

BAB I PENDAHULUAN

1.1. Latar Belakang

Pada masa saat ini *microblogging* sudah jadi layanan komunikasi yang benar-benar populer di kalangan penggiat internet. *Microblogging* ialah suatu wadah media sosial yang memungkinkan pengguna untuk mempublikasikan pesan pendek bermacam opini tentang bermacam topik, mengulas isu, menuliskan pembahasan suatu layanan, menulis kabar dalam jumlah karakter terbatas (kurang dari 200 karakter). Di mana jutaan pesan yang timbul setiap hari di dalam situs terkenal yang menyediakan layanan *microblogging* seperti *Twitter*, *Facebook*, serta *Tumblr* (Yenni Tresnawati, 2017).

Twitter merupakan salah satu situs *microblogging* dengan pengguna lebih dari 29,5 juta pengguna *Twitter* di Indonesia (Pamungkas, 2018). *Twitter* membolehkan penggunanya berbagi pesan pendek yang disebut *tweet*. *Twitter* bisa jadi sumber pendapat dan data sentimen masyarakat yang bisa digunakan untuk memasarkan suatu layanan ataupun produk.

Riset komputasi berasal dari opini-opini pengguna di internet, emosi serta sentimen melalui entitas serta atribut yang dipunyai serta diekspresikan dalam bentuk bacaan disebut *sentiment analysis* ataupun *opinion mining*. Analisis sentimen dilakukan untuk memandang komentar atau kecenderungan opini terhadap suatu permasalahan atau objek oleh seseorang, analisis sentimen akan mengelompokkan polaritas dari teks ke dalam kalimat ataupun dokumen untuk mengenali pendapat yang dikemukakan dalam bentuk teks ataupun kalimat tersebut berupa *positif*, netral, ataupun *negatif* (Pamungkas, 2018).

Dengan langkah manual analisa sentimen juga bisa dilakukan. Akan tetapi untuk data *tweet*, langkah manual tidak mungkin bisa dilakukan sebab jumlah data yang sangat besar serta terus meningkat. Di sinilah peran *preprocessing* yang secara otomatis dapat mengelola kata. Pada *preprocessing* terdapat beberapa proses yaitu *tokenizing*, *stopword*, *stemming*, serta pembobotan kata. Setelah melaksanakan *preprocessing*, maka dibutuhkan normalisasi (Yenni Tresnawati, 2017). Setelah melakukan normalisasi, dilakukan penggolongan atau *clustering* pada tiap *tweet*, salah satunya menggunakan metode *Naive Bayes Classifier*.

Pada penelitian ini dapat dilakukan analisis sentimen terhadap provider Indihome dengan menggunakan algoritma *Naive Bayes Classifier*. Data yang diambil berasal dari *Twitter* yang dipisahkan menjadi data *training* dan data *testing*. Data *training* digunakan untuk mendapatkan parameter yang bisa digunakan untuk proses pengklasifikasian memanfaatkan metode *Naive Bayes Classifier* sedangkan data *testing* digunakan untuk mengevaluasi hasil dari pengklasifikasian.

Berdasarkan penelitian yang telah tersedia sebelumnya, penelitian ini mencoba melakukan analisis sentimen dengan klasifikasi data *Twitter* berbahasa Indonesia. Data tersebut akan diproses dengan *text-preprocessing* untuk mengetahui data yang tidak sempurna.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan di atas, maka permasalahan yang akan dibahas dalam penelitian ini yaitu:

1. Bagaimana pendekatan metode *Naive Bayes Classifier* mampu mengelompokkan sentimen setiap *tweet* dengan baik?
2. Berapa akurasi yang diperoleh dari metode *Naive Bayes Classifier* untuk melakukan klasifikasi *tweet* pada data *Twitter*?

1.3. Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Metode yang digunakan untuk penelitian ini adalah *Naive Bayes Classifier* dan tidak menggunakan metode pembandingan lain.
2. Data yang digunakan adalah *tweet* dari pengguna *Twitter* yang mengandung kata kunci Indihome.
3. Data yang digunakan hanya dari *Twitter* tentang *tweet* opini terhadap layanan provider Indihome berbahasa Indonesia.
4. *Tweet* yang digunakan hanya berupa teks tidak mengandung gambar.
5. Sistem menggunakan bahasa pemrograman *Python*.

1.4. Tujuan

Tujuan yang hendak dicapai dalam penelitian ini adalah sebagai berikut:

1. Dapat melakukan analisis sentimen pada *Twitter* tentang *tweet* opini masyarakat mengenai layanan provider Indihome menggunakan metode *Naive Bayes Classifier*.
2. Mengetahui akurasi yang diperoleh dari metode *Naive Bayes Classifier* untuk melakukan klasifikasi *tweet* pada *Twitter*.
3. Mengetahui sentimen publik terhadap layanan Provide Indihome di media sosial *Twitter*.
4. Memvisualisasikan dari hasil analisis sentimen dapat ditampilkan dalam bentuk grafik pada aplikasi, sehingga hasil analisis sentimen dapat dibaca oleh pihak perusahaan.

1.5. Manfaat

Berdasarkan penelitian ini, diharapkan mempermudah pengambilan data secara otomatis. Selain itu analisis sentimen dapat memberi informasi mengenai *tweet* opini masyarakat untuk pihak Indihome agar dapat dijadikan bahan pertimbangan untuk memperbaiki layanannya ke depan.

1.6. Luaran

Luaran yang diharapkan pada penelitian ini berupa suatu sistem yang secara otomatis dapat melakukan analisis sentimen *tweet* pada *Twitter*.