

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

*Twitter* merupakan sosial media yang sampai saat ini sangat digemari dan menjadi penyebar informasi yang sangat cepat. Informasi yang beredar juga sangat banyak mulai dari berita, opini, komentar, kritik dan semuanya ada yang bersifat positif, negatif, dan netral. Menurut data yang dilansir Statista berdasarkan hasil penelitian PeerReach, Indonesia termasuk pengguna *twitter* yang paling aktif ke 3 di dunia dibawah Amerika Serikat dan Jepang. Sentimen analisi ini bertujuan untuk menentukan sentimen publik terhadap objek tertentu yang disampaikan di *twitter* dengan bahasa Indonesia yang dapat membantu usaha untuk melihat opini publik terkait dengan objek tersebut sebagai bahan pertimbangan pengambilan keputusan.

Internet yang sudah menjadi kebutuhan saat ini dan disertai dengan kemudahan aksesnya menjadikan informasi akan tersebar dengan cepat dan luas. Hampir setiap individu masyarakat sudah terikat dengan yang namanya smartphone. Smartphone menjadi salah satu alat pendukung untuk menjadikan internet sebagai kebutuhan utama. Penggunaan media sosial menjadi pemicunya. Masyarakat dapat dengan mudah memberikan opininya melalui media sosial. Salah satu media yang sampai saat ini masih efektif adalah *Twitter*. *Twitter* menjadi salah satu media sosial yang banyak digunakan di Bandung dimana kota Bandung sempat menjadi kota peringkat 6 dunia jumlah pengguna *twitter*. Dilihat dari aktifnya pengguna *twitter* sendiri akan memudahkan untuk pengumpulan opini dibanding dengan melakukan *survey* ataupun kuisisioner. Dengan banyaknya penyampaian opini, maka akan banyak opini yang terkumpul yang kemudian bisa dimanfaatkan sebagai bahan penilaian dan evaluasi mengenai Kota Bandung.

Untuk menghasilkan informasi dari opini yang ada, maka dilakukan pengolahan data dengan analisis sentimen yang dapat dijadikan salah satu solusi

permasalahan diatas yaitu dengan memisahkan opini ke dalam kelas sentimen positif, negative dan netral, serta pengambilan kesimpulan tentang faktor apa saja yang sering dibahas dalam opini-opini tersebut. Sebagai pejabat publik akun @ridwankamil merupakan akun yang sangat aktif. Isi dari *tweet* baik berupa informasi mengenai Kota Bandung dan tidak sedikit menjawab pertanyaan ataupun keluhan publik. Cukup banyak interaksi pada *tweet* @ridwankamil. *Tweet* yang digunakan adalah mention ke akun @ridwankamil. Ini nantinya semoga dapat membantu Ridwan Kamil untuk melihat umpan balik dari masyarakat terhadap kinerja pemerintah kota. Masyarakat juga dapat langsung mengeluhkan permasalahan kota Bandung melalui *twitter*.

Dari latar belakang untuk permasalahan diatas maka penulis pada penelitian ini mengambil judul “ Sentimen Analisis Twitter Berbahasa Indonesia” untuk melakukan klasifikasi opini dengan menggunakan metode *lexicon based* dan *naïve bayes*. Pemilihan penggunaan metode *lexicon based* dan *naïve bayes* ini mempertimbangkan pada peneltian-penelitian sebelumnya, penerapan metode *naïve bayes* yang menggunakan teori probabilitas walaupun perhitungan yg sederhana tetapi dapat menghasilkan nilai akurasi yang tinggi. Digabungkan dengan penerapan metode berbasis kamus berisikan daftar kata positif, negatif dan netral untuk membuktikan apakah penggunaan dua metode ini efektif untuk melakukan klasifikasi dalam menghasilkan nilai akurasi tinggi.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka dapat dirumuskan beberapa permasalahan sebagai berikut :

1. Bagaimana menganalisis sentimen bahasa indonesia dengan *Lexicon based* dan *Naïve Bayes*?
2. Bagaimana bagaimana tingkat akurasi metode *Lexicon based* dan *Naïve Bayes* pada sentimen analisis *tweet* berbahasa Indonesia?

### 1.3 Tujuan dan atau Manfaat Penelitian

Penyusunan laporan tugas akhir ini bertujuan untuk :

1. Menganalisis dan mengklasifikasikan sentimen pada sebuah *tweet* dengan menggunakan metode *Lexicon based* dan *Naïve Bayes* untuk mendapatkan kategori sentimen yang sesuai.
2. Menguji akurasi yang dihasilkan oleh metode *Lexicon based* dan *Naïve Bayes* pada *tweet* berbahasa Indonesia.

### 1.4 Batasan Masalah

Dari rumusan masalah yang ada maka dapat diberikan batasan sehingga pembahasan lebih terarah. Batasan-batasan yang menjadi acuan dalam tugas akhir ini adalah sebagai berikut :

1. *Tweet* yang digunakan adalah *tweet* yang berasal dari *Twitter* dengan kata kunci @ridwankamil dan telah diklasifikasikan secara manual (berdasarkan logika penulis) terlebih dahulu ke dalam tiga kelas, yaitu positif, negatif, dan netral.
2. *Tweet* yang menjadi dataset tidak melalui proses subjectivity classification (menentukan kalimat yang merupakan opini atau tidak)
3. *Tweet* yang digunakan hanya *tweet* berbahasa Indonesia
4. *Tweet* berupa teks
5. Metode yang digunakan adalah berbasis kamus (*Lexicon-Based*) berdasarkan Hu dan Liu opinion lexicon yang telah di terjemahkan kedalam Bahasa Indonesia dan *Naïve Bayes* berdasarkan perhitungan James Hennessey.

### 1.5 Sistematika Penulisan

Sistematika penulisan dari tiap bab dalam laporan akhir ini adalah sebagai berikut:

**BAB I PENDAHULUAN** ; Bab ini membahas tentang latar belakang masalah, identifikasi masalah, maksud dan tujuan pelaksanaan tugas akhir, kegunaan tugas akhir, serta sistematika penulisan.

**BAB II LANDASAN TEORI** ; Bab ini menjelaskan tentang teori-teori yang berkaitan dengan permasalahan yang di bahas pada tugas akhir.

**BAB III METODOLOGI** ; Bab ini akan membahas mengenai metodologi penelitian yang digunakan dalam tugas akhir ini.

**BAB IV ANALISIS DAN PERANCANGAN** ; Bab ini menjelaskan tentang perancangan yang diperlukan dalam pembuatan aplikasi dan mengenai kebutuhan-kebutuhan dan dalam melakukan penelitian.

**BAB V IMPLEMENTASI DAN PENGUJIAN** ; Bab ini membahas mengenai tahapan-tahapan apa saja yang dilakukan dalam proses implementasi. Bab ini juga membahas mengenai hasil dari proses implementasi dan analisis.

**BAB VI PENUTUP** ; Bab ini dituliskan kesimpulan dan saran terhadap penelitian yang telah dilakukan.