingat permasalahan yang dikaji sangat luas, agar penyajian lebih terarah dan mencapai sasaran yang ditentukan, maka diperlukan suatu pembatasan masalah atau ruang lingkup kajian yang meliputi hal-hal sebagai berikut:

1. *Tweet* yang akan dianalisis mengandung kata nama partai politik peserta PEMILU 2014 setelah dilakukan pengambilan data menggunakan Twitter *Streaming* API.
2. Opini yang dianalisis adalah opini terhadap 12 partai politik yaitu Partai Nasional Demokrat, Partai Kebangkitan Bangsa, Partai Keadilan Sejahtera, Partai Demokrasi Indonesia Perjuangan, Partai Golongan Karya, Partai Gerakan Indonesia Raya, Partai Demokrat, Partai Amanat Nasional, Partai Persatuan Pembangunan, Partai Hati Nurani Rakyat, Partai Bulan Bintang, dan Partai Keadilan dan Persatuan Indonesia.
3. Metode untuk klasifikasi *tweet* menggunakan pohon keputusan yang dibentuk algoritma C4.5.
4. Aplikasi ini akan dibangun dengan menggunakan pemodelan terstruktur dengan tools *Data Flow Diagram* (DFD) sebagai analisis kebutuhan fungsional dan *Entitiy Relationship Diagram* (ERD) sebagai analisis basis data.
5. Aplikasi ini dibangun dengan menggunakan bahasa pemrograman PHP dan database MySQL.

1.5 Metode Penelitian

Metode penelitian dalam penyusunan tugas akhir ini dibagi menjadi dua tahap yaitu, tahap pengumpulan data dan tahap pembangunan perangkat lunak.

1.5.1 Metode Pengumpulan Data

a) Studi Literatur

Pengumpulan data dengan cara mencari dan mempelajari bermacam-macam literatur yang dibutuhkan, baik dari buku, jurnal, *paper* dan bacaan-bacaan yang ada kaitannya dengan judul penelitian.

b) Observasi

Pengumpulan data dengan cara melakukan pengamatan langsung terhadap *tweet* yang bermunculan yang membahas tentang pemilihan umum 2014 dan partai politik pesertanya.

1.5.2 Metode Pembangunan Perangkat Lunak

Teknik analisis data dalam pembuatan perangkat lunak menggunakan paradigma perangkat lunak secara *waterfall*, yang meliputi beberapa proses diantaranya:

a. *System/Information Engineering*

Merupakan tahap yang terbesar dalam pengerjaan aplikasi, dimulai dengan mengumpulkan *tweet* menggunakan Twitter Streaming API.

b. *Analisis*

Merupakan tahap menganalisis hal-hal yang diperlukan dalam pelaksanaan pembuatan perangkat lunak. Pada tahap ini juga dilakukan analisis terhadap pohon keputusan yang dibangun algoritma C4.5.

c. *Design*

Design adalah tahap menerjemahkan dari keperluan-keperluan yang dianalisis ke dalam bentuk yang lebih mudah dimengerti oleh pemakai.

d. *Coding*

Coding adalah hasil perancangan sistem diterjemahkan ke dalam kode-kode dengan menggunakan bahasa pemrograman PHP.

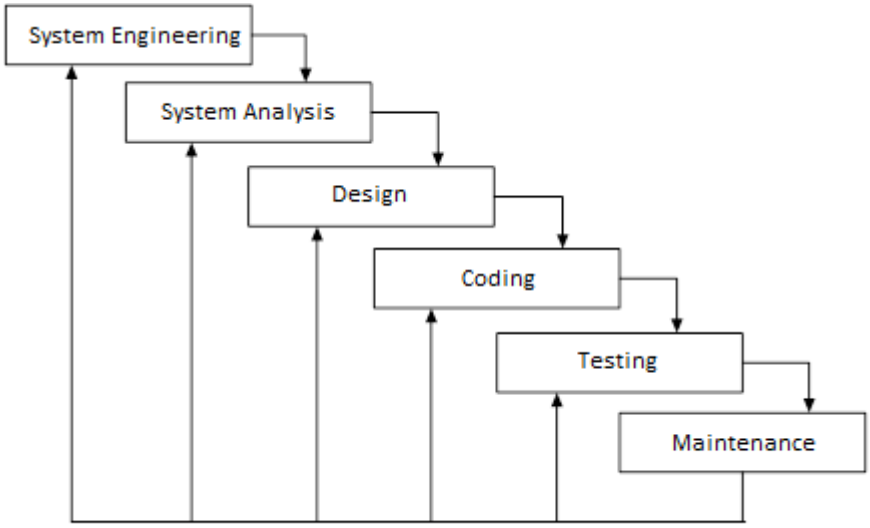
e. *Testing*

Testing adalah tahap pengujian terhadap program yang telah dibuat. Pengujian ini dilakukan untuk memastikan tidak ada kesalahan dalam melakukan klasifikasi terhadap *tweet* dan memberikan hasil yang sesuai dengan yang diinginkan.

f. Maintenance

Tahap akhir dimana suatu perangkat lunak yang sudah selesai dapat mengalami perubahan–perubahan atau penambahan sesuai dengan

permintaan *user*. Pada program yang dibangun ini tidak dilakukan tahap *maintenance*.



Gambar 1.1 Skema Model *Waterfall* [14]

1.6 Sistematika Penulisan

Sistematika penulisan penelitian ini disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah, rumusan masalah, maksud dan tujuan, metodologi penelitian, dan sistematikan penulisan.

BAB II LANDASAN TEORI

Bab ini berisi landasan teori yang membahas tentang *Text Mining*, *Text Pre-processing*, Klasifikasi, Algoritma C4.5, Basis data, *Database Management System*(DBMS), Pemodelan Analisis, ERD, DFD, MySQL, PHP, HTML yang digunakan untuk membangun aplikasi ini.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini berisi analisis sistem, analisis masalah, perancangan sistem, pengenalan aplikasi yang dibangun, definisi kebutuhan perangkat lunak, perancangan sistem, perancangan antarmuka.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi implementasi sistem yang meliputi kebutuhan perangkat keras, kebutuhan perangkat lunak, implementasi basis data, implementasi antarmuka. Pengujian alpha dan pengujian beta.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang didapat selama penulisan laporan skripsi dan juga saran untuk pengembangan dan penyempurnaan aplikasi ini di masa akan datang.