

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada saat ini penggunaan internet di Indonesia berkembang secara luar biasa cepatnya dan internet memegang peran penting di dalam aspek kehidupan manusia. Pertumbuhan pengguna internet berdasarkan data dari Asosiasi Penyelenggara Jasa Internet Indonesia pada tahun 2018 telah mencapai angka 171,17 juta pengguna yang mengalami kenaikan dari tahun sebelumnya yaitu dengan angka pengguna internet mencapai 143,26 juta. Penggunaan internet yang menempati urutan pertama paling sering digunakan adalah layanan untuk berkomunikasi lewat pesan yang mencapai 24,7% dan selanjutnya yang menempatkan urutan kedua adalah untuk sosial media yang mencapai 18,9% dengan menggunakan (*Facebook, Twitter, Instagram*) dan lain-lain [1].

Sosial media seperti *Twitter* merupakan media komunikasi yang populer di kalangan pengguna internet di Indonesia. Menurut data yang telah dirilis pada tahun 2012 Indonesia menjadi negara peringkat kelima pengguna *Twitter* terbesar di dunia [2]. *Twitter* sering digunakan untuk berbagi informasi dengan pengguna lain dengan cara membuat pesan singkat yang sering disebut dengan *tweet*.

Twitter seringkali digunakan untuk mengungkapkan emosi mengenai suatu hal, baik dalam hal memuji ataupun mencela dengan bentuk emosi. Emosi yang timbul itu sendiri dapat dikategorikan menjadi lima emosi dasar yaitu senang, cinta, sedih, takut dan marah. Emosi senang dan cinta termasuk ke dalam emosi positif. Emosi sedih, takut dan marah termasuk ke dalam emosi negatif [3].

Analisis opini dan sentimen (*opinion analysis and sentiment analysis*) dilakukan berdasarkan *tweet* dari pengguna *Twitter*. Analisis tersebut dilakukan untuk melihat pendapat atau kecenderungan opini pengguna terhadap sebuah masalah atau objek oleh seseorang, apakah beropini positif atau negatif [4]. Dengan banyaknya jumlah *tweet* yang dipublikasikan pengguna melalui *Twitter*, *tweet* yang mengandung opini mengenai sebuah objek, objek tersebut dapat berupa kejadian disekitar masyarakat seperti suatu produk atau layanan. Hal ini membuat perusahaan menggunakan *Twitter* sebagai media untuk menyebarkan informasi [5]. Contohnya adalah perusahaan jasa *Internet Service Provider* (ISP) seperti Biznet.

Biznet merupakan perusahaan ISP yang mengalami lonjakan permintaan berlangganan baru mencapai 40% pada masa pandemi corona saat ini. Dengan semakin meningkatnya *traffic* internet maupun jumlah pelanggan, Biznet akan terus menjaga kualitas koneksi internet yang saat ini telah menjadi kebutuhan utama [6]. Melalui *Twitter*, pengguna dapat memberikan opini mengenai kritik, saran ataupun kepuasan mereka terhadap layanan internet dari Biznet. Untuk melakukan hal itu, dibutuhkan sebuah sistem yang dapat menganalisis sentimen berdasarkan *tweet* pengguna menggunakan salah satu fungsi dari *text mining*, dalam hal ini adalah klasifikasi sentimen.

Text mining adalah salah satu teknik yang dapat digunakan untuk melakukan klasifikasi dokumen. Ada banyak model klasifikasi saat ini seperti *Support Vector Machine*, *Naïve Bayes Classifier*, *Artificial Neural Network*, *Decision Tree* dan metode lainnya. Namun metode yang akan digunakan dalam penelitian ini adalah *Support Vector Machine*. Metode *Support Vector Machine*

akan mengklasifikasikan apakah suatu opini akan termasuk dalam sentimen positif atau negatif dengan mencari *hyperlane* terbaik dari kedua kelas klasifikasi dan memiliki akurasi yang tinggi [7]. Sehingga dengan melakukan penelitian menggunakan metode *Support Vector Machine* diharapkan dapat mengetahui tingkat akurasi dan sentimen terhadap layanan internet Biznet berdasarkan emosi.

Penelitian mengenai analisis sentimen yang sudah pernah dilakukan yaitu pada jasa transportasi *online* yang menggunakan metode *Naïve Bayes Classifier*. Pada penelitian tersebut data yang digunakan didapat dengan *crawling* pada *Twitter*. Akurasi yang didapat menggunakan *Naïve Bayes Classifier* sebesar 86,80% [8].

Faktor yang mempengaruhi agar fitur klasifikasi memberikan hasil yang maksimal ada pada tahap *text-pre-processing* bagian dimana data *tweet* yang sudah didapat selanjutnya diolah dengan tahapan-tahapan *case folding*, *tokenizing*, *filtering* dan *stemming* [8].

Berdasarkan penelitian [9] *Support Vector Machine* memiliki kinerja yang lebih baik dibandingkan dengan *Naïve Bayes* dalam melakukan klasifikasi yang terdiri dari dua label kelas, yaitu positif dan negatif. Hasil klasifikasi sentimen pada penelitian ini menunjukkan nilai akurasi yang dihasilkan algoritma *Support Vector Machine* lebih tinggi dari *Naïve Bayes*. Sehingga penelitian ini akan terfokus pada analisis sentimen pengguna *Twitter* terhadap layanan internet Biznet dengan menggunakan algoritma *Support Vector Machine*.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah diuraikan diatas, maka rumusan masalah yang dibuat sebagai berikut :

1. Apakah metode *Support Vector Machine* mampu melakukan klasifikasi analisis sentimen pada setiap *tweet* pada *Twitter* khususnya sentimen analisis layanan internet Biznet ?
2. Seberapa baik performa dari metode *Support Vector Machine* dalam melakukan klasifikasi analisis sentimen *tweet* pada *Twitter* khususnya sentimen analisis layanan internet Biznet ?

1.3 Batasan Masalah

Pada penelitian ini diberikan batasan-batasan masalah antara lain :

1. *Tweet* yang diambil dan di analisis hanya *tweet* berbahasa Indonesia.
2. Menggunakan metode *Support Vector Machine*.
3. Menggunakan 2 label kelas klasifikasi yaitu positif dan negatif.
4. Untuk mengambil data *Twitter* menggunakan media *Twitter API*.
5. Pembuatan sistem menggunakan bahasa pemrograman *Python*.

1.4 Tujuan Penelitian

Tujuan diadakan penelitian ini yaitu untuk :

1. Membuktikan bahwa metode *Support Vector Machine* dapat digunakan dalam pengklasifikasian analisis sentimen pengguna *Twitter* terhadap layanan internet Biznet.

2. Mengetahui akurasi yang diperoleh dari metode *Support Vector Machine* dalam pengklasifikasian analisis sentimen pengguna *Twitter* terhadap layanan internet Biznet.

1.5 Manfaat Penelitian

Manfaat yang diberikan dari penelitian ini yaitu :

1. Membantu memberikan *feedback* mengenai layanan internet Biznet.
2. Membantu menganalisis sentimen pada *Twitter* khususnya sentimen analisis layanan internet Biznet dengan metode *Support Vector Machine*.
3. Menjadi referensi bagi penelitian berikutnya yang relevan dengan kasus analisis sentimen *Twitter*.

1.6 Luaran

Luaran yang diharapkan dari penelitian ini berupa suatu sistem yang dapat mengelompokkan *tweet* berisi opini positif dan negatif.

1.6.1 Metode Pengumpulan Data

Peneliti menggunakan metode pengumpulan data yang dilakukan dengan studi literatur dan studi pustaka, yaitu dengan mencari buku-buku, jurnal, skripsi, tutorial dan berbagai informasi yang berkaitan sebagai referensi serta melakukan observasi dengan streaming data dari Twitter.

1.6.1.1 Metode Analisis

Pada tahap ini dilakukan analisis terhadap studi literatur untuk mendapatkan pemahaman mengenai metode yang akan digunakan, yaitu metode

Support Vector Machine (SVM) untuk menyelesaikan masalah dalam analisis sentimen.

1.6.1.2 Metode Perancangan Desain Sistem

Pada tahap ini dilakukan perancangan berdasarkan spesifikasi yang telah ditentukan sebelumnya. Perancangan terfokus pada pengumpulan data, algoritma yang digunakan dan antarmuka atau interface untuk pengujian. Proses perancangan dilakukan berdasarkan hasil dari analisis studi literatur yang telah diperoleh.

1.7 Sistematika Penulisan

Pada bagian ini terdiri dari pokok-pokok permasalahan yang dibahas dan dibagi menjadi beberapa bagian, yaitu :

BAB I Pendahuluan

Pada bab ini berisikan gambaran umum penelitian dari latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, luaran, batasan masalah, metode penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini berisikan tentang penjelasan dan teori-teori yang dibutuhkan. Landasan teori yang berkaitan dengan sentimen, *text-processing* dalam pengambilan data *Twitter* dan algoritma *Support Vector Machine*.

BAB III METODOLOGI PENELITIAN

Pada bab ini berisikan tentang metode yang digunakan dalam penelitian meliputi metode pengambilan data, metode pengembangan sistem dan lain-lainnya.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini berisikan tentang penjabaran hasil dari penelitian, nilai akurasi yang didapat dari algoritma yang digunakan dan berbagai macam analisis hasil dari sistem yang telah dibuat.

BAB V KESIMPULAN

Pada bab ini berisikan tentang uraian kesimpulan berdasarkan penelitian yang dilakukan dan saran bagi pembaca untuk pengembangan penelitian yang telah dilakukan agar menjadi lebih baik.

