

BAB I

PENDAHULUAN

1.1 Latar Belakang

Virus Corona merupakan keluarga besar virus yang dapat menyerang manusia dan hewan yang biasanya menyebabkan penyakit infeksi saluran pernapasan, mulai dari flu hingga penyakit serius seperti MERS (*Middle-East respiratory syndrome*) pertama kali muncul di Timur Tengah pada tahun 2012 dan SARS (*Severe Acute Respiratory Syndrome*) pertama kali mewabah di China pada tahun 2002. (Hanati, 2020). *Coronavirus Disease 2019* disingkat COVID-19 adalah penyakit menular yang disebabkan oleh SARS-CoV-2. Gejala yang ditunjukkan penderita COVID-19 yaitu mengalami demam tinggi, batuk kering dan kesulitan bernafas yang berujung pada *pneumonia* dan kegagalan multiorgan. Virus ini menyebar melalui percikan (*droplets*) dari saluran pernapasan yang dikeluarkan saat sedang batuk atau bersin juga dari sentuhan pada benda - benda mati.(Informasi Tentang Virus Corona, 2020).

Pada akhir tahun 2019, dunia dikejutkan dengan munculnya COVID-19 di Wuhan, Ibu Kota Provinsi Hubei Tiongkok, China. Menurut data pemerintah yang dilihat oleh SCMP, pria berusia 55 tahun dari Provinsi Hubei kemungkinan menjadi orang pertama yang terinfeksi COVID-19 pada 17 November. Di Indonesia sendiri, kasus COVID-19 tercatat pertama kali pada 2 Maret 2020 di Depok, Jawa Barat yaitu dua WNI yang merupakan seorang ibu (64 tahun) dan putrinya (31 tahun). Keduanya diduga tertular virus corona karena melakukan kontak langsung dengan warga negara Jepang yang datang ke Indonesia. Hingga kini penyebaran virus corona di Indonesia meningkat sangat cepat. Pada 20 Maret 2020, tercatat pasien yang diperiksa sebanyak 1898 orang, pasien positif COVID-19 sebanyak 369 orang, pasien sembuh sebanyak 17 orang, pasien meninggal sebanyak 32 orang, pasien negatif sebanyak 1570 orang dan sedang dalam proses pemeriksaan sebanyak 21 orang (Indonesia.com, 2020).

Dilihat dari cepatnya peningkatan kasus pasien positif COVID-19 di Indonesia, tentunya meningkatkan keresahan masyarakat sehingga terjadinya peningkatan harga masker dan *hand sanitizer* yang tidak normal, adanya sentimen negatif yang terkesan mempolitisasi dalam pemberitaan, dikeluarkannya kebijakan – kebijakan pemerintah yang

dinilai masyarakat terlalu santai dan masih kurang tanggap dalam penanganan penyebaran virus, tidak transparannya pemberitaan saat awal kasus terjadi, banyaknya pernyataan dari tokoh publik yang dinilai tidak pantas dalam kondisi ini, juga tidak tertibnya masyarakat terhadap kebijakan yang sudah ditetapkan pemerintah terkait *social distancing* sehingga menimbulkan semakin banyaknya kasus terjadi dan kepanikan masyarakat yang semakin menjadi-jadi. (Perwitasari, 2020)

Untuk melaksanakan penelitian ini diperlukannya sebuah sistem penganalisa opini baik itu opini positif atau negatif atau biasa disebut analisis sentimen dari banyaknya pengguna media sosial khususnya Twitter yang memberikan pengaruh besar bagi masyarakat dalam penilaian berupa pendapat atau komentar terhadap penyebaran Virus Corona di Indonesia terkait apa yang sedang dilakukan, dirasakan, percakapan, penyebaran informasi, pelaporan berita yang juga bisa menjadi tolak ukur pemerintah terhadap kebijakan yang akan diambil berdasarkan aspirasi masyarakat yang bisa dilakukan dengan *Text Mining*.

Sentimen analisis merupakan suatu ilmu yang mencari emosi ataupun opini masyarakat baik berupa teks, maupun dari vokal. Sentimen analisis merupakan salah satu cabang dari *text mining* yang sedang marak dilakukan saat ini. Pada dasarnya sentimen analisis memiliki dua tipe pendekatan yaitu *machine learning* dan *lexicon based*. Beberapa metode klasifikasi *machine learning* seperti *naïve bayes*, *support vector machine*, dan *lexicon-based* sering digunakan untuk mendapatkan hasil yang terbaik.

Oleh karena itu, penelitian ini mencoba menganalisis *tweet* berbahasa Indonesia yang membicarakan tentang virus corona di Indonesia. Analisis dilakukan dengan melakukan klasifikasi *tweet* yang berisi sentimen masyarakat tentang pemberitaan mengenai penyebaran virus corona di Indonesia dan kebijakan pemerintah. Metode klasifikasi yang digunakan dalam penelitian ini adalah *Naive Bayes Classifier*. Metode ini dikombinasikan dengan fitur untuk dapat mendeteksi negasi dan pembobotan menggunakan *term frequency* serta TF-IDF. Klasifikasi *tweet* pada penelitian ini diperoleh berdasarkan kombinasi antara kelas sentimen dan kelas kategori. Hasil pengujian pada tools RStudio dalam tahap *preprocessing* memperlihatkan bahwa akurasi dengan *term frequency* memberikan hasil akurasi yang lebih baik daripada akurasi dengan fitur TF-IDF. Pada *Lexicon Based*, seleksi fitur tidak selalu meningkatkan akurasi pada metode

klasifikasi. Terdapat banyak kata tidak baku yang tidak terdeteksi sehingga menurunkan nilai evaluasi. Metode *Support Vector Machine* menghasilkan akurasi performans yang sama baiknya dengan metode *Naive Bayes* baik dalam klasifikasi sentimen maupun dalam klasifikasi kategori. Namun demikian, secara keseluruhan penggunaan metode *Support Vector Machine* dan *Naive Baye* sama-sama memiliki performansi yang cukup baik untuk melakukan klasifikasi tweet. (Fachril Rozy Saputra Rangkuti, 2018) Tetapi karena *Naive Bayes* memiliki waktu pemrosesan yang cepat serta mudah diimplementasikan dengan tingkat efektifitas yang tinggi. Dengan bahasa yang lebih sederhana, *Naive Bayes Classifier* mengasumsikan bahwa keberadaan maupun ketidakberadaan sebuah fitur dalam sebuah kelas tidak memiliki keterkaitan dengan keberadaan maupun ketidakberadaan fitur lainnya. Sehingga pada penelitian ini, penulis menggunakan metode *Naive Bayes Classifier* untuk melakukan pembobotan pada setiap kemunculan kata dan mengklasifikasikan sentimen dari opini dokumen *tweets* yang tersedia. (Nurrun Muchamad Shiddieqy Hadna, 2016).

1.2 Rumusan Masalah

1. Bagaimana pandangan atau pendapat masyarakat khususnya pengguna twitter terhadap pemberitaan mengenai COVID-19 di Indonesia?
2. Bagaimana metode *Naive Bayes Classifier* dapat digunakan dalam pengklasifikasian Analisis Sentimen COVID-19 di Indonesia dan seberapa baikkah performanya dalam proses pengklasifikasian tersebut?

1.3 Tujuan Penelitian

1. Membersihkan dokumen *tweet* dari atribut yang tidak dibutuhkan.
2. Mengetahui opini masyarakat khususnya pengguna twitter terkait pemberitaan mengenai COVID-19 di Indonesia.
3. Membuktikan bahwa metode *Naive Bayes Classifier* dapat digunakan dalam menghitung bobot dokumen *tweet*.

1.4 Batasan Masalah

1. Data yang dianalisis adalah data *tweets* dari media sosial Twitter menggunakan Bahasa Indonesia. Data *Tweets* yang diambil tidak melebihi dari 100 *tweets* dan hanya mencakup 7 hari terakhir dibuatnya *tweet* secara *realtime*.

2. Daftar kata pada perhitungan bobot terbatas.
3. Metode yang digunakan untuk klasifikasi adalah *Naïve Bayes Classifier*.
4. Target pengguna ditujukan kepada pihak terkait yang akan menjadikan hasil analisis ini sebagai tolak ukur contohnya pemerintah dan admin sebagai pengolah halaman sistem.
5. Bahasa pemrograman yang digunakan dalam proses pembuatan yaitu PHP, CSS dan JavaScript.
6. *Platform* yang digunakan pada penelitian ini adalah berbasis web dan basis data MySql.

1.5 Luaran yang Diharapkan

1. Memperoleh hasil *pre processing* kata pada data *tweet* yang sudah terstruktur semaksimal mungkin.
2. Menunjukkan kecendrungan opini masyarakat terhadap pertimbangan dalam permasalahan – permasalahan terkait Corona Virus yang terjadi di Indonesia seperti penerapan *social distancing* yang masih kurang diperhatikan oleh masyarakat, larangan untuk mudik dan pulang kampung kepada masyarakat, pembagian sembako yang menuai banyak pro dan kontra serta banyaknya penimbunan bahan pokok serta masker yang meresahkan.
3. Memberikan Informasi sentimen positif dan negatif terhadap pesebaran informasi terkait kebijakan pemerintah mengenai penanganan Virus Corona di Indonesia sebagai acuan pemerintah dan badan berkepentingan lainnya agar mempertimbangkan aspirasi masyarakat dalam mengeluarkan sebuah kebijakan sehingga apa yang akan diberikan kepada masyarakat boleh tepat sasaran dan sejalan dengan kinerja pemerintah.

1.6 Sistematika Penulisan

Sistematika penulisan dalam tugas akhir ini, disusun sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah, rumusan masalah, Batasan masalah, dan tujuan penulisan, serta sistematika penulisan

BAB II TINJAUAN PUSTAKA

Bab ini berisi penelitian terkait, metode yang digunakan, dan Langkah-langkah metode

BAB III PERANCANGAN SISTEM

Bab ini berisi desain arsitektur sistem, desain algoritma metode, flowchart sistem, dan diagram alur.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi hasil penelitian sampai saat ini, kendala-kendala yang dihadapi, rencana penyelesaian masalah

BAB V PENUTUP

Bab ini berisi kesimpulan dan saran dari hasil pengujian sistem yang telah dilakukan