

BAB III METODOLOGI PENELITIAN

3.1 Populasi dan Sampel Penelitian

Populasi penelitian berupa masyarakat atau calon pemilih presiden 2024. Sampel penelitian adalah komentar bakal calon presiden pada media sosial TikTok.

3.2 Alat dan Bahan

Perangkat keras dan perangkat lunak yang digunakan untuk membuat analisis sentimen capres dan cawapres 2024 berdasarkan komentar publik TikTok adalah sebagai berikut :

a. Perangkat Keras

- Laptop
- *Processor* intel pentium 3
- *RAM* 4Gb
- *Harddisk* 80Gb

b. Perangkat Lunak

- Sistem operasi *windows* 10 64 bit
- *Visual Code*
- *Xampp*
- *MySQL connector*
- *Python*
- *JDK*
- *Pentaho PDI*
- *Pentaho Designer*

3.3 Sketsa *Input*, *Proses*, dan *Output*: Analisis Sentimen Calon Presiden Indonesia 2024 Berdasarkan Komentar di TikTok

Sistem dibuat tidak lepas dari *input*, *proses*, dan *output*. Adapun sketsa dari *input*, *proses*, dan *output* Analisis Sentimen Calon Presiden Indonesia 2024 Berdasarkan Komentar di TikTok adalah sebagai berikut :

a. *Input*

Data Komentar TikTok terkait Calon Presiden 2024. Data akan diambil menggunakan *crawler* khusus untuk TikTok untuk pengguna akun resmi tvOneNews dan Metro TV.

b. *Proses*

1. Pengumpulan Data

a) *Crawler* akan digunakan untuk mengambil komentar-komentar terkait dengan calon presiden Indonesia 2024 di *platform* TikTok pada akun resmi tvOneNews dan Metro TV dengan #capres 2024 #elektabilitas #lembagasurvey #capres2024 #prabowo #ganjarpranowo #aniesbaswedan #calonpresiden2024 #pemilu2024.

b) Data akan terdiri dari teks komentar, nama pengguna, dan waktu komentar.

2. *Preprocessing* Data

a) Membersihkan dan menghilangkan karakter khusus dan *link* yang tidak diperlukan dari teks komentar.

b) Melakukan normalisasi teks seperti pengubahan huruf besar ke huruf kecil.

3. Analisis Sentimen

Pada tahap ini dibedakan antara komentar positif dan negatif pada akun TikTok dengan tagar tersebut.

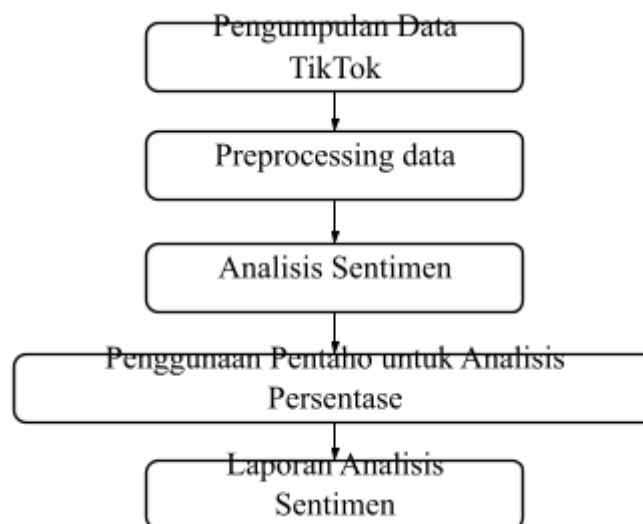
4. Penggunaan Pentaho untuk Analisis Persentase

Menggunakan *platform* Pentaho untuk menghitung dan memvisualisasikan persentase komentar positif dan negatif dari masing-masing kandidat calon presiden 2024.

c. *Output*

Laporan Analisis Sentimen Calon Presiden Indonesia 2024. Persentase komentar positif dan negatif dari seluruh data masing-masing kandidat calon presiden 2024 berdasarkan yang diambil dari pentaho.

Berikut adalah blok diagram yang menggambarkan cara kerja analisis sentimen terhadap calon presiden Indonesia 2024 berdasarkan komentar di TikTok seperti gambar 3.1.



Gambar 3.1 Blok Diagram Diusulkan

3.4 Pengumpulan Data

Pengumpulan data dalam penelitian ini menggunakan teknik *scraper* yang dibuat dengan menggunakan bahasa pemrograman *python*. Proses pengambilan data membutuhkan token dari TikTok dimulai dari memasukkan halaman *url* atau nama halaman *fanpage* untuk mengambil data komentar dari halaman tersebut. Halaman *fanpage* yang dipilih untuk analisis sentimen ini berupa halaman yang menampilkan video sentimen yang mencakup 3 (tiga) pasangan capres dan cawapres. Hal tersebut ditujukan agar komentar pada video yang ditampilkan tidak merujuk atau berat sebelah ke salah satu calon presiden saja (adil). Data komentar hasil *scraper* yang didapat random yang berarti tidak semua data

komentar pada halaman *fanpage* tersebut dapat diambil. Adapun url dan data komentar yang didapat dari teknik scraper terlihat pada tabel 3.1.

Tabel 3.1 Data Hasil *Crawling/Scraper*

No	Halaman Fanpage/URL	Jumlah Komentar
1	https://www.tiktok.com/@zeinchannel410/video/7267127149600017670	49
2	https://www.tiktok.com/@metro_tv/video/7306089644578360578	244
3	https://www.tiktok.com/@narasi/video/7281587109260201221	104
4	https://www.tiktok.com/@heerrzyy/video/7154497210330254618	260
5	https://www.tiktok.com/@undergreen.id/video/7269349287924993285	259
6	https://www.tiktok.com/@hanapijani/video/7293100060034764037	20
7	https://www.tiktok.com/@hanapijani/video/7291177260223089926	258
Total		1194

3.5 *Preprocessing Data*

Data yang diproses dalam tahap preprocessing hanya data komentar yang telah berhasil dikumpulkan. Adapun langkah-langkah preprocessing yang digunakan adalah tahap *cleansing*, *case folding*, *tokenizing*, dan *filtering*. Data yang digunakan untuk sampel berjumlah 22 komentar yang berasal dari halaman *fanpage* <https://www.tiktok.com/@zeinchannel410/video/7267127149600017670>.

a. Tahap *Cleansing*

Tahap *Cleansing* yaitu proses di mana menghapus atribut atau tanda baca yang tidak diperlukan pada proses analisis. Hasil *cleansing* terhadap data komentar yang digunakan terlihat pada tabel 3.2.

Tabel 3.2 Tahap *Cleansing* Data

No	<i>Input</i>	<i>Output</i>
1	Bpk prabowo subianto presiden ku 2024 RI	Bpk prabowo subianto presiden ku RI
2	Prabowo 🔥🔥	Prabowo 🔥🔥
3	pak Prabowo presiden .gaspoll pak Prabowo jangan kasih kendor	pak Prabowo presiden gaspoll pak Prabowo jangan kasih kendor
4	sebesar apapun kwalisimu tetep ganjar presiden	sebesar apapun kwalisimu tetep ganjar presiden
5	prabowo 2024	prabowo
6	Prabowo RI 1	Prabowo RI
7	Pak Prabowo Subianto pasti menang di pilpres 2024	Pak Prabowo Subianto pasti menang di pilpres
8	Anis mantap	Anis mantap
9	Pak ganjar 🤪👉👉	Pak ganjar 🤪👉👉
10	Prabowo-Ridwan kamil..... 😄😄	Prabowo Ridwan kamil 😄😄
11	@jjhrra._ ulah KTP mu, kita pilih Prabowo 🤪👍	ulah KTP mu kita pilih Prabowo 🤪👍
12	prabowo 🤪	prabowo 🤪
13	siap saya, dari, tangerang, banten, siap, pilih, pa, anis	siap saya dari tangerang banten siap pilih pa anis
14	pa ganjar 🤪	pa ganjar 🤪
15	Pak Ganjar Pilihanku	Pak Ganjar Pilihanku
16	masyarakat butuh perubahan	masyarakat butuh perubahan

No	Input	Output
17	prabowo di dukung sengkuni,bubar bubar 🌀	prabowo di dukung sengkuni bubar bubar 🌀
18	Mau pilih prabowo males ada golkar sama pan gara2 si artis pelakor shahnaz nyalon	Mau pilih prabowo males ada golkar sama pan gara si artis pelakor shahnaz nyalon
19	aku Klaten Jawa tengah dukung Anis Baswedan dan cak Imin menuju capres cawapres 2024 👍👍👍	aku Klaten Jawa tengah dukung Anis Baswedan dan cak Imin menuju capres cawapres 👍👍👍
20	BEKASI FULL PRABOWO	BEKASI FULL PRABOWO
21	Siapapun yg menang smoga amanah&bisa bikin maju Indonesia trutama memntingkan kpentingan rakyat bkn memjajah negri sndiri	Siapapun yg menang smoga amanah&bisa bikin maju Indonesia trutama memntingkan kpentingan rakyat bkn memjajah negri sndiri
22	saran aja jgn lihat janjinya... karena itu semua tidak akan di tepati...	saran aja jgn lihat janjinya karena itu semua tidak akan di tepati

b. Tahap *Case Folding*

Tahap *case folding* yakni mengubah semua huruf besar menjadi *lowercase* atau huruf kecil. Hasil *case folding* terhadap data komentar yang digunakan terlihat pada tabel 3.3.

Tabel 3.3 Tahap *Case Folding* Data

No	Input	Output
1	Bpk prabowo subianto presiden ku 2024 RI	bpk prabowo subianto presiden ku ri
2	Prabowo 🔥🔥	prabowo 🔥🔥
3	pak Prabowo presiden .gaspoll pak Prabowo jangan kasih kendor	pak prabowo presiden gaspoll pak prabowo jangan kasih kendor
4	sebesar apapun kualisimu tetep ganjar presiden	sebesar apapun kualisimu tetep ganjar presiden
5	prabowo 2024	prabowo
6	Prabowo RI 1	prabowo ri
7	Pak Prabowo Subianto pasti menang di pilpres 2024	pak prabowo subianto pasti menang di pilpres
8	Anis mantap	anis mantap
9	Pak ganjar 🤞🤞🤞	pak ganjar 🤞🤞🤞
10	Prabowo-Ridwan kamil..... 😄😄	prabowo ridwan kamil 😄😄
11	@jjhrra._ ulah KTP mu, kita pilih Prabowo 🤔👍	ulah ktp mu kita pilih prabowo 🤔👍
12	prabowo 😄	prabowo 😄
13	siap saya, dari, tangerang, banten, siap, pilih, pa, anis	siap saya dari tangerang banten siap pilih pa anis
14	pa ganjar 😄	pa ganjar 😄
15	Pak Ganjar Pilihanku	pak ganjar pilihanku
16	masyarakat butuh perubahan	masyarakat butuh perubahan

No	Input	Output
17	prabowo di dukung sengkuni,bubar bubar 🤔	prabowo di dukung sengkuni bubar bubar 🤔
18	Mau pilih prabowo males ada golkar sama pan gara2 si artis pelakor shahnaz nyalon	mau pilih prabowo males ada golkar sama pan gara si artis pelakor shahnaz nyalon
19	aku Klaten Jawa tengah dukung Anis Baswedan dan cak Imin menuju capres cawapres 2024 👍👍👍	aku klaten jawa tengah dukung anis baswedan dan cak imin menuju capres cawapres 👍👍👍
20	BEKASI FULL PRABOWO	bekasi full prabowo
21	Siapapun yg menang smoga amanah&bisa bikin maju Indonesia trutama memntingkan kpentingan rakyat bkn memjajah nagri sndiri	siapapun yg menang smoga amanah&bisa bikin maju indonesia trutama memntingkan kpentingan rakyat bkn memjajah negri sndiri
22	saran aja jgn lihat janjinya... karena itu semua tidak akan di tepati...	saran aja jgn lihat janjinya karena itu semua tidak akan di tepati

c. Tahap *Tokenizing*

Tahap *tokenizing* merupakan proses pemisahan kata berdasarkan tiap kata yang menyusunnya menjadi potongan tunggal. Hasil *tokenizing* terhadap data komentar yang digunakan terlihat pada tabel 3.4.

Tabel 3.4 Tahap *Tokenizing* Data

No	Input	Output
1	Bpk prabowo subianto presiden ku 2024 RI	“bpk” “prabowo” “subianto” “presiden” “ku” “ri”
2	Prabowo 🔥🔥	“prabowo” “🔥” “🔥”
3	pak Prabowo presiden .gaspoll pak Prabowo jangan kasih kendor	“pak” “prabowo” “presiden” “gaspoll” “pak” “prabowo” “jangan” “kasih” “kendor”
4	sebesar apapun kualisimu tetep ganjar presiden	“sebesar” “apapun” “kualisimu” “tetep” “ganjar” “presiden”
5	prabowo 2024	“prabowo”
6	Prabowo RI 1	“prabowo” “ri”
7	Pak Prabowo Subianto pasti menang di pilpres 2024	“pak” “prabowo” “subianto” “pasti” “menang” “di” “pilpres”
8	Anis mantap	“anis” “mantap”
9	Pak ganjar 🤪🤪🤪	pak” “ganjar” “🤪” “🤪” “🤪”
10	Prabowo-Ridwan kamil..... 😄😄	“prabowo” “ridwan” “kamil” “😄” “😄”
11	@jjhrra._ulah KTP mu, kita pilih Prabowo 🤪👍	“ulah” “ktp” “mu” “kita” “pilih” “prabowo” “🤪” “👍”
12	prabowo 🤪	“prabowo” “🤪”
13	siap saya, dari, tangerang, banten, siap, pilih, pa, anis	“siap” “saya” “dari” “tangerang” “banten” “siap” “pilih” “pa” “anis”
14	pa ganjar 🤪	“pa” “ganjar” “🤪”
15	Pak Ganjar Pilihanku	“pak” “ganjar” “pilihanku”

No	Input	Output
16	masyarakat butuh perubahan	“masyarakat” “butuh” “perubahan”
17	prabowo di dukung sengkuni, bubar bubar 🤔	“prabowo” “di” “dukung” “sengkuni” “bubar” “bubar” “🤔”
18	Mau pilih prabowo males ada golkar sama pan gara2 si artis pelakor shahnaz nyalon	“mau” “pilih” “prabowo” “males” “ada” “golkar” “sama” “pan” “gara” “si” “artis” “pelakor” “shahnaz” “nyalon”
19	aku Klaten Jawa tengah dukung Anis Baswedan dan cak Imin menuju capres cawapres 2024 👍👍👍	“aku” “klaten” “jawa” “tengah” “dukung” “anis” “baswedan” “dan” “cak” “imin” “menuju” “capres” “cawapres” “👍” “👍” “👍”
20	BEKASI FULL PRABOWO	“bekasi” “full” “prabowo”
21	Siapapun yg menang smoga amanah&bisa bikin maju Indonesia terutama memntingkan kpentingan rakyat bkn memjajah negri sendiri	“siapapun” “yg” “menang” “smoga” “amanah&bisa” “bikin” “maju” “indonesia” “trutama” “memntingkan” “kpentingan” “rakyat” “bkn” “memjajah” “negri” “sendiri”
22	saran aja jgn lihat janjinya... karena itu semua tidak akan di tepati...	“saran” “aja” “jgn” “lihat” “janjinya” “karena” “itu” “semua” “tidak” “akan” “di” “tepati”

d. Tahap *Filtering*

Tahap *filtering* yaitu tahap menghapus kumpulan dari kata yang tidak berpengaruh tanpa mengurangi informasi dari dokumen tersebut dan keberadaannya sering muncul. Dengan memanfaatkan *stopword* berbahasa Indonesia yang telah disusun dan dilakukan penambahan kata secara manual

karena disebabkan ada dari beberapa kata yang belum terdapat dalam *stopword*.

Tabel 3.5 Tahap *Filtering Data*

No	<i>Input</i>	<i>Output</i>
1	Bpk prabowo subianto presiden ku 2024 RI	prabowo subianto presiden
2	Prabowo 🔥🔥	prabowo 🔥
3	pak Prabowo presiden .gaspoll pak Prabowo jangan kasih kendor	prabowo gaspoll
4	sebesar apapun kualisimu tetep ganjar presiden	ganjar presiden
5	prabowo 2024	prabowo
6	Prabowo RI 1	prabowo
7	Pak Prabowo Subianto pasti menang di pilpres 2024	prabowo subianto menang
8	Anis mantap	anis mantap
9	Pak ganjar 🤪🤪🤪	ganjar 🤪🤪
10	Prabowo-Ridwan kamil..... 😄😄	prabowo 😄😄
11	@jjhrra._ ulah KTP mu, kita pilih Prabowo 🤪👍	pilih prabowo 🤪👍
12	prabowo 🍷	prabowo 🍷

No	<i>Input</i>	<i>Output</i>
13	siap saya, dari, tangerang, banten, siap, pilih, pa, anis	pilih anis
14	pa ganjar 🍷	ganjar 🍷
15	Pak Ganjar Pilihanku	ganjar pilihanku
16	masyarakat butuh perubahan	perubahan
17	prabowo di dukung sengkuni,bubar bubar 🗳️	prabowo dukung sengkuni 🗳️
18	Mau pilih prabowo males ada golkar sama pan gara2 si artis pelakor shahnaz nyalon	pilih prabowo males
19	aku Klaten Jawa tengah dukung Anis Baswedan dan cak Imin menuju capres cawapres 2024 🍷🍷🍷	dukung anis baswedan 🍷
20	BEKASI FULL PRABOWO	full prabowo
21	Siapapun yg menang smoga amanah&bisa bikin maju Indonesia trutama memntingkan kpentingan rakyat bkn memjajah negri sndiri	siapapun amanah
22	saran aja jgn lihat janjinya... karena itu semua tidak akan di tepati...	janjinya tidak tepati

3.6 Pelabelan Kelas Sentimen

Tahap pelabelan kelas sentimen adalah proses pemberian label atau kategori pada teks atau data untuk mengidentifikasi sentimen atau opini yang terkandung di dalamnya. Proses ini merupakan langkah penting dalam pengembangan model

analisis sentimen, yang bertujuan untuk mengenali apakah suatu teks atau dokumen memiliki sentimen positif, negatif, dan netral.

Tabel 3.6 Pelabelan Kelas Sentimen

No	Komentar	Pasangan Calon Presiden		
		Anies Baswedan	Prabowo Subianto	Ganjar Pranowo
1	prabowo subianto presiden		Positif	
2	prabowo 🏽		Positif	
3	prabowo gaspoll		Positif	
4	ganjar presiden			Positif
5	prabowo		Positif	
6	prabowo		Positif	
7	prabowo subianto menang		Positif	
8	anis mantap	Positif		
9	ganjar 🤪 🙌			Positif
10	prabowo 😁 😂		Positif	
11	pilih prabowo 🤪 👍		Positif	
12	prabowo 🥰		Positif	
13	pilih anis	Positif		
14	ganjar 🥰			Positif
15	ganjar pilihanku			Positif
16	perubahan	Positif		
17	prabowo dukung sengkuni 🤪		Negatif	
18	pilih prabowo males		Negatif	
19	dukung anis baswedan 👍	Positif		
20	full prabowo		Positif	
21	siapapun amanah	Netral	Netral	Netral
22	janjinya tidak tepati	Netral	Netral	Netral

3.7 Analisis Sentimen Menggunakan Algoritma *Naive Bayes Classifier*

Proses ini melibatkan analisis teks untuk menentukan apakah komentar bersifat positif, negatif, atau netral terhadap calon presiden 2024 dengan menentukan kata-kata yang bersifat negatif maupun kata-kata yang bersifat positif ataupun netral. Untuk menentukan hasil sentimen komentar positif, negatif, dan netral digunakan algoritma *naive bayes classifier* dengan dataset atau data *training* yang digunakan terlihat pada tabel 3.7.

Tabel 3.7 Data *Training*

No	Komentar	Pasangan Calon Presiden		
		Anies Baswedan	Prabowo Subianto	Ganjar Pranowo
1	prabowo subianto presiden		Positif	
2	prabowo 🗲		Positif	
3	prabowo gaspoll		Positif	
4	ganjar presiden			Positif
5	prabowo		Positif	
6	prabowo		Positif	
7	prabowo subianto menang		Positif	
8	anis mantap	Positif		
9	ganjar 💪👉			Positif
10	prabowo 😁😄		Positif	
11	pilih prabowo 🗲👍		Positif	
12	prabowo 🍷		Positif	
13	pilih anis	Positif		
14	ganjar 🍷			Positif
15	ganjar pilihanku			Positif
16	perubahan	Positif		

No	Komentar	Pasangan Calon Presiden		
		Anies Baswedan	Prabowo Subianto	Ganjar Pranowo
17	prabowo dukung sengkuni 🤔		Negatif	
18	pilih prabowo males		Negatif	
29	dukung anis baswedan 👍	Positif		
20	full prabowo		Positif	
21	siapapun amanah	Netral	Netral	Netral
22	janjinya tidak tepati	Netral	Netral	Netral

3.8 Penggunaan Pentaho untuk Analisis Sentimen

Setelah diketahui pengelompokan kata dari kata yang bersifat positif atau negatif, maka data tersebut akan diolah menggunakan *platform* Pentaho untuk menghitung dan memvisualisasikan persentase komentar positif dan negatif. Hasil analisis sentimen akan diubah menjadi laporan visual yang memberikan gambaran yang jelas tentang sentimen publik dalam bentuk digram pie atau diagram batang. Pada tahap ini akan dijelaskan pada bab IV.

3.9 Hasil/Laporan Analisis Sentimen

Hasil akhir dari proses ini adalah laporan analisis sentimen yang mencakup persentase komentar positif dan negatif dari seluruh data dalam bentuk digram pie atau diagram batang pada masing-masing kandidat calon presiden 2024. Hasil laporan dijelaskan pada bab IV.