

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pemerintah resmi menetapkan kenaikan cukai hasil tembakau (CHT) sebesar 23%. Keputusan tersebut ditetapkan usai Presiden Joko Widodo menerima laporan perencanaan kenaikan cukai rokok dari Menteri Keuangan Sri Mulyani Indrawati di istana kepresidenan pada hari Jumat, 13 September 2019. Menteri Keuangan Sri Mulyani mengatakan penetapan itu mulai berlaku pada tanggal 1 Januari 2020 mendatang. Tak hanya mengatur kenaikan tarif cukai hasil tembakau saja, pemerintah juga mengatur harga jual eceran (HJE) rokok sebesar 35%. Sri Mulyani mengatakan ada tiga faktor utama yang menjadi alasan pemerintah menaikkan tarif cukai rokok yaitu, upaya mengurangi konsumsi rokok yang cenderung meningkat terutama dikalangan remaja dan perempuan, mengatur industri rokok, dan menjaga penerimaan negara [1]. Dengan adanya penetapan kenaikan harga rokok yang bisa dikatakan cukup signifikan itu, disambut oleh respon masyarakat dari berbagai kalangan mulai dari masyarakat biasa, pengamat ekonomi, sampai politikus pun banyak yang mengutarakan pendapatnya terkait hal tersebut. Opini masyarakat terhadap kebijakan pemerintah menaikkan harga rokok pun beraneka ragam, ada yang merespon dengan baik dan ada juga yang meresponnya dengan kurang baik. Mayoritas masyarakat menyampaikan opininya melalui media sosial agar opininya bisa dilihat oleh orang banyak dan tentunya berharap pemerintah bisa mendengar aspirasinya.

Media sosial merupakan media yang dimanfaatkan sebagai sarana pergaulan sosial secara *online* di internet. Di media sosial, para penggunanya dapat saling berkomunikasi, berinteraksi, berbagi informasi, berpendapat, dan berbagai kegiatan lainnya. Media sosial menggunakan teknologi berbasis website atau aplikasi yang dapat mengubah suatu komunikasi ke dalam bentuk dialog interaktif. Salah satu contoh media sosial yang banyak digunakan adalah Twitter. Jumlah pengguna twitter secara global diklaim menaik. Pada laporan finansial twitter ke-3 tahun 2019, pengguna aktif harian di platform twitter dicata meningkat 17 persen, ke angka 145 juta pengguna, dan uniknya Indonesia diklaim menjadi salah satu negara yang pertumbuhan pengguna aktif harian twitter-nya paling besar. Klaim ini disampaikan oleh Dwi Adriansah selaku *Country Industry Head Twitter Indonesia* saat ditemui Kompas Tekno di kawasan SCBD Jakarta pada Rabu, 30 Oktober 2019 [2]. Kebanyakan orang sering kali menggunakan twitter sebagai media untuk menyampaikan opininya. Maka tidak heran ketika ada kejadian, pemberitaan, informasi, ataupun sesuatu hal yang menarik perhatian masyarakat hal tersebut akan menjadi salah satu trending topik di twitter, sebagai contohnya tentang kebijakan pemerintah menaikkan harga rokok. Banyak masyarakat yang menyampaikan opininya terkait hal tersebut, baik opini positif maupun negatif. Dengan adanya opini masyarakat yang disampaikan melalui twitter, merupakan sebuah hal positif apabila data tersebut dimanfaatkan sebagai bahan penelitian untuk menghasilkan sebuah informasi sehingga dapat digunakan sebagai bahan evaluasi. Namun untuk melakukan sebuah penelitian menggunakan data tersebut perlu adanya teknik analisis yang tepat salah satunya adalah analisis sentimen.

Sentiment analysis atau opini mining adalah riset komputasional dari opini, sentiment, dan emosi yang diekspresikan secara tekstual [3]. Dengan menggunakan analisis sentimen polaritas teks akan dikelompokkan untuk mengetahui teks tersebut bernilai positif atau negatif. Dalam melakukan sebuah analisis sentimen perlu adanya sebuah algoritme yang digunakan. Ada banyak algoritme yang bisa digunakan di dalam analisis sentimen salah satunya adalah algoritme *naive naves classifier*. *Xhemali Daniela dkk (2009)* mengatakan bahwa *Naive Bayes Classifier* memiliki tingkat akurasi yang lebih baik dibanding model classifier lainnya [4]. Sehingga pada penelitian ini akan menggunakan algoritme *naive bayes classifier*. Opini-opini masyarakat yang berada di twitter nantinya akan diklasifikasikan menggunakan algoritme *naive bayes classifier* sehingga menghasilkan sebuah klasifikasi opini positif dan negatif.

Berdasarkan masalah yang telah dipaparka, maka penulis ingin melakukan penelitian dengan judul “Analisis Sentimen Terhadap Kenaikan Harga Rokok Pada Media Sosial Twitter Menggunakan Algoritme Naive Bayes Classifier”.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan di atas, maka permasalahan yang akan dibahas dalam penelitian yaitu.

1. Bagaimana algoritme *naive bayes classifier* dapat mengklasifikasi data *tweet* dengan baik?

2. Berapa tingkat akurasi yang dihasilkan dari analisis sentimen terhadap kenaikan harga rokok pada media sosial twitter menggunakan algoritme *naive bayes classifier*.

1.3 Batasan Masalah

Berdasarkan rumusan masalah yang ada, terdapat batasan-batasan masalah dari penelitian ini antara lain.

1. Tweet yang dianalisis hanya tweet yang menggunakan bahasa Indonesia saja.
2. Tweet yang digunakan hanya yang berupa teks tidak berupa gambar
3. Penelitian hanya menggunakan data tweet dengan hastag yang berhubungan dengan kenaikan harga rokok saja seperti 'hargarokok', 'hargarokoknaik', 'rokokmahal'.
4. Analisis sentimen menggunakan data yang sudah ditraining.
5. Algoritme yang digunakan adalah *naive bayes classifier*.
6. Bahasa pemrograman yang digunakan adalah python.
7. Klasifikasi data dikategorikan menjadi 2 yaitu positif dan negatif.
8. Pengujian akurasi rata-rata keberhasilan sistem dilakukan dengan menggunakan metode *confusion matrix*.

1.4 Maksud dan Tujuan Penelitian

1.4.1 Maksud Penelitian

Adapun maksud dari penelitian ini adalah sebagai berikut.

1. Mengumpulkan tweet sentimen masyarakat yang berkaitan dengan kenaikan harga rokok.
2. Membuat system yang dapat melakukan analisis terhadap data training sentiment twitter yang berkaitan dengan kenaikan harga rokok.
3. Mengimplementasikan algoritme *naïve bayes classifier* pada proses analisis sentimen terhadap kenaikan harga rokok.

1.4.2 Tujuan Penelitian

Sedangkan tujuan dari penelitian ini adalah sebagai berikut.

1. Menganalisis tweet sentiment masyarakat terkait kenaikan harga rokok
2. Mengetahui tingkat akurasi dari kinerja algoritme *naive bayes*.

1.5 Manfaat Penelitian

Adapun manfaat yang dihasilkan dari penelitian ini adalah.

1. Dapat dijadikan rujukan penelitian selanjutnya yang berhubungan dengan analisis sentimen terutama yang menggunakan algoritme *naïve bayes classifier*.
2. Dapat dijadikan sumber informasi untuk mengetahui kinerja dari algoritme *naive bayes*.

1.6 Metode Penelitian

Dalam suatu penelitian perlu adanya sebuah metode yang digunakan untuk mendapatkan data atau informasi yang dapat menyelesaikan masalah-masalah dalam penelitian. Adapun metode yang digunakan penulis dalam penelitian ini antara lain.

1.6.1 Metode Pengumpulan Data

Pada metode pengumpulan data, penulis menggunakan dua metode yaitu metode deskriptis dan studi pustaka.

1. Metode Deskriptis

Dalam metode ini penulis mengumpulkan data dengan *crawling* menggunakan API twitter dengan hastag 'hargarokok', 'hargarokoknaik', 'hargarokokmahal' pada twitter.

2. Metode Studi Pustaka

Studi pustaka yang dilakukan yaitu mengumpulkan referensi baik dari buku, jurnal, paper, artikel, makalah, dan situs internet yang membahas analisis sentiment, algoritme *naïve bayes classifier*, serta referensi lain yang dapat membantu penelitian ini.

1.6.2 Metode Analisis

Metode analisis digunakan untuk menganalisis data yang sudah didapatkan pada proses pengumpulan data. Data *tweet* yang sudah didapatkan kemudian akan dianalisis melalui serangkaian proses *preprocessing text*. Kemudian data yang sudah diproses akan dianalisis menggunakan algoritme *naïve bayes classifier* untuk

mengetahui kinerjanya dalam mengklasifikasikan *tweet* positif dan negatif. Setelah proses klasifikasi berhasil, hasil dari analisis akan diuji menggunakan metode *confusion matrix*.

1.6.3 Metode Perancangan

Metode perancangan yang digunakan dalam penelitian ini yaitu dengan membuat diagram alur. Hal ini perlu dilakukan sebagai sebuah gambaran untuk memudahkan orang lain dalam memahami analisis sentimen ini. Selain itu juga perlu dibuat rancangan *interface* menggunakan python 3.8 untuk menampilkan halaman proses dan juga menampilkan hasil dari proses analisis sentimen.

1.6.4 Metode Implementasi

Setelah data sudah dianalisis serta rancangan analisis sudah diselesaikan, maka langkah selanjutnya yaitu mengimplemetasikan semua yang sudah terdapat pada rancangan penelitian.

1.6.5 Metode Pengembangan

Metode pengembangan yang diterapkan pada penelitian ini yaitu.

1. Crawling API twitter untuk mengambil sentiment masyarakat yang berhubungan dengan kenaikan harga rokok.
2. Melakukan perhitungan manual pada sample data sentiment.
3. Membuat *interface*.
4. Membuat koneksi antara data sentiment dengan *interface* program.

1.6.6 Metode Testing

Metode *testing* dilakukan untuk mengetahui apakah dalam melakukan analisis sentimen ini berjalan dengan baik atau tidak. Hasil dari penelitian juga perlu dilakukan pengujian apakah hasil analisis sentimen sesuai dengan tujuan dari penelitian atau tidak. Metode yang digunakan untuk mengetahui rata-rata keberhasilan sistem adalah *confusion matrix*.

1.7 Sistematika Penelitian

Sistem penulisan ditunjukkan kepada pembaca agar pembaca lebih mudah dalam memahami isi dari penelitian ini. Adapun sistematika dari penelitian ini terdiri dari beberapa bab sebagai berikut.

BAB I PENDAHULUAN

Pada bab ini terdapat penjelasan mengenai latar belakang masalah, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, manfaat penelitian, metode penelitian serta sistematika penelitian.

BAB II LANDASAN TEORI

Pada landasan teori menjelaskan hal hal yang menjadi dasar penelitian berupa teori-teori yang berkaitan dengan analisis sentiment menggunakan algoritme *naïve bayes classifier* dan beberapa hal yang mendukung proses analisis sentimen.

BAB III ANALISIS DAN PERANCANGAN

Dalam bab ini akan menjelaskan proses analisis data, analisis kebutuhan, serta analisis sistem serta perancangan sistem, program dan juga *interface*.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Pada bab ini akan menampilkan implementasi sistem, program, dan *interface* dalam analisis sentiment menggunakan algoritme *naïve bayes classifier* serta proses pengujian dari hasil penelitian menggunakan metode *confusion matrix*.

BAB V KESIMPULAN

Dalam bab terakhir akan dijelaskan kesimpulan dari hasil penelitian dan saran dari penulis untuk penelitian selanjutnya.

