

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pemilihan umum (Pemilu) di Indonesia diselenggarakan oleh Komisi Pemilihan Umum (KPU), sebuah lembaga independen yang mengatur dan mengawasi proses pemilihan umum. KPU dan Badan Pengawas Pemilihan Umum (Bawaslu) memiliki peran penting dalam memastikan Pemilu berjalan secara langsung, umum, bebas, rahasia, jujur, dan adil (LUBER), serta terjamin kebebasannya dari intervensi kekuasaan lain untuk menjaga masa depan demokrasi negara (Hananto Widodo & Dicky Eko Prasetyo, 2021). Untuk memberikan gambaran awal cepat hasil pemilihan sebelum pengumuman resmi, *Quick Count* (Hitung Cepat) digunakan untuk memperkirakan hasil suara kandidat presiden pada sampel suara dari sejumlah TPS. Metode ini tidak menghitung seluruh suara, melainkan sebagian kecil sampel yang diperkirakan mencerminkan hasil akhir nasional yang sering menjadi topik pembicaraan Masyarakat dan media (Nugraha et al., 2019). Hasil *Quick Count* sering dianggap perkiraan awal yang cukup akurat dan penting bagi kepercayaan publik terhadap proses pemilihan, meski bukan hasil resmi, namun pengumuman hasil yang tidak sesuai harapan dapat memicu ketegangan politik, protes, atau kerusuhan (Rhima Indria Saraswati, 2020).

Media sosial memiliki dampak signifikan dalam kehidupan masyarakat dengan kemudahan berkomunikasi jarak jauh dan bersosialisasi secara virtual. Namun, media sosial juga memiliki konsekuensi seperti penyebaran bahasa sarkasme atau ungkapan kasar oleh para netizen yang menggunakan media sosial untuk mengekspresikan identitas dan eksistensi diri mereka, bahkan dengan cara yang melanggar norma kesantunan berbahasa (Andi Saadillah et al., 2023). Pada pemilihan presiden sebelumnya, penggunaan media sosial berperan penting dalam kampanye politik, di mana para kandidat menyampaikan visi dan misi mereka melalui platform media sosial (Ulfa et al., 2020). Dengan berkembangnya platform digital yang mampu mewadahi opini masyarakat, pro maupun kontra tidak hanya bermunculan di layar kaca melalui

liputan berita, tetapi juga banyak bermunculan di media sosial X Twitter dan platform digital (Zhafira et al., 2021). Media sosial seperti X Twitter telah menjadi platform penting bagi masyarakat untuk mengekspresikan pandangan dan opini terkait isu politik termasuk pemilihan presiden (Wulandari, 2024).

Urgensi penelitian terletak pada pentingnya mengklasifikasikan teks terkait *Quick Count* Pemilihan Presiden Indonesia 2024 di media sosial X Twitter untuk mengidentifikasi respons masyarakat terhadap perkembangan politik yang terjadi. Menurut artikel dari Tirto.id oleh Irfan Amin, pengguna media sosial di Indonesia sudah mencapai 167 juta orang pada tahun 2023 berdasarkan catatan dari *We Are Social*, dengan sebagian besar pengguna berasal dari kalangan muda yang mendominasi lebih dari 50 persen pemilih pada Pemilu. Ketertarikan pengguna media sosial terhadap informasi politik cukup tinggi, dengan 46,07 persen pengguna menunjukkan ketertarikan yang signifikan terhadap informasi politik (Amin, 2023).

Proses ekstraksi informasi teks di media sosial adalah *text mining*, melibatkan pemantauan tren, pola, dan pembobotan kata untuk mendapatkan informasi berharga dari teks yang luas (Sosiawan & Wibowo, 2020). Klasifikasi adalah teknik dalam *text mining* yang bertujuan mengelompokkan objek-objek dengan karakteristik serupa ke dalam beberapa kelas untuk menganalisis pendapat, penilaian, sikap, dan perasaan orang (Rizki et al., 2023). *Text mining* memiliki definisi menambang data berupa teks, bertujuan mencari kata-kata yang dapat mewakili isi dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen dengan menggunakan algoritma *machine learning*. Salah satu teknik dalam *text mining* adalah klasifikasi, yang bertujuan mengelompokkan objek-objek dengan karakteristik serupa ke dalam beberapa kelas, seperti mengkaji pendapat dan sikap terhadap teks (Nurmalasari & Ribut Yuliantoro, 2022). Metode *Naive Bayes* dan TF-IDF telah terbukti efektif dalam klasifikasi teks studi terhadap kenaikan harga bahan pokok di X Twitter, metode ini mencapai akurasi tinggi setelah melalui *preprocessing* menyeluruh (Muslimin & Lusiana, 2023).

Penelitian sebelumnya telah menerapkan pembobotan kata TF-IDF dan klasifikasi dengan algoritma *Naive Bayes classifier* pada ulasan aplikasi Ruang Guru. Hasil penelitian menunjukkan bahwa prediksi sentimen lebih dominan pada sentimen positif dengan nilai presisi 71%, recall 69%, F1-score 69%, dan akurasi 69% (Novitasari et al., 2022). Penelitian lainnya menggunakan *Naive Bayes* dan ekstraksi fitur TF-IDF, dengan hasil sentimen positif 70,20%, negatif 29,80%, dan akurasi 65,39% (Salim & Solichin, 2022). Penelitian terdahulu telah menerapkan kombinasi pembobotan TF-IDF dan klasifikasi *Naive Bayes* pada berbagai bidang. Namun, belum ada penelitian yang secara khusus menerapkan pendekatan ini untuk mengklasifikasi teks *Quick Count* Pemilihan Presiden Indonesia 2024 di X Twitter.

Penelitian ini mengadopsi serangkaian pendekatan, dimulai dengan dengan pengumpulan data melalui pencarian Twitter dan web crawling menggunakan Python. Setelah data terkumpul, dilakukan labeling data oleh ahli dalam pelabelan data, tahap *preprocessing* teks yang mencakup koreksi kata tidak standar dengan menggunakan kamus yang disusun oleh peneliti, serta penerapan pembobotan kata dengan metode *Term Frequency Inverse Document Frequency* (TF-IDF). Data yang telah diolah kemudian dibagi menjadi data pelatihan dan data pengujian. Klasifikasi teks dilakukan dengan *Naive Bayes classifier*. Evaluasi akhir dilakukan dengan mengukur akurasi yang dihasilkan. Melalui pendekatan ini, diharapkan klasifikasi teks *Quick Count* Pemilihan Presiden 2024 di X Twitter mampu membantu para pihak terkait seperti lembaga survei dan pemangku kepentingan lainnya untuk mengidentifikasi respons masyarakat terhadap perkembangan politik yang terjadi. Dengan demikian, penelitian ini memberikan kontribusi akademis dalam bidang klasifikasi teks dan memiliki dampak praktis dalam mengidentifikasi respons masyarakat di media sosial.

1.2 Rumusan Masalah

Rumusan masalah penelitian ini adalah seberapa akurat metode *Naive Bayes* dengan pembobotan TF-IDF dalam mengklasifikasikan teks *Quick Count* Pemilihan Presiden Indonesia 2024 di media sosial X Twitter?

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah mengukur akurasi metode *Naive Bayes* dengan pembobotan TF-IDF dalam mengklasifikasikan teks *Quick Count* Pemilihan Presiden Indonesia 2024 di media sosial X Twitter.

1.4 Manfaat Penelitian

Manfaat Penelitian ini adalah:

1. Teoritis: Memberikan kontribusi dalam bidang text mining dan klasifikasi teks, khususnya dalam konteks politik di media sosial. Penelitian ini dapat memperkaya literatur tentang penggunaan metode TF-IDF dan Naive Bayes dalam klasifikasi teks.
2. Praktis: Membantu lembaga survei dan pemangku kepentingan lainnya mengidentifikasi respons masyarakat terkait teks *Quick Count* Pemilihan Presiden Indonesia 2024 di media sosial X Twitter, sehingga dapat merumuskan strategi komunikasi yang lebih efektif.
3. Metodologis: Menyediakan studi kasus yang dapat dijadikan referensi dalam penerapan metode TF-IDF dan Naive Bayes untuk klasifikasi teks di berbagai konteks lain, baik dalam ranah politik, pemasaran, maupun bidang lainnya yang membutuhkan analisis teks di media sosial.