

BAB I

PENDAHULUAN

1.1. Latar Belakang

Virus corona yang ada di Indonesia ini menjadi sebuah masalah yang membuat masyarakat di seluruh Indonesia geram dengan ada hal ini. Dikarenakan belum adanya vaksin/obat untuk mematikan virus corona tersebut. Maka dari itu pemerintah melakukan tindakan seperti lockdown/psbb terhadap masyarakat untuk stay dirumah. Sehingga total kasus virus corona di indonesia perhari ini menjadi 42,762 per 17 Juni 2020, hingga kamis (18/6/2020) sore, terdapat penambahan sebanyak 1,331 orang, hal ini di sampaikan oleh juru bicara pemerintah untuk penanganan Covid-19, Achmad Yurianto. Dalam konferensi pers di Gedung BNPB, Jakarta. menurut (sumber: tribunternate.com).

Twitter adalah layanan jejaring sosial dan mikroblog daring yang memungkinkan pengguna untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter akan tetapi pada tanggal 07 November 2017 bertambah hingga 280 karakter yang dikenal dengan sebutan kicauan (*tweet*). Twitter didirikan pada bulan Maret 2006 oleh Jack Dorsey, *twitter* bertujuan untuk melayani percakapan publik. Kekerasan, pelecehan dan perilaku sejenis lainnya membuat orang tidak berani mengekspresikan diri dan mencederai esensi percakapan publik secara global. Pengguna twitter di indonesia, informasi berdasarkan data PT Bakrie Telecom, memiliki 19,5 juta pengguna di indonesia dari total 500 juta pengguna global. Pengguna twitter di indonesia menempati peringkat 5 pengguna Twitter terbesar di dunia. Dan

kebanyakan pengguna twitter di indonesia adalah konsumen, yaitu dikarenakan tidak memiliki blog atau tidak pernah mengupload video di Youtube namun sering update status di Twitter, Menurut (sumber: KOMINFO). Sedangkan twitter itu sendiri adalah layanan jejaring sosial dan mikroblog daring yang memungkinkan penggunanya untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter akan tetapi pada tanggal 7 November 2017 bertambah hingga 280 karakter yang dikenal dengan sebutan kicauan (*tweet*). Menurut (sumber: wikipedia).

Analisis sentimen adalah bagian dari *opinion mining*. Analisis sentimen dilakukan untuk melihat pendapat terhadap sebuah masalah atau dapat juga digunakan untuk identifikasi kecenderungan hal di luar. Sentimen yang dapat menjadi solusi permasalahan yang telah dipaparkan sebelumnya, yaitu dengan memisahkan opini ke dalam kelas bersentimen positif atau negatif, serta dilakukan pengambilan kesimpulan faktor apa yang sering dibahas dalam opini-opini tersebut. Sentimen yang dibahas yaitu mengenai ekstraksi pendapat, evaluasi dan emosi orang tentang suatu topik tertentu yang tertulis menggunakan teknik pemrosesan bahasa alami.

Data Mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Turban dkk. 2005).

Text mining yang dikenal juga dengan *text data mining* atau pencarian pengetahuan di basis data tekstual adalah sebuah proses yang semi otomatis melakukan ekstraksi dari pola yang ada di *database*.. Untuk pemanfaatan dari *text mining* secara nyata sangatlah luas. Areanya seluas data tekstual yang terbentuk seperti di area pemasaran dengan data komentar konsumen, teknologi dengan data arsip paten, keuangan dengan data laporan triwulan, hukum dengan data putusan pengadilan. Sehingga, hasil dari klasifikasi tersebut bisa menjadi media alternatif bagi masyarakat untuk mengetahui data isu komentar twitter yang positif dan negatifnya.

Metode *support vector machine* yaitu dapat digunakan untuk mengklasifikasikan data opini berdasarkan atribut penilaian yang dimiliki, untuk dapat dipisahkan apakah opini termasuk ke dalam kelas bersentimen positif atau negatif. Oleh karena itu, dalam penelitian ini digunakan metode *support vector machine* (SVM) yang diharapkan sesuai dengan permasalahan. Dengan dilakukannya analisis sentimen ini, diharapkan dapat diketahui permasalahan yang terdapat pada isu vaksin covid-19 di Indonesia, sehingga dapat ditentukan isu tersebut positif atau negatif bahwa virus corona masuk ke Indonesia. Dalam penelitian ini penulis akan melakukan penelitian dengan memanfaatkan algoritma *support vector machine* dengan data tweet yang penulis ambil dari *twitter* serta mengklasifikasikan sentimen positif dan negatifnya dengan topik terkait isu vaksin covid-19 di Indonesia. Berdasarkan permasalahan diatas penulis akan membuat tugas akhir yang berjudul **“Penerapan Metode Support Vector Machine Dalam**

Analisis Sentimen Media Sosial Twitter Terkait Isu Vaksin Covid-19 Di Indonesia”

1.2. Rumusan Masalah

Berdasarkan latar belakang diatas, dapat dirumuskan permasalahan dalam penelitian ini adalah menganalisis, mengelola data opini pengguna twitter terkait isu vaksin covid-19 di Indonesia dan kemudian mengklasifikasikan sentimen data *tweet* tersebut dengan menggunakan metode *support vector machine* serta menghitung tingkat akurasi yang dihasilkan oleh metode *support vector machine*.

1.3. Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini adalah data *tweet* yang menggunakan Bahasa Indonesia, jumlah data *tweet* yang digunakan adalah 500 data yang sudah difilter, dan diambil pada bulan Januari 2021
2. Pengumpulan data menggunakan *Twitter API* dengan tool *R Studio*. Data *tweet* yang akan di nilai sentimennya adalah *tweet* yang memiliki salah satu *keyword* berikut : #corona #vaksin
3. Data sentimen yang diambil untuk klasifikasinya adalah positif dan negatif
4. Tools yang digunakan adalah *Kaggle.com* dengan berbasis Bahasa pemrograman *Python*
5. Data yang sudah terkumpul di input ke *Kaggle.com*
6. Metode yang digunakan untuek mengklasifikasian dalam penelitian ini adalah MetodeiSupportiVectoriMachine (SVM)

1.4. Tujuan

Tujuan dari penelitian ini adalah mengimplementasikan algoritma *Support Vector Machine* dalam analisis sentimen pengguna sosial media twitter dengan topik vaksin covid-19 di indonesia.

1. Mengetahui isu apa yang terjadi dengan isu vaksin covid-19 di Indonesia.
2. Menentukan hasil akurasi dengan menggunakan metode *support vector machine*.

1.5. Manfaat Penelitian

Penelitian ini dilakukan dengan harapan dapat memberikan manfaat, diantaranya:

1. Bagi peneliti, untuk memberikan informasi gambaran mengenai sentimen positif dan negatif terkait isu vaksin covid-19 di Indonesia dan paham dalam membangun model malukan klasifikasi tweet berdasarkan metode *Support Vector Machine* (SVM).
2. Bagi Lembaga, Dapat membantu sarana pengenalan dan tambahan informasi referensi Perpustakaan Universitas Stikubank Semarang (Unisbank) dan dapat dikembangkan lebih lanjut dengan menganalisis sentimen pada *twitter* atau pada permasalahan lain dengan metode yang berbeda.
3. Bagi masyarakat, dapat mengetahui opini masyarakat terkait isu vaksin covid-19 di Indonesia dan mendapatkan informasi hasil akurasi dari implementasi metode *support vector machine* secara publikasi.

1.6. Metodologi Penelitian

Menurut sumber skripsi dari (Gieffari Satria Abdilah, 2019) gambaran umum mengenai metode penelitian diharapkan dapat memperoleh hasil sebagai berikut :

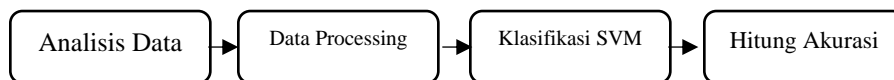
1.3.1. Perumusan objek penelitian

Penelitian ini menggunakan data tweet dari media sosial *twitter* yang terkait dengan isu vaksin covid-19 di Indonesia dengan cacah data sebanyak 500.

1.3.2. Metode pengumpulan data

Data sentimen yang digunakan diambil dari media sosial *twitter* dengan menggunakan pencarian isu terkait vaksin covid-19 di Indonesia. Pencarian data *tweet* dikhususkan dengan topik isu vaksin covid-19 di Indonesia. Untuk proses pengambilan data ini perlu menggunakan API (*Application Programming Interface*) dari *twitter* lalu di simpan dalam file Ms. Excel.

1.3.3. Metode analisis data



Gambar 1.1. Analisis Data

Dalam tahapan ini yang akan dilakukan dalam penelitian, metodenya sebagai berikut:

1. Analisis data

Setelah data tweet terkumpul maka data tersebut akan diolah dengan menentukan sentimen tiap *tweet*. Data sentimen yang akan digunakan yaitu data positif dan negatif. Cara menentukan data sentimen positif dan negatifnya itu

secara manual dan berdasarkan makna yang terkandung dalam *tweet* apakah bersifat positif atau negatif.

2. Data *Processing*

Dalam tahap ini data akan diolah untuk menghilangkan beberapa kata atau karakter yang terdapat pada *tweet* dengan proses *cleaning*, normalisasi kalimat, *stopword* dan tokenisasi. Di proses ini akan mendapatkan data yang rapi agar mempermudah dalam proses analisis data *tweet*.

3. Klasifikasi SVM

Klasifikasi merupakan suatu teknik untuk menemukan *hyperplane* yang bisa memisahkan dua set data dari kelas yang berbeda (Vapnik. 1999). Untuk klasifikasi yang dilakukan dengan metode *support vector machine* dalam penelitian ini yaitu menggunakan kernel linier, karena itu data yang diklasifikasi akan menentukan pembagian sentimen positif dan negatif.

4. Hitung Akurasi

Dalam metode yang digunakan akan menghasilkan akurasi data dengan variabel jumlah data, perbandingan antara data latihan dan data uji, bobot perbandingan antara positif dan negatif pada tiap *tweet*. Data yang di uji adalah data yang telah diketahui sentimennya. Sehingga klasifikasi sentimen akan dilakukan perhitungan akurasi atau ketepatan sistem dalam melakukan klasifikasinya. (Yonathan Sari Mahardika, Eri Zuliarso, 2018).

1.7. Sistematika Penulisan

Sistematika penulisan ini disusun untuk memberikan gambaran secara umum tentang penelitian yang akan dijalankan. Untuk memahami lebih jelas laporan ini, maka materi-materi yang tertera pada Laporan Skripsi ini dikelompokkan menjadi beberapa sub bab dengan sistematika penyampaian sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisi tentang latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian serta metodologi penelitian.

BAB II TINJAUAN PUSTAKA

Bab ini berisi kajian tentang penelitian-penelitian terdahulu yang pernah dilakukan terkait dengan penelitian yang dilakukan oleh peneliti sebelumnya, teori yang berkaitan dengan penelitian.

BAB III METODE PENELITIAN

Bab ini berisi tentang bagaimana peneliti meneliti dan metode yang digunakan dalam penelitian ini. Data yang digunakan adalah data tweet mengenai isu korona virus yang ada di indonesia.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi tentang keseluruhan hasil dari penelitian yang telah dikerjakan, serta pembahasan masalah yang ada dalam penelitian tersebut.

BAB V PENUTUP

Bab ini berisi kesimpulan dan saran.