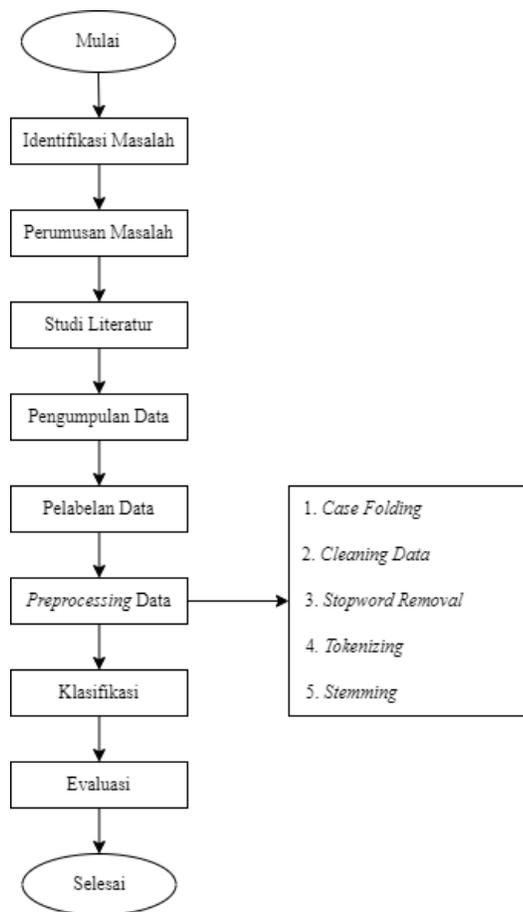


BAB III

METODOLOGI PENELITIAN

3.1 Kerangka Berpikir

Pada bab ini, memberikan penjelasan tentang langkah-langkah dan metode penulisan yang digunakan dalam penelitian ini. Dalam bab ini memberikan gambaran alur perancangan kerangka pikir yang akan digunakan sebagai acuan selama tugas akhir. Berikut adalah tahapan penelitian terdapat pada gambar 3.1.



Gambar 3. 1 Kerangka Berpikir

3.2 Identifikasi Masalah

Pada penelitian ini, hal pertama yang dilakukan peneliti adalah mengidentifikasi masalah. Dengan menggunakan metode *Decision Tree* berbasis SMOTE, penelitian ini menyelidiki *rating* pengguna layanan aplikasi Shopee di *Google Play Store*. Hal ini menjadi masalah karena setiap pengguna memiliki tingkat kepuasan yang berbeda dari aplikasi tersebut karena berbagai alasan.

3.3 Perumusan Masalah

Pada identifikasi masalah sebelumnya, perumusan masalah yaitu bagaimana menentukan hasil dari *rating* pengguna layanan aplikasi Shopee di *Google Play Store* dengan metode *Decision Tree* berbasis SMOTE.

3.4 Studi Literatur

Pada tahap studi literatur, diperlukan referensi yang relevan dengan permasalahan yang sedang diteliti dengan mencari sumber-sumber yang tersedia. Studi literatur dapat menjadi panduan untuk menemukan dan menyelesaikan masalah.

3.4 Pengumpulan Data

Pada penelitian ini, metode *web scraping* digunakan untuk mengumpulkan data dari *website* menggunakan Bahasa pemrograman *Python* di *Google Colab* dan mengumpulkan data tentang *rating* layanan aplikasi Shopee di *Google Play Store*. Tahap ini menghasilkan kumpulan data dalam bentuk CSV, yang akan digunakan untuk proses selanjutnya. Jumlah data yang digunakan sebanyak 8229 data.

3.6 Pelabelan Data

Hasil dari *scrapping* data disimpan dalam *file* bertipe CSV dan kemudian dilakukan pelabelan data untuk menentukan pendapat atau pandangan dari komentar yang diambil. Pada proses pelabelan data dibedakan menjadi dua kelas, yaitu kelas positif dan kelas negatif.

Pembagian pelabelan data dimulai dari *rating* bintang empat sampai dengan *rating* bintang lima diberikan label positif, *rating* bintang tiga diberikan label netral, dan *rating* bintang satu sampai dengan *rating* bintang dua diberikan label negatif.

Sebanyak 938 data *rating* bintang tiga dalam *dataset rating* aplikasi Shopee. Dalam penelitian ini, tidak menggunakan data label netral.

3.7 Preprocessing Data

Preprocessing data adalah tahap awal pemrosesan data teks komentar, dengan tujuan menghasilkan kalimat yang lebih sederhana yang mengandung sentimen. Tahap ini melibatkan pemilihan dan penghapusan kata-kata yang tidak diperlukan dari data teks komentar, yang berarti bahwa elemen tertentu dari data harus dihapus agar data dapat disaring dengan benar.

3.7.1 Case Folding

Proses *preprocessing* teks terdiri dari beberapa tahapan penting. Salah satu tahapan yang paling penting adalah *case folding*, di mana teks dilipat menjadi huruf kecil untuk memberinya bentuk standar. Selain itu, untuk membersihkan teks, angka dan simbol seperti tanda seru (!), koma (,), dan lainnya dihapus

3.7.2 Cleaning

Tahap *cleaning* melibatkan menghapus karakter tertentu dari teks, seperti *hashtag*, URL, *emoticon*, *username*, dan elemen lain yang dianggap tidak penting atau mengganggu saat menganalisis atau memproses teks.

3.7.3 Stopword Removal

Pada tahap ini, dilakukan proses penyaringan dan pembersihan teks dengan menghilangkan elemen-elemen yang tidak memiliki nilai semantik atau tidak relevan untuk analisis. Fokusnya adalah mengeliminasi kata-kata, simbol, atau karakter yang tidak berkontribusi pada makna inti dari teks.

3.7.4 Tokenizing

Tokenizing adalah proses yang memecah teks sebuah dokumen menjadi kata-kata yang berbeda, juga dikenal sebagai "token", dengan menggunakan spasi dan tanda baca untuk memisahkan kata-kata. Tujuannya adalah agar setiap kata dapat diperlakukan secara mandiri tanpa ketergantungan pada kata lainnya (Nata dan Yudiasra, 2017).

3.7.5 Stemming

Penguraian kata menjadi bentuk dasarnya dikenal sebagai *stemming*. Imbuhan kata awalan atau akhiran dihapus selama tahap *stemming*, sehingga hanya inti kata yang tersisa.

3.8 Klasifikasi

Pada tahapan ini, data akan dilakukan pembagian data menjadi data *training* dan data *testing*. Selanjutnya data akan melewati proses pembobotan kata menggunakan *CountVectorizer*. Setelah melakukan pembagian data dan pembobotan kata, dilanjutkan dengan pemodelan data menggunakan metode *Decision Tree* dan *Decision Tree* berbasis SMOTE.

3.9 Evaluasi

Pada tahap evaluasi dilakukan pengujian perbandingan metode *Decision Tree* tanpa smote dengan metode *Decision Tree* berbasis SMOTE. Pada tahap ini menggunakan *confusion matrix* yang berfungsi untuk mengevaluasi dan melihat hasil *accuracy*, *precision*, *recall*, dan *f1 score*.

3.10 Perangkat Pendukung Penelitian

Perangkat yang digunakan dalam analisis sentimen ulasan pengguna layanan aplikasi Shopee, yaitu:

1. Perangkat Keras

Perangkat yang digunakan dalam melakukan adalah laptop Asus Vivobook 14 dengan spesifikasi:

Processor : CPU Intel Core i3 1005G1, Intel core i5 1035G.

RAM : 4GB.

2. Perangkat Lunak

Perangkat lunak yang dipakai adalah:

- Sistem Operasi: Windows 10 Pro
- Software yang digunakan dalam penelitian ini:
 - Microsoft Word 2016

- Microsoft Excel 2016
- Browser Google Chrome (Google Colab)