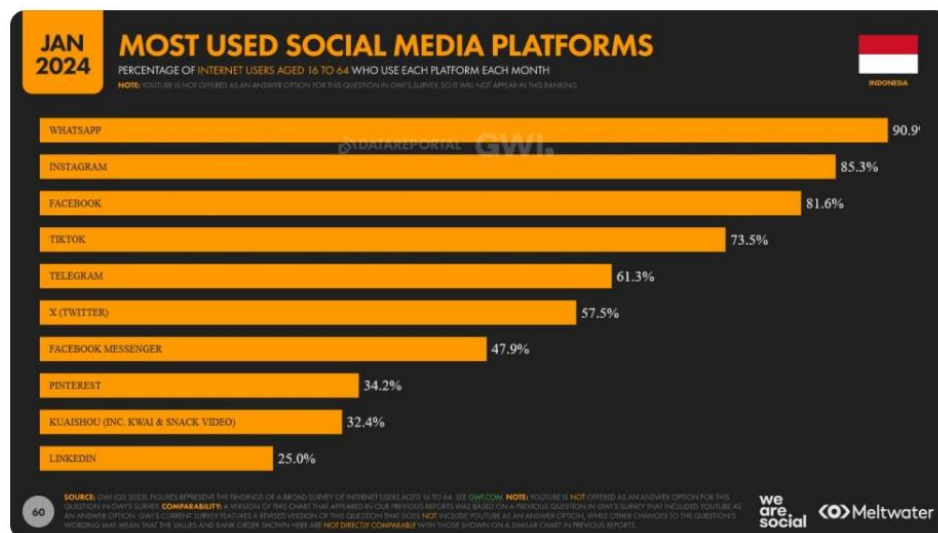


BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan media sosial telah mengubah cara masyarakat menyampaikan opini, pendapat, dan sikap mereka terhadap berbagai isu, menjadikan media sosial ini sebagai ruang diskusi publik yang semakin penting (Zempi et al., 2023). Media sosial tidak hanya berfungsi sebagai alat bantu komunikasi dengan interaksi antar sosial, tetapi juga menjadi ruang bagi masyarakat untuk terlibat dalam diskusi interaktif terkait isu-isu terkini. Di Indonesia, media sosial *X* sebelumnya dikenal sebagai *Twitter*, merupakan salah satu yang paling populer untuk berbagi informasi dan pendapat yang menduduki peringkat ke-6 pada tahun 2024 terlihat pada Gambar 1. 1 (Dwi Riyanto, 2024). Media sosial *X* mempunyai karakteristik yang cepat untuk menyebarkan informasi terkini dalam waktu singkat, salah satunya isu-isu politik khususnya dalam konteks kampanye Pemilihan Kepala Daerah (Cahya Herdiyani & Zailani, 2022).



Gambar 1. 1 *Most Used Social Media Media sosials in Indonesia (January 2024)*

(Sumber: (Dwi Riyanto, 2024))

Pemilihan Kepala Daerah (Pilkada) adalah peristiwa yang berdampak dalam dinamika politik yang menarik perhatian publik, terutama di era digital. Pilkada, termasuk di Jawa Timur, menjadi ajang persaingan kandidat sekaligus topik ramai di media sosial. Menurut KPU Jawa Timur, tiga pasangan calon Gubernur telah resmi mendaftar. Pada tanggal 23 September Tahun 2024, Komisi Pemilihan Umum (KPU) Jawa Timur telah menetapkan nomor urut calon gubernur. Kandidat nomor urut 1 adalah Luluk Nur Hamidah-Lukmanul Hakim, nomor 2 Khofifah Indar Parawansa-Emil Dardak, dan nomor 3 Tri Rismaharini-Zahrul Azhar Asumta. Setiap pasangan Gubernur ini tidak hanya mengandalkan kampanye konvensional melalui pertemuan langsung, tetapi juga memanfaatkan media sosial sebagai alat untuk memperkuat citra positif, membangun dukungan publik, dan menjangkau masyarakat yang lebih luas.

Antusiasme masyarakat Jawa Timur terhadap Pilkada 2024 terlihat melalui media sosial *X*, yang menjadi media sosial utama untuk menyuarakan pandangan, kekhawatiran, dan harapan terkait isu politik, termasuk Pilkada (Adib et al., 2024). Banyaknya sentimen publik yang disampaikan melalui media sosial menghasilkan data yang besar serta tidak terstruktur. Data ini mencakup berbagai sentimen positif, negatif, dan netral terhadap setiap kandidat gubernur beserta wakil gubernur. Sentimen yang bermunculan di media sosial *X* menjadi indikator penting dalam menggambarkan kepercayaan publik terhadap calon pemimpin dan dapat memengaruhi keputusan pemilih (Arifin et al., 2024). Oleh karena itu, diperlukan adanya teknologi analisis sentimen dalam mengklasifikasikan sentimen dengan teknik *text mining* untuk memahami dinamika opini yang berkembang, serta mengidentifikasi faktor-faktor yang memengaruhi keputusan pemilih dalam Pilkada (Afdal et al., 2022).

Menurut (Fahlevvi, 2022), Analisis sentimen adalah teknik untuk mengevaluasi pandangan atau perasaan seseorang terhadap suatu topik, biasanya melalui teks seperti ulasan produk, komentar media sosial, atau berita. Tujuannya untuk mengidentifikasi apakah pendapat tersebut positif, negatif, atau netral. Beberapa penelitian sebelumnya telah menguji berbagai metode dalam melakukan analisis sentimen di media sosial. Seperti penelitian yang dilakukan oleh (Ferdian Maulana Akbar et al., 2024) dalam memahami dampak isu kerangka motor ABC

terhadap reputasi merek Motor XYZ di *Twitter*, dengan mengumpulkan data dari media sosial tersebut dan menggunakan lima algoritma *machine learning*; algoritma *Naïve Bayes*, algoritma *Decision Tree*, algoritma *Support Vector Machine*, algoritma *Logistic Regression*, dan algoritma *Random Forest* dalam mengklasifikasikan sentimen publik. Hasil akhir penelitian mengindikasikan bahwa model *Random Forest* dengan penyesuaian *hyperparameter* memiliki performa terbaik, mencapai skor F1 sebesar 0,765.

Penelitian selanjutnya, seperti yang diterapkan oleh (Manullang et al., 2023) menganalisis sentimen publik terhadap calon presiden di media sosial *X* (sebelumnya *Twitter*) menggunakan metode *Lexicon Based* dan *Random Forest*. Pemilihan presiden adalah momen krusial dalam demokrasi, dan *Twitter* menjadi *media sosial* utama untuk mengungkap opini masyarakat mengenai calon yang bersaing. Dengan volume *tweet* yang besar, diperlukan metode analisis yang efisien. Hasil penelitian menunjukkan bahwa *Random Forest* mencapai akurasi 94%, sementara metode *Lexicon Based* kurang efektif dalam menganalisis teks kompleks. Hal ini membuktikan bahwa algoritma *Random Forest* tidak hanya lebih stabil, namun juga lebih mampu melakukan generalisasi.

Pada penelitian ini akan menganalisis sentimen publik berkaitan dengan Pemilihan Kepala Daerah (PILKADA) Provinsi Jawa Timur 2024 yang diungkapkan melalui media sosial *X*. Penelitian ini berfokus pada identifikasi persepsi publik yang berupa sentimen positif, netral, atau negatif terhadap masing-masing kandidat Gubernur beserta Wakil Gubernur pada Pilkada Provinsi Jawa Timur Tahun 2024. Dengan mengumpulkan ulasan dan opini masyarakat, penelitian ini bertujuan untuk memberikan gambaran yang lebih jelas mengenai pandangan publik terhadap setiap pasangan calon Gubernur dan Wakil Gubernur. Selain itu, penelitian ini juga akan mengevaluasi akurasi metode analisis yang digunakan, memastikan bahwa data yang diperoleh tidak hanya akurat tetapi juga mencerminkan secara tepat aspirasi dan opini masyarakat Jawa Timur terhadap masing-masing kandidat Gubernur beserta Wakil Gubernur pada Pilkada Provinsi Jawa Timur Tahun 2024.

Penelitian ini menggunakan metode *Random Forest* dalam mencapai tujuannya. Menurut (Pratama et al., 2021), *Random Forest* sering digunakan dalam analisis sentimen karena algoritma tersebut mampu menghasilkan prediksi yang akurat. Meskipun algoritma *Random Forest* telah banyak digunakan dalam berbagai penelitian analisis sentimen di media sosial, hingga saat ini, belum ada penelitian yang secara khusus menggunakan algoritma *Random Forest* untuk menganalisis sentimen publik terhadap kandidat Gubernur beserta Wakil Gubernur pada pelaksanaan Pilkada Provinsi Jawa Timur 2024. Selain itu, penggunaan algoritma *Random Forest* pada penelitian ini juga didorong oleh keunggulannya dalam menangani data berdimensi tinggi dan kompleks serta menghasilkan prediksi dengan akurasi tinggi (Wiraswendro & Soetanto, 2022).

Melalui penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai sentimen publik terhadap masing-masing kandidat Gubernur beserta Wakil Gubernur pada Pilkada Provinsi Jawa Timur Tahun 2024 di media sosial *X*. Dengan menerapkan metode *Random Forest*, analisis sentimen diharapkan dapat dilakukan secara lebih akurat, terutama dalam mengolah data besar dan tidak terstruktur. Hasil penelitian ini diharapkan dapat memberikan pemahaman yang lebih mendalam mengenai pandangan masyarakat terhadap setiap pasangan kandidat. Selain itu, hasilnya dapat menjadi bahan evaluasi yang bermanfaat dalam menyusun strategi kampanye pada masa mendatang, sehingga memungkinkan kandidat untuk memahami dan merespons aspirasi publik secara lebih tepat.

1.2 Rumusan Masalah

Berdasarkan dari penjelasan latar belakang yang telah dideskripsikan sebelumnya, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana pola sentimen analisis pengguna media sosial *X* terhadap Pilkada Provinsi Jawa Timur Tahun 2024 ?
2. Bagaimana performa model yang dikembangkan dalam mengklasifikasi sentimen terkait Pilkada Provinsi Jawa Timur Tahun 2024 di media sosial *X* ?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disusun, tujuan dari penelitian ini dapat adalah sebagai berikut:

1. Menganalisis sentimen publik terhadap Pilkada Provinsi Jawa Timur Tahun 2024 di media sosial *X* menggunakan algoritma *Random Forest*.
2. Mengevaluasi performa Menggunakan Confusion Matrix dan Teknik Balancing pada model *Random Forest* dalam mengklasifikasi sentimen terkait Pilkada Provinsi Jawa Timur Tahun 2024 di media sosial *X*

1.4 Batasan dan Asumsi Penelitian

Pada penelitian Analisis Sentimen Pemilihan Kepala Daerah (PILKADA) Provinsi Jawa Timur Tahun 2024 pada Media Sosial *X* menggunakan Algoritma *Random Forest* memiliki beberapa batasan yang ditetapkan agar fokus penelitian tetap terarah dan sesuai dengan tujuan. Batasan tersebut adalah sebagai berikut:

1. Penelitian yang akan dilakukan difokuskan pada data yang diambil dari unggahan di tanggal 01 Juli hingga tanggal 28 November Tahun 2024
2. Penelitian ini berfokus pada data berbahasa Indonesia yang diambil dari media sosial *X* dalam analisis sentimen publik terhadap Pemilihan Kepala Daerah (PILKADA) Provinsi Jawa Timur Tahun 2024.
3. Data yang diperoleh melalui proses *crawling* terbatas pada periode yang telah ditentukan, namun terdapat beberapa tanggal tertentu yang tidak tersedia dalam dataset. Ketidaktersediaan data ini disebabkan oleh pembatasan akses API atau kebijakan privasi media sosial media sosial *X*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi dalam berbagai aspek, baik praktis maupun akademis. Adapun manfaatnya dijelaskan sebagai berikut:

1. Bagi kandidat dan tim kampanye, hasil analisis sentimen ini dapat dijadikan acuan untuk mengevaluasi strategi komunikasi politik yang digunakan selama kampanye Pilkada. Dengan memahami bagaimana masyarakat merespons citra kandidat di media sosial *X*, tim kampanye dapat mengidentifikasi pendekatan yang efektif dan aspek yang perlu diperbaiki agar strategi komunikasi sesuai dengan aspirasi publik.

2. Bagi penulis, penelitian ini memperluas pemahaman penulis tentang respons masyarakat terhadap isu politik, kandidat, dan kampanye di ruang digital, sekaligus meningkatkan kemampuan dalam pengolahan data, penerapan algoritma *Random Forest*, dan analisis sentimen opini publik secara sistematis.
3. Bagi Telkom *University*, penelitian ini berkontribusi pada pengembangan kajian akademik di bidang analisis data, khususnya dalam konteks politik. Hasilnya dapat menjadi referensi bagi mahasiswa atau peneliti yang tertarik mengkaji opini publik di media sosial, serta mendorong penelitian lanjutan yang menggabungkan ilmu data dan ilmu sosial secara aplikatif.

1.6 Sistematika Penulisan

BAB 1 Pendahuluan

Pada Bab I ini berisi penjelasan mengenai latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan penelitian, sistematika penulisan, serta rencana jadwal kegiatan. Pembahasan difokuskan pada analisis sentimen pada Pemilihan Kepala Daerah (PILKADA) Provinsi Jawa Timur Tahun 2024.

BAB 2 Tinjauan Pustaka

Bab II menyajikan hasil kajian literatur yang mendukung penelitian ini, termasuk pembahasan teori, kerangka kerja, dan metode yang relevan dengan analisis sentimen pada Pemilihan Kepala Daerah Provinsi Jawa Timur Tahun 2024.

BAB 3 Metodologi Penelitian

Bab III ini menjelaskan proses penyelidikan terhadap masalah yang diangkat melalui tahapan penelitian yang dilakukan secara sistematis dan terstruktur untuk menganalisis sentimen pada Pemilihan Kepala Daerah Provinsi Jawa Timur Tahun 2024.

BAB 4 Analisis dan Perancangan

Bab IV ini memuat uraian mengenai implementasi metode yang digunakan untuk analisis sentimen opini masyarakat terhadap Pemilihan Kepala Daerah Provinsi Jawa Timur Tahun 2024 pada media sosial *X*.

BAB 5 Hasil Penelitian dan Pembahasan

Bab V ini menjelaskan hasil penelitian yang telah dilakukan, dengan analisis data terkait sentimen opini masyarakat terhadap Pemilihan Kepala Daerah Provinsi Jawa Timur Tahun 2024, yang diuraikan secara sistematis sesuai metode analisis yang digunakan.

BAB 6 Kesimpulan dan Saran

Bab VI ini berisi kesimpulan yang diperoleh dari hasil penelitian mengenai analisis sentimen opini masyarakat terhadap Pemilihan Kepala Daerah Provinsi Jawa Timur Tahun 2024, serta saran yang relevan dengan objek penelitian atau aspek terkait lainnya.

1.7 Rencana Jadwal Penelitian

Penelitian ini dirancang untuk diselesaikan dalam waktu 9 bulan melalui tahapan-tahapan yang tersusun secara sistematis. Setiap tahap direncanakan dengan cermat untuk memastikan pencapaian tujuan penelitian. Rencana jadwal kegiatan disajikan pada

Tabel 1. 1 yang menggambarkan kegiatan yang dilakukan pada setiap bulannya.

Tabel 1. 1 Rencana Jadwal Penelitian

No.	Kegiatan	Waktu Penelitian								
		1	2	3	4	5	6	7	8	9
1.	Pengumpulan Data									
2.	Preprocessing Data									
3.	Pelabelan Sentimen									
4.	Implementasi Metode									
5.	Visualisasi Sentimen									
6.	Penyusunan Buku TA									