

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Di Era globalisasi ini, Indonesia mengalami perkembangan teknologi yang sangat pesat, dimana dalam perkembangannya Indonesia mampu menyetarai sistem informasi yang semakin bermacam ragamnya (A. Elsa, A. Kartika, Z. Rafika, 2022). Kebanyakan warga negara memanfaatkan kemajuan teknologi informasi ini untuk berkomunikasi. Komunikasi di Indonesia sendiri sekarang lebih mudah dan banyak ragamnya salah satunya menggunakan akun media *Twitter*. *Twitter* diciptakan oleh Jack Dorsey di tahun 2006 dan pertama meluncur di dunia maya saat Juli 2006 dengan alamat <http://www.twitter.com> yang masih digunakan hingga saat ini. *Twitter* adalah salah satu media yang bagi pengguna media social di Indonesia dinyatakan keakuratannya (A. Elsa, A. Kartika, Z. Rafika, 2022). Biasanya pengguna media social memilih informasi dari *Twitter* untuk pengupdatean berita setiap harinya (A. Elsa, A. Kartika, Z. Rafika, 2022). Kemudian pada oktober 2022 Elon Musk mengutarakan rencananya untuk membuat aplikasi serba bernama X yang diakselerasi dengan mengakuisi *Twitter* dan mengubah nama perusahaan *Twitter Inc* menjadi *X Corp*. Pada bulan juli 2023 *twitter rebranding* menjadi X. *Twitter* tidak lagi menggunakan logo burung birunya yang kemudian digantikan oleh huruf X yang berwarna putih dengan background berwarna hitam. Elon Musk berkata diakunnya “

Dan segera kami akan ucapkan adieu (selamat tinggal) pada merk Twitter dan semua burung” pada senin 23 juli 2023 (CNN Indonesia, 2023).

Indonesia menepati posisi ketiga dalam penyumbang twit terbanyak dengan jumlah 1 miliar twit, dibawah Amerika Serikat (3,7 miliar) dan Jepang (1,8 miliar). Bahkan, Jakarta menjadi *Twitter Capital City*, yakni kota dengan jumlah twit terbanyak dan teraktif di dunia (MIT *Technology Review*, 2013). Adanya pengguna *twitter* yang banyak di Indonesia, membuat pengguna *Twitter* di negara lain membuat sebuah produk besar maupun kecil dan berlomba-lomba melakukan pemasaran via *Twitter* dan kemudian mereka menunggu *reply* atau *mention* dari pengguna lain untuk mengetahui bagaimana respon atas produk mereka.

Seiring dengan kemajuan teknologi dan informasi, kini tilang telah menggunakan sistem elektronik yang salah satunya dikenal dengan sistem *E-tilang*. Sistem *E-tilang* yang dimaksud dalam penelitian ini akan menggantikan sistem tilang manual yang menggunakan blanko/surat tilang, dimana pengendara yang melanggar akan dicatat melalui aplikasi yang dimiliki personil kepolisian. Sistem *E-tilang* mempunyai tujuan untuk mencegah praktik pungutan liar yang beberapa kali dilakukan oleh oknum polisi lalu lintas yang meresahkan masyarakat maupun masyarakat sendiri yang menawarkan suap kepada oknum polisi lalu lintas. Landasan hukum dari dilaksanakannya sistem *E-tilang* yaitu Undang-undang Nomor 22 Tahun 2009 tentang Lalu Lintas dan Angkutan Jalan, Peraturan Mahkamah Agung Nomor 12 Tahun 2016 tentang Tata Cara Penyelesaian Perkara Pelanggaran Lalu Lintas.

Banyaknya kabar mengenai system E-Tilang, membuat warganet salah satunya pengguna *Twitter* ramai memperbincangkan kabar tersebut. Cuitan atau *tweet* masyarakat pada *twitter* mempunyai nilai emosi positif, atau negative, dan juga bernilai netral. Komentar tersebut berupa teks selanjutnya akan diproses dengan salah satu teknik yang digunakan untuk memproses data teks yang biasa disebut *teks mining* yang salah satu bidang *Natural Language Processing* (NLP) karena *teks mining* mengolah data teks sehingga dapat dikomputasi dan menghasilkan informasi yang relevan.

Text mining sebenarnya merupakan bagian dari *data mining* dimana proses yang dilakukan utamanya adalah melakukan ekstraksi pengetahuan dan informasi pola-pola yang terdapat dalam sekumpulan dokumen teks menggunakan alat analisis tertentu (R. Feldman dan J.Sanger, 2007). Salah satu cara yang digunakan yaitu dengan Sentiment Analysis dimana dapat digunakan sebagai alat penentu kebijakan dari opini-opini yang diberikan oleh pelanggan pengguna akun *Twitter* yang berkomentar untuk system E-Tilang.

Sentiment Analysis merupakan perpaduan dari *data mining* dan *teks mining*, dimana Sentiment Analysis sendiri digunakan untuk mengolah berbagai macam opini yang telah diberikan oleh masyarakat atau para pakar melalui media yang ada, opini tersebut diberikan untuk sebuah produk, jasa maupun sebuah instansi (B.Liu, 2012). Pada sentiment Analysis terdapat 3 jenis opini yaitu opini negative, opini positif, dan opini netral. Dan juga ada beberapa algoritma yang dapat digunakan untuk metode sentiment ini, salah satunya yaitu algoritma *K-Nearest Neighbor* (K-NN).

Algoritma K-Nearest Neighbor (K-NN) adalah metode *lazy learning* yang mana tidak ada model yang dipelajari dari data pelatihan, yang ada hanya belajar bila contoh uji harus diklarifikasikan (Kartika Lidya Safitri, dkk, pp.1-8, 2015). *K-Nearest Neighbor* bertujuan untuk mengklarifikasikan objek berdasarkan atribut dan data *training*. Data klarifikasi yang tidak menggunakan data apapun untuk dicocokkan dan hanya berdasarkan memori. Didalam *query* terdapat titik yang akan ditemukan sejumlah k titik *training* yang posisinya paling dekat dengan titik *query*. Klarifikasi akan diketahui dengan adanya voting terbanyak antara klarifikasi k objek. Algoritma K-NN menggunakan klarifikasi ketentangan sebagai nilai prediksi dari sample uji yang baru (Kartika Lidya Safitri, dkk, pp.1-8, 2015).

Dari sini dapat diketahui jika Analisis Sentiment dapat berguna untuk menganalisis komentar-komentar masyarakat yang menggunakan social media terutama *Twitter*, dimana komentar-komentar tadi diterjemahkan menjadi sesuatu yang lebih bermakna. Sehingga komentar tadi bersifat membangun. Jadi, dari beberapa pemaparan data diatas peneliti ingin mengetahui mengenai bagaimana komentar warganet pengguna social media *Twitter* dalam menanggapi kasus E-Tilang. Maka didapat sebuah judul untuk penelitian ini yaitu Analisis Sentimen Data Komentar Sosial Media *Twitter* dengan Metode *K-Nearest Neighbor* (Studi Kasus Pada Penerapan E-Tilang Di Indonesia).

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan rumusan masalah untuk kasus ini yaitu bagaimana cara menganalisis sentimen data komentar sosial media *twitter* dengan metode *K-Nearest Neighbor* studi kasus pada penerapan E-Tilang di Indonesia.

1.3 Ruang Lingkup

Untuk memperjelas pembahasan ini diperlukan ruang lingkup yang jelas. Adapun batasan masalah dalam menganalisis sentiment data komentar social media *twitter* dengan metode *K-Nearest Neighbor* studi kasus pada e-tilang di Indonesia meliputi hal-hal sebagai berikut :

1. Data komentar yang dianalisis dibagi menjadi 3 kategori yaitu positif, negatif dan netral.
2. Data komentar yang dianalisis berasal dari Twitter masyarakat di Indonesia yang men-tweet mengandung kata E-Tilang.
3. Data mulai diambil dari bulan maret – desember 2021.
4. Data uji sebanyak 100 data dari jumlah data 1181 data.
5. Penelitian ini menggunakan modul *Python* untuk scraping data.

1.4 Tujuan Penelitian

Berdasarkan uraian diatas, tujuan penelitian ini adalah untuk menganalisis sentiment komentar social media twitter dengan metode K-NN studi kasus pada E-Tilang.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini diharapkan dapat digunakan sebagai bentuk informasi bagi masyarakat mengenai pemberlakuan e-tilang di Indonesia serta memberikan masukan kepada penegak hukum tentang pelaksanaan e-tilang.

1.6 Sistematika Penulisan

Penulisan skripsi ini dibagi atas lima bab dan masing-masing terdiri dari sub bab yaitu sebagai berikut :

BAB I : PENDAHULUAN

Bab ini berisikan uraian yang memuat tentang segala sesuatu yang melatar belakangi penulis melakukan penelitian dan yang mendasari permasalahan yang terdiri atas latar belakang masalah , rumusan masalah, ruang lingkup, tujuan penelitian dan sistematika penulisan.

BAB II : TINJAUAN PUSTAKA DAN DASAR TEORI

Bab ini menjelaskan mengenai sumber data yang digunakan sebagai referensi dalam kesuksesan penerapan analisis sentiment pada data komentar social media twitter

menggunakan metode K-NN (K- Nearest Neighbor) yang berisi teori tentang Text Mining, pembobotan TF-IDF, K-NN (K- Nearest Neighbor), Confusion Matrix.

BAB III : METODE PENELITIAN

Bab ini membahas tentang data dan peralatan yang diperlukan dalam penulisan dan perancangan system. Kebutuhan system yang meliputi kebutuhan system perangkat lunak dan kebutuhan perangkat keras, metode dan prosedur pengumpulan data, kebutuhan input dan output, serta membahas alur proses pengerjaan dan rancangan tampilan halaman.

BAB IV : IMPLEMENTASI DAN PEMBAHASAN SISTEM

Bab ini memuat langkah dan pembahasan mengenai pengambilan data , pengujian system, visualisasi system.

BAB V : KESIMPULAN

Bab ini berisi tentang kesimpulan dari penelitian yang dikerjakan dan saran yang diberikan pada penelitian selanjutnya.