

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Analisis Sentimen merupakan bidang studi yang menganalisis pendapat, sentimen, penilaian, sikap dan emosi dari orang-orang terhadap entitas dan atributnya yang diekspresikan melalui teks [1]. *Aspect Based Sentiment Analysis* (ABSA) merupakan salah satu dari tiga tingkat kedetailan pada analisis sentimen. Secara tradisional, Analisis sentimen menentukan sebuah opini tentang suatu produk hanya menentukan polaritasnya positif, negatif, atau netral, dibandingkan dengan Analisis Sentimen biasa, ABSA membantu memahami permasalahan pada analisis sentimen lebih baik dan memberikan peluang yang lebih baik dalam menganalisis sentimen karena ABSA menganalisis langsung dengan melihat opini dan target, serta menyadari pentingnya target dari opini [1], [2]. ABSA bertujuan untuk dapat mendeteksi polaritas dari sentimen teks berdasarkan aspek tertentu [3]. *Bidirectional Long Short-Term Memory (Bi-LSTM)* merupakan algoritma *Long Short Term-Memory* (LSTM) yang terdiri dari dua lapisan jaringan syaraf LSTM, yaitu *forward* dan *backward direction* untuk memodelkan masing-masing konteksnya (sebelum dan sesudahnya). Bi-LSTM menghubungkan dua *hidden layers* dari arah yang berlawanan menuju *output* yang sama. Dengan ini lapisan *neuron* dapat memperoleh informasi dari kondisi masa lalu (*past*) dan masa depan (*future*) secara serentak. Bi-LSTM sering digunakan untuk memodelkan informasi konteks di bidang *Natural Language Processing* (NLP) [4], [5].

Market place menyediakan fitur ulasan agar dapat digunakan oleh pengguna dalam memberikan opininya mengenai *market place* tersebut, namun banyaknya jumlah ulasan membuat para kandidat sulit untuk mengambil keputusan sehingga dibutuhkan analisis sentimen berbasis aspek untuk mengolah ulasan dan membedakannya berdasarkan polaritas dan aspeknya [6]. Beberapa tahun terakhir Analisis Sentimen Berbasis Aspek (ABSA) menjadi topik NLP yang banyak diteliti. Pada Penelitian [7] oleh Putra dkk. Menggunakan LSTM dengan GloVe

sebagai metode *word embedding* mendapatkan akurasi rata-rata 81%. Penelitian [8] oleh Xu dkk membandingkan metode Feature+SVM, LSTM, RNN dan CNN-CRF. Pada domain restoran Akurasi LSTM 66.45% lebih rendah dibandingkan Feature+SVM dan CNN-CRF, sementara pada domain laptop akurasinya 74.28% lebih tinggi dari CNN-CRF dan RNN. Lalu penelitian [9] oleh Khine & Aung membandingkan metode LSTM dengan SenticNet MA-LSTM, dimana LSTM akurasinya 82% sementara SenticNet MA-LSTM akurasinya 87.5% lebih tinggi dari LSTM. Dari hasil akurasi penelitian yang dijabarkan LSTM mendapatkan akurasi yang cukup baik, LSTM dapat menangkap teks sekuensial yang panjang, dan dapat menangani permasalahan penurunan gradien pada RNN [9]. Namun kekurangan LSTM yaitu tidak dapat menangkap *context semantic information*, dikarenakan informasi hanya ditransmisikan dari *past to future* [5], [10].

Dari permasalahan tersebut, solusinya adalah menggunakan algoritma Bi-LSTM, karena Bi-LSTM mampu menangani permasalahan pada LSTM yakni menangkap *context semantic information* lebih efektif dari dua arah (*past to future & future to past*) serta mendapatkan lebih baik representasi teks dari sentimen [10]. Metode *word embedding* menggunakan Word2Vec CBOW karena dapat juga membantu menangkap informasi semantik dari teks dan lebih baik dari Skip-Gram dan *One-hot Encoding* [4]. Berikut ini riset Bi-LSTM dalam bidang NLP. Pada penelitian [5] Bi-LSTM mendapatkan hasil akurasinya lebih besar yakni 85.69% dibanding CNN dan LSTM. Selanjutnya penelitian ujaran kebencian [4] hasil akurasi Bi-LSTM 96.93% lebih baik dari LSTM. Pada penelitian [11] kasus identifikasi *cyberbullying* Bi-LSTM mendapatkan hasil akurasi 95.24% lebih besar dibanding LSTM, lalu penelitian deteksi depresi dan kecemasan [12] Bi-LSTM mendapatkan hasil akurasi 94.12% lebih besar dibanding LSTM yang hanya 84.91%, Lalu penelitian [13] pada klasifikasi teks BI-LSTM mendapatkan akurasi 87.3% lebih tinggi dibandingkan LSTM yang akurasinya 85.2%.

Berdasarkan penjelasan latar belakang serta penelitian sebelumnya maka, akan dilakukan Analisis Sentimen Berbasis Aspek terhadap ulasan aplikasi *Market Place* menggunakan *Bidirectional Long Short-Term Memory* dengan metode *Word Embedding* Word2Vec CBOW.

1.2. Rumusan Masalah

Berdasarkan uraian yang disampaikan dalam latar belakang masalah, maka rumusan masalahnya adalah Bagaimanakah performansi dari metode *Bidirectional Long Short-Term Memory* dalam melakukan analisis sentimen berbasis aspek terhadap ulasan aplikasi *Market Place* pada Google Playstore.

1.3. Maksud dan Tujuan

Maksud dari penelitian ini adalah untuk implementasi *Bidirectional Long Short-Term Memory* pada kasus analisis sentimen berbasis aspek terhadap ulasan aplikasi *market place*. Berdasarkan rumusan masalah yang ada, maka tujuan penelitiannya adalah Mengetahui performansi dari metode *Bidirectional Long Short-Term Memory* pada Analisis Sentimen Berbasis Aspek terhadap ulasan aplikasi *Market place*.

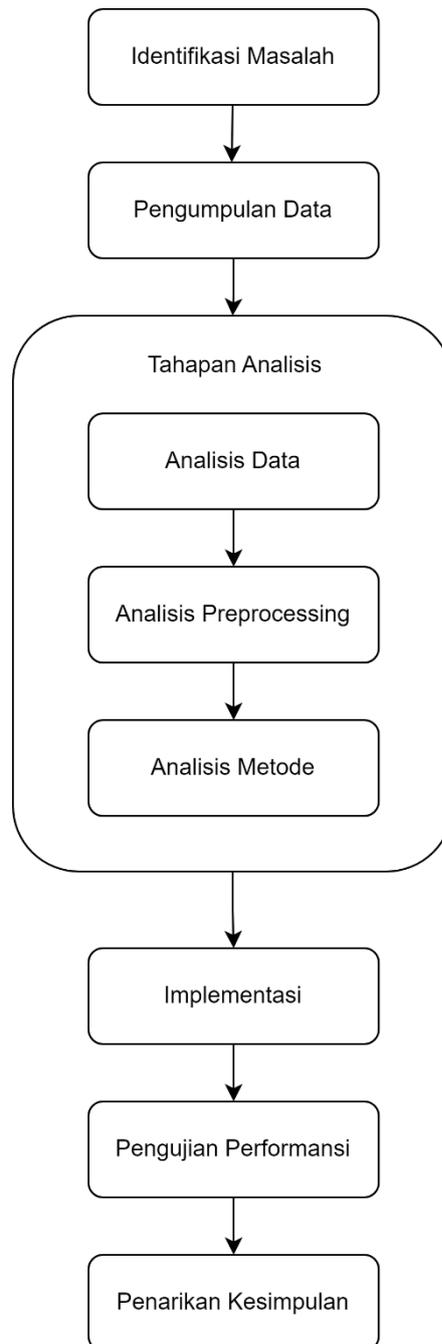
1.4. Batasan Masalah

Batasan masalah yang digunakan dalam penelitian adalah sebagai berikut:

1. Data yang digunakan dalam penelitian diambil dari ulasan aplikasi *Market Place* di Google Play Store berbahasa Indonesia data yang diambil kisaran Mei 2022 – Oktober 2022.
2. Dataset masukkan berupa format .csv.
3. Pengambilan data ulasan pada Google Play Store menggunakan *web scraping*.
4. Kelas label sentimen terdiri dari positif, negatif, dan non sentimen.
5. Aspek yang digunakan yaitu aspek Layanan, Fitur, Kebermanfaatan, Bisnis, dan Non Aspek.
6. Metode mengubah kata menjadi vektor menggunakan *Word2Vec* arsitektur *Continuous Bag of Words (CBOW)*.

1.5. Metode Penelitian

Metode penelitian yang diusulkan pada penelitian ini dapat dilihat pada Gambar 1.1.



Gambar 1.1. Metode Penelitian

1.5.1. Identifikasi Masalah

Langkah awal yang dilakukan adalah mengidentifikasi masalah. Terdapat pada latar belakang permasalahan yang diidentifikasi adalah Bagaimanakah performansi dari metode *Bidirectional Long Short-Term Memory* dalam melakukan

analisis sentimen berbasis aspek terhadap ulasan aplikasi *market place* pada Google Playstore.

1.5.2. Metode Pengumpulan Data

Metode Pengumpulan Data yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Studi Literatur, merupakan dengan cara mengumpulkan, jurnal, *paper*, buku, dan sumber literasi lainnya yang berkaitan dengan analisis sentimen berbasis aspek dan metode yang digunakan bertujuan untuk mendapatkan pemahaman.
2. Pengumpulan Dataset, dataset yang digunakan untuk penelitian ini adalah data yang diambil dari Google Playstore pada ulasan aplikasi *market place* berbahasa Indonesia dengan melakukan Teknik *web scraping*.

1.5.3. Metode Tahapan Analisis

Metode tahapan analisis yang dilakukan pada penelitian ini yaitu sebagai berikut:

1. Analisis Data, menganalisis struktur atau pola kalimat dari data yang didapatkan dari ulasan Aplikasi *Marketplace* pada Google Playstore.
2. Analisis Praproses, mentransformasikan data yang mentah menjadi data yang terstruktur serta mudah dipahami. Langkah-langkah yang dilakukan pada tahapan ini antara lain; *Case folding, filtering, word normalization, convert negation, tokenization, dan stopwords removal*.
3. Analisis Metode, pada tahapan ini dilakukan *word embedding* dengan metode Word2Vec CBOW serta dilakukan klasifikasi menggunakan metode *Bidirectional Long Short-Term Memory* pada kasus analisis sentimen berbasis aspek.

1.5.4. Implementasi

Tahapan implementasi ini adalah melakukan implementasi sistem berupa simulator, yang dimulai dari memasukkan dan membaca data ulasan yang telah dilabeli secara manual, kemudian melakukan implementasi *preprocessing* yaitu:

case folding, filtering, word normalization, convert negation, tokenization, dan stopword removal, setelah data masukkan dipraproses, dilanjut melakukan implementasi *word embedding* dengan *Word2Vec Continuous Bag of Words* (CBOW) memakai data yang telah dipraproses tadi, Data yang telah dilakukan proses *word embedding* akan menjadi vektor, namun sebelum melakukan pelatihan dan pengujian algoritma, dilakukan pembagian data, yaitu data latih dan data uji. Lalu implementasi algoritma membangun model Bi-LSTM pada proses pelatihan model Bi-LSTM akan dibagi menjadi pelatihan klasifikasi aspek dan klasifikasi sentimen. Pada proses pelatihan tadi dilakukan secara optimal agar hasilnya baik. Setelah pembangunan model Bi-LSTM, tahap selanjutnya yaitu pengujian yang akan menghasilkan prediksi, pengujian Bi-LSTM dalam ABSA yaitu pengujian klasifikasi aspek lalu pengujian klasifikasi sentimen.

1.5.5. Pengujian Performansi

Pada tahap ini melakukan pengujian dan evaluasi performansi seperti *accuracy, precision, recall, dan f1-score*. metode klasifikasi menggunakan *Bidirectional Long Short-Term Memory* dalam kasus analisis sentimen berbasis aspek dan mengevaluasi performansi model menggunakan *Confusion Matrix*.

1.5.6. Penarikan Kesimpulan

Pada tahap ini dilakukan penarikan kesimpulan dari hasil pengujian yang dilakukan dengan metode *Bidirectional Long Short-Term Memory* pada kasus analisis sentimen berbasis aspek apakah sesuai dengan tujuan penelitian.

1.6. Sistematika Penulisan

Sistematika penulisan ini dibuat agar dapat memudahkan dan memberikan gambaran mengenai penelitian yang dilakukan. Maka Sistematika penulisannya adalah sebagai berikut:

BAB 1 PENDAHULUAN

Pada Bab ini berisi latar belakang, rumusan masalah, maksud dan tujuan, Batasan masalah, metodologi penelitian, serta sistematika penulisan penelitian.

BAB 2 LANDASAN TEORI

Pada bab ini berisi teori-teori pendukung dalam penelitian, seperti teori *Natural Language Processing*, analisis sentimen berbasis aspek, tahapan *preprocessing*, *word embedding*, *deep learning*, algoritma *Bidirectional Long Short-Term Memory (BiLSTM)*, dan teori lainnya mengenai topik penelitian ini.

BAB 3 ANALISIS DAN PERANCANGAN

Pada bab ini berisi tentang analisis masalah, analisis data masukan, sistem yang dibangun, analisis *preprocessing*, analisis *word embedding*, dan analisis metode *Bidirectional Long Short-Term Memory*.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Pada bab ini berisi implementasi sistem dan pengujian berdasarkan analisis yang dilakukan pada bab sebelumnya, dan pengujian performansi algoritma *Bidirectional Long Short-Term Memory* dengan *confusion matrix* pada Analisis sentimen berbasis aspek ini.

BAB 5 KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dan saran dari hasil penelitian mengenai performansi analisis sentimen berbasis aspek menggunakan *Bidirectional Long Short-Term Memory* yang telah dilakukan, dan sebagai acuan dalam penelitian di masa mendatang mengenai Analisis sentimen berbasis aspek.