

## BAB I

### PENDAHULUAN

#### 1.1. Latar Belakang

Pada tahun 2016, pengguna internet di Indonesia telah menembus angka 102,8 juta (kominfo.go.id, 2014). Internet menjadi alat komunikasi yang sering digunakan masyarakat untuk berkomunikasi dengan pengguna internet lainnya, seperti melalui aplikasi *Chatting* (*Facebook*, *Twitter*, *Yahoo*, dan lain-lain) (Kompasiana, 2015).

Situs *microblogging* seperti *Twitter* telah menjadi alat komunikasi yang sangat populer di kalangan pengguna internet. Jutaan pesan muncul setiap hari di situs web populer ini (Alexa, 2013). Isi dari pesan yang kerap muncul setiap harinya kebanyakan berisi tentang kehidupan sehari-hari dari pengguna internet. Situs *microblogging* juga sering digunakan untuk berbagi informasi atau opini mereka terhadap suatu topik atau isu-isu yang sedang hangat diperbincangkan, ditambah lagi dengan didukungnya format pesan yang bebas dan aksesibilitas dari berbagai platform yang mudah, pengguna internet cenderung untuk beralih dari blog atau *miling list* ke layanan *microblogging* (Agarwal, 2011).

Twitter merupakan salah satu jejaring sosial yang memiliki pangsa pasar global sebesar 43% dibanding dengan jejaring sosial lainnya. Dengan rincian, pengguna aktif sebesar 21% dari pengguna internet global. Dapat dijabarkan pula ada 58% pengguna aktif *Twitter* secara

rutin mengakses Twitter (Wikanargo, 2016), sehingga Twitter menjadi salah satu jejaring sosial yang menyimpan banyak informasi atau biasa yang disebut dengan *big data*.

Untuk menganalisa data yang besar, diperlukan server yang mendukung penyimpanan data-data tersebut. Situs jejaring sosial sendiri mempunyai server yang dapat menyimpan *big data* dengan baik, segala sesuatu yang dilakukan oleh pengguna jejaring sosial, termasuk hal menyampaikan opini dalam jangka waktu yang cukup lama (Wikanargo, 2016). Dengan adanya dukungan penyimpanan data oleh server, maka memudahkan untuk dilakukannya proses selanjutnya seperti menganalisa data yang berisi informasi atau opini-opini yang telah tersimpan pada server. Data-data tersebut dapat diolah dengan menggunakan teknik *data mining* untuk mencari pola atau aturan tertentu dari *big data*, dan dapat dianalisis dengan menggunakan salah satu metode dari *data mining*, misalnya metode *Naïve Bayes Classifier*. Metode *Naïve Bayes Classifier* bekerja sangat baik dibanding dengan model *classifier* lainnya dan memiliki tingkat akurasi yang lebih baik dibanding model *classifier* lainnya (Xhemali, Hinde, & Stone, 2009). Metode *Naïve Bayes Classifier* akan menganalisis dengan asumsi independen atau naif yang kuat (Ganisaputra & Tan, 2013).

Menjelang pemilihan Gubernur DKI Jakarta 2017, ada ribuan percakapan per hari mengenai pemilihan Gubernur DKI Jakarta di media sosial, salah satunya melalui Twitter (Sari, 2016). Opini masyarakat Indonesia mengenai pemilihan Gubernur DKI Jakarta 2017 pun beragam, ada yang positif, dan ada yang negatif. Hal ini menjadi menarik

dan penting bagi pihak-pihak tertentu seperti pasangan calon Gubernur DKI Jakarta 2017 Basuki Tjahaja Purnama dan Djarot Saiful Hidayat atau bagi tokoh publik yang berperan di dalamnya, yang ingin mengetahui sentimen atau pendapat masyarakat Indonesia mengenai calon Gubernur Basuki Tjahaja Purnama menjelang pemilihan Gubernur DKI Jakarta 2017.

### **1.2. Rumusan Masalah**

Berdasarkan uraian latar belakang di atas, maka dapat dirumuskan masalah yaitu :

1. Bagaimana mengamati dan menganalisis opini masyarakat Indonesia mengenai calon Gubernur Basuki Tjahaja Purnama menjelang pemilihan Gubernur DKI Jakarta 2017 melalui *tweets* pada Twitter?
2. Bagaimana mengklasifikasikan sentimen pada *tweets* yang akan diuji?

### **1.3. Batasan Masalah**

Batasan masalah untuk penelitian ini memiliki ruang lingkup sebagai berikut :

1. Data yang digunakan untuk penelitian ini adalah data *tweets* yang menggunakan Bahasa Indonesia.
2. Calon Gubernur yang akan dikenakan analisis opininya adalah Ahok atau Basuki Tjahaja Purnama, terkait dengan opini positif dan negatif dari masyarakat Indonesia menjelang pemilihan Gubernur DKI Jakarta 2017.

3. Metode yang digunakan untuk pengklasifikasian dalam penelitian ini adalah metode *Naïve Bayes Classifier*.
4. Lama waktu pengambilan data dalam penelitian ini adalah dimulai dari tanggal 03 Oktober 2016 sampai dengan tanggal 25 Januari 2017.
5. Hasil keluaran pengklasifikasian sentimen pada penelitian ini dengan menggunakan metode *Naïve Bayes Classifier* berupa sentimen positif, negatif, dan netral.

#### **1.4. Tujuan Penelitian**

Tujuan dari penelitian ini adalah :

1. Mengamati dan menganalisis opini masyarakat Indonesia mengenai calon Gubernur Basuki Tjahaja Purnama menjelang pemilihan Gubernur DKI Jakarta 2017 melalui *tweets* pada Twitter.
2. Mengklasifikasikan sentimen pada *tweets*

#### **1.5. Metode Penelitian**

Metode yang digunakan dalam penelitian ini meliputi:

##### a. Studi Literatur

Studi literatur atau studi pustaka yaitu dengan mencari dan mempelajari sumber referensi dan literatur yang berkaitan dengan penelitian *Data Mining* dan *Naïve Bayes Classifier* dengan sumber data dari Twitter mengenai opini pemilihan Gubernur

DKI Jakarta 2017 terhadap calon Gubernur Ahok.

b. Metode Pembangunan Perangkat Lunak

Metode pembangunan perangkat lunak meliputi:

1. Analisis Kebutuhan Perangkat Lunak

Analisis kebutuhan perangkat lunak dilakukan dengan cara mencari data atau informasi terlebih dahulu mengenai opini pemilihan Gubernur DKI Jakarta 2017 terhadap calon Gubernur Basuki Tjahaja Purnama, sehingga menghasilkan informasi untuk dianalisis. Kemudian pada tahap ini akan dibentuk Spesifikasi Kebutuhan Perangkat Lunak (SKPL).

2. Perancangan Perangkat Lunak

Perancangan perangkat lunak dilakukan untuk mendapatkan deskripsi arsitektural perangkat lunak, deskripsi antarmuka, deskripsi data, dan deskripsi prosedural. Hasil perancangan berupa dokumen Deskripsi Perancangan Perangkat Lunak (DPPL).

3. Implementasi Perangkat Lunak.

Implementasi perangkat lunak adalah proses penulisan program yang merealisasikan rancangan sistem yang dikembangkan dengan menggunakan Bahasa pemrograman PHP dengan pengelolaan basis data menggunakan MySQL.

#### 4. Pengujian Perangkat Lunak

Pengujian dilakukan untuk menguji fungsionalitas perangkat lunak, apakah telah berjalan dengan baik atau belum.

#### **1.6. Sistem Penulisan Tugas Akhir**

Sistematika penulisan laporan tugas akhir adalah sebagai berikut :

BAB 1 PENDAHULUAN

BAB 2 TINJAUAN PUSTAKA

BAB 3 LANDASAN TEORI

BAB 4 ANALISIS DAN PERANCANGAN SISTEM

BAB 5 IMPLEMENTASI DAN PEGUJIAN SISTEM

BAB 6 KESIMPULAN DAN SARAN