

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada era modern saat ini banyak sekali teknologi yang terus berkembang dengan pesat. Salah satunya adalah media sosial yang sangat populer serta penggunaannya yang selalu bertambah seperti Twitter, Instagram, Facebook, Path dan lain sebagainya. Adapun pengguna media sosial sendiri mencapai angka 2,2 miliar dengan pengguna mobile mencapai 3,7 miliar. Menurut Kemenkominfo Republik Indonesia, di Indonesia sendiri 80 persen di antaranya terdiri dari remaja 15-19 tahun. Sedangkan pada kenyataannya, remaja menggunakan internet tidak hanya untuk kebutuhan edukasi saja, melainkan dipergunakan juga untuk hiburan, belanja, media sosial dan lain sebagainya[1]. Hal ini yang menjadi fokus utama dalam pelaksanaan penelitian ini adalah tingkat kepuasan dari masyarakat pengguna internet terhadap penyedia layanan internet tersebut.

Pada penelitian ini, media sosial merupakan sumber informasi atau data untuk dilakukannya analisis sentimen atau opini dalam bentuk teks berbahasa Indonesia serta media sosial yang digunakan adalah Twitter. Analisis sentimen adalah sebuah teknik untuk mendeteksi opini terhadap suatu obyek (misalnya individu, organisasi ataupun produk) dalam sebuah kumpulan data[2]. Dalam analisis sentimen hal mendasar yang dilakukan adalah pengelompokan polaritas kata yang berasal dari setiap kalimat atau pendapat pada dua sifat, yaitu positif dan negatif.

Jumlah pelanggan telepon seluler bergerak di Indonesia terus meningkat secara signifikan dari tahun ke tahun. Berdasarkan data ITU (International Telecommunication Union), pada tahun 2000 pelanggan telepon bergerak seluler Indonesia berjumlah sekitar 3,4 juta. Berdasarkan sumber yang sama, pada tahun 2012 angka tersebut meningkat secara signifikan menjadi 280 juta[3]. Dengan peningkatan jumlah pelanggan telepon seluler yang signifikan, apakah penyedia layanan operator seluler sudah memenuhi kepuasan dari para pelanggannya. Pada

penelitian ini akan terpaparkan penyedia layanan operator seluler mana yang lebih mengerti atau memenuhi kepuasan dari opini para penggunanya.

Terdapat beberapa metode *Machine Learning* yang dapat dilakukan untuk melakukan analisis sentimen. Namun pada penelitian ini dilakukan dengan menggunakan metode *K-Nearest Neighbor* (KNN). Pada penelitian lain metode *K-Nearest Neighbor* (KNN) digunakan untuk analisis sentimen terhadap tayangan televisi berdasarkan opini masyarakat dengan menghasilkan tingkat akurasi mencapai 80, 83%[4]. Pada penelitian lainnya, metode *K-Nearest Neighbor* (KNN) digunakan untuk analisis sentimen Twitter berbahasa Indonesia dengan tingkat akurasi mencapai 82,30%[5]. Dengan nilai keakurasian yang cukup tinggi, dapat dibilang metode ini sangat dapat dan cocok digunakan untuk melakukan analisis sentimen. Penulis juga menambahkan penggunaan *part-of-speech* (POS) *tagging* untuk mendukung keakuratan yang lebih tinggi. Dari hasil penelitian yang sudah dilakukan oleh penulis, maka didapatkan nilai rata-rata *precision* sebesar 92,21%, *recall* sebesar 93,74%, *F1-score* sebesar 92,20%, dan *Accuracy* sebesar 98,94%.

1.2 Rumusan Masalah

Adapun rumusan masalah berdasarkan latar belakang di atas adalah bagaimana menghasilkan informasi yang akurat terkait kepuasan pengguna dalam layanan data operator di Indonesia. Dan hasil dari informasi tersebut dapat dimanfaatkan oleh masyarakat sebagai tolak ukur untuk memilih operator dan dapat dimanfaatkan oleh pihak perusahaan penyedia layanan data operator untuk mengetahui opini masyarakat terhadap layanan yang mereka berikan.

1.3 Tujuan

Adapun tujuan yang ingin dicapai pada penelitian ini adalah dapat memberikan informasi kepada masyarakat Indonesia terkait kepuasan pengguna dalam layanan data operator seluler di Indonesia sehingga masyarakat dapat memilih operator seluler manakah yang baik digunakan untuk mereka masing-masing. Serta dapat memberikan informasi kepada perusahaan penyedia layanan data operator telekomunikasi untuk melihat opini konsumen atau masyarakat terhadap kualitas layanan yang sudah mereka berikan.

1.4 Batasan Masalah

Adapun batasan masalah dalam penyusunan tugas akhir ini adalah sebagai berikut.

- a. Metode atau algoritma yang digunakan adalah *K-Nearest Neighbor* (K-NN) dalam proses klasifikasi sentimen.
- b. Jumlah data yang digunakan adalah sebanyak 2.000 data *tweet* dari Twitter terkait dengan tingkat kepuasan masyarakat terhadap operator selular yang diinisialisasi sebagai *provider* A, B, C, D dan dari masing-masing *provider* diambil sebanyak 500 *tweets*.
- c. Proses pengambilan *tweet* terpisah dengan program utama dan disortir secara manual.
- d. Kamus Negatif, Positif serta Netral ditentukan secara manual.

1.5 Metodologi Penelitian

Metodologi penelitian yang dilakukan untuk dapat membantu pelaksanaan, pembuatan, dan penganalisaan tugas akhir ini adalah meliputi :

- a. Studi Literatur
Studi literatur dilakukan dengan cara mencari, mengumpulkan dan mempelajari referensi yang berasal dari berbagai sumber seperti buku, jurnal, paper, dan sumber lainnya sebagai dasar teori.
- b. Perancangan
Perancangan dilakukan dengan menggunakan software JetBrains Pycharm dalam bahasa pemrograman yang digunakan adalah Python.
- c. Pengumpulan Data dan Uji Coba
Pengambilan data diambil dengan menggunakan Twitter API untuk mendapatkan *tweet*. Pengujian program dilakukan dengan memasukan data *tweet* yang sudah terkumpul kedalam program, dalam melakukan analisis sentimen pada *tweet* yang sudah terkumpul.

d. Analisa Data

Berdasarkan data yang diperoleh, akan dibuat analisa tentang kinerja sistem dan tingkat kesalahan serta tingkat ketelitian.

e. Pembuatan Laporan Tugas Akhir

Pembuatan laporan tugas akhir dilakukan untuk mendokumentasikan penyelesaian tugas akhir dalam laporan tertulis.

1.6 Sistematika Penulisan

Tugas akhir ini terdiri dari beberapa bab yang berisi urutan secara garis besar dan kemudian dibagi lagi ke dalam sub-sub yang akan membahas dan menguraikan masalah yang lebih terperinci. Secara garis besar isi dari bab tugas akhir ini adalah sebagai berikut :

BAB I : Pendahuluan

Merupakan uraian mengenai latar belakang, rumusan masalah, batasan masalah, tujuan, metodologi penelitian, dan sistematika penulisan.

BAB II : Tinjauan Pustaka

Berupa uraian konsep dan teori dasar secara umum tentang analisis sentimen, preprocessing POS *Tagging*, serta konsep dasar metode yang akan digunakan yaitu *K-Nearest Neighbor*.

BAB III : Metodologi Penelitian

Pada bab ini dibahas mengenai perancangan sistem data mining. Program yang digunakan menggunakan bahasa python.

BAB IV : Implementasi dan Pengujian

Pada bab ini menguraikan hasil implementasi metode yang digunakan dan menganalisa dari tingkat keakuratan dan kesalahan yang terjadi.

BAB V : Kesimpulan dan Saran

Berisi tentang kesimpulan dan saran terhadap hasil yang diperoleh dari penelitian yang telah dilakukan serta saran untuk pengembangan sistem lebih jauh