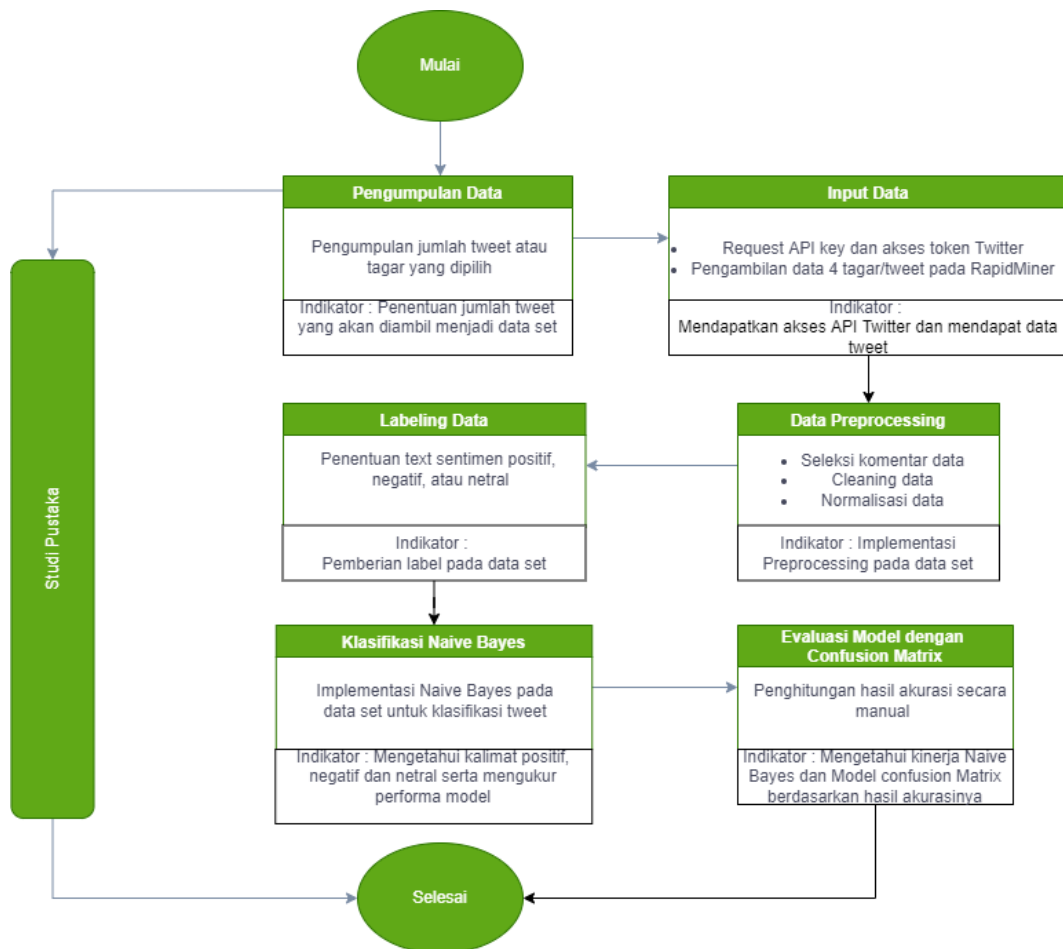


BAB III

METODOLOGI PENELITIAN

3.1 Metodologi Penelitian

Metodologi penelitian pada penelitian ini terdiri dari studi pustaka, pengumpulan data, input data, Data *pre-processing*, *Labeling data*, Klasifikasi dengan *Naive Bayes*, evaluasi model dengan *Confusion Matrix* yang ditunjukkan pada Gambar 3.1.



Gambar 3.1 Metodologi Penelitian

3.2. Studi Pustaka

Pada penelitian ini studi pustaka yang dilakukan diantaranya mengenai konsep *Data Mining*, *Naive Bayes*, dan *tools* yang digunakan dalam proses analisis sentimen.

3.3 Pengumpulan Data

Pengumpulan data dilakukan dengan studi literatur dan praktik untuk mendapatkan dataset. Data yang digunakan dalam penelitian ini diperoleh dengan input data dari media sosial Twitter. Spesifik yang dimaksud pada input data adalah data Twitter yang berisi kalimat opini dari masyarakat akan diambil secara acak dengan menggunakan metode pengumpulan *tweets* manual yaitu memasukkan berbagai macam kata kunci yang berhubungan dengan PSBB di DKI Jakarta.

3.4 Input Data

Input data dari Twitter dilakukan pengambilan secara acak dengan menggunakan metode pengumpulan *tweets* manual yaitu memasukkan berbagai macam kata kunci yang berhubungan dengan PSBB di Indonesia pada kolom pencarian Twitter dengan menggunakan *hashtag*.

3.5 Preprocessing Data

Terdapat beberapa metode yang digunakan pada tahap *preprocessing* seperti :

1. Seleksi Komentar Data yang telah dikumpulkan pada tahap pengumpulan data ternyata ada data yang berulang dikarenakan adanya *re-tweet* tanpa diberikan komentar.

2. *Cleaning* adalah proses pembersihan data teks yaitu dengan menghilangkan data yang tidak konsisten atau tidak relevan.
3. Normalisasi Data digunakan untuk menormalisasikan kalimat. Banyak pengguna dalam menuliskan kata menggunakan singkatan sehingga akan menghambat proses pemahaman teks, seperti contoh “BGT” jika dinormalisasi menjadi “BANGET”, kata perulangan “gara2” menjadi “gara-gara”.

3.6 Labeling Data

Proses labeling data dilakukan untuk menentukan sebuah dokumen teks mengandung sentimen positif, netral atau negatif.

3.7 Klasifikasi Dengan *Naïve Bayes*

Klasifikasi dengan metode *Naïve Bayes* dilakukan dengan menggunakan *tools* RapidMiner untuk mengetahui besar akurasi pembelajaran dengan *Naïve Bayes*.

3.8 Evaluasi Model

Untuk mengidentifikasi pola-pola menarik yang ditemukan dalam hal ini adalah tentang reaksi masyarakat tentang adanya PSBB di DKI Jakarta.