

BAB I

PENDAHULUAN

A. Latar Belakang

Pertumbuhan industri obat tradisional, farmasi, kimia, dan kosmetik tumbuh 5,59% pada kuartal pertama tahun 2020 [1]. Selain itu, data dari Badan Pengawas Obat dan Makanan (BPOM) menunjukkan bahwa jumlah perusahaan yang berada di industri kosmetik mengalami peningkatan pada periode 2021 hingga bulan Juli 2022 menjadi 913 industri dari 819 industri [2]. Pertumbuhan industri kosmetik juga didukung dengan kehadiran *e-commerce* di Indonesia. Pernyataan ini didukung dengan hasil survei yang dilakukan oleh Populix pada bulan Juli 2022. Survei tersebut menunjukkan bahwa 66% pembeli di Indonesia membeli produk kosmetik melalui *e-commerce* [3].

E-commerce merupakan suatu sistem yang memungkinkan konsumen untuk melakukan transaksi pembelian secara daring [4]. Salah satu fitur yang dimiliki oleh *e-commerce* adalah ulasan produk. Fitur ini memberikan pembeli sebuah akses untuk memberikan penilaian terhadap barang yang telah berhasil dibeli dan sampai ke alamat pengiriman dalam bentuk rentang nilai atau *rating* serta kalimat ulasan. Fitur ini memiliki pengaruh penting karena pembeli cenderung memberikan perhatian lebih ke ulasan negatif dibandingkan dengan ulasan positif [5]. Dalam beberapa kasus, pembeli memberikan nilai *rating* dan kalimat ulasan yang tidak sejalan. Hal ini dapat memengaruhi calon pembeli lain saat ingin membeli suatu produk melalui *e-commerce*.

Secara umum, setiap ulasan yang diberikan oleh pembeli memiliki nilai sentimen. Sentimen merupakan opini yang didasari oleh emosi, perasaan, dan sikap seseorang [6]. Sentimen dapat dibagi menjadi tiga, antara lain sentimen negatif, positif, dan netral. Sentimen negatif menggambarkan tentang kekecewaan, ketidakpuasan, dsb. Sentimen positif menggambarkan tentang kebahagiaan, kepuasan, dsb. Sentimen netral menggambarkan tentang opini yang tidak mengarah ke positif maupun negatif.

Sentimen dari suatu kalimat dapat dipelajari dengan menggunakan metode analisis sentimen. Analisis sentimen merupakan metode yang dapat digunakan untuk mengidentifikasi sentimen suatu kalimat. Analisis sentimen dapat dilakukan dengan memanfaatkan metode *deep learning*. *Deep learning* merupakan suatu metode yang dapat digunakan untuk mempelajari karakteristik dari data yang berukuran besar [7]. Perkembangan *deep learning* untuk melakukan tugas analisis sentimen dalam beberapa tahun ke belakang menunjukkan prospek positif [8]. Dengan demikian, penelitian ini menggunakan metode *deep learning* untuk melakukan analisis sentimen. Salah satu arsitektur *deep learning* yang dapat digunakan adalah *Long Short-Term Memory* (LSTM). Arsitektur ini memiliki *memory cell* yang dapat mengingat data berurutan. Dengan kemampuan tersebut, arsitektur LSTM cocok digunakan untuk mempelajari karakteristik suatu kalimat [9].

Berdasarkan uraian di atas, penelitian ini membahas tentang analisis sentimen terhadap ulasan pembeli produk kecantikan dari *e-commerce* menggunakan arsitektur *Long Short-Term Memory* (LSTM). *Dataset* yang dipakai dalam penelitian ini merupakan kumpulan ulasan dari pembeli produk kecantikan yang didapatkan dari beberapa *e-commerce*, antara lain Lazada, Sephora, Shopee, Sociolla, Tokopedia, dan Amazon. Pada penelitian ini, ulasan konsumen dipetakan menjadi sentimen negatif dan sentimen positif berdasarkan nilai *rating*. Ulasan dengan nilai *rating* satu sampai tiga pada penelitian ini lebih condong ke sentimen negatif dan ulasan dengan nilai *rating* empat sampai lima mengarah ke sentimen positif. Berdasarkan pertimbangan tersebut, pada penelitian ini ulasan dengan nilai *rating* satu sampai tiga diberikan label sentimen negatif, sedangkan ulasan dengan nilai *rating* empat dan lima diberikan label sentimen positif.

B. Rumusan Masalah

Berdasarkan penjelasan dari latar belakang, maka esensi permasalahan yang akan diselesaikan dari penelitian ini, yaitu:

1. Bagaimana tahapan dalam melakukan analisis sentimen positif dan negatif terhadap ulasan pembeli produk kecantikan menggunakan model dengan arsitektur LSTM?
2. Bagaimana performa model dengan arsitektur LSTM dalam melakukan analisis sentimen positif dan negatif terhadap ulasan produk kecantikan?

C. Batasan Masalah

Untuk mencegah pembahasan dalam penelitian ini menjadi terlalu melebar, penelitian ini menerapkan beberapa batasan masalah, yaitu:

1. *Dataset* yang digunakan berasal dari ulasan produk kecantikan pada beberapa *e-commerce*, seperti Lazada, Sephora, Shopee, Sociolla, Tokopedia, dan Amazon.
2. Sentimen ulasan dikelompokkan menjadi dua berdasarkan nilai *rating*, yaitu positif dan negatif. Ulasan dengan nilai *rating* satu sampai tiga dikelompokkan menjadi sentimen negatif, sedangkan ulasan dengan nilai *rating* empat dan lima dikelompokkan menjadi sentimen positif.
3. Arsitektur *deep learning* yang digunakan adalah *Long Short-Term Memory* (LSTM).
4. Jenis pengukuran evaluasi model yang digunakan adalah *specificity* dan *recall*.
5. Model dengan evaluasi paling optimal diimplementasikan ke dalam sistem web.

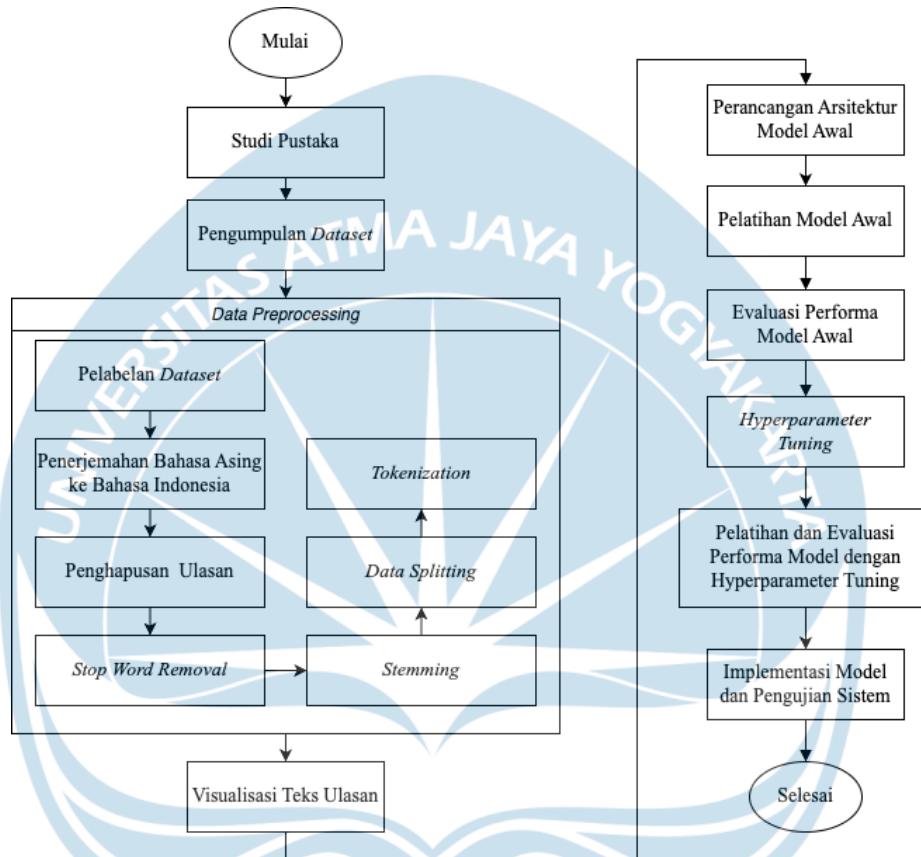
D. Tujuan Penelitian

Tujuan yang ingin dicapai dalam tugas akhir ini adalah sebagai berikut.

1. Mengetahui bagaimana prosedur analisis sentimen terhadap ulasan produk kecantikan menggunakan model berarsitektur LSTM.
2. Mengetahui performa model dengan arsitektur LSTM dalam melakukan analisis sentimen positif dan negatif terhadap ulasan pembeli produk kecantikan.

E. Metode Penelitian

Metode penelitian pada penelitian ini dapat dilihat pada Gambar 1.1 sebagai berikut.



Gambar 1.1 Diagram alir metode penelitian

1. Studi Pustaka

Pada tahap studi pustaka, studi dan literatur yang berkaitan dengan penelitian ini dikumpulkan dan dipelajari. Literatur yang digunakan berasal dari berbagai sumber, meliputi jurnal, buku, riset, dan artikel yang kredibel dari internet. Informasi spesifik yang dicari berkaitan dengan analisis sentimen, *deep learning*, dan *Long Short-Term Memory (LSTM)*.

2. Pengumpulan *Dataset*

Pada tahap ini, *dataset* dikumpulkan dengan menggunakan dua metode. Metode pertama adalah menggunakan *Application Programming Interface (API) scrapping* dari berbagai *e-commerce*, seperti Lazada, Sephora, Shopee,

Sociolla, Tokopedia. Metode kedua adalah mengunduh *dataset* ulasan produk kecantikan dari Amazon yang sudah tersedia di internet. Data yang diambil adalah teks ulasan dan nilai *rating* dari pembeli produk kecantikan. *Dataset* kemudian disimpan ke dalam suatu berkas dengan format *Comma Separated Values* (CSV) dengan tujuan mempermudah proses pembacaan data.

3. *Data Preprocessing*

Data preprocessing berfungsi untuk meningkatkan kualitas *dataset* dengan cara melakukan beberapa sub-tahapan, meliputi pelabelan *dataset*, penerjemahan teks ulasan dengan bahasa asing ke Bahasa Indonesia, penghapusan ulasan, *stop word removal*, dan *stemming*. *Dataset* yang telah melewati seluruh tahapan yang telah disebutkan kemudian dipecah menjadi tiga bagian pada tahap *data splitting*, yaitu: data latih, validasi, dan tes. Data latih merupakan data yang dipelajari oleh model. Data validasi merupakan data yang digunakan untuk mengevaluasi performa model saat proses pelatihan berlangsung. Data tes digunakan untuk mengukur performa model saat melakukan analisis sentimen terhadap data yang belum pernah dipelajari sebelumnya. Tahapan terakhir pada *data preprocessing* adalah *tokenization*. Pada tahap ini, setiap kalimat ulasan dipecah menjadi kumpulan token berbentuk *array*. Setiap token kemudian dirubah menjadi angka, sehingga karakteristik kalimat ulasan dapat dipelajari oleh model.

4. Visualisasi Teks Ulasan

Teks ulasan dengan sentimen negatif dan positif memiliki karakteristik yang berbeda-beda. Tahap ini berfungsi untuk memvisualisasikan karakteristik tersebut dengan menggunakan grafik *wordcloud*.

5. Perancangan Arsitektur Model Awal

Pada tahap ini, arsitektur model awal dirancang untuk dilatih dan dievaluasi pada tahap selanjutnya. Perancangan arsitektur model meliputi beberapa komponen, seperti *input layer*, *hidden layer*, dan *output layer*.

6. Pelatihan Model Awal

Setelah arsitektur model awal selesai dirancang, tahapan selanjutnya yang dilakukan adalah melatih model awal dengan menggunakan data latih dan validasi.

7. Evaluasi Performa Model

Pada tahap ini, model awal yang sudah dilatih dievaluasi performanya dengan menggunakan data tes. Jenis pengukuran evaluasi performa yang digunakan pada penelitian ini adalah *specificity* dan *recall*.

8. *Hyperparameter Tuning*

Setelah mendapatkan nilai *specificity* dan *recall* dari model awal, penelitian ini dilanjutkan ke tahap *hyperparameter tuning*. Tahap ini berfungsi untuk melakukan eksperimen terhadap kombinasi beberapa nilai *hyperparameter*, meliputi rasio *data splitting*, *input dimension*, *output dimension*, jumlah *unit* pada *hidden layer*, *activation function*, *batch size*, dan *epoch*.

9. Implementasi Model dan Pengujian Sistem

Tahap implementasi model dan pengujian sistem merupakan bagian akhir dalam penulisan laporan ini. Pada tahap ini, model dengan evaluasi performa yang paling optimal diimplementasikan ke dalam sistem web. Setelah itu, sistem web diuji untuk memastikan seluruh fungsionalitas berjalan dengan semestinya.

F. Sistematika Penulisan

Laporan tugas akhir ini menggunakan sistematika penulisan yang dideskripsikan sebagai berikut.

BAB I PENDAHULUAN

Bab I mendeskripsikan tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan dalam pembuatan laporan ini.

BAB II TINJAUAN PUSTAKA

Bab II berisi hasil tinjauan penelitian-penelitian sejenis yang sudah pernah dilakukan sebelumnya.

BAB III LANDASAN TEORI

Bab III menjelaskan tentang teori-teori yang berkaitan dengan tugas akhir ini.

BAB IV HASIL DAN PEMBAHASAN

Bab IV membahas tahapan-tahapan dalam melakukan analisis sentimen, antara lain: pengumpulan *dataset*, *data preprocessing*, visualisasi teks ulasan, perancangan arsitektur model awal, pelatihan model awal, evaluasi performa model awal, *hyperparameter tuning*, pelatihan dan evaluasi performa model dengan *hyperparameter tuning*, dan interpretasi hasil pelatihan model.

BAB V IMPLEMENTASI MODEL DAN PENGUJIAN SISTEM

Bab V membahas tentang bagaimana model yang telah dihasilkan diimplementasikan ke sistem web. Selain itu, bab ini juga berisi tentang interaksi sistem web dengan model supaya model tersebut dapat digunakan dalam sistem web. Poin terakhir yang dibahas adalah pengujian terhadap seluruh fitur yang ada pada sistem web.

BAB VI PENUTUP

Bab VI merupakan bab terakhir dalam penulisan laporan tugas akhir. Bab ini berisi kesimpulan dan saran dari penulis terkait dengan penelitian ini.