

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan dunia internet yang berkembang pesat pada saat ini dengan diiringi banyak muncul sosial media yang mudah untuk diakses seluruh kalangan masyarakat Indonesia, begitu pula dengan sebuah situs microblogging *Twitter*. Lebih dari enam juta *tweet* perhari ditulis oleh masyarakat Indonesia dalam akun *Twitter* mereka, mulai dari pandangan terhadap sebuah organisasi maupun personal. Situs *microblogging* tersebut dapat menjadi sumber daya yang sangat berharga untuk mendapatkan opini dari masyarakat.

Perusahaan dan organisasi membutuhkan analisis opini sebagai dasar manajemen merk (*brand management*), *corporate reputation*, *marketing*, sistem rekomendasi, dan intelijen. Merk telah dianggap sebagai salah satu *intangible asset* yang paling berharga dan manajemen merk adalah prioritas penting bagi pihak manajemen perusahaan atau organisasi. Di satu sisi, analisis sentimen memiliki banyak manfaat bagi organisasi, industri, dan individu, tetapi di sisi lain memiliki banyak tantangan dan masalah.

Salah satu masalah adalah variasi bahasa yang digunakan pengguna *Twitter* sangat tinggi, dengan kosa kata dan tata-bahasa yang tidak baku. Sebagai contoh, kata “tidak” memiliki 21 variasi seperti “tdk”, “gak”, “nggak”, “kagak”, “ga”, “nda”, “gag” dan seterusnya. Jumlah karakter yang dibatasi 140 membatasi pengguna untuk mengekspresikan dengan jelas apa yang diinginkan. Jumlah *tweet* yang sangat besar juga membuat pemrosesan secara manual akan membutuhkan tenaga dan waktu yang sangat besar.

Salah satu metode klasifikasi yang dapat digunakan adalah metode *Naïve Bayes* yang sering disebut dengan *Naïve Bayes Classifier*. Kelebihan metode *Naïve Bayes Classifier* adalah sederhana tetapi memiliki akurasi yang tinggi. Berdasarkan hasil eksperimen, *Naïve Bayes Classifier* pada penelitian terdahulu terbukti dapat digunakan secara efektif untuk mengklasifikasikan berita secara otomatis dengan akurasi mencapai 90.23%. Algoritma *Naïve Bayes Classifier* yang sederhana dan kecepatannya yang tinggi dalam proses pelatihan dan klasifikasi membuat algoritma ini menarik untuk digunakan sebagai salah satu metode klasifikasi.

Dari penjelasan diatas, selain kita menunjukkan seberapa besar opini dari masyarakat, kita juga akan menganalisa bagaimana cara untuk mengklasifikasikan opini positif, negatif, netral dan akan memasukkannya dalam sebuah peta. Dimana dari peta tersebut dapat melihat daerah mana user yang menyatakan respon positif , negatif maupun netral.

1.2 Perumusan Masalah

Beberapa masalah yang diteliti pada Tugas Akhir ini adalah :

- a) Bagaimana memanfaatkan data *Twitter* untuk melakukan analisis sentimen terhadap produk Telkomsel.
- b) Bagaimana membuat aplikasi yang berguna bagi sebuah perusahaan untuk mengetahui lokasi user yang menyatakan opini positif dan opini negatif terhadap produk.

1.3 Tujuan Dan Manfaat

Tujuan dari penelitian tugas akhir ini adalah :

- a) Sistem dapat menghimpun data dari situs microblogging *Twitter* dengan *crawler* sehingga memudahkan perusahaan untuk melakukan pemasaran.
- b) Sistem analisis sentimen menggunakan data opini *Twitter* diharapkan dapat membantu sebuah perusahaan untuk menganalisa pemasaran dan tanggapan konsumen terhadap produk.

Manfaat dari penelitian Tugas Akhir ini adalah memberikan solusi berupa perangkat lunak yang dapat digunakan untuk mendapatkan opini dari masyarakat tentang kepuasan terhadap produk sebuah produk atau perusahaan dan mengetahui daerah mana yang paling banyak mendapat respon positif dan negatif untuk menunjang pemasaran melalui opini yang diambil dari *twitter*. Agar perusahaan dapat mengetahui pada daerah mana pihak manajemen perlu mengadakan pemasaran produknya, kalau pada suatu daerah banyak mendapat respon negatif.

1.4 Batasan Masalah

Penelitian Tugas Akhir ini menggunakan fasilitas API dari *Twitter* berdasarkan keyword produk Telkomsel untuk mendapatkan opini berbahasa Indonesia yang bekerja secara *real time*. Ruang lingkup penelitian dibatasi klasifikasi positif, negatif pada Produk

Telkomsel. Analisis klasifikasi opini akan dilakukan dengan menggunakan metode *Naive Bayes Classifier* serta akan menampilkan daerah mana yang banyak mendapatkan respon positif atau negatif dengan menggunakan tampilan peta berdasarkan perprovinsi.

1.5 Metodologi Penelitian

Dalam penyelesaian tugas akhir ini dipergunakan metodologi sebagai berikut :

1. Studi Pustaka

Pada tahap ini dilakukan pengumpulan informasi yang diperlukan untuk pembuatan perangkat lunak. Informasi yang diperlukan diperoleh dengan mempelajari literatur-literatur mengenai permasalahan pengkategorian dokumen secara umum, untuk pembobotan menggunakan model *tf-idf*, serta penerapan metode *Naive Bayes Classifier* untuk pembuatan perangkat lunak berbasis web pengklasifikasian komentar dan opini tentang produk pada twitter, serta menampilkannya dalam sebuah peta Negara Indonesia.

2. Perancangan Sistem

Tahap ini meliputi analisis dan desain sistem yang akan dikembangkan dengan mengacu pada hasil pemahaman sistem dan studi pustaka serta data yang telah diperoleh sebelumnya. Pada tahapan ini, desain model data, dan desain proses didefinisikan sehingga dapat memenuhi kebutuhan sistem. Dalam melakukan penelitian ini. Penulis akan melakukan beberapa tahapan yang sudah penulis rancang. Berikut adalah rancangan yang akan penulis lakukan :

1. Pengumpulan data, disini penulis akan melakukan *crawling* tweet untuk mendapatkan tweet dalam jumlah banyak. Kemudian penulis akan melakukan seleksi bahasa pada tweet tersebut. Tweet yang berbahasa Indonesia lah yang akan penulis teliti. Kemudian dari tweet-tweet berbahasa Indonesia tersebut akan dicari lagi tweet yang mengandung produk-produk tertentu. Misal penulis akan meneliti tweet-tweet berbahasa Indonesia yang mengandung kata “Telkomsel”, “Telkomsel Flash”, “Kartu As”, ”Simpati”, “Kartu Hallo”. Karena penulis ingin mengetahui bagaimana sentimen konsumen terhadap produk tersebut.

2. *Preprocessing*, Pada proses ini, penulis akan melakukan penghapusan *stopword* atau *Filtering*, *Tokenizing* atau memisahkan kalimat menjadi kata-kata, memproses emoticon, mengubah *errastic casing*, mengubah kata-kata yang terkapitalisasi, dan menentukan tanda baca yang benar di dalam tweet-tweet berbahasa Indonesia yang mengandung kata “Telkomsel”. Selain proses yang ada diatas, dilakukan pula proses *deep parsing*.
 3. Pembobotan, menghitung tiap kata-kata apakah terkandung dalam sebuah tweet. Dalam tahap ini penulis akan melakukan pembobotan dengan menggunakan *tf-idf*.
 4. Klasifikasi, disini penulis memilih metode naïve bayes classifier untuk proses klasifikasi. Penulis akan mencoba mengklasifikasikan mana tweet yang mengandung sentimen positif dan mana tweet yang mengandung sentimen negatif terhadap produk “Telkomsel”.
 5. Validasi dan Evaluasi, setelah melakukan ujicoba pertama, penulis akan melakukan *cross validation* untuk memperoleh hasil yang lebih optimal.
 6. Visualisasi, hasil dari analisis tweet tersebut akan penulis visualisasikan agar bisa lebih mudah dipahami untuk pengguna aplikasi yang akan penulis buat.
3. **Implementasi Sistem**
Pada tahap ini dibuat implementasi dari desain yang telah dibuat. Mulai dari meng*crawler* data, *preprocessing*, pengklasifikasian dan memasukkan data hasil klasifikasi ke dalam peta Negara Indonesia sesuai daerah berdasarkan provinsi.
 4. **Evaluasi Sistem**
Pada tahap ini akan dilakukan pengujian terhadap perangkat lunak menggunakan data yang sudah dipersiapkan sebelumnya. Ujicoba dan evaluasi perangkat lunak dilakukan untuk mencari masalah yang mungkin timbul, mengevaluasi jalannya program, dan mengadakan perbaikan jika ada kekurangan.
 5. **Dokumentasi**
Pada tahap terakhir ini dilakukan penyusunan buku sebagai dokumentasi dan laporan dari pelaksanaan secara keseluruhan, dari tahap awal hingga tahap akhir pembuatan tugas akhir.

1.6 Sistematika Penelitian

Laporan tugas akhir ini dibagi menjadi lima bagian pembahasan yaitu :

BAB I PENDAHULUAN

Pada bagian ini dipaparkan mengenai hal-hal yang melatarbelakangi penelitian, rumusan dan batasan permasalahan dalam penelitian, serta tujuan dan relevansi atau manfaat penelitian terhadap perkembangan solusi dari permasalahan, serta sistematika penulisan dalam laporan tugas akhir ini.

BAB II TINJAUAN PUSTAKA

Dalam bab ini diuraikan teori-teori penunjang dalam penulisan laporan tugas akhir ini. Pustaka yang digunakan baik dari buku, diktat, maupun materi dari internet.

BAB III PERANCANGAN SISTEM

Dalam bagian ini diuraikan secara rinci rancangan system dari aplikasi tugas akhir ini.

BAB IV IMPLEMENTASI SISTEM

Pada bagian ini diuraikan secara rinci implementasi program serta hasil analisa yang diperoleh.

BAB V KESIMPULAN

Pada bab ini dijelaskan hasil dari implementasi, yang termaktub di dalam kesimpulan. Dan juga saran untuk mengembangkan hasil analisis penelitian tugas akhir ini agar dapat dikembangkan ke penelitian berikutnya.

[HALAMAN INI SENGAJA DIKOSONGKAN]