

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan komunikasi telah memberi fasilitas pertukaran informasi secara masif melalui media sosial, seperti platform X. Hal ini menjadikan isu terkait Ibu Kota Nusantara sebagai topik yang ramai dibicarakan. Pemindahan Ibu Kota Negara ke IKN, berdasarkan Undang-Undang Nomor 3 Tahun 2022, menjadi isu strategis nasional yang menghadirkan berbagai tantangan, seperti konflik agraria dan dampak sosial terhadap masyarakat adat serta lokal. Tingginya perdebatan publik di media sosial menegaskan pentingnya analisis sentimen untuk memahami persepsi masyarakat, khususnya pandangan pengguna media sosial X terhadap IKN [1]. Analisis sentimen adalah upaya untuk mengetahui bagaimana perasaan atau emosi seseorang atau sekelompok orang terhadap suatu hal. Ini dapat berupa hal-hal seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, atau atribut yang terkait [2].

Penelitian sebelumnya telah mengeksplorasi berbagai metode untuk menganalisis sentimen, termasuk penggunaan algoritma seperti *Support Vector Machine* [2]-[4], *K-Nearest Neighbors* [5]-[7], *Naïve Bayes* [5], *Artificial Neural Networks* [8], *Decision Tree* [9], [10], *Random Forest* [9], [11], dan *Adaptive Boosting* [12]. *Support Vector Machine* adalah algoritma *supervised learning* yang bekerja dengan menentukan hyperplane. Hyperplane yang ideal terletak di tengah antara kedua kelas, sehingga memaksimalkan jarak ke data-data terluar dari masing-masing kelas [6]. *K-Nearest Neighbors* adalah satu metode pembelajaran yang proses klasifikasinya didasarkan pada kualitas dataset pembelajaran. KNN sering digunakan karena memiliki rumus yang mudah dan tahan terhadap data pembelajaran yang penuh dengan noise [7]. *Artificial Neural Networks* adalah jaringan saraf tiruan yang mirip dengan otak manusia, terdiri dari unit proses yang disebut neuron. ANN memiliki struktur dan perilaku yang mirip dengan saraf alami,

dengan satu keluaran (sinapsis melalui akson) dan beberapa masukan (dendrit) [8]. *Decision Tree* adalah flowchart berbentuk pohon dengan setiap node (node internal) memiliki atribut, dan cabang-cabangnya menunjukkan hasil tes atau nilai atribut. Pengenalan pola statistik juga biasanya dilakukan dengan *Decision Tree* [10]. *Random Forest* adalah algoritma yang digunakan untuk klasifikasi data dalam jumlah besar. Proses klasifikasi dalam Random Forest menggabungkan beberapa pohon (tree) dan melakukan pelatihan pada sampel data yang tersedia. Semakin banyak pohon yang digunakan, semakin baik akurasi [11]. *Adaptive Boosting* adalah teknik boosting yang menggabungkan berbagai pendekatan pembelajaran mesin. Algoritma ini menilai setiap kasus secara berulang, dengan kasus yang salah diklasifikasikan menerima nilai yang lebih tinggi. Mengurangi kesalahan dan variasi dalam *supervised learning* adalah tujuan algoritma ini. Selain itu, algoritma ini memiliki kemampuan untuk menangani data yang kompleks dan interaksi antar fitur [12]. Studi-studi ini telah membuktikan keefektifan metode-metode tersebut dalam hal analisis sentimen umum, meskipun belum secara khusus pada sentimen terkait Ibu Kota Nusantara.

Salah satu masalah adalah kurangnya perbandingan langsung antara berbagai teknik klasifikasi untuk menganalisis sentimen, terutama yang berkaitan dengan Ibu Kota Nusantara di media sosial X. Selain itu, sangat penting untuk memilih algoritma klasifikasi yang tepat untuk menganalisis data tanggapan masyarakat Indonesia di media sosial X. Masing-masing algoritma, seperti *Support Vector Machine*, *K-Nearest Neighbors*, *Naïve Bayes*, *Artificial Neural Networks*, *Decision Tree*, *Random Forest*, dan *Adaptive Boosting*, memiliki kelebihan dan kekurangannya sendiri dalam menangani data teks. Oleh karena itu, penelitian perlu dilakukan untuk membandingkan akurasi, efektivitas, dan skalabilitas dari berbagai algoritma dalam menangani data besar dan dinamika perubahan sentimen yang cepat di media [13].

Penelitian membandingkan beberapa algoritma, yaitu *Support Vector Machine*, *K-Nearest Neighbors*, *Naïve Bayes*, *Artificial Neural Networks*, *Decision Tree*, *Random Forest*, dan *Adaptive Boosting*, dalam analisis sentimen terkait Ibu

Kota Nusantara di platform X. Fokus utama adalah menguji dan membandingkan akurasi prediksi dari ketujuh metode tersebut, serta mengevaluasi kemampuan masing-masing algoritma dalam menghadapi tantangan seperti variasi atau perbedaan penggunaan bahasa informal di media sosial X. Hasil dari penelitian ini diharapkan dapat memberikan panduan yang lebih baik dalam memilih algoritma yang paling efektif dan efisien untuk analisis sentimen terkait Ibu Kota Nusantara.

1.2 Rumusan Masalah

Melihat latar belakang masalah yang telah disampaikan, rumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana perbandingan kinerja dan akurasi algoritma *Support Vector Machine*, *K-Nearest Neighbors*, *Naïve Bayes*, *Artificial Neural Networks*, *Decision Tree*, *Random Forest*, dan *Adaptive Boosting* dalam menganalisis sentimen masyarakat terhadap Ibu Kota Nusantara di sosial media X ?

1.3 Batasan Masalah

1. Sumber data diambil dari media sosial X dengan memanfaatkan API.
2. Pengambilan data hanya mencakup kata kunci "Ibu Kota Nusantara" dan "IKN".
3. Bahasa yang digunakan dalam data adalah bahasa Indonesia.
4. Algoritma yang digunakan adalah *Support Vector Machine*, *K-Nearest Neighbors*, *Naïve Bayes*, *Artificial Neural Networks*, *Decision Tree*, *Random Forest*, dan *Adaptive Boosting*.
5. Memiliki 3 label sentimen, positif, negatif dan netral.
6. Menggunakan Bahasa pemrograman *python*.

1.4 Tujuan Penelitian

1. Mengetahui nilai perbandingan langsung terkait tingkat akurasi antara algoritma *Support Vector Machine*, *K-Nearest Neighbors*, *Naïve Bayes*, *Artificial Neural Networks*, *Decision Tree*, *Random Forest*, dan *Adaptive Boosting* dalam menangani klasifikasi sentimen Ibu Kota Nusantara.

1.5 Manfaat Penelitian

1. Memberikan wawasan tentang pandangan pengguna media sosial X terhadap Ibu Kota Nusantara.
2. Menambah pengetahuan atau data terkait algoritma *Support Vector Machine*, *K-Nearest Neighbors*, *Naïve Bayes*, *Artificial Neural Networks*, *Decision Tree*, *Random Forest*, dan *Adaptive Boosting*.
3. Sebagai pedoman atau sumber informasi bagi pembaca dan peneliti berikutnya dalam menerapkan algoritma yang lebih efisien.

1.6 Sistematika Penulisan

Berisi sistematika penulisan skripsi yang dilakukan secara urut pada penelitian ini:

BAB I PENDAHULUAN

Pendahuluan mencakup latar belakang permasalahan, rumusan masalah, batasan masalah, tujuan penelitian, dan sistematika penelitian.

BAB II TINJAUAN PUSTAKA

Landasan teori mencakup penjelasan dari penelitian sebelumnya mengenai analisis sentimen serta penerapan algoritma *Support Vector Machine*, *K-Nearest Neighbors*, *Naïve Bayes*, *Artificial Neural Networks*, *Decision Tree*, *Random Forest*, dan *Adaptive Boosting*.

BAB III METODE PENELITIAN

Membahas mengenai urutan langkah-langkah yang dilakukan dalam penelitian ini, serta alur yang diikuti untuk mencapai tujuan penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memuat temuan penelitian yang diperoleh dari kegiatan yang telah dilaksanakan, serta analisis mengenai penilaian terhadap hasil yang diperoleh dari proses penelitian.

BAB V PENUTUP

Bagian ini mencakup kesimpulan dan saran atau rekomendasi terkait penelitian yang telah dilaksanakan.

