

BAB III

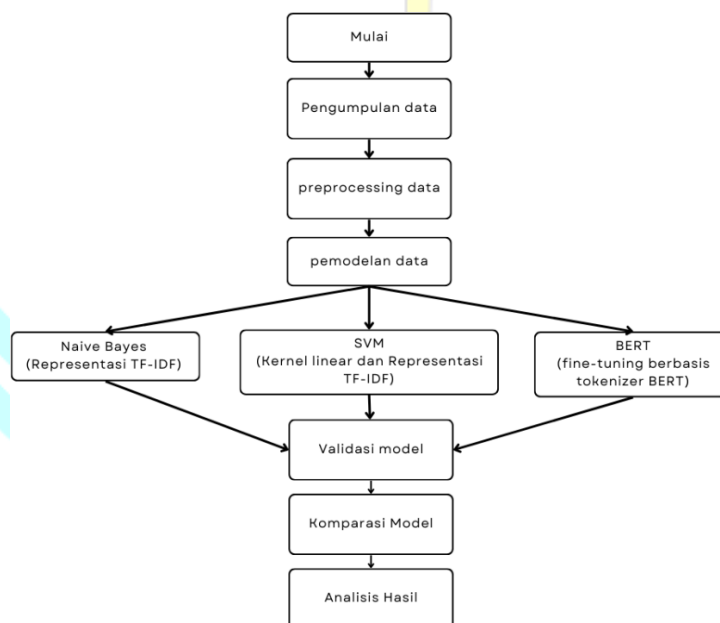
METODE PENELITIAN

3.1. Objek Penelitian

Penelitian ini menggunakan pendekatan machine learning untuk mengklasifikasikan ulasan pengguna aplikasi Yummy ke dalam kategori positif dan negatif. Objek penelitian meliputi data ulasan yang ada pada google playstore mulai dari bulan Mei 2019 hingga Desember 2024 yang diperoleh melalui scrapping menggunakan google colab

3.2. Prosedur Penelitian

Tahapan Prosedur penelitian ini dirancang dalam beberapa tahapan yang digambarkan melalui diagram alur pada Gambar 3.1 di bawah ini. Setiap tahapan akan dijelaskan secara rinci sebagai berikut:



Gambar 3. 1 Diagram Alur Penelitian

Gambar diatas merupakan diagram alur penelitian untuk pengolahan data ulasan aplikasi YUMMY menggunakan pendekatan *machine learning naïve bayes*, *support vector machine* dan *bidirectional encorder representation from transformers*. Tahapan dalam alur penelitian ini adalah sebagi berikut:

3.2.1. Mulai

merupakan awal proses penelitian

3.2.2. Pengumpulan Data

Proses ini bertujuan untuk memperoleh data ulasan pengguna dari aplikasi Yummy. Dataset berupa teks ulasan yang telah diberi label sebagai ulasan positif atau negatif. Dataset ini akan digunakan sebagai data masukan untuk proses selanjutnya. Sumber Data berupa ulasan aplikasi Yummy dalam format CSV

3.2.3. Preprocessing Data

Tahapan ini bertujuan untuk mempersiapkan data agar siap digunakan oleh algoritma. Langkah-langkah *preprocessing* meliputi:

1. Pembersihan Data: Menghapus karakter khusus, tanda baca, URL, dan angka yang tidak relevan.
2. Tokenisasi: Memecah teks menjadi kata-kata individual.
3. Stopword Removal: Menghapus kata-kata umum yang tidak memiliki makna signifikan seperti “*dan*”, “*ke*”, “*di*”, dan lainnya.
4. Stemming: Mengubah kata menjadi bentuk dasarnya menggunakan algoritma *Porter Stemmer* atau *Sastrawi*.
5. Tools yang digunakan dalam tahapan ini adalah Python dengan pustaka NLP seperti NLTK, SpaCy, dan Sastrawi.

3.2.4. Pemodelan Data

Pemodelan Data Pada tahap ini, data yang telah diproses akan digunakan sebagai masukan untuk tiga metode klasifikasi, yaitu:

1. Model Naive Bayes
Melatih model Naive Bayes menggunakan representasi TF-IDF.
2. Model SVM
Melatih model SVM dengan kernel linear dan representasi TF-IDF.
3. Model BERT
Melakukan fine-tuning pada model BERT menggunakan data ulasan Yummy, dengan tokenisasi berbasis tokenizer BERT.

Setiap metode akan diimplementasikan dengan menggunakan pustaka *scikit-learn*, *TensorFlow*, dan *HuggingFace*.

3.2.5. Validasi Model

Validasi model dilakukan dengan pengujian tiap metode untuk mengevaluasi performa dari ketiga metode menggunakan beberapa metrik evaluasi, yaitu:

1. Akurasi: Mengukur persentase prediksi yang benar.
2. Precision: Menilai ketepatan prediksi positif terhadap data positif yang sebenarnya.
3. Recall: Mengukur sejauh mana model dapat mendeteksi ulasan positif.
4. F1-Score: Rata-rata harmonis antara precision dan recall.

3.2.6. Komparasi model

Bertujuan untuk mengumpulkan hasil dari ketiga metode (Naive Bayes, SVM, dan BERT). Lalu membandingkan kinerja berdasarkan metrik evaluasi dan waktu komputasi. Yang dapat memberikan wawasan efektivitas pada tiap modelnya dengan menggunakan dataset ulasan yummy.

3.2.7. Analisis Hasil

Mengamati kinerja metode yang dibandingkan untuk memberikan penjelasan mendalam tentang kelebihan dan kekurangan masing-masing model dalam studi kasus ulasan aplikasi YUMMY serta menyimpulkan metode yang paling optimal dan memberikan rekomendasi untuk implementasi praktis.