

BAB 3

METODE PENELITIAN

Penelitian dengan cara menganalisis tentang sentimen positif serta negatif yang ada pada media sosial X. Penelitian ini mengambil data dari media sosial X yang berisi *tweet* berkaitan dengan program makan siang gratis menggunakan metode *Naïve Bayes Classification*. Setelah mendapatkan data dari media sosial X maka langkah selanjutnya adalah *preprocessing data* untuk menemukan hasil yang diharapkan. Kumpulan data yang diperoleh dari *tweet* pengguna media sosial X terkait program makan siang gratis akan menghasilkan informasi. Sehingga informasi yang didapat masih relevan mengenai program makan siang gratis.

Berikut ini adalah bahan penelitian yang dimiliki, alat penelitian yang harus digunakan, serta jalan penelitian yang dibuat untuk melakukan analisis sentimen tentang pasangan calon presiden dan wakil presiden terkait program makan siang gratis adalah sebagai berikut:

3.1 BAHAN PENELITIAN

Bahan penelitian diperoleh dari *tweet*, *re-tweet*, serta komentar di media sosial X yang masih berkaitan dengan program unggulan pada kampanye yang dilakukan. Data yang dikumpulkan dari 28 November 2023 hingga 10 Februari 2024 dari media sosial X berupa *tweet* terkait program unggulan yang ditawarkan dengan kata kunci “makan siang gratis” berhasil mendapatkan data *tweet* sebanyak 400 data.

3.2 ALAT PENELITIAN

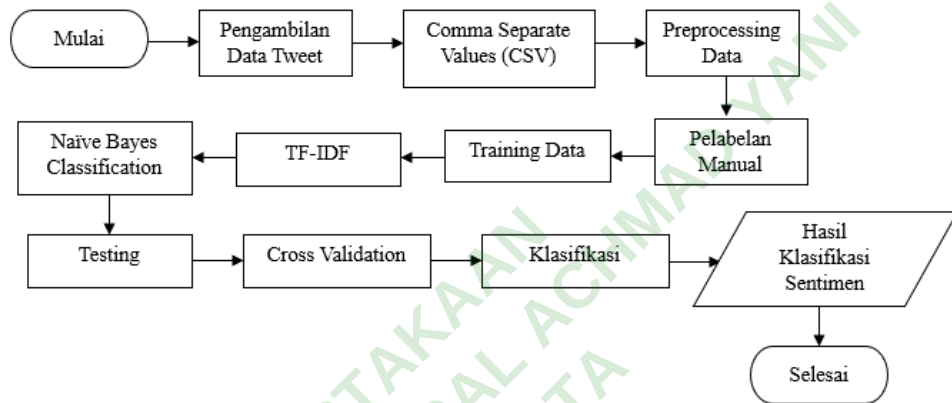
Komputer yang ada harus memiliki koneksi internet yang stabil agar dapat menjalankan sistem operasi dan program aplikasi untuk melakukan analisis merupakan alat yang digunakan pada penelitian ini. Alat-alat tersebut meliputi:

1. Sistem Operasi Windows 11
2. Google Colaboratory
3. Library Tweet Harvest

4. Microsoft Office Excel 2021

3.3 JALAN PENELITIAN

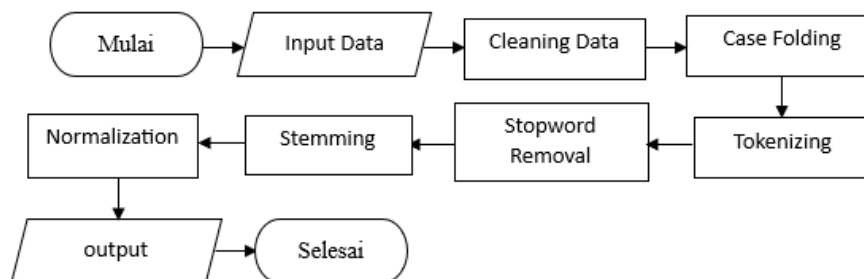
Jalan penelitian ini menggunakan *Google Colaboratory* untuk mengolah data *tweet* mengenai program makan siang gratis dari website X. Kemudian ditampilkan pada *Microsoft Office Excel*, seperti yang tertera pada Gambar 3.1 yang menjelaskan alur dari penelitian ini.



Gambar 3. 1. Flowchart Alur Penelitian

Penelitian ini menggunakan pelabelan manual yang dilakukan dengan cara menandai secara manual suatu kata pada dokumen. Merujuk pada penelitian yang dilakukan pada tahun 2022 lalu oleh Tanggraeni et al. bahwa pemberian nilai 1 untuk label yang bersifat positif sedangkan nilai 0 untuk label yang bersifat negatif.

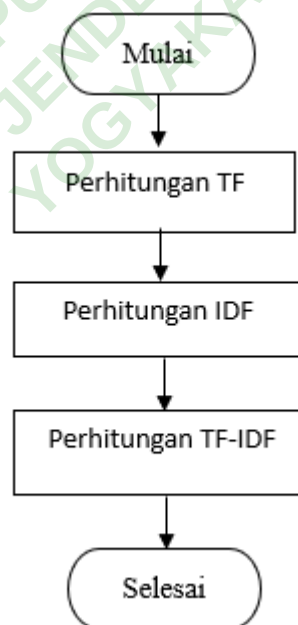
Cleaning data, case folding, tokenizing, stopword removal, stemming, serta normalization merupakan subproses dari *preprocessing data* seperti yang ditampilkan pada Gambar 3.2 di bawah ini:



Gambar 3.2 Flowchart Preprocessing Data

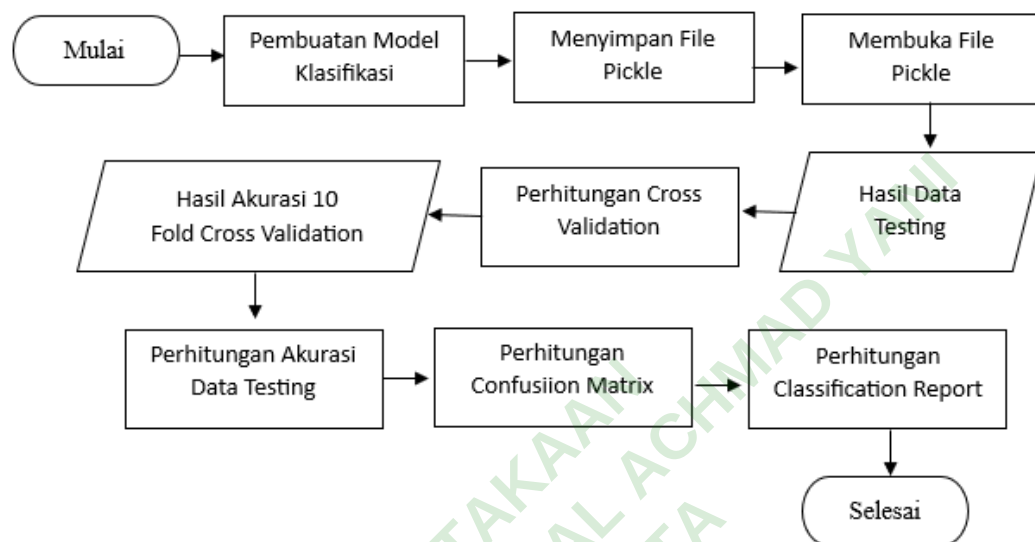
Untuk memulai tahapan *preprocessing data*, data yang sudah didapatkan akan diinput untuk menjalankan proses selanjutnya yaitu *cleaning data* yang bertujuan menghilangkan link url, tanda baca serta angka yang tidak memiliki arti pada data yang diambil. *Case folding* dikenal dengan teknik mengubah huruf kapital menjadi huruf kecil. Kemudian *tokenizing* yaitu proses yang dilewati untuk membagi teks menjadi beberapa bagian yang dikenal dengan *token*. Selanjutnya teknik menghapus kata-kata umum dari teks yang tidak memiliki informasi penting dikenal sebagai *stopword removal*, dilanjut dengan *stemming* yang bertujuan untuk mengurangi kata-kata ke bentuk dasarnya sehingga variasi kata yang terdapat kesamaan dianggap sebagai kata yang sama dalam analisis teks. Terakhir yaitu tahapan *normalization* yang bertujuan untuk menyeragamkan *term* yang terdapat kesalahan pada penulisan untuk diubah menjadi bentuk standar sehingga dapat dilakukan analisis.

Subproses perhitungan TF-IDF masuk ke dalam proses *training data* seperti yang ditampilkan pada Gambar 3.3:



Gambar 3.3 Flowchart TF-IDF

Terdapat Gambar 3.4 yang menunjukkan tahap perhitungan dengan metode *Naïve Bayes Classification*.



Gambar 3.4 Flowchart Perhitungan Naïve Bayes Classification