

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pemilu merupakan proses memilih kandidat untuk jabatan politik strategis di lembaga eksekutif dan legislatif, baik tingkat pusat maupun daerah. Dalam studi politik Indonesia, istilah pemilu mencakup pemilihan presiden (pilpres) hingga pemilihan kepala daerah (pilkada) (Indrawan & IP, 2022). Pilkada sendiri didasari oleh landasan filosofis dan yuridis Pasal 18 Ayat 4 UUD 1945 yang mengamanatkan Gubernur, Bupati, dan Walikota dipilih secara demokratis.

Sutrisno (2017) menyatakan bahwa pilkada adalah mekanisme demokratis yang memungkinkan masyarakat memilih pemimpin daerah. Dalam hal ini masyarakat aktif berpartisipasi dalam hal suara untuk memilih siapa pemimpin baik tingkat pusat atau daerah selama periode yang ditentukan.

Perkembangan teknologi dan era informasi yang dinamis melalui internet, menurut Andi Saadillah dkk. (2023), warganet atau *netizen* yaitu orang yang aktif menggunakan internet banyak bercerita tentang opini, pengalaman, dan hal lain. Manusia berbicara dengan berbagai bahasa, bahasa mencerminkan perasaan positif, negatif, atau netral (Yusof & Aziz Fadzillah, 2022). Beredarnya informasi yang semakin beragam, *netizen* pun menjadi pencari informasi yang andal.

Fleksibilitas komunikasi dari media sosial seperti Facebook, X, Instagram, TikTok, dan LinkedIn memudahkan *netizen* mengakses informasi dengan cepat. Di antaranya, X menjadi platform yang paling dinamis, dengan *tweet* berupa pesan informasi mikro blog hingga 280 karakter yang terus mengalir (Stracqualursi & Agati, 2024). Pak & Paroubek (2010) menegaskan X sebagai sumber informasi utama untuk peristiwa terkini karena kontennya singkat, aktual, dan relevan.

Berita dan opini *netizen* saling memengaruhi hingga membentuk kesamaan pandangan yang mengubah sentimen pribadi menjadi sentimen publik terbentuk atas pertukaran informasi melalui media sosial X. Dengan begitu informasi berserta

sentimen yang tersebar terhadap pilkada dibutuhkan pendekatan analisis yang dihasilkan oleh pengguna media sosial untuk mengolah data teks tersebut.

*Data science* memainkan peran penting dalam penelitian *web science research*, memberikan *insight* dalam pengembangan komunitas, perilaku orang secara daring, dan penyebaran informasi melalui *big data analysis* (Phethean dkk., 2016). Pendapat Ullman (2021) mengemukakan bahwa *data science* lebih dari sekadar *machine learning* atau statistik, melainkan evolusi alami dari teknologi pengelolaan data berskala besar untuk memecahkan masalah pada bidang ilmiah. Penelitian analisis sentimen dalam *data science* terus berkembang dengan cepat.

Dalam analisis data, *Machine Learning-based sentiment analysis* menjadi pendekatan populer dan efektif (Prastyo dkk., 2020), dengan klasifikasi sentimen positif dan negatif. Teknik ini memanfaatkan *Natural Language Processing (NLP)* untuk memahami bahasa manusia, termasuk tugas seperti peringkasan, penerjemahan, dan analisis (Agarwal, 2019). Salah satu metode yang diimplementasikan adalah *Random Forest*, algoritma konsisten dan adaptif untuk klasifikasi maupun regresi (Scornet dkk., 2015).

Dalam literatur, pemilu banyak dianalisis menggunakan analisis sentimen. Chaudhry dkk. (2021) menyebutkan Pemilu Amerika 2020 menuai kontroversi karena tuduhan kecurangan, sementara analisis sentimen mampu mengungkap opini *netizen* dari berbagai partai dan kandidat. Manullang dkk. (2023) menegaskan pada Pilpres Indonesia 2024, analisis sentimen berperan dalam memprediksi opini publik dan reaksi masyarakat terhadap calon presiden, sehingga membantu pihak berwenang mengambil keputusan lebih tepat. Peristiwa ini menjadi sorotan karena memengaruhi kebijakan dan informasi penting bagi masyarakat.

Implementasi sistem informasi dalam pemilihan pemimpin daerah penting untuk menyediakan informasi aktual dan berimbang, sehingga *netizen* dapat mengungkap opini, isu, dan regulasi berbasis data. Balázs dkk. (2021) menegaskan kualitas data opini dan berita krusial bagi kredibilitas dan keberlanjutan, sehingga diperlukan pendekatan terbaik guna menjamin validitas dan keandalan, khususnya terkait paslon ideal dalam pilkada.

Melihat informasi yang ada, judul yang diajukan yaitu “Penerapan Metode Random Forest Pada Analisis Sentimen Pemilihan Kepala Daerah Jawa Barat” yang dapat menjadi referensi untuk memahami pandangan dan sentimen publik terkait dengan Pilkada Jabar 2024 serta isu-isu yang muncul. Penelitian ini bertujuan menyajikan pandangan mendalam tentang proses demokrasi serta pengaruh analisis sentimen publik terhadap hasil pilkada. Hasilnya diharapkan mendukung pemilihan yang lebih informatif, membantu masyarakat mengambil keputusan politik yang bijak, serta menjadi acuan penelitian lanjutan di bidang *data science*.

### **1.2 Identifikasi Masalah**

Beberapa poin masalah dapat diidentifikasi berdasarkan latar belakang masalah yaitu;

1. Masyarakat pengguna internet di Jawa Barat membutuhkan sistem berbasis *data science* yang mampu memberikan informasi calon pemimpin daerahnya sehingga pemilih bisa mendapatkan gambaran secara nyata dari sentimen-sentimen yang ada di media sosial.
2. Belum terdapat evaluasi yang mendalam mengenai efektivitas dan kemudahan penggunaan sistem analisis sentimen yang dibangun oleh masyarakat pengguna internet, khususnya dalam konteks pemilihan kepala daerah di Jawa Barat.

### **1.3 Rumusan Masalah**

Permasalahan yang diidentifikasi dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimana membangun sistem untuk mengklasifikasi sentimen positif, negatif dan netral di media sosial X untuk masyarakat pengguna internet yang akan memilih kepala daerah Jawa Barat?
2. Bagaimana mengimplementasikan metode *Random Forest* pada sistem yang dibangun?
3. Bagaimana menguji sistem yang dibuat sehingga dapat digunakan oleh masyarakat pengguna internet?

#### 1.4 Tujuan Penelitian

Penelitian ini tidak hanya menentukan identifikasi dan rumusan masalah, tetapi juga menentukan tujuan penelitian sehingga hasilnya sesuai dengan rencana. Berikut adalah beberapa tujuannya:

1. Untuk membangun sistem yang dapat mengklasifikasi sentimen positif, negatif dan netral di media sosial X untuk masyarakat pengguna internet yang akan memilih kepala daerah Jawa Barat.
2. Untuk mengimplementasikan metode *Random Forest* pada sistem yang dibangun.
3. Untuk menguji dan memvalidasi sistem yang dibuat sehingga bisa dipakai oleh masyarakat pengguna internet.

#### 1.5 Manfaat Penelitian

Kesimpulan dari penelitian ini diharapkan dapat memberikan manfaat, yaitu:

1. Hasil penelitian ini diharapkan berkontribusi untuk masyarakat dalam menentukan pilihan pemimpinnya dalam pemilihan umum baik daerah maupun nasional.
2. Hasil penelitian ini diharapkan berkontribusi kegiatan pemilihan umum dan dapat diperluas kegunaannya pada tingkat nasional ataupun internasional.
3. Hasil penelitian ini diharapkan dapat berkontribusi dalam pengembangan dan implementasi *data science* untuk kepentingan masyarakat.

#### 1.6 Ruang Lingkup

Untuk memastikan penelitian ini memiliki batasan yang jelas dan terfokus, inilah ruang lingkup yang akan dibahas dalam penelitian ini:

1. Data yang digunakan dalam *dataset* ini diperoleh melalui proses *crawl* dari media sosial X dengan kata kunci *tweet*: “pilgub jabar”, “cagub jabar” dalam rentang waktu 27-08-2024 sampai dengan 27-11-2024.
2. *Tweet* dalam bahasa daerah, campuran bahasa, singkatan, atau penggunaan *emoticon* mungkin tidak terklasifikasi secara akurat.
3. Pelabelan *dataset* terdiri dari sentimen positif, negatif dan netral.

4. Pelabelan sentimen otomatis menggunakan *library TextBlob* dengan menerjemahkan *tweet* bahasa Indonesia ke bahasa Inggris, sehingga berpotensi menimbulkan pergeseran makna asli.
5. Analisis tidak secara khusus mendeteksi sarkasme, satir, atau bentuk ironi lain yang dapat menyebabkan perbedaan antara makna literal dan makna sebenarnya dari suatu *tweet*/opini yang dijadikan *dataset*.
6. Opini yang diperoleh dari media sosial X tidak selalu merepresentasikan opini masyarakat luas, karena terdapat potensi bias demografis, keberadaan akun anonim, *bot account* / *buzzer*, maupun propaganda politik.
7. Hasil analisis sentimen hanya merepresentasikan opini publik di media sosial X, bukan cerminan langsung dari dukungan atau perolehan suara masyarakat nyata.

### **1.7 Sistematika Penulisan**

Sistematika penulisan yang digunakan untuk memberikan tinjauan umum mengenai subjek penelitian ini sebagai berikut:

#### **BAB I Pendahuluan**

Bab bagian awal berisi pendahuluan suatu rangkuman permasalahan yang akan dibahas. Pendahuluan terdiri dari subbab yang melingkupi latar belakang, identifikasi masalah, dan rumusan dari masalah penelitian, selanjutnya subbab tujuan, manfaat penelitian, dan yang subbab akhir sistematika penulisan.

#### **BAB II Studi Pustaka**

Pembahasan mengenai ulasan garis besar dari studi-studi relevan terkait dengan penelitian ini. Tinjauan literatur memberikan pemahaman tentang berbagai teori serta konsep dasar terkait dengan referensi analisis sentimen berdasarkan teknik dan tahapan seperti pengambilan data, *preprocessing* data teks, ekstraksi fitur, metode klasifikasi, uji model, evaluasi model dan analisis model.

#### **BAB III Metode Penelitian**

Menyajikan penjelasan menyeluruh perihal metode dan proses penelitian dan mendapatkan bukti yang cukup untuk mendukung tujuan penelitian ini, serta tahapan yang dilakukan dalam saat penelitian.

#### BAB IV Hasil dan Pembahasan

Uraian dalam bab ini menggambarkan hasil disertai pembahasan penelitian yang sudah dilakukan sesuai tahapan penelitian yang dijelaskan pada bab sebelumnya. Serta memberikan analisis data, mempresentasikan visualisasi data berdasarkan informasi yang telah dipaparkan sebelumnya dan memuat interpretasi hasil yang diperoleh dan kesimpulan yang diambil.

#### BAB V Kesimpulan dan Saran

Bab ini memaparkan kesimpulan dari hasil penelitian, menguraikan relevansi temuan dengan tujuan studi, serta menyertakan rekomendasi untuk penelitian di masa mendatang.

