

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Di era teknologi media sosial yang semakin berkembang, informasi dapat tersebar dengan mudah sehingga mempengaruhi cara pandang dan gaya hidup masyarakat. Salah satu media sosial yang banyak digunakan masyarakat saat ini yaitu *Twitter*. Menurut Rahma, dkk., (2021) *twitter* adalah media sosial untuk berbagi tentang kehidupan sehari-hari atau informasi secara *real time*. Ada berbagai macam fitur yang tersedia pada media sosial *Twitter* salah satunya yaitu fitur *trending topic* yang dapat memberikan informasi terpopuler yang sedang diperbincangkan oleh banyak pengguna (Rufaidha & Irhandayaningsih, 2022).

Salah satu berita yang sedang diperbincangkan oleh pengguna *Twitter* pada saat ini ialah tentang calon pemilihan Presiden 2024. Pemilihan Presiden di Indonesia merupakan momen penting dimana masyarakat Indonesia dapat melalui proses demokrasi (Saputra, dkk., 2022). Dikarenakan masa jabatan Presiden Jokowi saat ini telah mencapai 2 periode yaitu sejak tahun 2014 hingga 2024 mendatang serta berdasarkan ketentuan pada pasal 7 UUD 1945 dalam amandemen pertama disebutkan bahwa “Presiden dan Wakil Presiden memegang jabatan selama lima tahun, dan sesudahnya dapat dipilih kembali dalam jabatan yang sama, hanya untuk satu kali masa jabatan”. Oleh sebab itu, banyak isu terhadap tokoh-tokoh yang ingin mencalonkan diri sebagai presiden salah satunya ialah Ganjar Pranowo.

Ganjar Pranowo adalah Gubernur Jawa Tengah dua periode yang menjabat sejak 23 Agustus 2013. Beliau juga salah satu politisi yang memanfaatkan media sosial untuk mengkampanyekan diri, membentuk citra maupun membangun komunikasi politik kepada masyarakat. Oleh sebab itu, penting untuk menganalisis sebuah *tweet* seputar isu-isu yang diutarakan pengguna *twitter* dengan menerapkan teknik analisis sentimen.

Analisis sentimen adalah proses meninjau opini publik terhadap suatu topik tertentu yang datanya bersumber dari internet atau media sosial (Kurniawan & Waluyo, 2022). Menurut Classifier & Bayes (2021) analisis sentimen itu perlu

dilakukan karena penggunaan media sosial yang tumbuh untuk mempengaruhi perkembangan opini publik. Pada analisis sentimen ini, penulis menggunakan tahapan ekstraksi fitur TF-IDF. TF-IDF adalah metode yang berguna untuk menghitung nilai bobot dalam sebuah teks serta dapat mengurangi ukuran teks sehingga menghindari ruang fitur yang besar (Suparyanto & Rosad, 2020). Jadi, untuk mengetahui performa dalam analisis sentimen tersebut, selain menggunakan ekstraksi fitur TF-IDF penulis juga menerapkan algoritma *Multivariate Bernoulli*.

Menurut Suparyanto & Rosad (2020) algoritma *Multivariate Bernoulli* adalah suatu metode perkembangan dari *Naïve Naves*. Metode ini memiliki hipotesis yang kuat dalam mengklasifikasikan suatu kejadian analisis. Metode *Multivariate Bernoulli* memiliki akurasi yang cukup baik dalam pengklasifikasiannya. Sebagaimana penelitian yang dilakukan oleh (Pratama, dkk., 2022) yang berjudul “Pengklasifikasian Kanker Payudara dan Kanker Paru-Paru dengan Metode *Gaussian Naïve Bayes*, *Multinomial Naïve Bayes*, dan *Bernoulli Naïve Bayes*”. Hasil dari penelitian tersebut menunjukkan performa rata-rata metode *Bernoulli Naïve Bayes* lebih unggul dengan hasil rata-rata *accuracy* 93,25%, rata-rata *precesion* 94,23%, dan rata-rata *recall* 94,69%. Dari penelitian yang telah disebutkan, hal ini menunjukkan bahwa algoritma *Multivariate Bernoulli* memiliki akurasi yang cukup tinggi dalam mengklasifikasikan sebuah analisis sentimen. Dikarenakan terdapat pro dan kontra mengenai isu pencalonan Ganjar Pranowo sehingga tak luput dari banyaknya opini serta adu argumen antar warganet. Oleh karena itu, penelitian ini berfokus untuk menganalisis *tweet* terhadap opini publik dengan menerapkan proses klasifikasi analisis sentiment dengan hasil positif dan negatif yang berjudul “Implementasi Algoritma *Multivariate Bernoulli* pada Analisis Sentimen *Tweet* terhadap Opini Publik atas Isu Pencalonan Ganjar Pranowo Sebagai CAPRES 2024”. Dengan demikian, peneliti dapat mengetahui akurasi, presisi, dan *recall* dari metode *Multivariate Bernoulli* serta jumlah sentimen pada *tweet* yang mengacu pada opini publik atas isu pencalonan Ganjar Pranowo sebagai CAPRES 2024.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang, maka dapat dituliskan rumusan masalah sebagai berikut:

1. Berapakah akurasi, presisi, dan *recall* yang dihasilkan oleh algoritma *Multivariate Bernoulli* pada analisis sentimen *tweet* terhadap opini publik atas isu pencalonan Ganjar Pranowo sebagai CAPRES 2024?
2. Berapakah jumlah sentimen positif dan negatif pada analisis sentimen *tweet* terhadap opini publik atas isu pencalonan Ganjar Pranowo sebagai CAPRES 2024?

## 1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengetahui akurasi, presisi, dan *recall* yang dihasilkan oleh algoritma *Multivariate Bernoulli* pada analisis sentimen *tweet* terhadap opini publik atas isu pencalonan Ganjar Pranowo sebagai CAPRES 2024.
2. Mengetahui jumlah sentimen positif dan negatif pada analisis sentimen *tweet* terhadap opini publik atas isu pencalonan Ganjar Pranowo sebagai CAPRES 2024.

## 1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Mengetahui performa metode algoritma *Multivariate Bernoulli* dalam mengklasifikasikan pada analisis sentimen *tweet* terhadap opini publik atas isu pencalonan Ganjar Pranowo sebagai CAPRES 2024.
2. Hasil penelitian ini dapat digunakan sebagai referensi oleh penelitian selanjutnya di bidang analisis sentimen dengan menggunakan metode *Multivariate Bernoulli*.

## 1.5 Batasan Penelitian

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Menggunakan algoritma *Multivariate Bernoulli*.

2. Pengambilan data pada *Twitter* menggunakan API (Application Programming Interface).
3. Memperoleh data sebanyak 600 data. Data diambil pada tanggal 01 September 2022 – 31 Januari 2023.
4. Pencarian data menggunakan kata kunci “Capres Ganjar Pranowo”.
5. Klasifikasi akhir dibagi menjadi 2 kelas yaitu positif dan negatif.
6. Menggunakan (TF-IDF) *Term Frequency-Inverse Document Frequency* sebagai pembobotan kata.
7. Menggunakan bahasa pemrograman *Phyton*.

