

BAB I

PENDAHULUAN

1.1 Latar Belakang

Salah satu Negara yang kurang berkembang adalah Negara Indonesia. Hal ini ditandai dengan penanganan masalah pandemi *Covid19* yang terjadi pada saat ini. Kondisi dimana terjadinya penyebaran virus corona yang pertama kali dari Negara Cina lebih tepatnya di kota Wuhan pada saat itu pada tahun 2019. Lalu wabah virus corona menyebar sangat luas sampai ke berbagai negara. Di Negara Indonesia sendiri korban terkena virus semakin hari semakin bertambah kematian juga semakin banyak. Penyebab penularan yang terjadi secara cepat karena berinteraksi langsung dengan orang yang positif corona. Maka pemerintah mengeluarkan undang undang yang berisi untuk semua kegiatan beraktifitas di luar dan berinteraksi dengan banyak orang sementara harus ditiadakan. Semua pekerja dilakukan secara online dan pembelajaran dilakukan secara daring. Tetapi keadaan masyarakat Indonesia banyak dari kalangan menengah kebawah, maka undang-undang tidak berjalan sesuai perintah. Pandemi yang terjadi berkelanjutan kurang lebih 1.5 tahun. Korban meninggal bertambah hari bertambah banyak, orang sakit yang berhubungan dengan pernafasan dianggap corona bahkan dikabarkan politik dalam rumah sakit. Pihak pemerintah

mengalami hutang yang sangat besar dan krisisnya ekonomi pada masyarakat. Indonesia terima bantuan dari luar negeri berupa suntik pencegah virus corona yang disebut vaksin, sedikit membantu kebebasan masyarakat dalam beraktifitas kembali. Setelah keadaan yang terjadi masyarakat mengekspresikan opininya pada media sosial. Salah satu media sosial yang populer di kalangan pengguna internet saat ini adalah *Twitter*.

Twitter merupakan media sosial yang dibuat oleh Jack Dorsey pada tahun 2006. Pada tahun 2013 Berdasarkan *press-release* *Twitter* ada 500 juta *tweet* atau kicauan oleh pengguna *twitter* per harinya. Sebanyak 500 juta *tweet* digunakan untuk mem-post hal tentang diri pengguna dan berbagi informasi. Isi *tweet* juga dapat mengekspresikan perasaan. Opini melalui *tweet* inilah yang dapat dimanfaatkan untuk melihat bagaimana sentimen yang dimunculkan salah satunya adalah mengenai opini seseorang terhadap keadaan pandemi *covid19*.

Penentuan polaritas positif atau negatifnya suatu opini dapat dilakukan secara manual, tetapi seiring bertambahnya sumber opini menjadi semakin banyak tentunya waktu dan usaha yang dibutuhkan untuk mengklasifikasikan polaritas opini tersebut akan semakin banyak terpakai. Maka dari itu diajukan penerapan metode pembelajaran mesin untuk melakukan klasifikasi opini dari sumber yang banyak berdasarkan komentar

masyarakat tentang *covid19*. Untuk melakukan itu bias menggunakan salah satu fungsi dari text mining.

Terdapat beberapa teknik klasifikasi dokumen, diantaranya *Naïve Bayes Classifier*, *K-Nearest Neighbor* dan *Support Vector Machines*. *Naïve bayes* adalah memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya. Hal ini dibuktikan pada jurnal *Xhemali, Daniela, Chris J. Hinde, and Roger G. Stone* mengatakan bahwa “*Naïve Bayes* Klasifier memiliki tingkat akurasi yang lebih baik dibanding model *classifier* lainnya”(2009). Alasan inilah tujuan penulis menggunakan algoritma tersebut supaya mendapatkan hasil akurasi yang tinggi untuk menentukan suatu opini negatif ataupun positif pada data *twitter*. Pada kasus ini digunakan bahasa pemrograman *python* dan software *kaggle*.

Penelitian ini penting dilakukan untuk memberi informasi kedepannya kepada para ilmuwan tentang kondisi masyarakat pandemi *covid19*. Sehingga dalam menyelesaikan permasalahan tersebut dilakukan “Analisis Sentimen Pada Media Sosial Twitter Menggunakan Algoritma *Naive Bayes Classifier*”, terhadap kondisi pandemi *Covid!9*.

1.2 Rumusan Masalah

Dari latar belakang yang telah dijelaskan, dapat ditarik beberapa permasalahan antara lain :

1. Bagaimana mengetahui pendapat masyarakat terhadap komentar *covid19* berdasarkan analisis sentiment dengan sumber data dari media sosial *twitter*?
2. Bagaimana cara melakukan implementasi dan visualisasi data pada penelitian analisis sentiment pandemic *covid19*?
3. Bagaimana hasil akurasi yang diperoleh dari penerapan algoritma *Naïve Bayes Classifier*?

1.3 Batasan Permasalahan

Dalam penelitian ini memerlukan batasan-batasan sebagai berikut:

- a. Pengambilan data melalui media *Twitter API* dengan 500 sampel data yang menggunakan bahasa Indonesia.
- b. Studi kasus yang digunakan berupa data *tweet* masyarakat terhadap keadaan yang terjadi selama pandemi *covid19* dengan kata kunci pencarian *#corona* dan *#vaksin*.

1.4 Tujuan dan Manfaat Penelitian

Tujuan yang ingin dicapai dari penelitian:

- a. Melakukan analisis sentiment menggunakan pemrograman *python* dengan *Twitter API* sebagai akses data *crawling*.
- b. Melakukan implementasi dan visualisasi data dengan *python* dan *package* yang terdapat pada pemrograman.
- c. Menguji tingkat akurasi sistem analisis sentimen menggunakan metode klasifikasi *Naïve Bayes*.

Manfaat yang ingin di capai dari penelitian ini adalah sebagai berikut:

1. Bagi Penulis

Dapat menerapkan teori dan praktik yang sudah di dapatkan selama menjalani pendidikan di Universitas Stikubank UNISBANK Semarang khususnya berupa pembuatan aplikasi text mining dengan metode analisa sentimen.

2. Bagi pembaca

Memberikan informasi akurat tentang penggunaan algoritma *Naïve Bayes Classifier* untuk mengelompokkan opini masyarakat terhadap pandemi *covid19* yang diambil dari pengguna *twitter*.

3. Bagi Universitas

Untuk melengkapi hasil penelitian yang nanti selanjutnya bisa digunakan sebagai acuan bagi peneliti lainnya.

1.5 Metode Penelitian

Metode penelitian yang digunakan dalam penulisan ini sebagai berikut:

1.5.1 Objek penelitian

Penelitian menggunakan opini masyarakat pada twitter terhadap keadaan pandemi *covid19* dengan kata kunci #corona dan #vaksin

1.5.2 Metode Pengumpulan Data

Pada penelitian ini, metode yang digunakan untuk memperoleh informasi dan menyelesaikan masalah adalah sebagai berikut:

1. Studi Kepustakaan

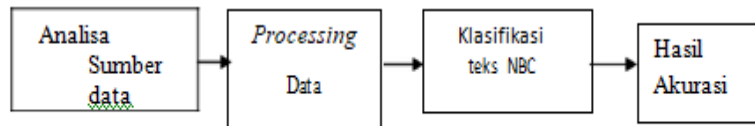
Terkait dengan permasalahan pandemi *covid19*, peneliti ini mengumpulkan data dengan melakukan studi pustaka seperti mengakses dokumen jurnal, paper, maupun situs internet sebagai acuan yang berkaitan dengan materi yang dibahas yaitu dengan algoritma *Naïve Bayes Classifier* .

2. Observasi

Melakukan pengamatan langsung pada obyek untuk mendapatkan gambaran yang lengkap mengenai data yang diproses.

1.5.3 Metode Analisa Data

Metode analisis data adalah proses yang digunakan dalam menganalisa setiap tahapan tahapan text mining data dari tweet sehingga diketahui kelas tersebut sentimen positif atau negatif. Berikut adalah gambaran penyelesaian masalah penelitian ini dapat dilihat pada gambar1.1



Gambar1. 1 Alur Sistem Analisa Data

Berikut adalah gambaran penyelesaian klasifikasi analisis sentimen pada gambar 1.1

1. Analisa Sumber Data

Penelitian ini menggunakan *API Twitter* untuk *mengcrawling* data dari *twitter* dan diunduh di dalam file Ms.Excel dalam format CSV, langkah selanjutnya yaitu menentukan analisis secara terhadap data yang sudah ada. Data berjumlah 500 data *tweet* yang terdiri dari 350 data *training* dan 150 data *testing*.

2. *Processing Data*

Dalam proses data Pada tahap ini, yaitu meliputi: *case folding*, normalisasi *tweet*, *stopword removal*, dan tokenisasi.

3. Klasifikasi *Naïve Bayes Classifier*

Sebuah metode untuk proses klasifikasi yang berakar pada *teorema Bayes* menggunakan algoritma *Naïve Bayes Classifier* dengan cara menghitung probabilitas pada setiap *tweet* yang akan menentukan opini negatif atau positif.

4. Hasil Akurasi

Proses terakhir adalah menghitung tingkat akurasi, yang akan dilakukan pengujian data yang sudah diketahui, sehingga setelah proses klasifikasi, selanjutnya menghitung tingkat akurasi sistem dalam pengklasifikasian menggunakan metode pengukuran *confusion matrix*.

1.6 Sistem Penulisan

Sistematika penulisan penelitian ini adalah sebagai berikut :

BABI PENDAHULUAN

Bab ini berisi dari Latar Belakang Penelitian, Perumusan Masalah, Tujuan dan Manfaat Penelitian, Metodologi Penelitian, dan Sistematika Penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini berisi uraian sistematis tentang informasi hasil penelitian yang telah dilakukan sebelumnya dan menghubungkannya dengan masalah penelitian yang sedang diteliti.

BAB III METODOLOGI PENELITIAN

Bab ini berisi tentang analisis pada komentar mengenai pelayanan kesehatan lalu di implementasi sistem yang sudah dirancang dalam bab sebelumnya.

BAB IV HASIL DAN PEMBAHASAN

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

BAB V PENUTUP

Bab ini berisi kesimpulan dan Saran.