

BAB I

PENDAHULUAN

1.1 Latar Belakang

Klasifikasi sentimen merupakan salah satu teknik dalam analisis teks atau pemrosesan bahasa alami yang bertujuan untuk mengekstrak informasi mengenai sikap atau pandangan seseorang terhadap suatu isu atau kejadian. Teknik ini berperan penting dalam mengelompokkan polaritas atau sentimen dari teks menjadi beberapa kategori, seperti positif atau negative [1]. Dalam era digital saat ini, klasifikasi sentimen menjadi semakin relevan karena banyaknya data teks yang tersedia di berbagai platform, terutama media sosial. Media sosial merupakan salah satu platform yang banyak digunakan untuk berkomunikasi dan mengakses informasi dengan cepat. Informasi yang ada di media sosial sangat bervariasi mulai dari komentar kritik hingga opini [2].

Twitter, dengan jutaan pengguna aktif di Indonesia, merupakan salah satu platform utama di mana opini publik sering kali diekspresikan. Ada 18,45 juta pengguna Twitter di Indonesia pada tahun 2022, menurut We Are Social [3]. Pengguna Twitter secara rutin membagikan pandangan, kritik, dan dukungan mereka terhadap berbagai isu, termasuk topik politik. Analisis sentimen terhadap tweet-tweet ini dapat memberikan wawasan mendalam tentang bagaimana masyarakat merespon suatu isu atau tokoh tertentu [4].

Untuk mengklasifikasikan sentimen dalam teks yang kompleks dan panjang, metode *Long Short-Term Memory (LSTM)* menjadi pilihan yang unggul. *LSTM* adalah jenis jaringan saraf rekurensial (*RNN*) yang dirancang untuk mengatasi masalah vanishing gradient dan mampu mempertahankan informasi jangka panjang [5]. Hal ini membuat *LSTM* sangat cocok untuk tugas klasifikasi sentimen, terutama pada data yang berasal dari media sosial seperti Twitter, di mana teks sering kali bersifat informal dan memiliki struktur yang bervariasi.

Beberapa penelitian sebelumnya telah menunjukkan efektivitas *LSTM* dalam analisis sentimen di media sosial. Contohnya, studi oleh Miftahul Ikhsan dkk pada

tahun 2023 menemukan bahwa *LSTM* memberikan akurasi yang lebih tinggi dibandingkan dengan metode lain seperti *Naïve Bayes* dan *SVM* dalam klasifikasi sentimen terkait vaksin COVID-19 di Twitter. Parameter yang paling optimal diperoleh leaky-relu 104 dengan akurasi 64% pada validasi data dan 66% pada data uji [6]. Penelitian oleh Jovita Nurvania dkk pada tahun 2021 menggunakan *LSTM* untuk mengklasifikasikan ulasan pengguna tentang dampak COVID-19 terhadap pariwisata Bali, yang juga menunjukkan hasil yang memuaskan, Model yang dibangun memiliki akurasi 71,67% [7].

Dalam konteks Pemilihan Presiden Indonesia 2024, analisis sentimen menggunakan *LSTM* dapat memberikan pemahaman yang lebih mendalam tentang pandangan masyarakat terhadap para calon presiden. Dengan menganalisis data dalam jumlah besar yang diperoleh dari media sosial, khususnya Twitter, dapat mengidentifikasi tren, pola, dan perubahan sentimen yang mungkin terjadi seiring berjalannya waktu [3]. Penelitian ini diharapkan mampu menghasilkan gambaran yang lebih akurat, objektif, dan representatif mengenai opini publik, yang nantinya dapat menjadi dasar yang kuat untuk mendukung proses pengambilan keputusan oleh masyarakat, Tim kampanye, dan pemangku kepentingan lainnya.

Penelitian ini tidak hanya bertujuan untuk memahami opini masyarakat terhadap calon presiden, tetapi juga menekankan pentingnya klasifikasi sentimen sebagai alat analisis yang efektif dan efisien dalam memahami pandangan publik di media sosial yang cepat, dinamis, dan terus berubah. Dengan menerapkan *LSTM*, penelitian ini berupaya meningkatkan akurasi prediksi sentimen, sehingga hasil yang diperoleh dapat menjadi referensi yang kredibel dan dapat diandalkan dalam berbagai konteks, termasuk strategi komunikasi politik, penyusunan kebijakan, dan perencanaan kampanye yang lebih tepat sasaran.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, maka dapat dipaparkan rumusan masalah sebagai berikut:

1.2.1 Permasalahan

Bagaimana mengklasifikasikan opini publik di Twitter mengenai calon Presiden Indonesia 2024 menggunakan *LSTM (Long Short-Term Memory)*.

1.2.2 Metode Usulan

Dalam studi ini mengusulkan klasifikasi sentimen terhadap opini publik terkait calon Presiden di Twitter dengan menggunakan *LSTM (Long Short-Term Memory)*.

1.2.3 Pertanyaan Penelitian

Di bawah ini menjelaskan pertanyaan penelitian yang menjadi fokus utama dalam penelitian ini.

1. Berapa tingkat akurasi model *LSTM* dalam mengklasifikasikan opini publik di Twitter mengenai tiga calon presiden?
2. Berapa nilai hyperparameter yang memberikan hasil terbaik ketika diuji dengan metode *LSTM* melalui hyperparameter tuning?

1.3 Tujuan Dan Manfaat

Terdapat sebuah tujuan dan manfaat terhadap penelitian tentang klasifikasi sentimen pengguna twitter terhadap calon presiden menggunakan *LSTM*.

1.3.1 Tujuan Penelitian

Tujuan penelitian ini dirumuskan sebagai berikut:

1. Untuk Menentukan tingkat akurasi model *LSTM* dalam mengklasifikasikan opini publik di Twitter mengenai tiga calon Presiden.
2. Untuk menentukan nilai hyperparameter yang optimal dalam metode *LSTM (Long Short-Term Memory)* melalui proses tuning hyperparameter, guna memperoleh hasil terbaik saat diuji.

1.3.2 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Membantu dalam memahami seberapa baik model *LSTM* dapat mengklasifikasikan opini publik di Twitter, memberikan wawasan yang lebih akurat tentang pandangan masyarakat terhadap tiga calon Presiden.
2. Mengidentifikasi nilai hyperparameter yang optimal untuk metode *LSTM* melalui proses tuning hyperparameter, sehingga dapat meningkatkan akurasi dan kinerja model dalam berbagai aplikasi praktis.

1.4 Batasan Masalah

Untuk memudahkan penelitian dan untuk mengatasi masalah yang diangkat oleh penulis, maka dibuat beberapa batasan masalah sebagai berikut:

1. Pengambilan data diperoleh dari twitter dengan *keyword* “Anies”, “Prabowo” dan “Ganjar”.
2. Data diambil pada tanggal 16 Agustus 2023 sampai tanggal 31 Desember 2023
3. Data twitter yang diambil hanya ulasan berbahasa Indonesia.
4. Data sentimen yang digunakan yaitu hasil ulasan sentiment positif dan negatif.
5. Data ulasan didapatkan dengan menggunakan Teknik *Scraping* menggunakan *Instant Data Scraper*.
6. Metode yang digunakan ketika tahap klasifikasi sentimen menggunakan *Long Short-Term Memory (LSTM)*, beserta menerapkan *Hyperparameter Tuning* untuk mencari nilai parameter terbaik.

1.5 Sistematika Penulisan

Penulisan skripsi disusun secara sistematis dan dibagi menjadi beberapa BAB, seperti berikut:

BAB I PENDAHULUAN

Bab 1 pendahuluan meliputi latar belakang penelitian, rumusan masalah penelitian termasuk permasalahan, metode usulan, pertanyaan penelitian, tujuan dan manfaat penelitian, batasan masalah, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab II Menjelaskan tentang tinjauan pustaka dengan kata lain landasan teori yang meliputi *Text Preprocessing*, *Klasifikasi Sentimen*, *Fitur dan Pembobotan*, *Confusion Matrix*, *Long Short Term Memory (LSTM)*, dan *Penelitian Terkait*. Diharapkan Semua literatur ini dapat memudahkan pemahaman terhadap dasar-dasar teori yang diterapkan dalam penelitian.

BAB III METODE PENELITIAN

Bab III metode penelitian mencakup berbagai aspek seperti peta penelitian, pengumpulan data, arsitektur model dan metode pengujian.

BAB IV HASIL DAN PEMBAHASAN

Bab IV menjelaskan tentang implementasi sistem yang dikerjakan sesuai dengan rancangan yang dibuat pada bab selanjutnya. Dimana terdapat Lingkungan uji coba, persiapan data termasuk pengumpulan data scraping, pelabelan data, text preprocessing. Pembobotan kata *TF-IDF*, klasifikasi menggunakan *LSTM* hasil evaluasi dan rasio jumlah sentiment positif.

BAB V PENUTUP

Bab V kesimpulan dari hasil penelitian yang meliputi hasil akurasi dari metode LSTM dalam mengklasifikasikan opini publik di Twitter mengenai tiga calon presiden dan nilai hyperparameter terbaik ketika diuji menggunakan metode LSTM.