

1. Pendahuluan

Latar Belakang

Analisis Sentimen adalah teknik dalam pemrosesan bahasa alami yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini publik yang diungkapkan dalam teks sebagai nilai positif, negatif, atau netral [9]. Analisis sentimen berfokus pada ekstraksi, kategorisasi, pemahaman, dan penilaian opini dalam teks yang dapat membantu dalam pengambilan keputusan di berbagai sektor seperti bisnis, keuangan, politik, pendidikan, dan pelayanan [4] [5].

Pembangunan Ibu Kota Nusantara (IKN) di Kalimantan Timur adalah kebijakan strategis untuk mengurangi ketimpangan pembangunan dan beban Jakarta [1]. IKN berstatus pemerintahan daerah khusus setingkat provinsi sesuai UU No.3 Tahun 2022 [2]. Pembangunan IKN menghadapi tantangan pertanahan, terutama hak masyarakat hukum adat, kepemilikan tanah, dan mata pencaharian [3]. Hal ini memunculkan berbagai reaksi dari masyarakat terhadap pembangunan Ibu Kota Nusantara (IKN), karena pemindahan Ibu Kota Negara yang baru merupakan isu sensitif sehingga banyak diperbincangkan di media sosial, terutama Twitter (X).

Twitter adalah salah satu platform media sosial yang dapat berbagi cuitan dan berinteraksi sesama pengguna. Platform ini digunakan untuk berbagai tujuan, seperti penyebaran berita, diskusi politik, hiburan, serta analisis sentimen publik. Topik yang sedang hangat atau sering diperbincangkan oleh masyarakat, seperti pemindahan Ibu Kota di Jakarta ke Kalimantan Timur. Dengan berbagai macam opini di Twitter, dapat menunjukkan bagaimana tanggapan masyarakat terhadap pembangunan Ibu Kota Nusantara (IKN).

Penelitian ini akan berfokus pada sentimen publik twitter yang membahas pembangunan IKN. Data diperoleh menggunakan *Scraper* Twitter. Data yang diperoleh dari beberapa tweet akan menyimpang karena bertepatan dengan masa periode pemilihan calon presiden dan wakil presiden 2024.

Untuk melakukan analisis sentimen pembangunan Ibu Kota Nusantara (IKN), diperlukan algoritma yang dapat melakukan klasifikasi data dengan menggunakan perbandingan algoritma *Logistic Regression* dan *Naive Bayes*. Penelitian ini diharapkan memberi kontribusi untuk membantu menganalisa opini masyarakat terkait pembangunan Ibu Kota Nusantara (IKN) dengan cara menemukan algoritma *Machine Learning* yang tepat untuk melakukan analisis sentimen terkait Pembangunan IKN.

Topik dan Batasannya

Berdasarkan latar belakang sebelumnya, didapatkan topik dan batasan dari penelitian ini sebagai berikut:

1. Bagaimana mengklasifikasi analisis sentimen terkait Pembangunan IKN?
2. Bagaimana performansi dalam mengklasifikasi analisis sentimen terkait Pembangunan IKN dengan menggunakan *Logistic Regression* dan *Naive Bayes*?

Adapun beberapa batasan masalah terhadap sistem yang dibangun dalam penelitian ini adalah sebagai berikut.

1. Dataset diambil dari platform Twitter.
2. Data yang diambil menggunakan kata kunci “ikn”, “IKN”, dan “Ibu Kota Nusantara”
3. Pengambilan data dari bulan Januari – Maret 2024 yang bertepatan dengan masa periode pemilihan calon presiden dan wakil presiden 2024.
4. Pelabelan dataset dengan kategori nilai “positif” atau “negatif”
5. Metode yang digunakan untuk *feature extraction* adalah TF-IDF.
6. Metode yang digunakan untuk klasifikasi yaitu *Logistic Regression* dan *Naive Bayes*.

Tujuan

Pada penelitian ini bertujuan untuk menjelaskan proses dalam melakukan *preprocessing* untuk data twitter pada analisis sentimen terkait Pembangunan IKN, menjelaskan pengimplementasian dan menganalisis performansi metode *Logistic Regression* dan *Naive Bayes* dalam pengklasifikasian analisis sentimen terkait Pembangunan IKN.

Organisasi Tulisan

Pada penelitian ini memiliki urutan sebagai berikut: Bagian 1 menjelaskan latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, serta metode penelitian. Bagian 2 menjelaskan studi literatur dari landasan teori yang digunakan. Bagian 3 menjelaskan sistem yang dibangun yaitu sistem klasifikasi analisis sentimen terkait Pembangunan IKN dengan menggunakan metode *Logistic Regression* dan *Naive Bayes*. Bagian 4 menjelaskan hasil yang didapat melalui algoritma dan sistem yang sudah dibangun. Bagian 5 menjelaskan hasil dari kesimpulan yang didapat dan saran yang bisa dikembangkan untuk penelitian selanjutnya.

