

BAB I PENDAHULUAN

1.1 Latar Belakang

Sejak masuknya era digital, telah terjadi peningkatan signifikan dalam jumlah pengguna internet di seluruh dunia. Peningkatan ini sejalan dengan pertumbuhan jumlah data yang tersedia di internet. Ketersediaan data yang melimpah, terutama data teks, telah mendorong banyak penelitian dalam bidang *text mining* dan *Natural Language Processing (NLP)*. Salah satu bidang penelitian *NLP* yang populer adalah analisis sentimen. Analisis sentimen adalah proses komputasi untuk mengklasifikasikan data teks sesuai dengan sentimen yang terkandung di dalamnya. Popularitas penelitian analisis sentimen juga didorong oleh kebutuhan untuk memahami opini publik tentang topik tertentu. Proses analisis sentimen dipengaruhi oleh kualitas dataset yang digunakan. Untuk dataset yang terdiri dari kalimat-kalimat panjang, penanganannya memerlukan pendekatan yang berbeda dan lebih kompleks [1]. Aplikasi seluler telah menjadi tulang punggung bagi berbagai layanan dan aktivitas dalam era digital. Salah satu contohnya adalah aplikasi *Grab*, yang telah menjadi salah satu layanan berbasis teknologi utama di sektor transportasi dan pengiriman makanan. Orang sering kali mengungkapkan opini dan ulasan mereka tentang layanan aplikasi seperti *Grab* melalui berbagai media, termasuk *Google Play Reviews*. Data ulasan ini dapat digunakan sebagai sumber utama dalam penelitian analisis sentimen [2].

Ojek *online* telah menjadi pilihan populer karena kemudahannya dalam mengakses layanan transportasi melalui aplikasi tanpa perlu datang ke tempat konvensional untuk memesan. Orang sering kali mengungkapkan opini dan ulasan mereka tentang layanan ojek *online* melalui berbagai media, termasuk *Google Play Reviews*. Melalui platform ini, pengguna dapat memberikan ulasan tentang kelebihan dan kekurangan layanan ojek *online*. Ulasan pengguna aplikasi ojek *online* dapat bervariasi, mulai dari yang sangat positif hingga sangat negatif, tergantung pada pengalaman masing-masing pengguna. Selain memberikan ulasan,

pengguna juga dapat memberikan rating berupa skor bintang dan menyatakan apakah mereka menyukai aplikasi tersebut atau tidak. [3].

Analisis sentimen adalah salah satu solusi yang dapat digunakan untuk menentukan tingkat kepuasan pengguna dari data yang tidak terstruktur. Dalam konteks ini, *Deep Learning* muncul sebagai metode yang efektif untuk mengimplementasikan analisis sentimen. *Deep Learning* bekerja dengan mengekstraksi data menggunakan jaringan syaraf tiruan, di mana model belajar dengan menyesuaikan nilai error. Salah satu pendekatan *Deep Learning* yang umum digunakan untuk klasifikasi teks adalah *Long Short-Term Memory (LSTM)*. Dengan kemampuannya memahami konteks dan hubungan antar kata dalam teks, menjadi metode yang ideal untuk menganalisis sentimen ulasan pelanggan terhadap layanan Grab [4].

Variasi seperti *Stacked LSTM* dan *Bidirectional LSTM* juga menunjukkan potensi untuk meningkatkan performa dalam klasifikasi teks. Pada penelitian sebelumnya, algoritma *BiLSTM* menunjukkan performa yang lebih baik dibandingkan dengan *LSTM* dalam analisis sentimen. Namun, hasil penelitian tersebut masih memiliki keterbatasan, terutama dalam mengoptimalkan hubungan antar kata pada teks yang kompleks [4]. Untuk mengatasi keterbatasan tersebut, penelitian ini menambahkan *Multi-Head Attention* pada kedua model yang akan dibandingkan, yaitu *Stacked LSTM* dan *Bidirectional LSTM*. Dengan pendekatan ini, diharapkan performa model dapat ditingkatkan, khususnya dalam memahami konteks yang lebih luas dan menangani kompleksitas data teks secara lebih efektif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dikaji diatas, dapat disimpulkan beberapa rumusan masalah seperti berikut :

1. Bagaimana performa dari model *Stacked LSTM* dan *Bidirectional LSTM* dalam menganalisis sentiment ulasan dari data aplikasi Grab?
2. Apakah penambahan *Multi-Head Attention* pada model *Stacked LSTM* dan *Bi-Directional* dapat meningkatkan performa model dalam melakukan

analisis sentimen ulasan pengguna aplikasi grab?

3. Berdasarkan hasil evaluasi akhir terkait akurasi, loss, precision, recall dan F1-score, model manakah yang memiliki performa terbaik untuk analisis sentiment pada ulasan pengguna aplikasi grab?

1.3 Batasan Masalah

Permasalahan dalam penelitian ini dibatasi berdasarkan batasan-batasan berikut:

- 1) Penelitian ini hanya berfokus pada perbandingan performa model *Stacked LSTM* dan *Bi-Directional LSTM* dengan penambahan *Multi-Head Attention* sebagai upaya untuk meningkatkan hasil kinerja model.
- 2) Penelitian ini menggunakan 2000 data ulasan aplikasi Grab yang diambil pada Google Play Store.
- 3) Analisis hanya difokuskan pada kategori sentimen positif dan negatif, tanpa mempertimbangkan kategori sentimen netral.
- 4) Penelitian ini menggunakan google collab untuk pengambilan data dengan bahasa *python*.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi *Grab* di platform *Google Play*, dengan fokus pada perbandingan performa dari model *Deep Learning*, khususnya *Long Short-Term Memory (LSTM)*, dalam menganalisis tingkat kepuasan pengguna, penelitian ini akan membandingkan *Stacked LSTM* dan *Bidirectional LSTM* yang akan di tambahkan dengan mekanisme *Multi Head Attention* untuk menentukan hasil terbaik dalam mengidentifikasi sentimen pengguna, baik yang bersifat positif maupun negatif. Selain itu, penelitian ini juga akan mengevaluasi performa masing-masing model berdasarkan akurasi, *loss*, *precision*, *recall*, dan *f1-score*, untuk memberikan gambaran tentang seberapa efektif model *LSTM* dalam menangkap sentimen pengguna terhadap layanan yang diberikan oleh pengguna pada aplikasi *Grab*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis maupun praktis sebagai berikut:

1) Manfaat Teoritis

Penelitian ini diharapkan dapat berkontribusi pada pengembangan ilmu pengetahuan dalam bidang pemrosesan bahasa alami (*NLP*), terutama dalam hal penerapan model *Deep Learning* seperti *Stacked LSTM* dan *Bidirectional LSTM* yang akan dilengkapi dengan *Multi-Head Attention*. Selain itu penelitian ini juga bertujuan untuk meningkatkan pengetahuan tentang analisis sentiment dengan dataset berbahasa Indonesia. Penelitian ini juga diharapkan menjadi acuan bagi penelitian selanjutnya yang berfokus pada pembuatan model *Deep Learning* untuk analisis sentiment dan klasifikasi teks lainnya.

2) Manfaat Praktis

Penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam terkait pendapat dan perasaan pelanggan tentang layanan yang diberikan, sehingga Perusahaan dapat memahami dan mengambil Kesimpulan terkait kebutuhan pengguna untuk meningkatkan kualitas layanan. Selain itu penelitian ini juga diharapkan memberikan informasi yang berguna untuk keputusan strategis, seperti meningkatkan fitur aplikasi atau meningkatkan kualitas layanan driver.

1.6 Sistematika Penulisan

Penelitian ini disusun secara sistematis untuk memudahkan pembaca dalam memahami alur penelitian. Berikut adalah ringkasan isi dari keseluruhan bab:

BAB I PENDAHULUAN

Bab ini menguraikan dasar penelitian, termasuk latar belakang yang menekankan pentingnya analisis sentimen pada aplikasi *Grab*, rumusan masalah yang berfokus pada evaluasi kinerja model *Stacked LSTM* dan *Bi-Directional LSTM* dengan *Multi-Head Attention*, serta tujuan, batasan, dan manfaat penelitian.

BAB II TINJAUAN PUSTAKA

Bab ini mencakup teori-teori yang menjadi dasar penelitian, seperti konsep *Natural Language Processing (NLP)*, analisis sentimen, *Long Short-Term Memory*

(*LSTM*), dan *Multi-Head Attention*. Selain itu, dibahas juga metode evaluasi kinerja model dan ulasan terhadap penelitian sebelumnya yang relevan.

BAB III METODE PENELITIAN

Bab ini menjelaskan metode yang digunakan, termasuk proses pengumpulan data melalui *scraping* dari *Google Play Store*, pelabelan, *preprocessing*, penyeimbangan data menggunakan teknik *SMOTE*, serta perancangan arsitektur *Stacked LSTM* dan *Bi-Directional LSTM* dengan *Multi-Head Attention*. Metode evaluasi melibatkan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil penelitian dan membandingkan performa model berdasarkan metrik evaluasi. Penjelasan hasil disajikan secara mendetail menggunakan tabel, grafik, dan analisis, disertai pembahasan mengenai kelebihan dan kekurangan setiap model yang diuji.

BAB V PENUTUP

Bab terakhir ini berisi kesimpulan yang dirumuskan berdasarkan hasil penelitian, serta saran untuk penelitian mendatang. Kesimpulan menyoroti model terbaik dalam analisis sentimen, sementara saran mengusulkan pengembangan lebih lanjut pada metode atau penerapan penelitian ini.