

BAB III

METODOLOGI PENELITIAN

3.1 Subjek dan Objek Penelitian

Subjek dalam penelitian ini adalah menyelidiki pendapat masyarakat tentang isu resesi tahun 2023 di Indonesia melalui media sosial YouTube. Objek penelitian adalah persepsi masyarakat mengenai isu resesi tahun 2023 di Indonesia berdasarkan data komentar pada berita isu resesi 2023 yang diunggah oleh *channel* YouTube CNN Indonesia pada tanggal 10 Oktober 2022 dan diambil dari tanggal 10 Oktober 2022 sampai 27 Oktober 2022.

3.2 Alat dan Bahan Penelitian

Alat :

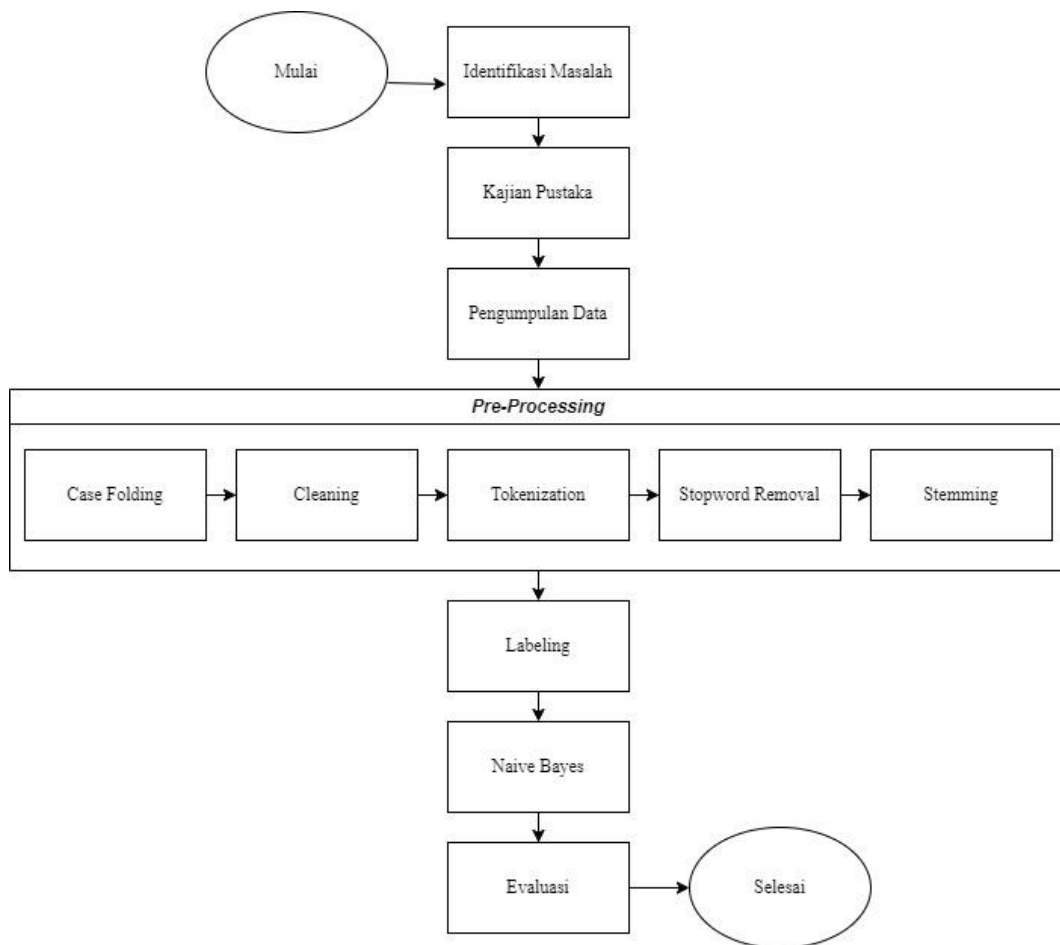
1. Laptop (CPU Intel i7 gen 9, RAM 8GB, GPU GTX 1660Ti)
2. Notebook Google Colab
3. Google Spreadsheet

Bahan :

1. Data komentar YouTube

3.3 Diagram Alir Penelitian

Diagram alir merupakan serangkaian langkah yang dilakukan dalam proses penelitian dengan tujuan sesuai dari tahap awal perencanaan penelitian dibuat. Gambar 3.1 akan menunjukkan diagram alir penelitian.



Gambar 3.1 Diagram Alir Penelitian

3.3.1 Identifikasi Masalah

Tahap identifikasi masalah pada penelitian ini dilakukan karena isu resesi yang akan terjadi di Indonesia tahun 2023. Topik ini diambil dari sumber terpercaya yaitu CNN Indonesia, selanjutnya peneliti menentukan tujuan dan manfaat dari penelitian ini.

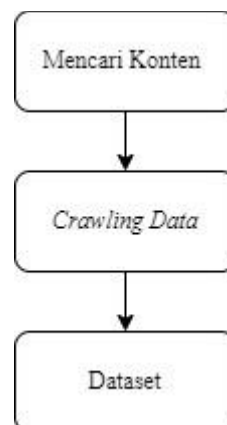
3.3.2 Kajian Pustaka

Peneliti mencari referensi dari penelitian sebelumnya yang relevan sebagai pedoman dalam penyusunan proposal. Referensi tersebut dicari

melalui jurnal, artikel. Tujuan dikumpulkannya referensi ini adalah untuk memperkuat permasalahan yang akan diangkat.

3.3.3 Pengumpulan Data

Melakukan pengambilan data yang nantinya akan digunakan dalam penelitian. Peneliti menggunakan data komentar dari *channel* YouTube CNN Indonesia yang diunggah pada tanggal 10 Oktober 2022 dan data diambil dari tanggal 10 Oktober 2022 sampai 27 Oktober 2022. Proses pengumpulan data dilakukan menggunakan bantuan *tools* Google Spreadsheet dan menggunakan ekstensi Apps Script.



Gambar 3.2 Diagram Pengumpulan Data

Dalam penelitian ini, peneliti mengumpulkan data seperti pada Gambar 3.2. Peneliti mencari berita mengenai isu resesi 2023 pada media YouTube dan mendapatkan berita dari *channel* CNN Indonesia. Selanjutnya peneliti mendapatkan alamat dari konten tersebut dan mendapatkan API untuk tahap *crawling data* guna mendapatkan data komentar YouTube menggunakan API YouTube dari layanan Google Spreadsheet dengan bantuan Apps Script. Kemudian peneliti mendapatkan file yang telah didapatkan kedalam satu file dengan format .csv.

3.3.4 Pre-Processing

Setelah melakukan pengumpulan data, Langkah selanjutnya adalah *pre-processing* data.

3.3.4.1 Case Folding

Case folding berguna untuk mengubah semua huruf dalam sebuah dokumen teks menjadi huruf kecil atau *lowercase*. Contoh *case folding* ada pada Tabel 3.1.

Tabel 3.1 Tabel *Case Folding*

Sebelum <i>Case Folding</i>	Sesudah <i>Case Folding</i>
Semoga aja gak terjadi ya Allah semoga yang terjadi lebih baik dari yang dipikirkan	semoga aja gak terjadi ya allah semoga yang terjadi lebih baik dari yang dipikirkan
INSYAALAH TERANG BENDERANG	insyaalah terang benderang

3.3.4.2 Pembersihan Data

Pada proses ini, peneliti menghapus atribut yang tidak dibutuhkan dan tidak memiliki makna, seperti *hashtag*, *mention*, *retweet*, *whitespace*, dll. Berikut Tabel 3.2 adalah contoh pembersihan data pada salah satu dataset.

Tabel 3.2 Tabel Pembersihan Data

Sebelum <i>Cleaning</i>	Sesudah <i>Cleaning</i>
I N F bukan ay en ef... Indonesia apa alien sih	I n f bukan ay en ef Indonesia apa alien sih
PREDIKSI PERTENGAHAN 2025 ..	prediksi pertengahan

3.3.4.3 Tokenisasi

Proses *tokenization* adalah proses untuk memecah dokumen tersebut menjadi bagian-bagian kata yang disebut token. Tabel 3.3 menjelaskan bagaimana tokenisasi bekerja.

Tabel 3.3 Tabel Tokenisasi

Sebelum Tokenisasi	Sesudah Tokenisasi
Semoga aja gak terjadi ya Allah semoga yang terjadi lebih baik dari yang dipikirkan	[semoga, aja, gak, terjadi, ya, allah, semoga, yang, terjadi, lebih, baik, dari, yang, dipikirkan]
gak ngerti gw mksd mba nya ini membuat berita	[gak, ngerti, gw, mksd, mba, nya, ini, membuat, berita]

3.3.4.4 Stopword Removal

Proses *stopword removal* ini adalah mengambil kata-kata yang dianggap penting dari hasil *tokenization* atau membuang kata-kata yang dianggap tidak terlalu mempunyai arti penting. Tabel 3.4 merupakan contoh dari *stopword removal*.

Tabel 3.4 Tabel *Stopword*

Sebelum <i>Stopword</i>	Setelah <i>Stopword</i>
[semoga, aja, gak, terjadi, ya, allah, semoga, yang, terjadi, lebih, baik, dari, yang, dipikirkan]	['semoga', 'terjadi', 'allah', ['semoga', 'terjadi', 'baik', 'dipikirkan']]
[gak, ngerti, gw, mksd, mba, nya, ini, membuat, berita]	['ngerti', 'gw', 'mksd', 'mba', 'berita']

3.3.4.5 Stemming

Tahapan selanjutnya adalah *stemming* untuk mentransformasikan kata menjadi kata dasarnya atau *root word*

dengan menghilangkan semua imbuhan kata. Tabel 3.5 adalah contoh dari penggunaan *stemming*.

Tabel 3.5 Tabel *Stemming*

Sebelum <i>Stemming</i>	Setelah <i>Stemming</i>
['semoga', 'terjadi', 'allah', ['semoga', 'terjadi', 'baik', 'dipikirkan']	['semoga', 'terjadi', 'allah', ['semoga', 'terjadi', 'baik', 'pikir']

3.3.5 Labeling

Dalam klasifikasi sentimen, ada satu hal yang paling penting yaitu proses pelabelan atau labeling. Proses pelabelan adalah proses pemberian label pada dataset ulasan. Tabel 3.6 merupakan contoh dari labeling.

Tabel 3.6 Tabel Labeling

Komentar	Labeling
cinta produk indonesia penuh ges	Positif
bismillah indonesia bertahanbraamiin aamiin ya robbal alaamiin	Netral
bias gelap gilir tengah atas gelap gulita	Negatif

3.3.6 Naïve Bayes

Dalam analisis sentimen *text mining*, Naïve Bayes adalah metode klasifikasi yang memungkinkan klasifikasi, presisi, dan perhitungan data [20]. Contoh perhitungan manual Naïve Bayes:

Dokumen	Komentar	Label
1	cinta produk indonesia penuh ges	Positif
2	agenda elite global	Negatif
3	channel elite global illuminati dajjal	Negatif
4	tani palawan indonesia	Netral
5	tani sulit makan	Netral

Dokumen	Komentar	Label
6	tani ternak jaya	Netral
7	kaum samir elite global	?

- Label Positif

$$P(\text{Positif}) = \frac{N_{\text{positif}}}{\text{jumlah label}} = \frac{1}{6} = 0,167$$

- Label Negatif

$$P(\text{Negatif}) = \frac{N_{\text{negatif}}}{\text{jumlah label}} = \frac{2}{6} = 0,333$$

- Label Netral

$$P(\text{Netral}) = \frac{N_{\text{netral}}}{\text{jumlah label}} = \frac{3}{6} = 0,5$$

- Probabilitas dokumen ke-7

$$P(\text{kaum}|\text{positif}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{5 + 19} = \frac{1}{24} = 0,0416$$

$$P(\text{samir}|\text{positif}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{5 + 19} = \frac{1}{24} = 0,0416$$

$$P(\text{elite}|\text{positif}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{5 + 19} = \frac{1}{24} = 0,0416$$

$$P(\text{global}|\text{positif}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{5 + 19} = \frac{1}{24} = 0,0416$$

$$P(\text{kaum}|\text{negatif}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{8 + 19} = \frac{1}{27} = 0,037$$

$$P(\text{samir}|\text{negatif}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{8 + 19} = \frac{1}{27} = 0,037$$

$$P(\text{elite}|\text{negatif}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{2 + 1}{8 + 19} = \frac{3}{27} = 0,1111$$

$$P(\text{global}|\text{negatif}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{2 + 1}{8 + 19} = \frac{3}{27} = 0,1111$$

$$P(\text{kaum}|\text{netral}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{9 + 19} = \frac{1}{28} = 0,0357$$

$$P(\text{samir}|\text{netral}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{9 + 19} = \frac{1}{28} = 0,0357$$

$$P(\text{elite}|\text{netral}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{9 + 19} = \frac{1}{28} = 0,0357$$

$$P(\text{global}|\text{netral}) = \frac{\text{count}(W_iC) + 1}{\text{count}(c) + |v|} = \frac{0 + 1}{9 + 19} = \frac{1}{28} = 0,0357$$

- Naïve Bayes

$$P(\text{Positif}|d7) = 0,167 * (0,04167 * 0,04167 * 0,04167 * 0,04167) \\ = 0,0278$$

$$P(\text{Negatif}|d7) = 0,333 * (0,037 * 0,037 * 0,1111 * 0,1111) \\ = 5,6269$$

$$P(\text{Netral}|d7) = 0,5 * (0,0357 * 0,0357 * 0,0357 * 0,0357) \\ = 0,0714$$

D7 merupakan kalimat negatif karena nilai probabilitas negatif lebih tinggi dibandingkan probabilitas positif atau netral.

3.3.7 Evaluasi

Tahapan terakhir yaitu melakukan evaluasi data, peneliti melakukan evaluasi untuk mendapatkan *precision*, *recall*, *f1-score*, dan akurasi seperti pada Gambar 3.3.

	precision	recall	f1-score	support
Positive	0.88	0.59	0.71	39
Neutral	0.94	0.97	0.95	397
Negative	0.50	0.56	0.53	18
accuracy			0.92	454
macro avg	0.78	0.70	0.73	454
weighted avg	0.92	0.92	0.92	454

Gambar 3.3 Hasil Evaluasi