

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis sentimen merupakan cabang dari ilmu *text mining* yang merupakan metode klasifikasi dengan teknik menganalisis data teks tentang suatu topik, peristiwa, atau individu yang bertujuan untuk mengekstrak informasi sentimen, penilaian, sikap, evaluasi, dan emosi milik seseorang dari kalimat opini. Data sentimen yang sudah didapatkan kemudian akan dikategorikan berdasarkan polaritasnya, apakah positif atau negatif [1]. Sentimen positif menyatakan menyajikan nilai yang baik sedangkan sentimen negatif menyatakan memberikan nilai buruk pada data yang digunakan yang mana data tersebut berbentuk teks.

Di era *big data* saat ini dengan perkembangan yang cukup pesat, penelitian data di bidang analisis sentimen sangat penting dan diperlukan. Analisis sentimen cukup sering dimanfaatkan oleh beberapa pelaku usaha yang mengeluarkan produk berupa penyedia layanan untuk menganalisis ulasan dari konsumen terhadap produk tersebut [2]. Satu diantara contoh dari pengaplikasian metode analisis sentimen di dunia bisnis adalah analisis sentimen terhadap ulasan suatu aplikasi. Salah satu penelitian yang sudah dilakukan yaitu analisis sentimen dengan objek ulasan aplikasi Tokopedia pada tahun 2019 dengan menggunakan metode *naïve bayes* [3].

Selain dalam dunia bisnis, analisis sentimen juga dipergunakan dalam pengkategorian opini positif dan negatif dari khalayak umum terhadap sebuah kebijakan atau issue yang sedang ramai dibicarakan. Seperti akhir-akhir ini banyak dibicarakan tentang pemilu tahun 2024. Dalam pelaksanaan pemilu, isu sensitif seringkali muncul. Salah satu contohnya yaitu isu terkait dinasti politik ataupun kecurangan lainnya, dan pendapat publik yang cukup bermacam terhadap calon presiden dan wakil presiden atau pada suatu partai politik tertentu. Pemilu seharusnya dijalankan dengan standar yang lebih baik agar terjaminnya persaingan yang sehat, partisipatif, serta memiliki mekanisme yang dapat dipertanggungjawabkan dengan jelas. Dengan demikian, untuk memastikan adanya jaminan pesta demokrasi yang sehat, pemilihan umum yang akan dilaksanakan

perlu memiliki kualitas yang baik. Setiap pemilih harus memiliki peran yang setara dan dihormati, sementara setiap peserta pemilu juga harus mendapatkan kesempatan yang sama serta adil untuk bersaing. Hal ini penting untuk memastikan bahwasannya hasil dari pemilihan merefleksikan kehendak dan juga aspirasi mayoritas masyarakat. [17].

Pada penelitian terdahulu, telah dilakukan analisis sentimen dengan objek tentang opini publik terhadap kebijakan yang dibuat pemerintah yaitu terkait tentang revisi KPK pada tahun 2019 yang mana *dataset* yang didapat diambil dari media sosial *X/Twitter*. Dengan melakukan komparasi terhadap tiga metode klasifikasi yang diusulkan pada jurnal ini memperoleh nilai akurasi sebesar 81.70% menggunakan metode *Support Vector Machine*, kemudian sebesar 80.90% dengan metode *Naïve Bayes* dan 74,55% dengan *Decision Tree* [4].

Didasari oleh penelitian yang sudah dilaksanakan dapat dijabarkan bahwa penngaplikasian metode *Support Vector Machine* pada klasifikasi ini sangat baik dalam mengklasifikasikan data teks. Hal ini dapat dilihat dengan membandingkan metode *Naïve Baiyes* dan *Decision Tree*, dengan *Support Vector Machine* yang mana hasil akurasinya lebih baik. SVM adalah salah satu macam dari metode klasifikasi yang mana menggunakan *metode machine learning (supervised learning)* dengan cara kerja memprediksi kelas yang didasarkan pada pola dari hasil proses 2 pelatihan yang dicetuskan oleh Vladimir Vapnik pada tahun 1992 [5]. Cara kerja klasifikasi pada metode SVM diawali dengan garis *hyperplane* yaitu garis pembagi yang memisahkan antara kelas bernilai opini positif dan dengan kelas bernilai opini negatif. Metode klasifikasi ini merupakan satu di antara beberapa metode yang memiliki tingkat nilai akurasi yang baik [6].

Berdasarkan penjelasan dari penelitian sebelumnya, menunjukkan bahwa penggunaan *Support Vector Machine* pada klasifikasi ini sangat baik dalam mengklasifikasikan data teks hal yang lebih baik dilihat dari hasil akurasi dibandingkan dengan metode *Naïve Bayes*. Selain itu dikarenakan beberapa isu kecurangan yang ada di proses pemilihan umum di Indonesia. Seperti penggelembungan suara, kampanye hitam, dan politik uang merupakan isu kecurangan yang sering terjadi. Dalam konteks ini ada beberapa pandangan rakyat

Indonesia yang berupa tanggapan positif ataupun negatif. [18] Oleh karena itu, dilakukan penelitian analisis sentimen dengan menggunakan metode *Support Vector Machine* pada opini publik terhadap Pemilu 2024.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Berdasarkan penjelasan pada latar belakang diatas, maka permasalahan pada penelitian ini adalah kurangnya akurasi yang dihasilkan dalam mengklasifikasikan opini publik menggunakan metode klasifikasi seperti *naïve bayes* berdasarkan penelitian terdahulu.

1.2.2 Metode Penelitian

Berdasarkan permasalahan diatas, dari penelitian ini yaitu digunakannya teknik klasifikasi menggunakan metode *Support Vector Machine*. Sedangkan pengambilan data dilakukan menggunakan metode *Crawling* data X/Twitter.

1.2.3 Pertanyaan Penelitian

Berdasarkan permasalahan dan solusi yang akan digunakan untuk melakukan klasifikasi sentimen didapatkan pertanyaan yaitu, Berapa nilai akurasi klasifikasi data sentimen pada opini publik terhadap Pemilu 2024 dengan menggunakan metode *Support Vector Machine*?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan Penelitian

Berdasarkan penjelasan pada latar belakang diatas, tujuan dari penelitian ini adalah untuk memperoleh hasil akurasi optimal dalam mengklasifikasi data sentimen untuk mengklasifikasikan data dengan kelas positif dan negatif

1.3.2 Manfaat Penelitian

Manfaat dengan adanya penelitian analisis ini yakni sebagai berikut :

1. Mengetahui hasil kinerja dari metode klasifikasi *Support Vector Machine* dalam melakukan klasifikasi data teks.
2. Memberikan informasi atau gambaran tentang pandangan masyarakat terhadap Pemilu 2024.

3. Menjadi acuan bagi peneliti berikutnya yang akan melakukan penelitian sejenis.

1.4 Batasan Masalah

Penelitian ini hanya untuk analisis sentimen pada *Tweet* di media sosial *X/Twitter* yang kemudian dibatasi dengan beberapa hal berikut:

1. *Tweet* yang digunakan menggunakan Bahasa Indonesia
2. Hanya *tweet* yang mengandung kata kunci “Prabowo”, “Gibran”, “Ganjar”, “Mahfud MD”, “Anies Baswedan”, “Muhaimin”, dan “Pemilu 2024” saja yang digunakan pada penelitian ini
3. Data diambil sejak tanggal 18 Oktober - 24 Oktober 2023 sebanyak 2015 *tweet*.
4. Pembagian data latih dan data uji menggunakan metode *k-fold cross validation*.
5. Kumpulan *tweet* yang sesuai dengan topik pada penelitian ini akan diklasifikasikan ke dalam kelas negatif dan positif.
6. Perhitungan dan analisis menggunakan metode algoritma *Support Vector Machine*.
7. Pengujian data menggunakan aplikasi yang dibuat.

1.5 Sistematika Penelitian

Sistematika pada penelitian ini adalah sebagai berikut :

1. Bab I Pendahuluan

Bab ini merupakan bagian awal dari penelitian ini yang menjelaskan tentang latar belakang penelitian, permasalahan, tujuan penelitian, batasan masalah, dan sistematika penulisan.

2. Bab II Kajian Pustaka

Pada bab ini berisi teori dan tinjauan pustaka yang menjadi dasar penunjang yang berkaitan dengan penelitian ini penelitian. Teori yang dimuat ialah teori yang berhubungan pada bidang *Text Mining* dengan menggunakan metode algoritma *Support Vector Machine*. Serta beberapa penelitian sebelumnya yang dipakai menjadi rujukan dalam mengerjakan penelitian ini.

3. Bab III Metode Usulan

Pada bab ini berisikan penjelasan proses atau langkah-langkah yang digunakan dalam penyelesaian permasalahan menggunakan metode yang telah ditentukan.

4. Bab IV Hasil dan Pembahasan

Bab ini menjelaskan hasil klasifikasi sentimen yang dihasilkan dengan menggunakan metode *Support Vector Machine*.

5. Bab V Penutup

Bab ini ialah bagian terakhir dari penelitian yang memuat tentang kesimpulan yang didapatkan dari penelitian yang sudah dikerjakan serta juga saran untuk pengembangan penelitian lebih lanjut.