

BAB I PENDAHULUAN

1.1. Latar Belakang

Aplikasi *mobile banking* telah menjadi bagian penting dari kehidupan masyarakat. Aplikasi *mobile banking* menawarkan berbagai fitur dan layanan yang memudahkan pengguna untuk melakukan transaksi keuangan. Salah satu aplikasi *mobile banking* tersebut ialah Brimo. Brimo merupakan aplikasi *mobile banking* dari bank BRI guna memenuhi layanan nasabah secara online.

Bank Rakyat Indonesia (BRI) adalah salah satu bank terbesar milik pemerintah Indonesia. Dengan jaringan yang luas terutama di daerah pedesaan, BRI memungkinkan akses layanan perbankan bagi masyarakat terpencil, termasuk yang belum terjangkau oleh bank lain. BRI menawarkan berbagai produk, termasuk program *corporate social responsibility* (CSR) untuk pembangunan kesejahteraan masyarakat. Bank ini juga dikenal fokus pada pengembangan usaha mikro, kecil, dan menengah (UMKM), dengan banyak nasabah dari kalangan pelaku UMKM. Dengan banyaknya nasabah, keluhan, saran, dan kritik terhadap layanan atau produk BRI tidak dapat dihindari. Sebagian besar umpan balik ini disampaikan melalui ulasan pengguna aplikasi *mobile banking* Brimo. Ulasan pengguna dapat memberikan informasi tentang pengalaman mereka menggunakan aplikasi, serta aspek-aspek seperti fitur dan fungsi produk, kemudahan penggunaan, kualitas layanan pelanggan, dan nilai produk atau layanan [1].

Maka dari itu diperlukan anotasi sentimen mengenai ulasan pengguna aplikasi Brimo agar pihak BRI bisa mempercepat penyelesaian khusus terkait spesifik masalah yang disebutkan pengguna, serta pengguna atau nasabah BRI mendapat solusi dari apa yang mereka sampaikan.

Anotasi sentimen adalah proses pemberian label sentimen, seperti positif, negatif, atau netral pada teks [1]. Anotasi sentimen dapat digunakan

untuk menganalisis ulasan pengguna aplikasi *mobile banking*. Terdapat beberapa model yang biasa digunakan untuk anotasi sentimen antara lain, model yang berbasis leksikon dan model yang berbasis semantik.

Model anotasi sentimen berbasis leksikon adalah model anotasi sentimen yang mengenali jenis sentimen berdasarkan daftar kata atau kamus yang dikaitkan dengan sentimen positif, negatif, atau netral [2]. Model ini bekerja dengan cara menghitung frekuensi kemunculan kata-kata atau frasa tersebut dalam teks yang akan dianotasi. Sedangkan model semantik bekerja dengan cara menggunakan informasi kontekstual dari teks untuk menentukan maknanya. Informasi kontekstual ini dapat berupa tata bahasa, leksikon, atau pengetahuan dunia [1]. Dalam pemodelan lebih lanjut, model semantik dapat juga berupa model anotasi sentimen berbasis *machine learning*. Model anotasi berbasis *machine learning* adalah model anotasi sentimen yang menggunakan algoritma pembelajaran mesin untuk memprediksi sentimen dari teks [2].

Di antara model anotasi sentimen berbasis *machine learning* yang sudah teruji untuk *task sentiment annotation* antara lain model *pretrained* IndoBERT [3], [4], [5], [6], [7] dan model BiLSTM [8], [9], [10], [11], [12]. Model anotasi sentimen IndoBERT merupakan model *pretrained* yaitu model yang telah dilatih dengan dataset yang besar untuk spesifik *task* tertentu. Model *pretrained* IndoBERT telah menunjukkan kemampuan yang baik dalam memahami konteks dan makna kata dalam bahasa Indonesia karena model IndoBERT telah dilatih ulang menggunakan dataset berbahasa Indonesia, IndoBERT terbukti efektif untuk berbagai tugas *natural language processing* (NLP) seperti klasifikasi sentimen, terjemahan, dan *summarization* [13]. Sedangkan LSTM memiliki kemampuan yang baik dalam mengingat informasi jangka pendek dan jangka panjang [14]. Namun LSTM memiliki satu kelemahan yaitu hanya mampu menangkap konteks kalimat dari kata sebelumnya, sehingga akan dilakukan percobaan menggunakan BiLSTM yang terbukti mampu untuk memahami konteks secara dua arah urutan kata yang berbeda yakni dari kiri

ke kanan dan dari kanan ke kiri[15]. Penggunaan model *pretrained* IndoBERT dipilih untuk membuktikan kemampuan IndoBERT dalam melakukan anotasi sentimen pada data yang berbentuk ulasan. Sedangkan, BiLSTM diujikan sebagai pembanding karena model ini mampu untuk menangkap makna kata dari kiri ke kanan dan dari kanan ke kiri, model dibuat berdasarkan penelitian [16]. Dengan membatasi model yang digunakan, maka penelitian ini dapat fokus untuk membandingkan kinerja kedua model tersebut.

1.2. Rumusan Masalah

Pretrained model IndoBERT menggunakan data korpus teks berbentuk artikel dengan total 220 juta kata yang terdiri dari 74 juta kata dari wikipedia, 55 juta kata dari kompas, tempo, dan liputan 6, 90 juta kata dari *Indonesian Web Corpus*, sedangkan dataset yang dianotasi adalah komentar dari google play store, maka untuk mengetahui performa model, dilakukan pemodelan lain menggunakan model BiLSTM, sehingga akan diperoleh hasil perbandingan akurasi antara *pretrained* model IndoBERT yang telah dilatih dengan BiLSTM.

1.3. Pertanyaan Penelitian

1. Apakah ada perbandingan akurasi yang signifikan antara *pretrained* IndoBERT dan BiLSTM dalam melakukan anotasi sentimen pada ulasan aplikasi *mobile banking* brimo?
2. Bagaimana bentuk arsitektur anotasi sentimen model BiLSTM yang mampu mengklasifikasi sentimen pada ulasan aplikasi *mobile banking* Brimo?

1.4. Batasan Masalah

1. Data yang akan digunakan adalah teks ulasan pada aplikasi *mobile banking* Brimo pada layanan *google play store*, pada rentang waktu Maret sampai Mei 2024.

2. Data yang digunakan hanya sejumlah 20.000 baris ulasan yang dibagi menjadi data latih sebanyak 10.000 ulasan, 9.000 ulasan untuk data *testing*, dan 1.000 validasi.
3. Data hanya akan diklasifikasi menjadi 2 jenis sentimen yaitu positif dan negatif.

1.5. Tujuan Penelitian

1. Mengetahui perbandingan kinerja IndoBERT dan BiLSTM dalam melakukan anotasi sentimen pada ulasan aplikasi *mobile banking* Brimo.
2. Mengetahui arsitektur BiLSTM yang akurasinya tinggi pada anotasi sentimen ulasan aplikasi *mobile banking* Brimo.

1.6. Manfaat Penelitian

1. Manfaat teoritis, secara teoritis tulisan ini dapat menambah ilmu pengetahuan dan perkembangan di bidang *artificial intelligence* khususnya NLP (*natural language processing*). Serta penelitian ini dapat dijadikan referensi mengenai analisis sentimen atau klasifikasi teks dengan metode IndoBERT dan BiLSTM.
2. Manfaat praktis bagi peneliti, menambah pengetahuan mengenai anotasi sentimen dalam klasifikasi teks.
3. Bagi Institut Teknologi Telkom Purwokerto dan mahasiswa, dapat menambah referensi dan koleksi mengenai penggunaan *pretrained* IndoBERT dan BiLSTM yang bermanfaat bagi Institut Teknologi Telkom Purwokerto dan mahasiswa informatika lainnya.