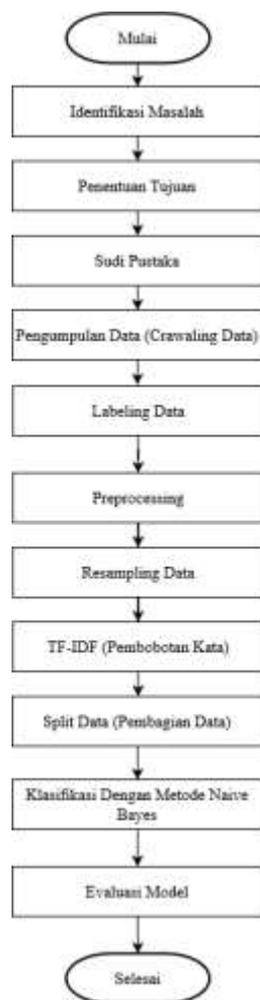


BAB III

METODOLOGI PENELITIAN

Pada bab ini, akan dijelaskan tentang metodologi penelitian Analisis sentimen masyarakat terhadap polri tentang kasus penembakan brigadir J menggunakan Naïve Bayes. Metodologi penelitian digunakan untuk acuan atau pedoman dalam agenda penelitian yang akan dilakukan agar penulis dapat melakukan penelitian secara terstruktur dan dapat menyelesaikan penelitian tepat pada waktunya, juga agar penelitian dapat berjalan sesuai dengan yang diharapkan. Metodologi penelitian dalam penelitian ini disajikan pada Gambar 3.1.



Gambar 3. 1 Metodologi Penelitian

Berikut ini adalah penjelasan langkah-langkah metodologi penelitian pada tugas akhir ini.

3.1 Identifikasi Masalah

Mengamati dan menemukan permasalahan yang terjadi tentang kasus yang sedang viral yaitu penembakan brigadir J yang menjadi sorotan bagi masyarakat. Oleh karena itu dilakukan analisis sentimen masyarakat terhadap penembakan brigadir J menggunakan metode Naïve Bayes.

3.2 Penentuan Tujuan

Penentuan tujuan berfungsi untuk memperjelas kerangka tentang apa saja yang menjadi sasaran dari penelitian ini. Tujuan dari penelitian ini adalah mengetahui pengaruh kasus penembakan brigadir J terhadap tingkat kepercayaan masyarakat kepada Polri dengan cara melakukan analisis sentimen menggunakan metode Naïve Bayes.

3.3 Studi Pustaka

Studi Pustaka merupakan langkah awal dalam metode pengumpulan data, studi pustaka merupakan metode pengumpulan data yang di arahkan kepada pencarian data dan informasi melalui dokumen dokumen baik tertulis, foto-foto, gambar, maupun dokumen elektronik yang dapat mendukung dalam proses penulisan. Diantaranya mengenai konsep, *NAÏVE BAYES* dan *tools* yang digunakan untuk proses *analisis sentimen*, studi pustaka merupakan dapat memengaruhi kredibilitas hasil penelitian yang dilakukan.

3.4 Pengumpulan Data (Crawling Data)

Sebagain umum di artikan sebagai proses, cara atau produser yang digunakan untuk memecahkan suatu masalah. Metode yang digunakan untuk mengumpulkan data dalam penelitian ini di input dari media sosial berupa Twitter. Untuk data yang di cari atau spesifikasi nya adalah data yang terdiri dari komentar masyarakat tentang Analisis sentimen masyarakat terhadap polri tentang kasus brigadir J menggunakan naïve bayes. Pengumpulan data dengan menginput hastag (#ferdisambo dan #brigadirj). Data diambil menggunakan teknik crawling.

Kategori data yang pakai pada penelitian ini dibagi menjadi 3 kategori (*class*) *tweet* yaitu kategori *tweet* positif, kategori *tweet* negatif dan kategori *tweet* netral.

3.5 Labeling Data

Labeling data dilakukan setelah data sudah dikumpulkan melalui teknik crawling. Data tersebut akan dimasukan ke tahapan selanjutnya yaitu pemberian label secara manual yang terdiri dari label positif, negatif, dan netral. Pelabelan manual dilakukan di microsoft excel.

3.6 Preprocessing

Tahap preprocessing merupakan tahap awal untuk mengolah data teks menjadi analisa sentimen. Tahap Preprocessing dilakukan pada jupyter notebook dengan bahasa pemograman python. Terdapat beberapa metode yang digunakan pada tahap preprocessing yaitu *cleansing*, *tokenizing* dan *case folding*. *Cleansing* adalah pembersihan data, *tokenizing* merupakan pemisahan dokumen menjadi kata dan terakhir adalah *case folding* yang merupakan tahap mengubah karakter huruf pada text menjadi kecil.

3.7 Resampling Data

Pada tahap ini, data yang sudah melewati tahap preprocessing akan diproses dengan membuat data baru dari data atau populasi yang sudah ada pada dataset. Keunggulan menggunakan teknik resampling adalah membuat dataset menjadi lebih bagus sehingga meningkatkan akurasi data dalam pembentukan model klasifikasi.

3.8 TF-IDF (Pembobotan Kata)

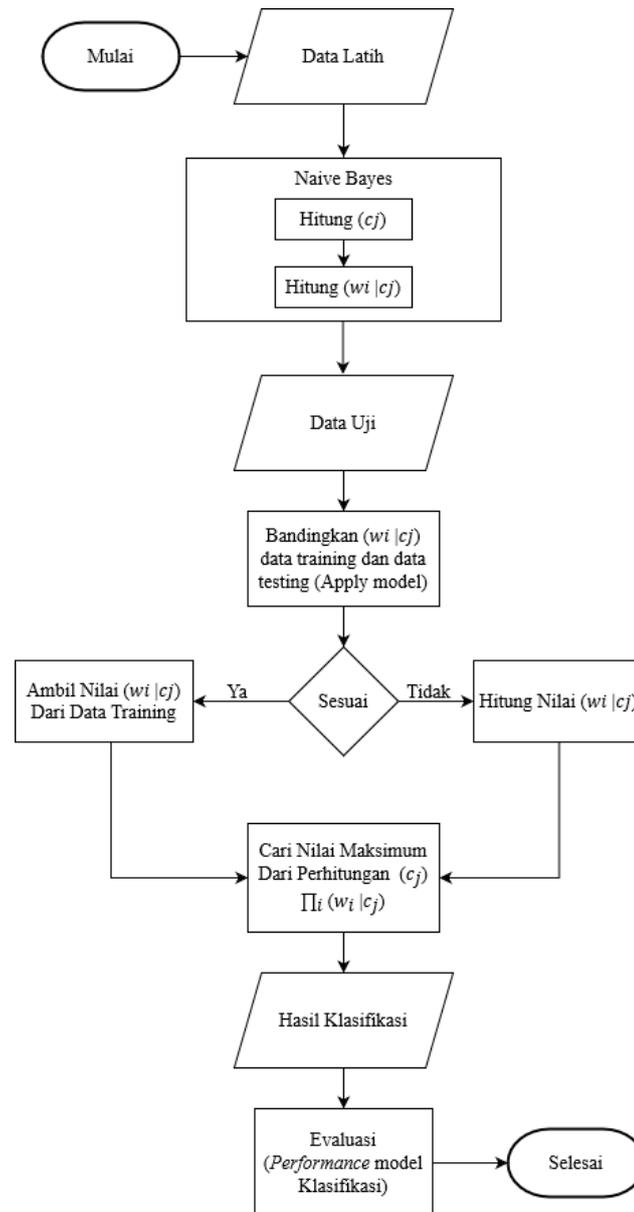
Tahap selanjutnya pada penelitian ini yaitu pemecahan kalimat menjadi beberapa kata serta memberikan bobot pada setiap kata dengan menggunakan TF-IDF. Tahap ini merupakan Statistic Numeric yang digunakan untuk mengungkapkan tingkat kepentingan sebuah kata untuk dokumen dalam suatu koleksi.

3.9 Split Data (Pembagian Data)

Split data atau pembagian data merupakan proses dimana dataset yang sudah ada akan dibagi menjadi data latih dan data uji. Data latih merupakan data yang digunakan untuk membuat model machine learning. Dalam proses pembuatan model, data latih diinputkan ke dalam algoritma machine learning untuk mempelajari pola dan hubungan antara input dan output. Sedangkan data uji merupakan data yang akan diuji untuk dilihat bagaimana performa atau akurasi dari model yang sudah dibuat. Pembagian data pada penelitian ini adalah dengan rasio 80% untuk data latih dan 20% untuk data uji.

3.10 Klasifikasi Naïve Bayes

Klasifikasi dilakukan dengan menggunakan metode *Naïve Bayes*. Metode ini di sebut simple dan efektif untuk di terapkan dalam analisis sentimen, dengan menggunakan bahasa pemograman python pada jupyter notebook untuk mengetahui besar akurasi pembelajaran dengan naïve bayes, untuk tahapannya data lebih dulu di bagi menjadi 80% untuk data latih dan 20% untuk data uji. Flowchart klasifikasi dengan metode Naïve Bayes dapat dilihat pada Gambar 3.2.



Gambar 3. 2 Flowchart Metode *Naive Bayes*

3.11 Evaluasi Model

Evaluasi model merupakan untuk desain evaluasi yang di kembangkan oleh para ahli evaluasi, yang biasanya dinamakan sama dengan pembuatannya atau tahap evaluasinya, dan untuk mengidentifikasi hal-hal yang menarik untuk di temukan tentang reaksi masyarakat tentang adanya tingkat kepercayaan terhadap Polri.