

BAB I

PENDAHULUAN

Bab satu pendahuluan berisi mengenai latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, sistematika penulisan sebagai berikut.

1.1. Latar Belakang

Virus Corona atau yang lebih dikenal dengan nama *Severe Acute Respiratory Syndrome Coronavirus 2* (SARS-CoV-2) merupakan virus baru yang menyerang dan menginfeksi sistem pernafasan orang yang terjangkit, umumnya virus ini dikenal sebagai covid-19 [1] Dan virus covid 19 adalah salah satu jenis virus yang sedang mewabah saat ini. Virus ini berasal dari kota Wuhan Cina pada bulan Desember 2019. Penyebarannya sudah hampir ke semua negara yang ada di dunia termasuk Indonesia. Menurut data pada website WHO (*World Health Organization*) Secara global, pada 17:31 CEST, 18 Juni 2021, ada 177.1108.695 kasus COVID-19 yang dikonfirmasi, dengan jumlah kasus kematian sebanyak 3.840.223. Karena jumlah kasus yang semakin bertambah, beberapa negara melakukan upaya dan membuat kebijakan untuk menangani dan memutus penyebaran covid 19 ini.

Di Indonesia sendiri pemerintah mengeluarkan dan menerapkan beberapa kebijakan guna untuk meminimalisir penyebaran covid-19 diantaranya yaitu penetapan PSBB yang ditetapkan pada surat edaran Nomor 21 Tahun 2020, penetapan protokol kesehatan perjalanan Internasional pada surat edaran Nomor 2 tahun 2021 dan kebijakan tentang peniadaan mudik hari raya idul fitri tahun 2021 pada surat edaran Nomor 13 tahun 2021 [2] .

Salah satu kebijakan yang sedang diterapkan pada saat ini yaitu larangan untuk mudik. Presiden Republik Indonesia Joko Widodo membuat larangan kepada masyarakat Indonesia untuk mudik tahun ini guna melakukan pencegahan penularan virus covid 19. Hal ini ditetapkan berdasarkan hasil dari keputusan SKB (Surat Keputusan Bersama) 3 Menteri dan peraturan yang telah disahkan oleh pemerintah mengenai hari libur dan larangan untuk mudik 2021. Larangan mudik diberlakukan selama 12 hari yaitu pada tanggal 06 Mei sampai 17 Mei 2021. Penetapan kebijakan ini menimbulkan berbagai persepsi di kalangan masyarakat dikarenakan pada awalnya Menteri Perhubungan RI Budi Karya Sumadi mengatakan bahwa pemerintah tidak melarang pelaksanaan mudik 2021, sehingga pernyataan tersebut memberikan harapan bagi masyarakat untuk merayakan hari raya idul fitri bersama keluarga di kampung halaman. Namun, pada tanggal 26 Maret 2021 keputusan mengenai perjalanan mudik 2021 diputuskan dan menyatakan bahwa mudik 2021 dilarang. Sehingga keputusan tersebut menimbulkan kekecewaan pada masyarakat

yang sudah mengharapakan adanya mudik 2021.

Berbagai pendapat dari masyarakat mengenai kebijakan larangan mudik ini di sampaikan secara langsung maupun melalui media sosial. Media sosial merupakan sebuah *platform* yang digunakan untuk penyampaian informasi dan dilakukan secara *online* [3]. Salah satu media sosial yang digunakan sebagai wadah untuk menyampaikan opini publik adalah *twitter*. *Twitter* merupakan sebuah *platform* media sosial yang memiliki banyak pengguna. Berdasarkan data dari Global Digital Statistic pada tahun 2019 jumlah pengguna media sosial di Indonesia tercatat sekitar 150 juta lebih pengguna dan *twitter* merupakan salah satu media sosial yang paling banyak digunakan yang mencakup lebih dari 52 persen pengguna. *Twitter* juga dapat menghasilkan data atau informasi yang berguna untuk di analisis, karena data yang dihasilkan dapat di ekstrak menjadi informasi penting berupa pendapat masyarakat tentang peluncuran produk atau tren tertentu.

Pendapat masyarakat mengenai kebijakan larangan mudik 2021 akan menjadi sumber data pada penelitian ini yang di dapatkan melalui *crawling* data pada *website twitter*. Dari hasil *crawling* akan dilakukan sebuah model analisis dengan menggunakan metode klasifikasi serta uji akurasi. Hal ini dilakukan untuk menguji sebuah model terhadap dataset mudik. Salah satu model analisis yang dapat dilakukan adalah analisis sentimen.

Analisis sentimen atau yang sering disebut dengan opinion mining merupakan studi komputasi untuk mengenali serta mengekspresikan opini, penilaian, perilaku, emosi, subjektivitas atau pandangan yang terdapat dalam suatu teks dengan mengelompokkan polaritas pada teks kemudian ditentukan termasuk kalimat positif, negatif atau netral [4]. Ada beberapa metode yang ditemukan untuk analisis sentimen diantaranya adalah *Support Vector Machine* (SVM), *Naïve Bayes*, KNN, dan *Artificial Neural Evolution*. Metode *Naïve bayes* digunakan karena metode klasifikasinya sederhana untuk sekumpulan dataset yang besar dan memiliki tingkat akurasi yang tinggi serta metode klasifikasi yang paling banyak digunakan [5]. Namun metode ini masih memiliki beberapa kekurangan salah satunya yaitu sangat sensitif terhadap pemilihan fitur. Jumlah fitur yang banyak dapat mengurangi kinerja klasifikasi. Oleh karena itu diperlukan sebuah proses pemilihan fitur yang dapat membuat proses klasifikasi lebih efisien dengan mengurangi jumlah data yang akan di analisis. Ada beberapa metode yang dapat digunakan untuk pemilihan fitur, diantaranya adalah *Document Frequency*, *Information gain*, dan *Chi-Square*. Penggunaan metode *Information gain* lebih unggul dari metode yang lain karena cara kerja *Information gain* mengukur seberapa banyak kehadiran dan ketidakhadiran suatu kata. *Information gain* merupakan salah satu pemilihan fitur yang sukses dalam klasifikasi teks dan dapat meningkatkan hasil akurasi [6]

Terdapat beberapa penelitian sebelumnya yang menggunakan metode *Naive bayes* sebagai model untuk klasifikasi diantaranya yaitu penelitian Analisis Sentimen yang dilakukan oleh [6] yang menggunakan metode *naive bayes* dengan seleksi fitur *information gain*. Pengklasifikasian sentimen yang digunakan yaitu untuk menentukan ulasan pada maskapai penerbangan kedalam kategori positif dan negatif. Hasil pengujian yang didapatkan menunjukkan bahwa penambahan pemilihan fitur *Information gain* mendapatkan hasil yang lebih baik jika dibandingkan dengan sebelum penambahan *information gain*. Sedangkan penelitian yang dilakukan oleh [7] menggunakan metode *Naive Baye* dengan seleksi fitur *Information gain* terhadap kebijakan merdeka belajar yang dibagi kedalam kelas kelas positif dan negatif. Dari hasil pengujian diperoleh hasil terbaik menggunakan jumlah fitur sebanyak 40% dengan nilai akurasi 81.48%.

Berdasarkan latar belakang di atas, Peneliti membuat sebuah sistem yang dapat digunakan untuk menganalisa pendapat masyarakat tentang kebijakan pemerintah mengenai larangan mudik 2021 menggunakan metode *Naive bayes* dengan *Information gain* sebagai metode pemilihan fitur.

1.2. Rumusan Masalah

Untuk mengetahui persepsi masyarakat mengenai kebijakan larangan mudik 2021, diperlukan informasi yang diberikan melalui opini masyarakat. *Twitter* merupakan salah satu *platform* yang menjadi wadah masyarakat untuk menyampaikan opini. *Twitter* dapat digunakan sebagai salah satu *platform* untuk menggali informasi mengenai opini masyarakat terhadap kebijakan larangan mudik 2021. Tetapi, *twitter* tidak mempunyai fitur untuk mengagregasi opini tersebut menjadi sebuah kesimpulan. Berdasarkan pernyataan tersebut, maka dalam penelitian ini permasalahan yang akan dirumuskan yaitu sebagai berikut:

1.2.1. Permasalahan

1. Bagaimana mengklasifikasikan ulasan masyarakat yang terdapat pada *twitter* terhadap kebijakan larangan mudik 2021.
2. Bagaimana performa algoritma *naive bayes* dalam mengklasifikasikan ulasan dari masyarakat mengenai kebijakan larangan mudik 2021.

1.2.2. Metode Usulan

Analisis Sentimen masyarakat mengenai kebijakan larangan untuk mudik 2021 menggunakan metode usulan *Naive bayes* dengan seleksi fitur *information gain*. Metode *Naive bayes* digunakan karena metode klasifikasinya sederhana untuk sekumpulan dataset yang besar dan memiliki tingkat akurasi yang tinggi serta metode klasifikasi yang paling banyak digunakan. Penggunaan seleksi fitur *information gain* yaitu untuk mengurangi penggunaan jumlah fitur agar

dapat meningkatkan hasil akurasi pada model.

1.2.3. Pertanyaan Penelitian

1. Bagaimana tingkat akurasi yang didapatkan dengan menggunakan metode klasifikasi *Naive bayes* dalam mengklasifikasikan opini masyarakat mengenai kebijakan larangan mudik 2021?
2. Bagaimana tingkat akurasi yang didapatkan menggunakan metode klasifikasi *Naive bayes* dengan seleksi fitur *Information gain* pada pengklasifikasian analisis sentimen pada dataset mudik 2021?
3. Apakah metode klasifikasi *naive bayes* menghasilkan akurasi yang tinggi dalam mengklasifikasikan opini masyarakat mengenai kebijakan larangan mudik 2021?

1.3. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Mengetahui hasil penerapan metode *Naive bayes* dalam mengklasifikasikan opini masyarakat mengenai kebijakan larangan mudik 2021.
2. Mengetahui efisiensi metode *Naive bayes* dengan seleksi fitur *Information gain* pada dataset mudik 2021?
3. Mengetahui tingkat akurasi yang didapatkan oleh metode *naive bayes* dalam mengklasifikasikan opini masyarakat mengenai kebijakan larangan mudik 2021.

1.4. Batasan Masalah

Pada penelitian ini, peneliti memberikan batasan masalah agar tidak terjadi kesalahan pada saat penelitian dilakukan. Berikut adalah batasan masalah pada proses penelitian yaitu:

1. Opini yang digunakan adalah opini dalam bahasa Indonesia.
2. Data yang digunakan berasal dari *tweets hashtag* #janganmudik, #laranganmudik2021, #lebaran2021, #mudik, #mudik2021, #Mudikdilarang, #mudiklebaran, #MudikLebaran2021.
3. Data yang didapatkan merupakan *tweet* pada tanggal 30 April-07 Mei 2021
4. Jumlah data *tweet* yang didapatkan sebanyak 2251 *tweet*.
5. Klasifikasi atau sentimen yang dilakukan dibagi menjadi 3 kategori yaitu : positif, negatif dan netral.

1.5. Manfaat Penelitian

Manfaat yang didapatkan dari hasil penelitian ini adalah:

1. Dapat menambah wawasan serta kemampuan berpikir dalam menerapkan teori yang telah didapat selama perkuliahan kedalam penelitian yang sebenarnya.
2. Mengetahui klasifikasi opini pengguna *twitter* terhadap kebijakan mudik 2021.
3. Sebagai acuan atau referensi untuk pengembangan penelitian selanjutnya.

1.6. Sistematika Pembahasan

Penulisan skripsi ini terdiri dari 5 bab dengan masing-masing bab akan dijelaskan secara singkat sebagai berikut:

Bab I Pendahuluan

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

Bab II Landasan Teori

Bab ini berisi tentang teori-teori pendukung skripsi yang digunakan untuk memahami permasalahan yang di bahas pada penelitian yaitu tentang analisis sentimen, *preprocessing text*, *Twitter*, *Scrapping Twitter*, *Naive bayes*, *k-fold cross validation*, *Confusion matrix* dan penelitian terdahulu.

Bab III Analisis dan Perancangan

Bab ini akan dijelaskan mengenai proses analisis *text mining* pada analisis sentimen masyarakat tentang kebijakan mengenai larangan mudik 2021, metode penelitian, dan metode analisis yang digunakan yaitu metode *Naive bayes* berdasarkan tahapan implementasi.

Bab IV Hasil dan Pembahasan

Pada Bagian ini menjelaskan tentang hasil pengolahan data, pengujian dari hasil proses klasifikasi menggunakan metode *Naive bayes* dengan Seleksi Fitur *Information gain* pada dataset kebijakan mudik 2021.

Bab V Penutup

Bab ini membahas tentang kesimpulan yang di dapatkan dari hasil permasalahan penelitian berdasarkan proses yang sudah di lakukan. Serta saran yang akan disampaikan untuk kebutuhan pengembangan penelitian selanjutnya.