

BAB I PENDAHULUAN

1.1 Latar Belakang Masalah

Akhir tahun 2019 dunia digemparkan oleh kemunculan virus baru yang diberi nama COVID-19. Virus ini menjadi momok yang besar karena tingkat penyebarannya yang sangat pesat. Menurut data Worldometers, ada sebanyak 198 negara yang telah mengonfirmasi kasus positif Covid-19. Sementara, angka kasusnya tercatat 467.520 kasus yang terkonfirmasi, dengan 21.174 orang meninggal dunia, dan 113.808 pasien sembuh. benar saja hingga tanggal 08 Juni 2020 sudah terdapat 198 Negara, 467.520 Orang Terinfeksi (Aida, 2020). Pemerintah di berbagai negara telah melakukan segala upaya untuk meredam persebaran virus COVID-19 tidak lain halnya dengan pemerintah Indonesia, salah satu negara yang terdampak virus COVID-19. Upaya yang dilakukan pemerintah Indonesia untuk menekan persebaran virus COVID-19 yaitu dengan kebijakan *lockdown*.

Pemerintah berkewajiban membuat kebijakan untuk menyejahterakan masyarakatnya dalam menanggulangi persebaran virus COVID-19. Maka dari pada itu perlu adanya evaluasi secara berkala terhadap kebijakan yang dikeluarkan oleh pemerintah, tapi sebelum itu perlu adanya sikap terbuka dari pemerintah untuk menerima kritik, saran, dan keluhan dari masyarakat. Pada masa pandemi virus COVID-19 pemerintah Indonesia hingga saat ini belum mengeluarkan kebijakan *lockdown*, banyak opini yang tersebar dikalang masyarakat terkait kebijakan yang dikeluarkan pemerintah baik itu opini positif maupun opini negatif. Berdasarkan hal inilah perlu dilakukannya analisis sentimen terhadap kebijakan *lockdown* yang seharusnya dikeluarkan pemerintah untuk mengetahui sentimen yang beredar di masyarakat Indonesia.

Pada saat ini masyarakat sudah sangat dimudahkan dalam menyampaikan sebuah opini terkait suatu hal. Berdasarkan data statistik terdapat 160 juta pengguna aktif media sosial di Indonesia (*detik.com*). dari hal tersebut dapat kita ketahui bahwa media sosial merupakan perantara yang paling banyak diminati masyarakat. Twitter merupakan media sosial *microblogging* yang sangat digandrungi masyarakat Indonesia saat ini. *Microblogging* merupakan suatu layanan media sosial yang memungkinkan pengguna *me-publish* berupa opini, komentar, berita dalam karakter yang terbatas (kurang dari

200 karakter) [5]. *Twitter* sering kali digunakan untuk mengungkapkan emosi terkait suatu hal. Dari sini dapat kita lihat bahwa *twitter* sangatlah berpotensi untuk melihat sentimen yang ada di kalangan masyarakat terkait kebijakan *lockdown* yang dikeluarkan pemerintah.

Analisis sentimen sendiri sebenarnya dapat dilakukan secara manual, tetapi mengingat jumlah data *tweet* yang sangat banyak dan terus bertambah hal ini tidaklah mungkin dilakukan. Disinilah peran *text mining* diperlukan. *Text mining* adalah proses memeriksa sekumpulan dokumen untuk menemukan informasi baru atau membantu menjawab pernyataan penelitian tertentu (Linguamatics.com). *Text mining* memiliki berapa tahap yaitu *tokenizing*, *filtering*, *stemming*, *tagging*, *analyzing*. Kemudian dari tahap tersebut dilakukan penggolongan pada setiap *tweet*, salah satunya dengan metode *support vector machine*.

Penelitian mengenai sentimen analisis sudah dilakukan sebelumnya yaitu komparasi algoritma untuk analisa sentimen review produk pada twitter. Pada penelitian ini menggunakan algoritma *baseline*, *nive bayes*, dan *maximum entropy*. Dari penelitian tersebut didapat *maximum entropy* menghasilkan akurasi yang lebih tinggi dibandingkan *baseline* dan *naive bayes*. Pada penelitian lainnya tentang analisis sentimen menggunakan *support vector machine* dan *maximum entropy* didapatkan hasil algoritma *support vector machine* memiliki akurasi lebih tinggi sebesar 75% dibandingkan algoritma *maximum entropy* [11]. Berdasarkan penelitian yang telah ada sebelumnya maka pada penelitian ini menggunakan algoritma *support vector machine* dalam melakukan sentimen analisis.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan diatas maka dirumuskan masalah dalam penelitian ini, yaitu :

1. Bagaimana mengimplementasikan algoritma *support vector machine* terhadap sentiment analis twitter ?
2. Bagaimana menghitung akurasi data uji dan data latih terbaik untuk algoritma *support vector machine* ?
3. Bagaimana sentiment rakyat Indonesia terhadap kebijakan *lockdown* ?

1.3 Batasan Masalah

1. Data yang dipakai hanya *tweet* berbahasa Indonesia dari media sosial *twitter*
2. *Tweet* yang diambil hanya yang mengandung kata *lockdown*, *COVID-19*, *#lockdown* dan *#dirumahaja*. Sejak tanggal 01 Maret 2020 sampai 16 April 2020

1.4 Tujuan Penelitian

1. Penelitian ini memiliki tujuan untuk dapat mengklasifikasikan informasi pada *tweet* di Indonesia terkait kebijakan *lockdown* positif atau negatif menggunakan algoritma *support vector machine* dan di visualisasikan dalam website.
2. Memberikan informasi kepada khalayak umum terkait sentimen yang beredar terkait kebijakan *lockdown*
3. Bisa menjadi bahan masukan untuk pemerintah terkait kebijakan *lockdown*

1.5 Manfaat Penelitian

1.5.1 Bagi Penulis

Mengaplikasikan ilmu-ilmu akademis yang didapat selama perkuliahan untuk identifikasi sentimen *twitter* terkait kebijakan *lockdown* yang dikeluarkan pemerintah Indonesia dengan algoritma *support vector machine*.

1.5.2 Bagi Universitas

1. Memberikan gambaran terhadap penerapan ilmu pengetahuan yang telah diterima selama kuliah
2. Menjadi sumbangan literatur karya ilmiah dalam disiplin ilmu teknologi khususnya bidang informatika
3. Mengukur tingkat kemampuan mahasiswa dalam ilmu akademis maupun non-akademis di lingkungan masyarakat

1.5.3 Bagi Masyarakat

1. Menambah wawasan pembaca mengenai penerapan algoritma *support vector machine* dalam sentimen analisis
2. Memberikan informasi tentang sentimen yang beredar di masyarakat Indonesia terkait kebijakan *lockdown*.

1.6 Metode penelitian

1.6.1 Metode Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data *tweet* dari *twitter* dengan menggunakan library python *twint*. Data yang diambil merupakan *tweet* yang mengandung kata *social distancing*, *physical distancing*, dan *#dirumahaja* dari tanggal 15 Maret 2020 sampai 31 Desember 2020

1.6.2 Metode Perancangan

Dalam penelitian ini, peneliti menggunakan *flowchart* untuk menunjukkan aliran di dalam program yang dibuat atau prosedur sistem secara logika dari tahap manipulasi data hingga proses klasifikasi.

1.6.3 Metode Analisis

Pada metode ini dilakukan beberapa tahapan :

1. Data Collecting

Data *tweet* akan dilakukan *crawling* menggunakan library python yaitu *twint*. Proses ini dilakukan dengan cara memasukkan keyword yang akan dicari, jumlah *tweet* yang akan diambil, dan tanggal *tweet*. *Tweet* yang telah diambil kemudian akan disimpan dalam format *csv*.

2. Preprocessing

Preprocessing merupakan tahap pembersihan data sehingga data siap digunakan untuk model klasifikasi. Terdapat beberapa tahap dalam *preprocessing* yaitu *cleansing*, *case folding*, *tokenisasi*, *normalization*, penghilang *stopword*, *stemming*.

3. Feature Extraction

Pada tahap ini, dilakukan pembobotan kata pada data yang sudah dikumpulkan sebelumnya menggunakan metode TF-IDF. Metode ini akan menentukan pembobotan kata berdasarkan frekuensi atau banyaknya kemunculan suatu kata dalam sebuah

dokumen.

Apabila presentase munculnya suatu kata semakin besar, maka nilainya akan semakin besar.

4. Klasifikasi

Pada tahap ini dilakukan klasifikasi terhadap *tweet* yang telah dilakukan *preprocessing* dan *feature extraction* sebelumnya. Metode yang digunakan dalam klasifikasi adalah SVM. SVM adalah salah satu metode *supervised learning* yang biasanya digunakan untuk klasifikasi dan regresi dengan cara mencari *hyperplane* terbaik.

1.6.4 Metode Perancangan

Dalam penelitian ini, peneliti menggunakan *flowchart* untuk menunjukkan aliran di dalam program yang dibuat atau prosedur sistem secara logika dari tahap manipulasi data hingga proses klasifikasi.

1.6.5 Metode Implementasi

Tahap implementasi merupakan penerapan dari metode analisis, seperti *data collecting*, *preprocessing*, *feature extraction*, dan *klasifikasi*.

1.6.6 Metode Pengujian

Model yang diperoleh pada proses pelatihan akan dilakukan pengujian *precision*, *recall*, dan *f-measure* yang akan disajikan dalam *confusion matrix* untuk mengetahui rata-rata keberhasilan dari model.

1.7 Sistematika Penelitian

Sistematika penulisan laporan Skripsi ini terdiri dari pokok-pokok permasalahan yang dibahas pada masing-masing yang akan diuraikan menjadi beberapa bagian.

BAB I PENDAHULUAN

Bab ini membahas tentang gambaran umum isi skripsi yang meliputi latar belakang masalah, rumusan masalah, batasan masalah, tujuan

penelitian, manfaat penelitian, metode penelitian, dan sistematika penelitian.

BAB II LANDASAN TEORI

Bab ini menguraikan tentang pengertian dan teori-teori yang dibutuhkan sebagai landasan atau dasar dalam penelitian yang akan dilakukan.

BAB III ANALISI DAN PERANCANGAN SISTEM

Bab ini menjelaskan mengenai metode yang digunakan dalam penelitian yaitu metode pengumpulan data, metode pengembangan sistem dan lain-lainnya.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Bab ini berisi tentang paparan implementasi dan analisis hasil uji coba program dari tahap analisis, desain, implementasi desain, hasil testing dan implementasinya, berupa penjelasan teoritik, baik secara kualitatif, kuantitatif, atau secara statistik.

BAB V PENUTUP

Bab ini berisi tentang uraian mengenai kesimpulan berdasarkan penelitian yang telah dilakukan dan tentang saran-saran bagi pembaca untuk pengembangan penelitian yang telah dilakukan agar menjadi lebih baik.