

BAB I

PENDAHULUAN

1.1 Latar Belakang

Media sosial merupakan alat komunikasi yang digunakan orang untuk melakukan kegiatan promosi ataupun bersosial media yang berbentuk text, foto dan video. Pada tahun 2019, pengguna internet di Indonesia saat ini mencapai 63 juta orang. Dari angka tersebut, 95 persennya menggunakan internet untuk mengakses jejaringan sosial.

Penelitian yang dilakukan oleh Mohammad Zoqi Sarwani dan Wayan Firdaus Mahmudy yang berjudul “ Analisis twitter untuk mengetahui karakter seseorang menggunakan algoritma naive bayes classifier” pada tahun 2015. Pada penelitian mereka melakukan analisis yang berkaitan dengan karakter seseorang yang diambil pada salah satu layanan sosial media yang paling populer saat itu adalah twitter. *Twitter* menghasilkan 110 juta tweet setiap hari dan telah memiliki 200 juta pengguna. Banyak pengguna media sosial secara tidak sadar memberikan informasi tentang kepribadiannya melalui tweets atau posting yang mereka buat dengan bahasa alami. Penelitian sebelumnya untuk mengetahui tentang kepribadian seseorang.

Dilakukan pengolahan data diperlukan proses *pre-processing* terlebih dahulu sebelum dilakukannya pengklasifikasian. Dalam penelitian ini, metode klasifikasi yang dipakai adalah metode klasifikasi *Naive Bayes Classifier*.

Menurut World Health Organization (2017) pada umumnya gangguan mental yang terjadi adalah gangguan kecemasan dan gangguan depresi.

Diperkirakan 4,4% dari populasi global menderita gangguan depresi, dan 3,6% dari gangguan kecemasan. Jumlah penderita depresi meningkat lebih dari 18% antara tahun 2005 dan 2015. Depresi merupakan penyebab terbesar kecacatan di seluruh dunia. Lebih dari 80% penyakit ini dialami orang-orang yang tinggal di negara yang berpenghasilan rendah dan menengah (WHO, 2017).

Gangguan jiwa dapat terjadi pada siapa saja dan kapan saja. Hasil analisis dari WHO sekitar 450 juta orang menderita gangguan jiwa termasuk skizofrenia. Skizofrenia menjadi gangguan jiwa paling dominan dibanding gangguan jiwa lainnya. Penderita gangguan jiwa sepertiga tinggal di negara berkembang, 8 dari 10 orang yang menderita skizofrenia tidak mendapatkan penanganan medis. Gejala skizofrenia muncul pada usia 15-25 tahun lebih banyak ditemukan pada laki-laki dibandingkan pada

perempuan (Ashturkar & Dixit, 2013).

Skizofrenia adalah salah satu jenis psikotik yang menunjukkan gejala halusinasi dan waham (Townsend, 2011). Pasien dengan skizofrenia mempunyai gejala salah satunya adalah halusinasi akibat cemas berkepanjangan yang tidak mampu dihadapi pasien menggunakan mekanisme koping dalam diri pasien. Pendapat lain menyebutkan bahwa halusinasi yang terjadi pada pasien skizofrenia halusinasi gangguan alam perasaan yang tidak menentu, isi kebesaran atau kejaran, sering bertengkar atau berdebat, dan perilaku cemas yang tidak menentu dan kemarahan (Hawari, 2014). Penyebab gangguan jiwa salah satunya adalah adanya tekanan yang berat dalam peristiwa hidup. Stres berasal dari lingkungan atau biologi ataupun bisa keduanya (Videback, 2008).

Indonesia yang paling banyak diakses adalah Facebook dan Twitter. Untuk di Indonesia pengguna Twitter menempati peringkat 5 terbesar di dunia. Posisi ini kalah dari USA, Brazil, Jepang, dan Inggris. Rata-rata pengguna aktif harian di twitter mencapai 126 juta pengguna.

Lebih kurang tiga miliar orang, sekitar 40% populasi dunia, menggunakan media sosial dan menurut sejumlah laporan, kita menghabiskan rata-rata dua jam setiap hari untuk membagikan, menyukai, menuliskan cuitan dan perbaharui

perangkat dan hampir setengah juta cuitan dan foto dibagikan setiap menitnya. Cuitan ini selain dilakukan oleh pengguna tertentu seringkali juga dilakukan oleh akun-akun palsu yang digunakan untuk menyerang seseorang dan kebanyakan orang tentunya ini sangat mengganggu dan berpengaruh untuk membuat kegaduhan.

Penelitian ini menggunakan media twitter untuk memprediksi pola tingkah laku sosial pada seseorang. Dalam proses menganalisis pola tingkah laku seseorang melewati twitter harus menggunakan metodologi tepat untuk hasil yang akurat.

Untuk mengatasi hal ini, maka telah memberikan solusi yang memungkinkan pengguna untuk dapat mengetahui tweet yang dianggap mengandung unsur berdampak pada tingkah laku sosial. namun tindakannya membutuhkan waktu yang lama. Maka dari itu perlu pembuatan suatu sistem untuk mengklasifikasikan cuitan pada tingkah laku berbahasa Indonesia.

Twitter atau posting pada *Tweet* merupakan kumpulan kata yang tidak baku sehingga dibutuhkan perlakuan khusus untuk mendapatkan data yang bisa diproses.

Oleh karena itu, dalam pengolahan data diperlukan proses *pre-processing* terlebih dahulu yang kemudian bisa dilakukan pengklasifikasian. Dalam penelitian ini, metode klasifikasi yang dipilih adalah metode klasifikasi *Naïve Bayes*. Ada berbagai metode untuk klasifikasi seperti *Naive Bayes*, *Decision Trees*, *Support Vector Machines* atau *K-Nearest Neighbor*. Metode yang digunakan dalam penelitian ini adalah *Naive Bayes*. Metode *Naive Bayes* menggunakan konsep probabilitas setiap kelas dalam proses pembelajaran klasifikasinya.

Metode ini diambil karena mudah di implementasikan dan tidak terlalu membutuhkan sumber daya komputer yang besar.

Berdasarkan uraian diatas maka penulis akan melakukan penelitian sebagai skripsi dengan judul **“Implementasi Algoritma Naive Bayes Untuk Mengklasifikasi Gangguan Kesehatan Mental Berdasarkan Cuitan di Media Sosial”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka dapat dirumuskan rumusan masalah sebagai berikut:

1. Bagaimana implementasi algoritma *Naive Bayes* untuk klasifikasi cuitan pada tingkah laku di media sosial Twitter?

2. Berapa jumlah presentase tingkat akurasi algoritma *Naive Bayes* pada pengklasifikasian cuitan pada tingkah laku di media sosial Twitter?

1.3 Batasan Masalah

Agar penelitian dapat terfokus, maka perlu adanya batasan untuk membatasi lingkup penelitian ini. Adapun batasan penelitian yang dilakukan antara lain:

1. Algoritma yang akan digunakan pada penelitian ini adalah *Naive Bayes*.
2. Sistem menggunakan media *website*.
3. Proses *stemming* menggunakan *Rapidminer*
4. Data di peroleh dari *twitter* tahun 2019 - 2020.
5. Tweetan yang akan di klasifikasikan adalah cuitan yang berbahasa Indonesia.
6. Cuitan akan dikategorikan ke dalam tiga kelas dalam bahasa Indonesia yaitu Kecemasan, Suasana Hati dan *Skizofrenia*?
7. Untuk Menghitung Hasil Menggunakan Excel

1.4 Maksud dan Tujuan Penelitian

Maksud dan tujuan penelitian ini adalah sebagai berikut:

1. Membangun model yang dapat mengklasifikasikan cuitan Kecemasan, Suasana Hati dan *Skizofrenia*.
2. Mengevaluasi hasil klasifikasi dari model yang telah ditentukan.

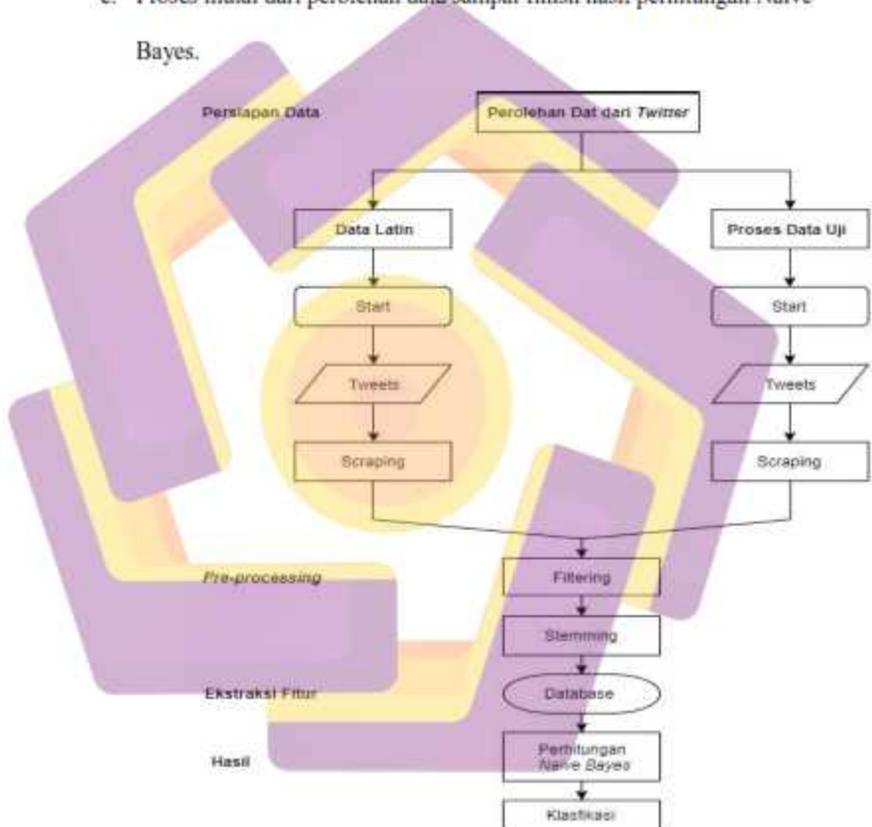


1.5 Metodologi Penelitian

Pada penelitian ini, penulis menggunakan beberapa metode penelitian, metode penelitian yang digunakan adalah sebagai berikut :

- a. Pengumpulan data diperlukan agar dapat informasi yang akurat dengan topik penelitian yang diambil. Adapun metode pengumpulan data yang diambil dari studi pustaka dalam proses pengumpulan bahan-bahan penelitian, antara lain buku, data cuitan, berbahasa Indonesia dari *Twitter*, jurnal-jurnal, skripsi maupun situs internet mengenai twitter dan algoritma *Naive Bayes* untuk membantu proses penelitian.
- b. Text mining masuk dalam text processing adalah sebuah metode untuk melakukan pemrosesan teks. Pertama, melakukan *ToLowerCase* yaitu mengubah seluruh data teks menjadi huruf kecil. Setelah itu melakukan pemisahan data berupa kalimat menggunakan delimiter spasi serta menghilangkan angka dan tanda baca lainnya yang tidak perlu pada suatu kata. Pada tahap ini menghilangkan kata yang dianggap *stopword*. Kemudian melakukan proses *stemming* untuk menghilangkan kata-kata yang memiliki imbuhan. *Stopword* adalah kata yang tidak memiliki arti. *Stemming* adalah proses yang dilakukan untuk melakukan pemetaan dan penguraian bentuk kata sehingga memiliki kata dasar.
- c. Perhitungan menggunakan *Naive Bayes* salah satu algoritma yang digunakan untuk mengklasifikasi teks serta merupakan metode *Machine Learning* yang menggunakan perhitungan probabilitas dan statistik yang dikemukakan oleh Thomas Bayes.

- d. Hasil Analisis masuk ketahap klasifikasi yang dilakukan dalam penelitian ini menggunakan beberapa pengguna *twitter* sebagai data testing. [7]
- e. Proses mulai dari perolehan data sampai finish hasil perhitungan Naive Bayes.



Gambar 1.1 Proses

1.6 Sistematika Penulisan

Penulisan skripsi ini di buat secara sistematis dalam beberapa bab, antara lain sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas tentang latar belakang penelitian, rumusan masalah, batasan masalah, maksud dan tujuan, metode penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini membahas tentang teori-teori dan tinjauan pustaka yang menjadi rujukan yang dijadikan dasar pada proses penelitian.

BAB III METODE PENELITIAN

Bab ini menjelaskan mulai dari persiapan penelitian, pengolahan data, alur penelitian dan melakukan analisis sistem menggunakan swot.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Bab ini menjelaskan secara rinci tentang program yang dibangun dan model yang diusulkan.

BAB V PENUTUP

Bab ini berisi kesimpulan dan saran dari penulis mengenai penelitian dan pengembangan lebih lanjut yang dapat dilakukan berkaitan dengan klasifikasi menggunakan algoritma *Naive Bayes*.

DAFTAR PUSTAKA

Daftar Pustaka berisi sumber rujukan penulis dalam melakukan penelitian.