

BAB I

PENDAHULUAN

1.1. Latar Belakang

Sentimen adalah ekspresi atau penilaian subjektif dari perasaan atau sikap seseorang terhadap suatu topik, peristiwa, atau situasi. Analisis sentimen adalah teknik menafsirkan dan pengolahan data tekstual untuk mengekstrak data opini dan menentukan sentimen yang ada dalam sebuah opini (Sari & Wibowo, 2019). Analisis sentimen memiliki banyak dampak positif yang dapat di raih salah satunya dalam penelitian dan pemahaman masyarakat. Analisis sentimen penting dalam penelitian sosial untuk memahami pandangan masyarakat terhadap topik tertentu (Liawati et al., 2023). Dalam melakukan analisis sentimen diperlukan metode untuk mengklasifikasikan sentimen tersebut. Di penelitian ini klasifikasi sentimen akan menggunakan metode *K-Nearest Neighbors* dan Metode *Naive Bayes Classifier*.

Metode *K-Nearest Neighbors* merupakan metode yang digunakan untuk melakukan klasifikasi data berdasarkan jarak terdekat terhadap objek data (Cholil et al., 2021). Metode *K-Nearest Neighbors* banyak digunakan di dalam penelitian sebelumnya dikarenakan modelnya yang mudah dipahami, dan memberikan performa yang wajar tanpa banyak penyesuaian (Raharjo, 2021). Contoh penelitian yang telah menggunakan metode *K-Nearest Neighbors* yaitu pada penelitian Ramadhani & Wahyudin (2022) kemudian pada penelitian Aris & Informatika (2022), dan penelitian Fudhail Ferio Supeli & Nusa Mandiri (2023). Dari ketiga penelitian di atas, peneliti menyimpulkan bahwa kelebihan dari algoritma *K-Nearest Neighbors* yaitu sangat mudah diimplementasikan jika dengan jumlah data yang tidak terlalu banyak. Akan tetapi jika datanya berjumlah sangat banyak akan sangat sulit untuk dihitung. Hal ini dapat dilihat dari tingkat akurasi yang semakin berkurang seiring dengan semakin banyaknya data.

Selanjutnya Metode *Naive Bayes Classifier* merupakan sebuah metode kategorisasi data yang merupakan turunan dari teorema Bayes yang menggunakan teknik statistik serta probabilistik (Felicia Watratan et al., 2020). Metode ini memiliki banyak kelebihan seperti dapat digunakan untuk data kualitatif dan kuantitatif, perhitungannya yang cepat dan efisien, serta mudah dipahami. Dengan banyaknya kelebihan ini banyak penelitian sebelumnya yang memakai metode ini, seperti penelitian oleh Afriansyah et al. (2024). Kemudian penelitian oleh Yuyun et al. (2021). Dan pada penelitian Asmiati et al. (2020). Dari ketiga penelitian di atas metode *Naive Bayes Classifier* memiliki akurasi yang cukup tinggi di angka lebih dari 70%.

Dari dua penjelasan di atas, peneliti memutuskan untuk melakukan perbandingan antara dua metode klasifikasi yaitu, metode *K-Nearest Neighbors* dan metode *Naive Bayes Classifier*. Data yang dikumpulkan berjumlah 670 *data tweet* dalam kurun waktu 1 bulan. Perbandingan yang akan dilakukan

adalah perbandingan akurasi dalam mengklasifikasikan data dari sentimen masyarakat terhadap Pemilu Presiden Tahun 2024. Dengan membandingkan akurasi diantara kedua metode tersebut diharapkan penelitian ini dapat memberikan informasi yang lebih baik dan akurat dalam klasifikasi sentimen masyarakat terhadap Pemilu Presiden Tahun 2024.

1.2. Rumusan Masalah

Bagaimana akurasi metode *K-Nearest Neighbors* dan metode *Naive Bayes Classifier* terhadap pengklasifikasian sentimen masyarakat terhadap Pemilu Presiden Tahun 2024?

1.3. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk membandingkan akurasi dalam klasifikasi sentimen masyarakat antara metode *K-Nearest Neighbors* dan metode *Naive Bayes Classifier*.

1.4. Ruang Lingkup Permasalahan

Ruang lingkup yang akan di bahas dalam penelitian ini yaitu Mencari sentimen atau opini masyarakat terhadap Pilpres 2024 melalui perbandingan metode *K-Nearest Neighbors* dan *Naive Bayes Classifier*. Sumber datanya berupa *tweet* pengguna dengan topik Pilpres 2024 di aplikasi *Twitter* selama 1 bulan sebelum sidang keputusan mahkamah konstitusi. Data teks yang diolah hanyalah *teks tweet* pada aplikasi *Twitter* yang telah diterjemahkan dari bahasa Indonesia ke bahasa Inggris dengan menggunakan klasifikasi emosi positif, negatif, dan netral. Keakuratan metode *K-Nearest Neighbors* dan *Naive Bayes Classifier* pada *dataset* yang diperoleh merupakan temuan yang perlu dibandingkan.

1.5. Manfaat Penelitian

Perbandingan dua metode klasifikasi sentimen masyarakat dapat memberikan pemahaman yang lebih baik tentang kelebihan dan kekurangan masing-masing metode, serta membantu dalam memilih metode yang paling sesuai untuk tujuan analisis tertentu. Berikut adalah beberapa manfaat dari melakukan perbandingan tersebut:

1. Mengetahui hasil akurasi tertinggi dari perbandingan metode *K-Nearest Neighbors* dan *Naive Bayes Classifier*
2. Memperoleh pemahaman yang lebih mendalam tentang bagaimana setiap metode bekerja, algoritma yang digunakan, dan pendekatan yang diambil untuk mengklasifikasikan sentimen.
3. Sebagai referensi penelitian selanjutnya yang berkaitan dengan analisis sentimen.