

BAB III

METODOLOGI PENELITIAN

3.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data primer. Data primer adalah informasi yang diperoleh langsung dari sumber datanya (Faisal et al., 2020). Data ulasan Google Classroom berasal dari situs Google Play Store dengan metode *web scraping* yang dilakukan dengan menggunakan *library* dari *python* yaitu *google-play-scraper* sebanyak 10.000 data ulasan tekstual bahasa inggris (Sunyoto, Setiabudi and Tjondrowiguno, n.d.). Berdasarkan 10.000 data ulasan tersebut kemudian diurutkan sehingga diperoleh kurun waktu data ulasan yang digunakan yaitu 1 Februari 2022 hingga 6 September 2022.

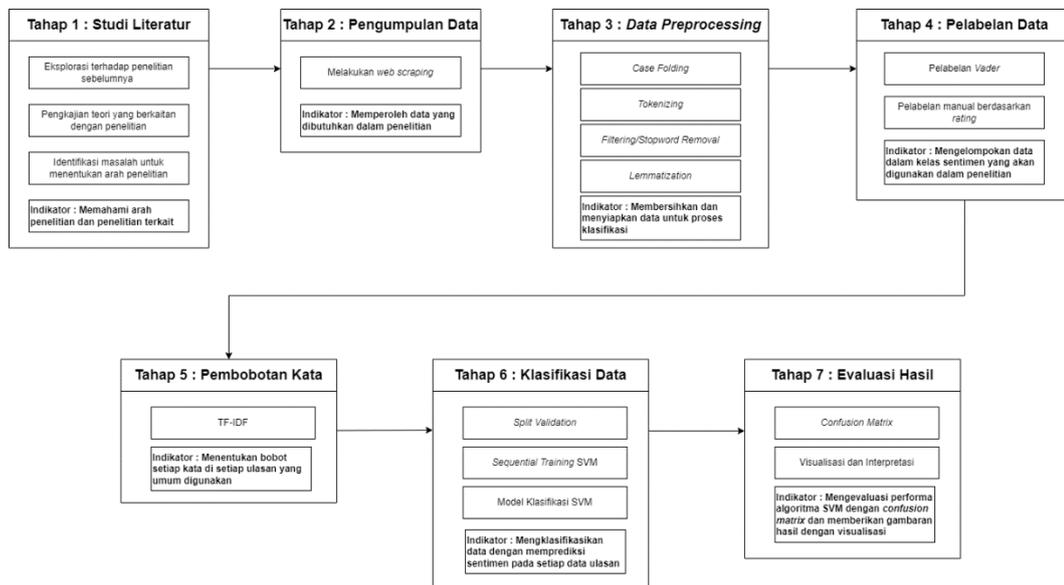
3.2 Variabel Penelitian

Variabel penelitian merupakan segala hal yang peneliti pilih untuk dipelajari dengan cara apapun dalam mengumpulkan data untuk membuat kesimpulan (Moto, 2019). Variabel yang digunakan pada penelitian ini terdiri dari 4 bagian, yaitu:

1. *Username* merupakan nama pengguna yang memberikan ulasan.
2. *Date* merupakan tanggal dibuatnya ulasan.
3. *Rating* merupakan tingkat kepuasan pengguna.
4. *Review* merupakan isi ulasan yang dibuat oleh pengguna.

3.3 Tahapan Penelitian

Berikut merupakan tahapan dari penelitian yang dimulai dari studi literatur, pengumpulan data, *data preprocessing*, pelabelan data, pembobotan kata, klasifikasi data dan evaluasi hasil seperti yang tertera pada Gambar 3.1



Gambar 3.1 Tahapan Penelitian

Berdasarkan Gambar 3.1 tahapan penelitian yang dilakukan dijelaskan sebagai berikut.

1. Studi Literatur

Studi literatur dilakukan dengan melakukan literatur terhadap beberapa penelitian terkait dengan mengkaji dari beberapa sumber seperti buku, jurnal, serta laporan penelitian yang berkaitan dengan ulasan Google Classroom, analisis sentimen, algoritma *Support Vector Machine* dan pelabelan VADER. Tahap ini bertujuan untuk mengidentifikasi permasalahan penelitian sehingga dapat terlihat arah penelitian yang dilakukan.

2. Pengumpulan Data

Pengumpulan data dilakukan dengan teknik *web scraping* pada situs Google Play Store. Pengumpulan data bertujuan untuk memperoleh informasi yang dibutuhkan dalam rangka mencapai tujuan penelitian.

3. *Data Preprocessing*

Data preprocessing dilakukan untuk membersihkan dan menyiapkan data ulasan untuk proses klasifikasi. *Data Preprocessing* dilakukan untuk membersihkan beberapa proses diantaranya *Case Folding*, *Tokenizing*, *Filtering/Stopword Removal* dan *Lemmatization*.

4. Pelabelan Data

Pelabelan data dilakukan untuk memberikan label pada setiap data ulasan dengan kelas sentimen. Data ulasan dapat memiliki kelas sentimen positif atau negatif. Metode pelabelan yang digunakan yaitu metode pelabelan VADER.

5. Pembobotan Kata

Pembobotan kata dilakukan untuk menentukan bobot kata pada setiap data ulasan. Nilai bobot suatu kata ditentukan dengan seberapa sering kata muncul dalam setiap ulasan. Metode yang digunakan untuk pembobotan kata yaitu metode ekstraksi fitur TF-IDF.

6. Klasifikasi Data

Klasifikasi data meliputi aktivitas *Split Validation* untuk membagi data menjadi dua bagian yaitu data latih dan data uji. Data latih digunakan untuk membangun dan melatih model klasifikasi melalui proses *Sequential Training*.

Data diuji digunakan untuk menguji model klasifikasi untuk memperoleh performa atau tingkat akurasi algoritma SVM.

7. Evaluasi Hasil

Evaluasi hasil meliputi proses menghitung nilai akurasi, presisi dan *recall* dari algoritma SVM menggunakan *confusion matrix* dan memberikan gambaran hasil klasifikasi sentimen dengan visualisasi dan interpretasi.