

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Analisis sentimen merupakan suatu proses klasifikasi dokumen teks yang berupa penilaian, pendapat atau sentimen ke dalam dua kelas yakni positif atau negatif[1]. Terdapat beberapa penamaan dari analisis sentimen, yaitu *opinion mining*, *sentiment mining*, dan *opinion extraction*. Analisis sentimen termasuk salah satu cabang ilmu dari *Artificial Intelligence*, *Natural Language Processing*, dan *Text Mining*. Istilah “analisis sentimen” pertama kali digagas oleh Nasukawa dan Yi pada tahun 2003[2]. Dalam garis besar, analisis sentimen merupakan suatu proses dalam mengolah suatu data untuk diklasifikasikan kedalam jenis sentimen. Adapun sentimen yang umum digunakan adalah sentimen negatif, sentimen positif, dan sentimen netral. Sehingga hasil dari pengklasifikasian itu dapat diolah menjadi suatu informasi yang dapat berguna untuk menjadi acuan dalam membuat suatu pertimbangan yang berkaitan dengan objek yang diteliti pada analisis sentimen.

Moderasi Beragama di Indonesia menuai beragam sentimen dari masyarakat. Sentimen masyarakat tersebut terpolarisasi menjadi dua, yakni sentimen positif dan sentimen negatif. Sentimen positif dapat berupa dukungan ataupun sikap dan pernyataan setuju, sedangkan sentimen negatif dapat berupa penolakan dan sikap tidak setuju terhadap program tersebut. Beragam survei mengenai tanggapan masyarakat terhadap isu yang berkaitan dengan moderasi beragama telah dilakukan sebelumnya, seperti riset yang dilakukan oleh Tim Kelompok Kerja Moderasi Beragama terhadap indikator kinerja penguatan moderasi beragama tahun 2024[3]. Selain itu terdapat pula riset Convey 2018-2019, dari berbagai objek penelitian diantaranya guru PAI, Buku Ajar di PTU, Survey Nasional, Ulama, Imam, Takmir, dan Khatib, Dosen PAI, dan Buletin Jum’at. Data-data tersebut perlu dianalisa lebih lanjut agar kedepannya dapat dijadikan acuan yang tepat dalam membuat keputusan yang berkaitan dengan Moderasi Beragama, dalam hal ini dibutuhkan analisis sentimen terhadap isu Moderasi Beragama di Indonesia.

Twitter merupakan salah satu media sosial yang paling banyak digunakan setelah *Facebook*[4]. Indonesia menduduki peringkat ke-lima sebagai negara

dengan penduduk yang paling banyak menggunakan *Twitter*[5]. *Twitter* merupakan wadah yang tepat bagi pengguna media sosial untuk mengekspresikan emosi berupa tanggapan ataupun pandangan dalam bentuk teks. Adapun dalam pandangan psikologi terbagi menjadi dua yakni emosi negatif dan emosi positif[6]. Dengan adanya fitur *trending* pada *Twitter*, Pengguna dapat membagikan tanggapan atau pandangan terhadap suatu hal yang sedang marak dibicarakan baik di lingkup internasional maupun di dalam negeri. Salah satu topik yang sedang hangat diperbincangkan adalah Moderasi Beragama. Oleh karena itu pada penelitian ini akan dilakukan analisis sentimen Moderasi Beragama pada media sosial *Twitter* dengan memanfaatkan algoritma *Naïve Bayes*. Pemilihan Algoritma *Naïve Bayes* dalam penelitian ini didasari beberapa berikut:

1. *Naïve Bayes Classifier* bekerja sangat baik dibandingkan dengan model *classifier* lainnya. Seperti yang disebutkan dalam buku karangan Olson Delen, 2008 yang berjudul “Advanced Data Mining Techniques”, Dibuktikan oleh Xhemali, Hinde Stone dalam Jurnal “*Naïve Bayes vs Decision Tree vs Neural Networks in Classification of Training Web Pages*”[7].
2. Tidak memerlukan spesifikasi perangkat keras yang tinggi untuk pemrosesan data[8].
3. diantaranya Algoritma *Naïve Bayes* cenderung membutuhkan sedikit data dalam pemrosesan dibandingkan algoritma pesaingnya[9].
4. Perhitungan yang cepat dan efisien[8].
5. kodifikasi yang sederhana, dan lebih cepat dan efisien dibandingkan algoritma pesaingnya[8].

Penelitian sebelumnya mengenai analisis sentimen telah dilakukan diantaranya oleh Gunawan (2018) dengan judul “Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode *Naive Bayes*”, dan Nurhuda (2013) dengan judul “Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari *Twitter* Menggunakan Metode *Naive Bayes Classifier*”[8]. Terdapat beberapa perbedaan antara penelitian ini dengan penelitian serupa yang sebelumnya telah dilakukan, yang pertama dari segi model dan rancangan, terletak pada proses

pelabelan data yang pada penelitian sebelumnya dilakukan secara manual, namun pada penelitian ini dilakukan secara otomatis dengan menggunakan *TextBlob*. Selain itu proses pengambilan data yang dilakukan pada penelitian ini berbeda dengan penelitian sebelumnya yang diambil langsung dengan metode *scrapping data* dari *API Twitter*, namun menggunakan bantuan platform *Netlytic.org*. Dari segi objek penelitian terletak pada topik pembahasan yang akan dianalisis. Penelitian ini mengangkat topik moderasi beragama untuk dijadikan objek sentimen, dimana belum terdapat penelitian dengan topik yang sama dan menggunakan metode yang sama dengan penelitian ini.

1.2 Rumusan Masalah Penelitian

Adapun rumusan masalah dari penelitian ini adalah sebagai berikut :

1. Berapa akurasi yang diperoleh dari Algoritma *Naïve Bayes* terhadap analisis sentimen klasifikasi *tweet* pada topik moderasi beragama dari beberapa variasi nilai rasio yang berbeda antara data latih dan data uji?
2. Bagaimana hasil analisis pandangan masyarakat Indonesia terhadap Moderasi Beragama yang diperoleh dari perhitungan analisis sentimen menggunakan Algoritma *Naïve Bayes*?

1.3 Tujuan Penelitian

Tujuan dan manfaat dilakukannya penelitian ini adalah sebagai berikut :

1. Mengetahui akurasi yang diperoleh dari Algoritma *Naïve Bayes* pada kasus Moderasi Beragama dari beberapa variasi nilai rasio yang berbeda antara data latih dengan data uji.
2. Mengetahui hasil analisis pandangan masyarakat Indonesia terhadap Moderasi Beragama yang diperoleh dari perhitungan analisis sentimen menggunakan Algoritma *Naïve Bayes*.

1.4 Batasan Masalah Penelitian

Untuk membatasi pembahasan agar fokus penelitian tidak keluar dari topik, maka dibutuhkan batasan masalah. Adapun batasan-batasan masalah yang ditetapkan sebagai berikut :

1. *Tweet* yang digunakan untuk dianalisis hanyalah *tweet* dengan bahasa Indonesia.

2. *Keywords* yang digunakan dalam pengambilan data *tweet* yakni kata “moderasi” dan “beragama”.
3. Algoritma yang digunakan pada penelitian ini adalah *Naïve Bayes*.
4. Sentimen *Twitter* yang dianalisis hanya yang berkaitan dengan topik Moderasi Beragama.
5. Penelitian ini dilakukan dengan menggunakan Metode CRISP-DM tanpa tahapan *deployment*.
6. Pengambilan data menggunakan *API Twitter*.

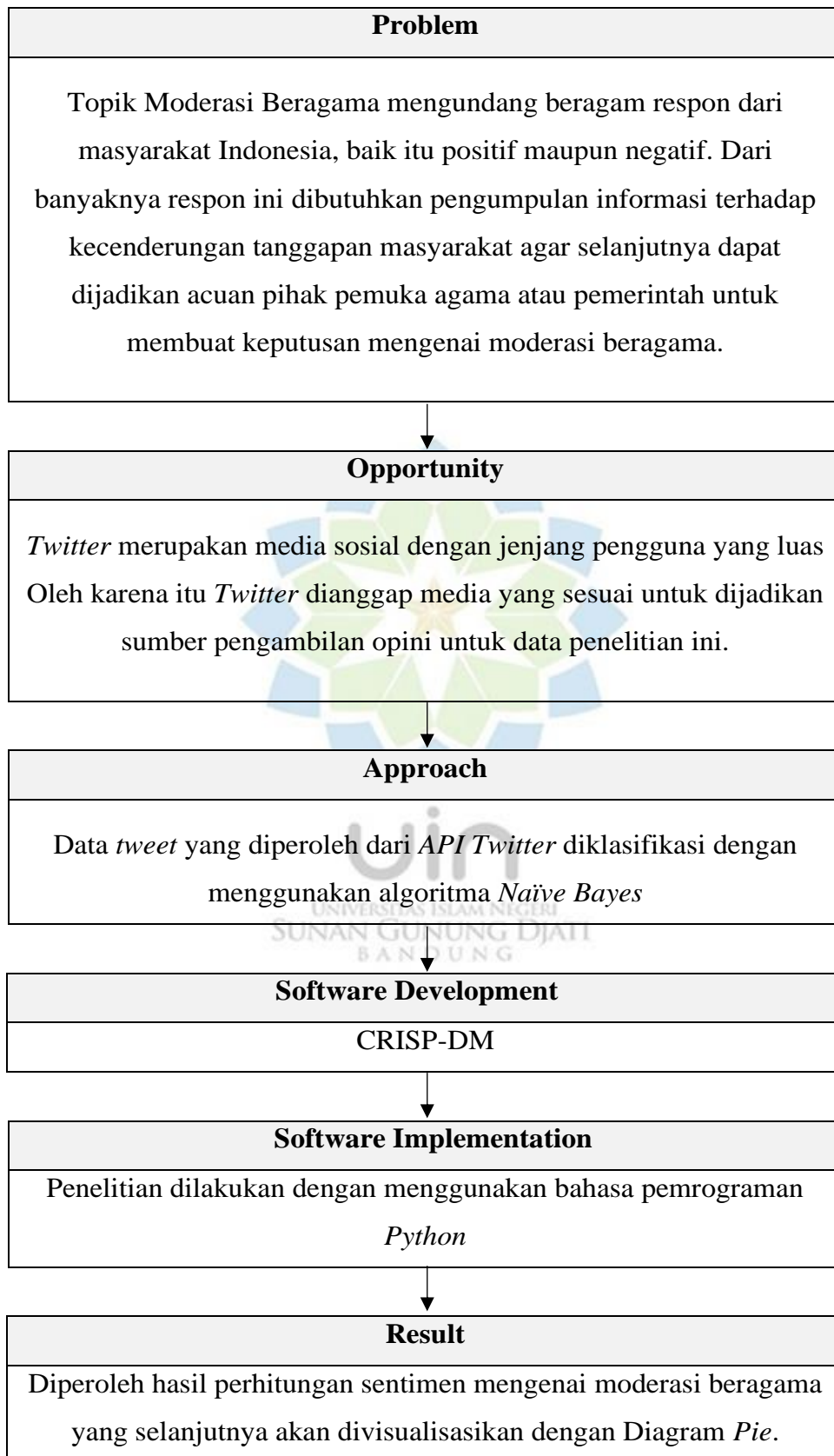
1.5 Kerangka Pemikiran Penelitian

Moderasi Beragama yang saat ini tengah ramai di sosialisasikan baik oleh kementerian agama maupun oleh instansi-instansi pendidikan seperti kampus dan sekolah mempunyai tujuan inti untuk mencegah radikalisme dan terorisme, namun demikian pandangan masyarakat terhadap konsep tersebut beragam dari mulai dukungan hingga penolakan.

Untuk mengetahui mayoritas pendapat berada di kubu mana, dibutuhkan data yang faktual dan terkini. Dalam mendapatkan data pandangan masyarakat di Indonesia tidak memungkinkan untuk membuat suatu kuisisioner, oleh karena itu pada penelitian ini digunakan media sosial *Twitter* sebagai sumber data. *Twitter* merupakan media sosial dengan jenjang pengguna yang luas dari mulai kalangan muda sampai tua, kalangan kelas ekonomi menengah sampai atas, dan kalangan pelajar, karyawan, hingga aparat negara. Oleh karena luasnya cakupan pengguna tersebut, *Twitter* dianggap media yang sesuai untuk dijadikan sumber pengambilan opini untuk data penelitian ini. pada umumnya dalam analisis sentimen yang bersumber data dari *Twitter*, pengambilan data dilakukan dengan cara *scrapping data* dengan memanfaatkan *API Twitter*, namun pada penelitian ini proses pengambilan data dilakukan dengan menggunakan bantuan *netlytic.org*. sehingga data yang diperoleh tersebut dapat dijadikan sampel gambaran dari keseluruhan suara penduduk di Indonesia. Solusi yang dilakukan untuk menganalisis kecenderungan masyarakat adalah dengan melakukan analisis sentimen terhadap topik tersebut.

Lebih jelasnya mengenai kerangka pemikiran penelitian ini diilustrasikan pada diagram 1.1 berikut:

Tabel 1. 1 Kerangka Pemikiran Penelitian



Dari segi data, data *Tweet* yang diperoleh akan diolah dengan menggunakan algoritma *Multinomial Naïve Bayes* dikarenakan alasan efisiensi dan kesesuaian dengan ukuran data yang telah disiapkan[10]. Penelitian dilakukan dengan metode *Cross-Industry Standard Process for Data Mining* atau CRISP-DM hingga tahapan *evaluation*, dengan mengimplementasikan bahasa pemrograman *Python*.

Hasil akhir dari penelitian ini adalah perhitungan sentimen mengenai moderasi beragama di Indonesia yang akan divisualisasikan dengan beberapa macam diagram, dan juga disertai dengan perhitungan akurasi, *precision*, *f1-score*, dan *recall* dari model yang telah dibuat.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Bab ini berisikan latar belakang penelitian, rumusan masalah penelitian, batasan masalah penelitian, tujuan penelitian dan manfaat penelitian, kerangka pemikiran, metodologi penelitian, dan sistematika penyusunan.

BAB II LANDASAN TEORI

Bab ini berisikan penjelasan mengenai teori-teori yang akan digunakan pada penelitian. Penjelasan tersebut mulai dari Moderasi Beragama hingga *Naïve Bayes Classifier* sebagai algoritma yang digunakan pada penelitian analisis sentimen moderasi beragama ini.

BAB III METODOLOGI PENELITIAN

Bab ini berisikan penjelasan mengenai metode yang digunakan dalam penelitian ini, dan deskripsi mengenai proses dari setiap tahapan pada metode yang dipilih.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisikan penjelasan dan laporan mengenai hasil penelitian, proses dan laporan penelitian, pembahasan akurasi hingga variabel terkait penelitian.

BAB V PENUTUP

Bab ini berisikan kesimpulan dari seluruh laporan yang menjawab masalah dari penelitian dan saran juga perbaikan untuk penelitian selanjutnya.