BAB III
METODE PENELITIAN

3.1 Objek Penelitian

Alfamart merupakan salah satu perusahaan ritel terbesar di Indonesia yang memiliki ribuan gerai yang tersebar di berbagai wilayah. Sebagai perusahaan yang bergerak di bidang perdagangan, Alfamart terus berinovasi untuk meningkatkan pelayanan kepada pelanggan, salah satunya dengan menghadirkan layanan berbasis digital. Alfamart menghadirkan aplikasi Alfagift sebagai solusi belanja yang lebih praktis dan efisien.

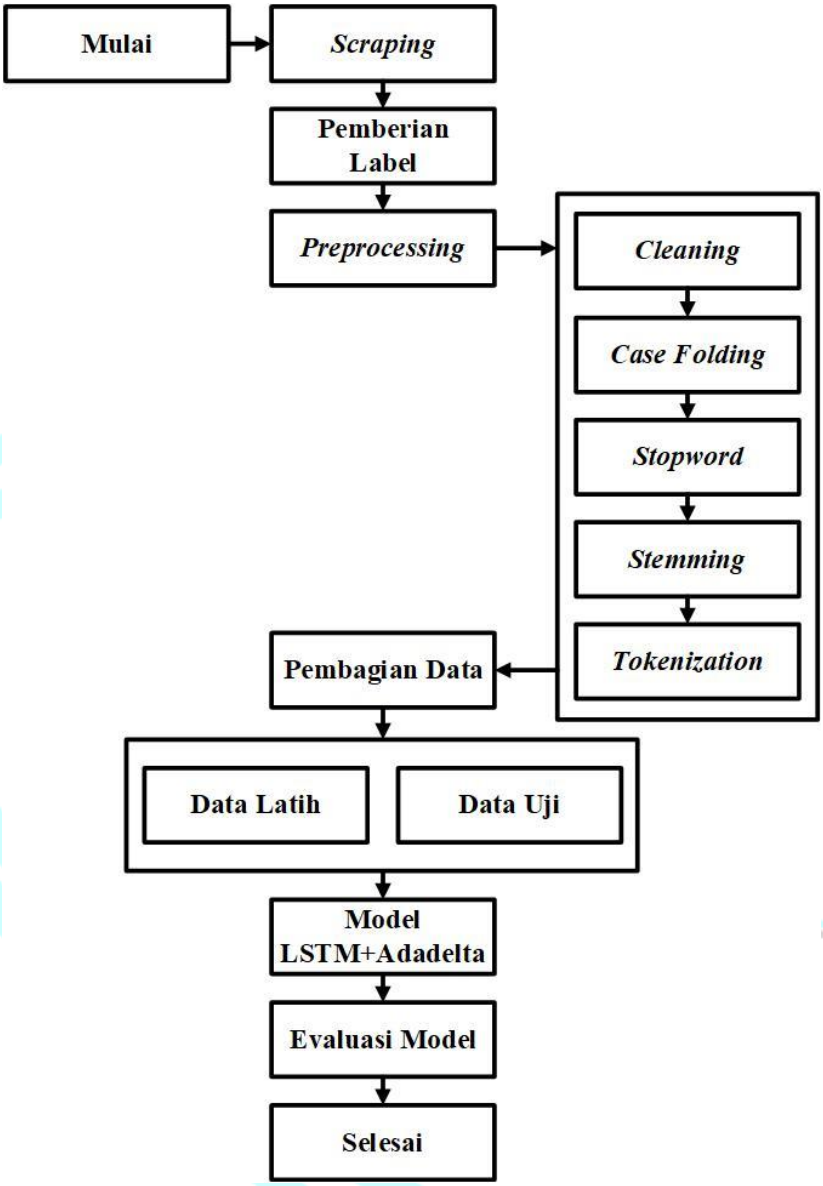
Objek penelitian ini adalah ulasan pengguna aplikasi Alfagift yang diperoleh dari *Google Play Store* dengan menggunakan teknik web *scraping*. Jumlah data yang digunakan dalam penelitian ini sebanyak 1000 data ulasan pengguna pada bulan November 2024 dengan versi aplikasi 4.33.0. Fokus penelitian meliputi mengklasifikasikan ulasan ke dalam beberapa kategori meliputi:

1. Sistem
2. Pelayanan
3. Transaksi
4. Keamanan
5. Promosi

Tujuan dari penelitian ini adalah untuk membuat model klasifikasi dan prediksi menggunakan algoritma LSTM dengan optimasi Adadelta. Algoritma LSTM dengan optimasi Adadelta digunakan karena dapat mengolah data teks dengan baik.

3.2 Prosedur Penelitian

Pada Gambar 3.1 merupakan prosedur penelitian yang dilakukan oleh peneliti mulai dari pengambilan data menggunakan metode *web scraping*, dilanjut dengan pemberian label pada ulasan, lalu melakukan *preprocessing* untuk membersihkan data ulasan. Membagi data menjadi data uji dan data latih, model LSTM dengan optimasi Adadelta digunakan untuk melatih model, terakhir melakukan evaluasi model menggunakan *confusion matrix*(Tukino et al., 2024).



Gambar 3.1 Alur Penelitian

3.2.1 Scraping Data

Pengambilan data ulasan dilakukan melalui proses *scraping* menggunakan *Google Colab*. Proses ini juga menggunakan library seperti *pandas, numpy,* serta *package google-play-scraper* dari bahasa pemrograman *Python*. (Rehatalanit et al., 2024b). Data dapat berupa terstruktur, semi-terstruktur, atau tidak terstruktur, termasuk teks, gambar, video, dan lainnya(Hilabi et al., 2025). Penelitian ini mengambil data berupa ulasan dari aplikasi Alfagift yang tersedia di Google Play Store.

3.2.2 Pemberian Label

Proses pemberian label dilakukan secara manual menggunakan aplikasi Microsoft Excel dengan memeriksa setiap ulasan satu per satu(Huda et al., 2024). Setelah itu, ulasan tersebut diberi label berdasarkan kategori transaksi, sistem, promosi, pelayanan, atau keamanan.

3.2.3 Preprocessing Data

Preprocessing merupakan proses persiapan dataset atau pembersihan data untuk digunakan dalam proses pelatihan data(Sujjada et al., 2024b). Pada penelitian ini, *preprocessing* bertujuan untuk meningkatkan kualitas data dan mengurangi *noise*. Ini terdiri dari beberapa tahap:

a. *Cleaning*, proses ini bertujuan untuk menghilangkan karakter seperti angka, tanda baca, simbol, link url, dan emoticon dari teks untuk mengurangi *noise*(Fremmuzar & Baita, 2023b).
b. *Case folding*, proses ini melibatkan pengubahan huruf kapital dalam kalimat atau kata diubah menjadi huruf kecil. Ini termasuk huruf dari "A"-"Z" dalam data menjadi “a”-“z” (Kholifatullah & Prihanto, 2023).
c. *Stopword*, proses ini merupakan *filtering*, yaitu pemilihan kata-kata penting yang digunakan untuk mewakili sebuah dokumen. (Zamsuri et al., 2023).
d. *Stemming*, proses ini digunakan untuk menghilangkan atau mengubah imbuhan, dari kata-kata yang memiliki imbuhan menjadi bentuk kata dasarnya(Maulana et al., 2023a).
e. *Tokenization*, proses ini bertujuan untuk membagi kata atau kalimat menjadi bagian-bagian kecil yang dikenal sebagai token(S et al., 2022).

3.2.4 Pembagian Data

Pembagian data ialah proses memisahkan dataset menjadi dua atau lebih subset yang berbeda, umumnya untuk keperluan pelatihan dan pengujian model *machine learning*(Putra et al., 2024). Pada penelitian ini, peneliti membagi data menjadi data latih untuk melatih model dan data uji untuk menguji atau mengevaluasi model.

3.2.5 Model Klasifikasi LSTM

Long Short Term Memory(LSTM) ialah sebuah jenis arsitektur jaringan saraf tiruan yang digunakan untuk memproses data berurutan atau *time series* untuk mengatasi masalah *gradient vanishing* dan *exploding*(Rolangon et al., 2023). Adadelta adalah ekstensi Adagrad yang berfungsi sebagai alternatif untuk mengurangi keagresifan Adagrad dan kecepatan belajar monoton. Ini juga lebih berkonsentrasi pada komponen kecepatan belajar(L. Azizah & Sukestiyarno, 2022). Dengan menggabungkan LSTM dan Adadelta, pemrosesan data menjadi lebih efektif karena LSTM mampu mengatasi permasalahan umum pada jaringan saraf tiruan, sementara Adadelta membantu dalam mengoptimalkan proses pembelajaran.

Tabel 3.1 Skenario Pelatihan

Skenario	***Learning Rate***	***Batch Size***	***Epoch***
1	1.0	16	100
2	1.0	32	100
3	1.0	64	100

3.2.6 Evaluasi Model

Data uji tidak sama dengan data latih yang digunakan dalam pelatihan model saat melakukan evaluasi. Untuk menilai seberapa baik kinerja metode klasifikasi, tahap ini menggunakan *confusion matrix*(Arthansa et al., 2024). Terdapat empat nilai yang dihasilkan oleh tabel *confusion matrix* diantaranya *True Positif, False Positif, False Negatif,* dan *True Negatif*(Anshari et al., 2023). Untuk mencari nilai dari *accuracy, precision, recall,* dan *f1-score* dapat menggunakan rumus berikut (Darmawan et al., 2023):

$$Accuracy = \frac{TP+TN}{Total} \times 100\% \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

$$F1\text{-}Score = \frac{2 \times Recall \times Precision}{Recall + Precision} \times 100\% \quad (4)$$

Keterangan:

TP = hasil dengan nilai prediksi positif dan nilai aktualnya positif.

TN = hasil dengan nilai prediksi negatif dan nilai aktualnya negatif.

FP = hasil dengan nilai prediksi positif dan nilai aktualnya negatif.

FN = hasil dengan nilai prediksi negatif dan nilai aktualnya positif