

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi telekomunikasi terutama dalam bidang telepon seluler berkembang semakin pesat hingga saat ini. Teknologi komunikasi tidak terlepas dari penyedia layanan telekomunikasi seluler. Dalam menunjang kelancaran dan kemudahan komunikasi diperlukan penyedia layanan telekomunikasi seluler yang sesuai dengan kebutuhan. Perusahaan layanan telekomunikasi seluler di Indonesia banyak yang berinovasi dalam meningkatkan mutu pada layanan produknya, seperti biaya bertelepon yang semakin murah, sinyal yang menjangkau ke semua wilayah, paket layanan data internet dengan kuota yang banyak dan koneksi tinggi, bahkan menyediakan teknologi terkini seperti teknologi 4G hingga 5G. Terdapat tiga penyedia layanan telepon seluler yang paling populer di Indonesia yaitu Telkomsel, XL, dan Indosat [1][2]. Dalam melayani penggunanya, perusahaan layanan telekomunikasi seluler memiliki keunggulan masing-masing yang dapat memenuhi kebutuhan penggunanya namun, tidak lepas juga dari kekurangan yang dapat membuat pengguna mengeluh dan mengakibatkan berpindahnya pengguna ke penyedia layanan telekomunikasi seluler yang lain. Kualitas layanan dari setiap perusahaan layanan telekomunikasi seluler memiliki pengaruh terhadap loyalitas penggunanya. Agar konsumen tidak berpindah ke penyedia layanan telekomunikasi yang lain perusahaan perlu memperhatikan keunggulan dan kelemahan produknya

[3].

Dua penyedia layanan telekomunikasi seluler yang populer di Indonesia adalah Indosat dan XL. Pada kuartal kedua tahun 2019 pengguna layanan Indosat dan XL mempunyai jumlah yang sama sebanyak 57 juta pengguna [2]. Jumlah pengikut (*followers*) di Twitter dari kedua penyedia layanan telekomunikasi tersebut juga cukup banyak dibandingkan dengan penyedia layanan telekomunikasi seluler yang lain. Pada bulan Oktober 2019, akun Indosat memiliki pengikut sebanyak 898,8 ribu, sedangkan akun XL memiliki 1,34 juta pengikut [1]. Dari data tersebut dapat disimpulkan bahwa kedua penyedia perusahaan tersebut merupakan penyedia layanan telekomunikasi seluler yang populer dan banyak pengguna yang mengakses maupun berinteraksi dengan akun Twitter dari Indosat maupun XL.

Untuk mengetahui bagaimana opini masyarakat terhadap kedua penyedia layanan telekomunikasi seluler tersebut diperlukan analisis sentimen melalui Twitter pada pengguna yang menggunakan layanannya. Twitter merupakan jejaring sosial dan layanan komunikasi *real-time* yang diluncurkan tahun 2006, digunakan oleh jutaan orang dan organisasi untuk membagikan maupun mencari tahu informasi secara cepat [4]. Fungsi lain Twitter selain digunakan sebagai media sosial, Twitter dapat dimanfaatkan untuk mengetahui sentimen masyarakat terhadap suatu hal [5].

Analisis sentimen pada penyedia layanan telekomunikasi seluler dapat digunakan untuk mengetahui opini masyarakat terhadap layanan yang diberikan. Analisis sentimen adalah suatu proses untuk memahami,

mengekstrak, dan mengolah data teks secara otomatis untuk mendapatkan informasi berupa sentimen yang terdapat dalam sebuah kalimat opini. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini seseorang terhadap suatu permasalahan atau objek, yang kemudian akan diketahui apakah opini tersebut bernilai positif ataupun negatif [6]. Twitter tidak dapat secara mandiri memberikan kesimpulan sentimen masyarakat terhadap suatu hal, oleh karena itu perlu dilakukan analisis sentimen terhadap layanan telekomunikasi seluler agar diketahui kecenderungan opini pengguna layanan telekomunikasi seluler di Indonesia.

Penelitian ini dilakukan menggunakan dua algoritma klasifikasi yaitu K-Nearest Neighbor dan Support Vector Machine. Algoritma K-Nearest Neighbor dipilih karena metode ini mudah diimplementasikan hanya dengan mengatur satu parameter k , bekerja secara lokal dengan menghitung sejumlah nilai k dari data masukan, tangguh terhadap data latih yang mempunyai banyak *noise* dan efektif apabila data latihnya besar [7][8]. Sedangkan algoritma Support Vector Machine dikenal sebagai metode yang mempunyai generalisasi data yang tinggi, mampu menghasilkan model klasifikasi yang baik meskipun data latihnya sedikit, paling kuat dan akurat diantara algoritma lainnya [8]. Dalam sebuah studi tentang 10 besar algoritma dalam *data mining*, K-Nearest Neighbor dikategorikan dalam peringkat delapan. Sementara Support Vector Machine berada di peringkat ketiga [9]. Dalam penelitian ini akan dilihat perbandingan akurasi dan kecepatan waktu proses dari dua jenis algoritma klasifikasi tersebut untuk mengetahui metode yang

mempunyai performa paling baik.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dijabarkan diatas, dapat disimpulkan rumusan masalah antara lain:

1. Kecenderungan opini masyarakat terhadap penyedia layanan telekomunikasi seluler belum diketahui.
2. Penentuan analisis sentimen menggunakan K-Nearest Neighbor dan Support Vector Machine untuk kecenderungan opini masyarakat banyak digunakan, tetapi penerapan pada penyedia layanan telekomunikasi belum diketahui.
3. Performa dan kecepatan waktu proses dari dua algoritma klasifikasi untuk analisis sentimen masyarakat terhadap penyedia layanan telekomunikasi belum diketahui.

1.3 Batasan Masalah

Berdasarkan perumusan masalah diatas agar dalam penyusunan penelitian ini tidak keluar dari pokok permasalahan, maka lingkup permasalahan dibatasi pada:

1. Penyedia layanan telekomunikasi seluler yang dijadikan objek penelitian adalah Indosat dan XL.
2. Sentimen yang dianalisis meliputi sentimen positif dan negatif.
3. Data berupa teks berbahasa Indonesia dan tidak menyertai gambar.
4. Proses pemberian label untuk data latih dilakukan secara manual dengan bantuan manusia.

5. Data yang digunakan dalam penelitian merupakan *tweet* yang didapat dalam rentang tanggal 04 November 2019 hingga 05 Januari 2020.
6. Bahasa pemrograman yang digunakan untuk membangun sistem merupakan Python 3.7.

1.4 Tujuan Penelitian

Penelitian ini dilakukan dengan tujuan sebagai berikut:

1. Menganalisis sentimen pengguna Twitter di Indonesia terhadap dua penyedia layanan telekomunikasi seluler yaitu Indosat dan XL menggunakan algoritma K-Nearest Neighbor dan Support Vector Machine sehingga dapat diketahui kecenderungan opini masyarakat apakah positif atau negatif.
2. Membandingkan akurasi dan kecepatan waktu proses dari dua algoritma yang digunakan.

1.5 Metode Penelitian

Tahapan-tahapan metode yang dipakai dalam penelitian ini adalah sebagai berikut:

1.5.1 Studi Pustaka

Studi pustaka dilakukan oleh penulis untuk menghimpun informasi dari penelitian-penelitian terkait untuk digunakan sebagai acuan penulis dalam melakukan penelitian.

1.5.2 Pengumpulan Data

Data *input* yang akan digunakan bersumber dari *tweet mention* ke akun Indosat dan XL, juga dari hasil pencarian dengan kata kunci seperti “indosat” dan “xl prabayar” untuk mencakup data-data *no mention*. Data yang sudah didapat kemudian diseleksi secara manual untuk menentukan mana *tweet* yang relevan dengan penelitian, lalu dibagi menjadi data latih dan data uji dengan perbandingan 80:20 sesuai dengan *pareto principle*, prinsip ini menyatakan bahwa dari 80% *output* disebabkan oleh 20% *input* [10]. Dalam penelitian ini 80% digunakan sebagai data latih dan 20% digunakan sebagai data uji. Data latih merupakan data *tweet input* yang sudah diberi label sentimen secara manual, sedangkan data uji dibiarkan tidak berlabel.

1.5.3 Analisis dan Perancangan

Dalam mengetahui kebutuhan sistem yang dibutuhkan maka perlu dilakukan analisis sistem. Kemudian akan dirancang sistem yang dapat menganalisis sentimen *tweet*. Tahapan dalam proses perancangan sistem ini antara lain perancangan *pre-processing*, perancangan pembobotan dengan metode TF-IDF, perancangan *training* dan *testing* algoritma K-Nearest Neighbor dan Support Vector Machine, serta perancangan evaluasi model menggunakan metode *K-fold cross-validation*.

1.5.4 Implementasi

Implementasi dilakukan dengan menggunakan rancangan sistem yang telah dibuat pada tahapan sebelumnya. Implementasi sistem pada penelitian ini menggunakan bahasa pemrograman Python 3.7, Anaconda3 2019.07, Twitter API, dan memanfaatkan library seperti Sci-kit Learn untuk *machine learning*, NLTK, dan PySastrawi untuk *pre-processing* data.

1.5.5 Pengujian

Pengujian dilakukan ke sistem yang telah diimplementasikan. Pengujian dilakukan untuk evaluasi model dan mengetahui nilai performa dari kedua algoritma yang digunakan. Teknik yang digunakan dalam penelitian ini adalah *K-fold cross-validation*.

1.5.6 Penulisan Laporan

Hasil dari penelitian kemudian dituliskan dalam sebuah laporan. Pada penulisan laporan penelitian akan ditarik suatu kesimpulan berdasarkan hasil pembahasan dan pengujian, serta diberikan saran untuk penelitian berikutnya.

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini berdasarkan pokok permasalahan agar lebih terarah diurutkan sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi tentang latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini memuat penelitian-penelitian dengan tema sejenis yang sudah pernah dilakukan sebelumnya dan membahas teori yang dijadikan landasan penelitian ini, antara lain Twitter, *text mining*, analisis sentimen, *term weighting*, klasifikasi, K-Nearest Neighbor, Support Vector Machine, *K-fold cross-validation*, dan evaluasi model.

BAB III ANALISIS DAN PERANCANGAN

Bab ini berisi tentang analisis permasalahan dan perancangan pembuatan sistem dalam analisis sentimen terhadap pengguna layanan telekomunikasi seluler.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Bab ini memaparkan pembahasan proses dan hasil dari implementasi dan pengujian yang dilakukan.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil analisis, perancangan, dan implementasi serta saran untuk pengembangan maupun untuk penelitian selanjutnya.