

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Media sosial di saat ini merupakan salah satu bagian penting dari kehidupan bersosial masyarakat. Hampir seluruh kalangan umur terkhususnya di Indonesia memiliki media sosial. Berdasarkan data yang dicatat oleh *We Are Social*, melansir bahwa total pengguna media sosial aktif yang ada di Indonesia sebanyak 167 juta orang pada Januari 2023, dan jika dibandingkan dengan jumlah populasi di Indonesia, angka tersebut setara dengan 60.4% dari populasi masyarakat Indonesia.

Twitter merupakan salah satu media sosial terkenal, dan banyak digunakan di Indonesia. Berdasarkan survei yang telah dilakukan oleh *Global Web Index (GWI)* pada kuartal 3 tahun 2022 berdasarkan persentase pengguna internet berusia 16-64 tahun, Twitter termasuk ke dalam peringkat ke-7 dengan pengguna terbanyak di Indonesia. Sebanyak 60.2% pemakai internet di Indonesia mengakses twitter, dengan jumlah sekitar 24 juta pengguna di Indonesia pada tahun 2023 (GWI, 2022).

Terdapat satu fitur yang andalan yang terdapat pada Twitter, yaitu *tweet*, para pengguna diberikan kebebasan untuk mem-*posting* apapun, baik dalam bentuk teks, foto, bahkan video. Berdasarkan data dari *Statista* terdapat 500 juta *tweet* diterbitkan tiap harinya, dan 350.000 *tweet* di-*posting* setiap menit. Topik *tweet* yang ada pada Twitter juga beragam, mulai dari hobi, opini, dan juga kritik terhadap suatu hal. Twitter Indonesia menghimpun topik yang paling ramai dibicarakan pengguna di Indonesia, dan tercatat bahwa politik menjadi salah satu pembahasan yang sedang hangat dan sering dibahas di kalangan pengguna Twitter di Indonesia.

Pemilu 2024 menjadi ajang tahun politik di Indonesia. Mulai dari tahun ini, banyak berseliweran di media sosial mengenai politik, baik berupa kampanye ataupun isu mengenai politik itu sendiri. Berdasarkan hal tersebut topik politik juga semakin marak dibicarakan di Twitter oleh masyarakat Indonesia. Salah satu topik politik yang sering dibicarakan adalah mengenai calon presiden.

Beberapa bakal calon presiden berdasarkan hasil survei yang sudah dilakukan menghasilkan tiga nama calon presiden pada pemilu 2024 nanti, antara lain Ganjar

Pranowo, Anies Baswedan, dan Prabowo Subianto. Nama-nama tokoh politik diatas nantinya akan dijadikan sebagai kata kunci dalam klasifikasi emosi dari para pengguna Twitter terhadap pilpres 2024, dan akan dijadikan acuan untuk melihat respon masyarakat.

Respon yang diberikan masyarakat Indonesia terhadap topik politik tersebut juga beragam, ada yang membenci dan juga mendukung. Tidak jarang adanya doa, ungkapan senang, dan bahkan ujaran kebencian ditemukan di Twitter mengenai bakal calon presiden yang terlibat dalam tahun politik. Beragamnya emosi dari masyarakat sangat erat kaitannya dengan tahun politik yang sedang berlangsung saat ini, bahkan fenomena ini sudah berlangsung dari tahun ke tahun.

Banyaknya konten dan juga respon dari masyarakat kepada suatu tokoh politik sering berdampak negatif terhadap media sosial di Indonesia, salah satunya yaitu munculnya konten *hoax*. Tidak jarang muncul konten-konten *hoax* dengan tujuan untuk menjatuhkan dan memberikan citra buruk terhadap suatu tokoh politik, dan hal ini sudah menjadi fenomena lumrah yang terjadi menjelang tahun politik. Hal yang sama juga terjadi pada tahun 2019 silam, di momen yang sama yaitu ketika tahun politik. Berdasarkan data dari Kementerian Komunikasi dan Informatika (Kominfo) pada tahun 2019 lalu, jumlah konten *hoax* yang teridentifikasi pada tahun 2018 – 2019 silam sebanyak 1.731 konten. Konten *hoax* yang memicu munculnya ujaran kebencian diidentifikasi sebanyak 486 unggahan selama bulan April 2019 (Kominfo, 2019).

Beragamnya respon para pengguna Twitter terhadap bakal calon presiden di pemilu 2024 mengakibatkan juga banyaknya jenis dari emosi “cuitan” para penggunanya. Emosi sendiri merupakan reaksi yang diberikan manusia sebagai respons terhadap suatu kondisi atau peristiwa. Reaksi yang diberikan melibatkan perubahan-perubahan fisik dan psikologi terhadap rangsangan yang diterima. Emosi juga dapat diartikan sebagai perasaan intens yang muncul terhadap seseorang atau sesuatu dan merupakan reaksi terhadap situasi tertentu (Neufeld, 1999).

Teks adalah jenis media yang dipakai dalam berkomunikasi dan memberikan informasi. Selain sebagai penyampai informasi, teks juga digunakan untuk mengekspresikan emosi (Hirat & Mittal, 2015). Emosi yang diberikan manusia

terbagi menjadi beberapa jenis, dalam penelitian ini, emosi yang akan digunakan dibagi menjadi ke dalam enam jenis emosi, yaitu emosi senang, sedih, marah, takut, terkejut, jijik, dan ditambah dengan satu emosi netral untuk data teks yang tidak memiliki emosi apapun. Keenam emosi ini ditemukan pertama kali oleh Paul Ekman, seorang psikolog yang menjadi perintis dalam studi emosi. Teori Paul Ekman telah banyak digunakan dalam bidang klasifikasi emosi, salah satu penelitian yang menerapkan teori ini adalah pengenalan emosi dari ekspresi wajah manusia (Septian *et al.*, 2020).

Klasifikasi emosi pada teks merupakan penerapan metode yang digunakan untuk dapat membagi jenis-jenis teks berdasarkan emosinya. Penerapan dari metode ini sendiri bertujuan untuk memahami perasaan yang disampaikan penulis, lalu juga untuk mengidentifikasi perasaan positif atau negatif seseorang terhadap suatu hal.

Algoritma Naïve Bayes adalah algoritma klasifikasi yang akan digunakan pada penelitian ini. Algoritma Naïve Bayes memiliki kelebihan pada bentuk model yang sederhana, akan tetapi algoritma ini masih dapat bersaing dengan model algoritma yang lainnya. Pada penelitian terdahulu Naïve Bayes adalah salah satu pengklasifikasian paling sederhana dalam bidang *machine learning*, namun algoritma ini masih mendukung mesin vektor (Rish, 2001)., Naïve Bayes sendiri juga memiliki kemampuan kecepatan yang efisien, karena hanya memerlukan waktu yang sedikit dalam melakukan pelatihan.

Algoritma *Naïve Bayes* memiliki keunggulan yaitu cepat dihitung dan akurat meskipun sederhana. Untuk memastikan perkiraan yang dibutuhkan dalam proses klasifikasi, *Naïve Bayes* tidak memerlukan banyak data pelatihan (*Data Training*). Jalur perhitungan yang lebih pendek membuat algoritma ini lebih mudah digunakan (Rosandy, 2016). Berdasarkan penelitian yang sudah dilaksanakan oleh (Ifon Purnama *et al.*, 2020). Metode *Naïve Bayes* menerapkan model *Confusion Matrix* dengan 210 jumlah data, 154 sebagai *data training* dan 56 *data testing* akurasi yang dihasilkan sebesar 82,14%.

Metode *Naïve Bayes* dapat diterapkan pada pengklasifikasian data yang berupa berbentuk teks. Dalam penelitian yang dilakukan sebelumnya oleh (Harijati, 2019), metode *Multinomial Naïve Bayes* dapat digunakan untuk pengolahan data,

terutama dalam pengklasifikasian teks. Penelitian yang sama juga dilakukan oleh (Aulia & Amelia, 2020). Metode *Naïve Bayes* lumrahnya sering digunakan dalam penelitian analisis sentimen. Biasanya pada penelitian analisis sentimen terdapat satu penggunaan metode penting praproses penghapusan kata (*stopword removal*) mendapatkan hasil yang efektif dalam proses pengklasifikasian analisis sentimen.

Hasil yang diharapkan dari penelitian ini adalah sebuah model yang dapat digunakan untuk klasifikasi emosi berdasarkan teks *tweet* para pengguna Twitter di Indonesia. Data emosi pengguna ini dapat digunakan sebagai acuan untuk mengidentifikasi emosi dan juga pendapat masyarakat Indonesia terhadap tokoh politik di pemilu 2024 nanti. Diharapkan nantinya setelah adanya penelitian ini dapat menjadi bahan rujukan dalam penelitian mengenai klasifikasi emosi dalam teks dengan menggunakan algoritma *Naïve Bayes*.

1.2 Identifikasi Masalah

Berdasarkan paparan latar belakang di atas, maka dapat diidentifikasi masalah dalam penelitian ini antara lain sebagai berikut:

- 1) Terdapat banyak *tweet* dari masyarakat Indonesia terhadap bakal calon presiden menjelang pemilu 2024 .
- 2) Beragamnya emosi pada teks dari *tweet* yang diberikan masyarakat Indonesia terhadap bakal calon presiden menjelang pemilu 2024.
- 3) Tidak terklasifikasinya jenis-jenis emosi dari keseluruhan *tweet* yang diberikan masyarakat Indonesia terhadap bakal calon presiden menjelang pemilu 2024.
- 4) Respon pengguna Twitter di Indonesia mempengaruhi citra publik para bakal calon presiden di pemilu 2024.

1.3 Batasan Masalah

Untuk memberikan ruang lingkup yang jelas dalam pembahasan, maka perlu dilakukan pembatasan masalah dalam penelitian ini sebagai berikut:

- 1) Data *tweet* yang digunakan untuk dataset adalah *tweet* yang berbahasa Indonesia terhadap 3 bakal calon presiden, yaitu Anies Baswedan, Ganjar Pranowo, dan Prabowo Subianto pada bulan Juli 2023 minggu pertama.

- 2) Data yang digunakan dalam penelitian ini sebanyak 5000 *tweet* tiap bakal calon presiden, dengan total data keseluruhan adalah 15.000 *tweet*.
- 3) Klasifikasi emosi hanya dilakukan terhadap *tweet* teks yang berisi opini atau pendapat pengguna Twitter mengenai bakal calon presiden pemilu 2024.
- 4) Kelas emosi yang digunakan sebagai data awal ditentukan dengan subjektif peneliti dengan referensi dari validator.
- 5) Emosi yang diklasifikasikan hanya dilakukan terhadap teks dengan emosi yang bersifat eksplisit.
- 6) Hasil dari *preprocessing data* terkadang tidak akurat, sehingga proses pelabelan dilakukan penyesuaian dengan narasi dan makna teks.

1.4 Rumusan Masalah

Berdasarkan identifikasi masalah yang telah dilakukan, maka dirumuskan masalah yang akan dibahas dalam penelitian ini sebagai berikut:

- 1) Bagaimana menyusun dan melatih model Naive Bayes untuk mengklasifikasi emosi dari teks pada pengguna Twitter terhadap bakal calon presiden di pemilu 2024?
- 2) Bagaimana mengevaluasi keakuratan model Naive Bayes dalam mengklasifikasi emosi dari teks pada pengguna Twitter bakal calon presiden di pemilu 2024?
- 3) Bagaimana memperoleh informasi mengenai citra publik masyarakat dan menjadi perbandingan terhadap para bakal calon presiden di pemilu 2024?

1.5 Tujuan Penelitian

Berdasarkan masalah yang telah diuraikan diatas maka tujuan penelitian yang akan dicapai adalah sebagai berikut:

- 1) Untuk menyusun dan melatih model Naive Bayes dalam mengklasifikasi emosi dari teks pada pengguna Twitter terhadap bakal calon presiden di pemilu 2024.
- 2) Untuk mengevaluasi keakuratan model Naive Bayes dalam mengklasifikasi emosi dari teks pada pengguna Twitter terhadap Twitter terhadap bakal calon presiden di pemilu 2024.

- 3) Untuk memperoleh informasi mengenai citra publik masyarakat dan menjadi perbandingan terhadap para bakal calon presiden di pemilu 2024.

1.6 Manfaat Penelitian

- 1) Menjadi pengembangan ilmu dan menambah pengetahuan tentang penerapan algoritma Naïve Bayes dalam klasifikasi emosi terhadap *tweet* dari pengguna Twitter di Indonesia.
- 2) Memberikan wawasan mengenai preferensi dan emosi pemilih terkait bakal calon presiden pada pemilu 2024 berdasarkan *tweet* di Twitter.
- 3) Menjadi acuan bagi para pemimpin politik dan partai politik dalam hal pengambilan kebijakan terkait bakal calon presiden di pemilu 2024 berdasarkan emosi masyarakat yang beredar di Twitter.

