# ANALISIS SENTIMEN TERHADAP PEMILU INDONESIA TAHUN 2024 MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

## SKRIPSI

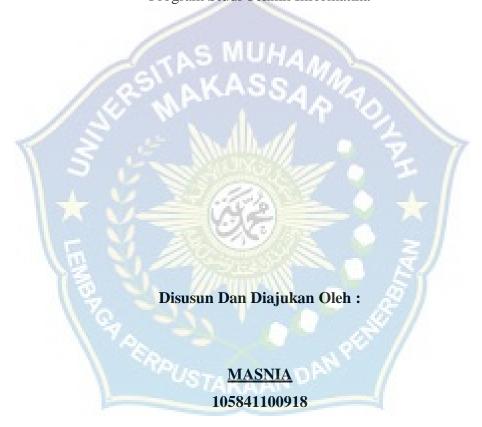
Diajukan Untuk Memenuhi Salah Satu Syarat Memperoleh Gelar Sarjana Komputer Program Studi Teknik Informatika



PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAKASSA
2024

## ANALISIS SENTIMEN TERHADAP PEMILU INDONESIA TAHUN 2024 MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

Diajukan Untuk Memenuhi Salah Satu Syarat Memperoleh Gelar Sarjana Komputer Program Studi Teknik Informatika



PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAKASSA
2024

# UNIVERSITAS MUHAMMADIYAH MAKASSAR FAKULTAS TEKNIK



**GEDUNG MENARA IQRA LT. 3** 

Jl. Sultan Alauddin No. 259 Telp. (0411) 866 972 Fax (0411) 865 588 Makassar 90221 Website: <a href="http://toknik.unismuh.makassar.ac.id">www.unismuh.ac.id</a>, e\_mail: <a href="http://toknik.unismuh.makassar.ac.id">unismuh@gmail.com</a> Website: <a href="http://toknik.unismuh.makassar.ac.id">http://toknik.unismuh.makassar.ac.id</a>

# PENGESAHAN

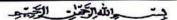
Skripsi atas nama Maania dengan nomor induk Mahasiswa 105 84 11009 18, dinyatakan ditorima dan disahkan oleh Panitia Ujian Tugas Akhir/Skripsi sesuai dengan Surat Keputusan Dekan Fakultas Teknik Universitas Muhammadiyah Makassar Nomor : 001/05/A.5-II/I/45/2024, sebagai salah satu syarat guna memperoleh gelar Sarjana Komputer pada Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar pada hari Sabtu tanggal 31 Januari 2024.

Panitla Ujian :	Makossar,	19 Rajab	1445 H
1. Pengawas Umum		31 Vanuari	2024 M
a. Rektor Universites Muhammadiyah Maka Prof. Dr. H. AMBO ASSE MAG	MUL		
b. Dekan Fakultas Teknik Universitas Hasa Prof. Dr. Eng. MUHAMMAD BRAN SAM	nitidity S	Ja	iva
2. Penguji	ully of E		
a. Kotue : Dr. Ir Zahir Zaihuddin M	Sc.	6	7
b. Sekertaris : Desi Anggrea V. S. Kom	Mc Simon		y-
3. Anggota : 1. Lukman, S. Kom My.		14)	<b></b>
2. Muhyiddin A M Hayal 3. Lukman Anas, S. Kom	NO TEXEN		\$.
Mon	getahur:		
Pembimbing I	Pembin	lyenida	
Ms		W	_
Titin Wahyuni S.Pd., MT.	Fahrim Irhashna	Rachman, S	S.Kom., MT
MAMMAGUERO	Fakultas Teknik		
	100		
	awaty, ST., MT., IPI 795 108	<u>VI</u>	

# UNIVERSITAS MUHAMMADIYAH MAKASSAR FAKULTAS TEKNIK

#### **GEDUNG MENARA IQRA LT. 3**

Jl. Sultan Alauddin No. 259 Telp. (0411) 866 972 Fax (0411) 865 588 Makassar 90221 Website: <a href="mailto:www.unismuh.ac.id">www.unismuh.ac.id</a>, e\_mail: <a href="mailto:unismuh@amail.com">unismuh@amail.com</a> Website: <a href="http://teknik.unismuh.makassar.ac.id">http://teknik.unismuh.makassar.ac.id</a>



#### HALAMAN PENGESAHAN

Tugas Akhir ini diajukan untuk memenuhi syarat ujian guna memperoleh gelar Sarjana Komputer (S.Kom) Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar.

Judul Skripsi : ANALISIS SENTIMEN TERHADAP PEMILU INDONESIA TAHUN
2024 MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER.

Nama

: MASNIA

Stambuk

: 105841100918

Makassar, 31 Januari 2024

Telah Diperiksa dan Disetujui Oleh Dosen Pembimbing;

Pembimbing I

Titin Wahyuni S.Pd., MT.

Pembimbing II

Fahrim Irhamna/Rachman, S.Kom., MT

mormatika

Mulvidin A. M. Hayat, S.kom., MT

#### **ABSTRACT**

Abstract. Sentiment Analysis Of The 2024 Indonesian Election Using The Naïve Bayes Classifier Method. General Elections or abbreviated as Elections are the process of activities carried out to elect leaders who will occupy leadership positions in the Executive and Legislative institutions in the Indonesian Government. Elections are a manifestation of a country that adheres to a democratic system. The election process determines leaders from Regency/City, Provincial, to National levels. Elections often use social media as a campaign tool because it is considered a solution by the public in sharing their opinions freely. Among these social media, Twitter is one of the popular social media platforms used (Karami et al., 2022). Because Twitter is considered to be an effective forum for use by Indonesian people, both individuals, corporations and even politicians, even in disseminating abstract information, the first thing to be written is the Indonesian version. The abstract indents with a distance of 1 cm each to the left and right. Italics. Abstracts are written in 1 paragraph, 100-150 words long.

Key Words: Anies President 2024, Ganjar President 2024, Prabowo President 2024, Election 2024

#### ABSTRAK

Abstract. Anaisis sentimen terhadap pemilu Indonesia tahun 2024 menggunakan metode Naive Bayes Classifier. Pemilihan Umum atau disingkat Pemilu adalah proses kegiatan yang dilakukan untuk memilih pemimpin yang akan menempati kursi kepemimpinan di lembaga Eksekutif maupun Legislatif di Pemerintahan Indonesia. Pemilu merupakan perwujudan dari sebuah negara yang menganut sistem demokrasi. Adapun proses pemilihan ini menentukan pemimpin mulai dari tingkatan Kabupaten/Kota, Provinsi, hingga Nasional Pemilu sering kali menggunakan media sosial sebagai alat kampanye karena dianggap sebagai solusi oleh public dalam membagikan pendapatnya secara bebas. Diantara media sosial tersebut twitter merupakan salah satu platform media sosial popular yang digunakan (Karami et al., 2022). Karena twitter dianggap menjadi wadah yang efektif untuk digunakan oleh masyarakat Indonesia baik perseorangan, corporate bahkan politikus sekalipun dalam penyebarluasan informasi abstrak yang pertama dituliskan adalah versi Bahasa Indonesianya. Abstrak menjorok kedalam dg jarak kiri dan kanan masing-masing 1 cm. Italic. Abstrak ditulis dalam 1 paragraf, dengan panjang 100-150 kata.

**Kata kunci:** Anies Presiden 2024, Ganjar Presiden 2024, Prabowo Presiden 2024, Pemilu 2024.

#### **KATA PENGANTAR**

### Assalamu'alaikum Warahmatullahi Wabaraktuh

Segala Puji Bagi Allah SWT yang telah memberikan rahmat serta hidayah Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul "Analisis Sentimen Publik Terhadap Calon Presiden Tahun 2024 Menggunakan Metode Naïve Bayes Classifier" dengan baik. Shalawat beserta salam semoga senantiasa tercurah kepada junjungan kita Baginda Rasulullah Shalallahu "Alaihi Wassalam beserta para keluarga, sahabat, dan para pengikut Beliau hingga akhir zaman.

Dalam pembuatan laporan skripsi ini, penulis banyak dibantu oleh berbagai pihak dengan memberikan banyak masukan dan petunjuk, serta mendukung dan menjadi motivasi bagi penulis. Maka dari itu, ucapan terima kasih penulis haturkan kepada:

- Terkhususnya kepada kedua orang tua penulis yang sudah mendukung dan memberikan semangat serta motivasi selama penyusunan Skripsi, dan untuk semua keluarga yang telah mendoakan dan memberikan motivasi, semangat serta bantuannya.
- 2. Ibu **Dr.Ir.Hj Nurnawati, S.T., M.T., I.P.M**, selaku Dekan Fakultas Teknik
- 3. Bapak **Muh. Syafaat S Kuba, S.T., M.T**, selaku Wakil Dekan Fakultas Teknik.
- 4. Bapak Muhyiddin AM Hayat S.Kom., M.T, selaku Ketua Prodi Informatika.
- 5. Ibu **Titin Wahyuni S.Pd., M.T**, selaku Dosen Pembimbing 1 proposal.
- Bapak Fahrim Irhamna Rachman S.Kom., M.T, selaku Dosen Pembimbing
   Proposal.
- 7. Dosen dan Staf Fakultas Teknik Universitas Muhammadiyah Makassar.
- 8. Teman-teman Khususnya Angkatan Mekanika 2018 Fakultas Teknik, Universitas Muhammadiyah Makassar, terima kasih atas dukungan dan doanya.

9. Teman – teman seperjuangan saya yang ada di lembaga kemahasiswaan Fakultas Teknik terima kasih atas dukungan dan doanya.

Akhir kata, penulis mengharapkan agar skripsi ini dapat bermanfaat, baik bagi penulis pribadi maupun pada pihak-pihak lain. Serta, semoga segala masukan baik berupa kritik maupun saran yang membangun yang ditujukan kepada penulis dapat menjadikan penulis menjadi lebih baik lagi untuk kedepan. Terima kasih.

Billahi fisabililhaq, fastabiqul khairat.

Wassalamualaikum Wr. Wb.



## **DAFTAR ISI**

HALAMAN PENGESAHAN	ii
KATA PENGANTAR	
DAFTAR ISI	
DAFTAR GAMBAR	
DAFTAR TABEL	
DAFTAR ISTILAH	xiv
BAB I PENDAHULUAN	1
A. Latar Belakang	1
B. Rumusan Masalah	3
C. Tujuan Masalah	3
D. Batasan Penelitian	4
E. Manfaat Penelitian	
F. Ruang Lingkup Penelitian	5
G. Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
A. LANDASAN TEORI	7
1. Analisis Sentimen	
2. Data Mining	8
3. Metode Naïve Bayes Classifier	
4 Flowchart	1.0

B.	Penelitian Terkait	. 12
BAB	III METODE PENELITIAN	.14
A.	Tempat dan Waktu Penelitian	. 14
B.	Bahan dan Alat	. 14
C.	Perancangan Sistem.	. 14
D.	Teknik Pengujian sistem	. 16
E.	Teknik Analisis Data	
BAB	IV HASIL DAN PEMBAHASAN	.20
A.	Deskripsi Data	. 20
В.	Data Preprocessing	. 21
C.	PelabelanPelabelan	. 21
D.	Proses Penerapan Metode Naive Bayes	. 22
E.	Teknik Pengujian Sistem	
BAB	V PENUTUP	.61
A.	Kesimpulan	. 61
В.	Saran	. 61
DAF	ΓAR PUSTAKA	.62

## **DAFTAR GAMBAR**

Gambar 1. Knowledge Discovery in Database	9
Gambar 2. Flowchart System	15
Gambar 3. import data set	22
Gambar 4. tampillan data	32
Gambar 5. proses cleaning	32
Gambar 6. pembagian data frame kolom	33
Gambar 7. jumlah data sesuai dengan pengimputan	33
Gambar 8. pembagian data training dan testing	34
Gambar 9. proses tokenisasi, vektorisasi menggunakan tf-idf	34
Gambar 10. proses uji coba	35
Gambar 11. proses uji coa kata secara manual	35
Gambar 12. proses uji coba naive bayes	36
Gambar 13. hasil prediksi	36
Gambar 14. ekpor data hasil prediksi	

## DAFTAR TABEL

Tabel 1. Daftar Simbol Flowchart Diagram (Jogiyanto, 2001)	11
Tabel 2. Tabel Ringkasan Data Mentah	20
Tabel 3. Tabel DataSet Bersih	21
Tabel 4. Tabel dataset yang sudah di label	22
Tabel 5. Dataset Hasil Prediksi	59



## DAFTAR LAMPIRAN

Lampiran 1. Deskripsi Data	66
Lampiran 2. Data yang Sudah Di Bersihkan	66
Lampiran 3. Data yang Sudah di Label	67
Lampiran 4. Tampilan Data	70
Lampiran 5. Source Code	
Lampiran 6. Grafik	



#### DAFTAR ISTILAH

#### **Analisis Sentimen**

Analisis sentimen dapat di sebut juga sebagai penambangan opini, seperti dalam bidang studi yang menganilisis pendapat, sentiment, evaluasi, penilaian, sikap dan emosi seseorang terhadap entitas seperti produk, layanan, organisasi,individu, masalah, peristiwa, topi dan atributnya.

#### Flowchart

Flowchart merupakan gambaran tentang langkah-langkah dan prosedur suatu aplikasi. Flowchart membantu analis dan pemrogram membuat suatu gambaran dari aplikasi yang dibuat dan memecahnya 13 menjadi segmen yang lebih kecil untuk analisis yang lebih mudah. Flowchart digunakan dalam desain aplikasi untuk menggambarkan logika aplikasi.

#### **Preprocessing**

Preprocessing merupakan suatu tahapan yang bertujuan untuk memudahkan dalam proses pengolahan data untuk diolah pada tahapan selanjutnya. Proses preprocessing dilakukan untuk menyeleksi data serta mengubahnya menjadi data yang lebih terstruktur.

## **Counfusion Matrix**

Teknik ini digunakan untuk memvisualisasikan dan menghitung jumlah klasifikasi yang benar dan salah untuk setiap kelas sentimen.Dari confusion matrix, Anda dapat menghitung akurasi, presisi, recall, dan F1-score untuk masing-masing kelas sentimen.

## **Data Mining**

Data Mining adalah analisis data pengamatan yang besar untuk menemukan hubungan yang sebelumnya tidak diketahui dan dua metode baru untuk meringkas data sehingga mudah dipahami dan penggunaannya sesuai dengan data pemilih.

#### **Cross Validation**

Teknik ini dapat digunakan untuk menguji stabilitas dan konsistensi model Naive Bayes. Dalam pengujian crossvalidation, dataset Anda dibagi menjadi beberapa subset yang saling bergantian sebagai data training dan testing.

## **Tranformation**

Merupakan proses integrasi data yang dipilih untuk membuat data cocok untuk proses penambangan data. Ini adalah proses yang sangat tergantung pada jenis atau pola data yang dicari dalam database.

## Pengujian Akurasi

Teknik ini bertujuan untuk mengukur sejauh mana model dapat mengklasifikasikan sentimen dengan benar. Hitung akurasi model dengan membandingkan hasil klasifikasi sentimen dari model dengan label sentimen yang sebenarnya pada data testing.

#### BAB 1

#### PENDAHULUAN

## A. Latar Belakang

Pemilihan Umum atau disingkat Pemilu adalah proses kegiatan yang dilakukan untuk memilih pemimpin yang akan menempati kursi kepemimpinan di lembaga Eksekutif maupun Legislatif di Pemerintahan Indonesia. Pemilu merupakan perwujudan dari sebuah negara yang menganut sistem demokrasi. Adapun proses pemilihan ini menentukan pemimpin mulai dari tingkatan Kabupaten/Kota, Provinsi, hingga Nasional. Begitupun juga pemilihan presiden hingga saat ini menerapkan sistem didalam Pemilu.

Menurut Komisi Pemilihan Umum (KPU) pada websitenya mengumumkan tahapan dan jadwal penyelenggaraan Pemilu tahun 2024 mendatang. Pencalonan presiden dan wakil presiden yaitu 19 oktober hingga 25 november 2023 (Admin, 2022b). Adanya hal tersebut membuat beberapa lembaga survey terus merilis risetnya tentang pencalonan presiden dan wakil presiden. Berdasarkan salah satu website lembaga survey yaitu Poltracking Indonesia (Admin, 2022a) merilis riset terbarunya tentang pilpres 2024. Hasil survey tersebut mengungkapkan bahwa tiga nama terkuat sebagai bakal calon presiden yaitu Ganjar Pranowo (Gubernur Jawa Tengah), Prabowo Subianto (Ketua Umum Partai Gerindra), dan Anies Baswedan (Gubernur DKI Jakarta).

Pemilu sering kali menggunakan media sosial sebagai alat kampanye karena dianggap sebagai solusi oleh public dalam membagikan pendapatnya secara bebas. Diantara media sosial tersebut twitter merupakan salah satu platform media sosial popular yang digunakan (Karami et al., 2022). Karena twitter dianggap menjadi wadah yang efektif untuk digunakan oleh masyarakat Indonesia baik perseorangan, corporate bahkan politikus sekalipun dalam penyebarluasan informasi. Menurut We Are Social, jumlah pengguna twitter di Indonesia pada tahun 2022 adalah mencapai angka 18,45 juta.

Pengembangan pengetahuan dan teknologi terus meningkat sehingga memunculkan ide dan gagasan baru pada bidangnya. Salah satu pemanfaatan teknologi yang terus didorong untuk digunakan sebagai sistem pengambilan kebijakan, yaitu analisis sentiment. Analisis sentimen merupakan riset komputasi dari sentimen, emosi, dan opini yang diluapkan kedalam teks. Jika diberikan suatu set dokumen teks yang berisi sentimen mengenai suatu objek, maka opinion mining bertujuan untuk mengekstrak atribut dan komponen dari objek yang telah dikomentasi pada setiap dokumen dan untuk menentukan apakah komentar tersebut bermakna positif atau negatif (Nurul et al., 2019).

Penentuan kelas sentimen memerlukan metode klasifikasi untuk mengukur tingkat keberhasilan kelas. Salah satu metode klasifikasi yang digunakan untuk klasifikasi teks yaitu Naïve Bayes. Naïve bayes telah terbukti menjadi algoritma yang terkenal dalam melakukan analisis sentimen dalam sebuah teks. Metode naïve bayes didukung library pada python dalam melakukan berbagai domain klasifikasi (Soesanto et al., 2023).

Berdasarkan berbagai penelitian yang telah dipaparkan, penelitian ini akan membahas analisis sentimen sebagai objek penelitian. Penentuan sentimen yang diberikan oleh masyarakat terhadap bakal calon presiden potensial berdasarkan cuitan pada twitter akan dilakukan analisis menggunakan metode klasifikasi naïve bayes.

Analisis sentimen adalah bidang dalam ilmu data untuk mencapai pandangan holistik yang lebih luas tentang kebutuhan dan harapan pengguna. Analisis sentimen bertujuan untuk mendefinisikan emosi pengguna melalui teks pada kalimat, dokumen atau aspek lainnya (Abayomi-alli et al., 2022). Karena opini publik di Indonesia berpotensi menjadi informasi berharga yang digunakan pada objek analisis sentimen (Hayatin et al., 2020). Data yang digunakan berasal dari platform media sosial twitter yang didapat menggunakan teknik crawling data sejumlah 15000 dataset.

Naïve bayes merupakan metode yang dipilih karena terbukti memberikan hasil akurasi yang baik, cepat dan dapat bekerja dengan baik dalam melakukan verifikasi analisis sentimen (Santoso et al., 2020). Tingkat akurasi pada sebuah metode klasifikasi juga dipengaruhi oleh penerapan bag of words, yaitu TF-IDF dengan tujuan untuk melakukan perbaikan konteks kalimat dalam sebuah metode klasifikasi teks. M. Liang (2022) dalam penelitiannya (Liang & Niu, 2022) mengungkapkan bahwa penggunaan TF-IDF merupakan pemberian nilai komperhensif yang sangat penting dalam pembobotan kata di sebuah klasifikasi teks.

Berdasarkan latar belakang di atas, maka peneliti melakukan penelitian dengan mengangkat judul tentang "Analisis Sentimen Pemilu Indonesia Tahun 2024 Menggunakan Metode Naïve Bayes Classifier". penelitian ini akan membahas sentimen dari pengguna Twitter dilakukan dengan menggunakan metode Naïve Bayes

#### B. Rumusan Masalah

Rumusan masalah pada penelitian ini adalah:

- a. Bagaimana opini masyarakat di media social twitter mengenai pemilihan presiden (calon presiden) tahun 2024? Apakah cenderung mendatang positif, negative, ataukah netral?
- b. Bagaimana proses pembobotan menggunakan metode Naïve Bayes Classifier?
- c. Bagaimana menguji tingkat keberhasilan penelitian?

## C. Tujuan Masalah

Penelitian ini memiliki tujuan sebagai berikut:

 Untuk mengetahui bagaimana opini masyarakat menanggapi pemilu di tahun 2024 mendatang, apakah cenderung positif, negative, ataukah netral.

- 2. Untuk melakukan transformasi teks ke dalam bentuk kumpulan kata menggunakan text Preprocessing.
- 3. Untuk melakukan implementasi penelitian menggunakan metode klasifikasi Naïve Bayes Classifier.
- 4. Untuk menguji tingkat keberhasilan berdasarkan akurasi menggunakan metode klasifikasi Naïve Bayes Classifier.

#### D. Batasan Penelitian

Batasan masalah pada penelitian ini adalah:

- 1. Dataset yang digunakan seputar pemilihan umum presiden 2024.
- 2. Dataset yang digunakan berasal dari media sosial twitter dengan rentang waktu
- 3. Pelabelan dataset dengan kategori kelas "anies presiden 2024", "ganjar presiden 2024", dan "prabowo presiden 2024".
- 4. Metode klasifikasi yang digunakan adalah Naïve Bayes Classifier (NBC).
- 5. Bahasa pemrograman yang digunakan untuk klasifikasi adalah python.
- 6. Eksistensi calon merupakan salah satu indikator penting untuk membaca potensi kemenangan dari kandidat yang diusung.

#### E. Manfaat Penelitian

Penelitian yang dilakukan diharapkan dapat memberikan manfaat sebagai berikut:

- Manfaat bagi peneliti, yaitu menambah wawasan peneliti dalam melakukan analisis sentiment menggunakan metode Naïve Bayesdan juga memberikan referensi kepada peneliti lain dengan topik penelitian serupa ataupun pengembangan dari penelitian ini.
- 2. Manfaat secara teoritis, yaitu memberikan masukan pada bidang ilmu pengetahuan & teknologi terhadap materi mengenai analisis sentimen dengan metode pembobotan TF-IDF dan metode klasifikasi Naïve Bayes.

## 3. Manfaat bagi praktisi, antara lain:

- a. Bagi partai politik, yaitu memberikan perhitungan potensi suara yang dimiliki oleh bakal calon presiden kepada partai politik pengusung sebagai bahan pertimbangan dalam mengambil keputusan.
- b. Bagi masyarakat, yaitu memberikan informasi bakal calon presiden potensial berdasarkan opini publik di twitter.

## F. Ruang Lingkup Penelitian

Berdasarkan latar belakang yang telah dipaparkan maka dapat Ruang lingkup sebagai berikut:

- Perlunya pemetaan untuk menentukan gerbong politik dalam memastikan peluang bakal calon Presiden terpilih menjadi Presiden Indonesia 2024.
- 2. Metode yang di gunakan adalah metode Naïve Bayes Clasiffier dan efesien pada twitter.

## G. Sistematika Penulisan

Sistematika penulisan dalam penyusunan laporan proposal penelitian ini adalah sebagai berikut :

## **BAB I PENDAHULUAN**

Pada bab ini menguraikan tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitia, ruang lingkup penelitian dan sistematika penulisan.

#### BAB II TINJAUAN PUSTAKA

Pada bab ini menguraikan mengenai studi pustaka terhadap teori — teori yang di gunakan dalam penelitian serta aspek — aspek yang mendukung dalam penelitian.

#### BAB III METODOLOGI PENELITIAN

Bab ini berisi penjelasan tentang langkah – langkah yang di gunakan dalam penelitian, di mulai dari pengidentifikasian masalah, pengumpulan data, analisa algoritma, dan alat ukur yang di gunakan dalam penelitian.

## BAB IV ANALISA DAN PENGUJIAN

Bab ini berisikan hasil desain sistem serta pembahasan terhadap desain tersebut

## BAB V KESIMPULAN DAN SARAN

Bab terakhir akan memuat kesimpulan isi dari keseluruhan uraian babbab sebelumnya dan saran-saran dari hasil yang diperoleh yang diharapkan dapat bermanfaat dalam pengembangan selanjutnya.

#### **BAB II**

#### TINJAUAN PUSTAKA

#### A. LANDASAN TEORI

#### 1. Analisis Sentimen

Analisis sentimen dapat di sebut juga sebagai penambangan opini, seperti dalam bidang studi yang menganilisis pendapat, sentiment, evaluasi, penilaian, sikap dan emosi seseorang terhadap entitas seperti produk, layanan, organisasi,individu, masalah, peristiwa, topi dan atributnya(Zhang & Liu 2016)

Selanjutnya, analisis sentiment dapat di katakan sebagai opinion mining yang dapat di katakana sebagai proses dalam memahami sampai kepada mengolah data secara sistem demi mendapatkan informasi yang berkaitan dengan sebuah opini dan pendapat. Analisis sentimen biasanya di lakukan dengan tujuan untuk mengetahui sifat sebuah opini yang sedang di bhasa, yang biasanya bersifat pada opini yang negative atau positif (Buntoro, Adji, & Purnama sari, 2014).

Berdasarkan beberapa pendapat tentang analisis sentiment dapat di ketahui bahwa pada dasarnya analisis sentiment merupakan cara bagaimana menganalisis pendapat, tanggapan, komentar. Opini dan sentiment dari masyarakat yang biasanya di curahkan ke dalam social media ataupun media online. Dengan opini yang bermacam macam, dapat di lihat opini yang bersifat positif, netral ataupun Negatif.

Seperti yang di lakukan oleh L.Lee dan B.Pang analisis sentiment di bagi menjadi 2 kategori, yaitu Coarse-grained sentiment analysis merupakan sebuah proses klasifikasi di lakukan berdasarkan orientasi sebuah dokumen secara keseluruhan . dalam orientasi tersebut di bagi menjadi tiga jenis yaitu negative, netral, positif. Sedangkan Fined —grained sentiment analisis menggunakan objek yang berupa sebuah kalimat melainkan bukan sebuah dokumen secara keseluruhan. Seperti contoh kalimat yang mengandung sifat positif "semoga debat pilpres dapat mempermudah masyarakat memilih presidennya" dan kalimat negative " debat pilpres tahun ini banyak menuai perpecahan masyarakat Indonesia". Pada dasarnya

tidak semua tweet menggunakan kata baku dan memiliki banyak kata yang artinya sama seperti "saya dan aku" dengan kata lan "gua. Gue, aq, ak, sa" dan sebagainya.

## 2. Data Mining

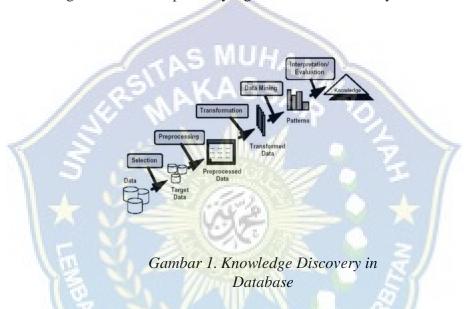
Menurut para ahli, *data mining* adalah analisis data pengamatan yang besar untuk menemukan hubungan yang sebelumnya tidak diketahui dan dua metode baru untuk meringkas data sehingga mudah dipahami dan penggunaannya sesuai dengan data pemilih (Jefri, 2013).

Sementara itu (Larose, 2005) *data mining* didefinisikan sebagai proses pencarian hubungan, pola, dan tren baru yang signifikan dengan menyaring data gudang yang sangat besar menggunakan teknik pengenalan pola seperti teknik statistik dan matematika.

Menurut (Benri Melpa Metisen, 2015) data mining terdiri dari beberapa langkah, antara lain:

- a. Pemilihan data (*Data Selection*): Buat set data target, pilih set data atau fokuskan pada subset variabel atau sampel data untuk membuat penemuan. Hasil pemilihan disimpan dalam file terpisah dari database operasional.
- b. *Pre-processing / cleaning* data: *Pre-processing* dan pembersihan data adalah operasi dasar, seperti pengurangan kebisingan. Proses pembersihan melibatkan penghapusan data duplikat, memeriksa data yang bertentangan dan memperbaiki kesalahan dalam data seperti kesalahan ketik. Data dapat diperkaya dengan data atau informasi eksternal terkait.
- c. Transformation: Merupakan proses integrasi data yang dipilih untuk membuat data cocok untuk proses penambangan data. Ini adalah proses yang sangat tergantung pada jenis atau pola data yang dicari dalam database.
- d. Pemodelan Data (*Data mining*): Pemilihan tugas *data mining* adalah pemilihan tujuan dari proses *knowledge discovery in database*, seperti karakterisasi, klasifikasi, regresi, pengelompokan, asosiasi, dll. Pemilihan tugas penambangan data adalah pemilihan target dari proses *knowledge discovery in database*. Karakterisasi, klasifikasi, regresi, pengelompokan,

- asosiasi, dll. Pemilihan teknik, metode atau algoritma yang tepat sangat bergantung pada tujuan dan keseluruhan proses *knowlegde discovery in database*.
- e. Evaluasi dan interpretasi (Interpretation/ Evaluation): Yakni, penerjemahan pola yang dihasilkan dari *data mining*. Model data yang dihasilkan harus disajikan dalam format yang mudah dipahami, pada tahap ini diperiksa apakah pola atau informasi yang ditemukan tidak sesuai dengan fakta atau hipotesis yang sudah ada sebelumnya.



Data mining dapat membantu mengidentifikasi model dan tren bisnis yang dapat digunakan untuk membuat keputusan yang lebih baik dan lebih efisien. Misalnya, penambangan data dapat membuat operasi lebih efisien, meningkatkan penjualan, meningkatkan layanan pelanggan, dan meningkatkan keamanan informasi. Namun, penggunaan penambangan data juga dapat menimbulkan masalah privasi, keamanan, dan etika. Oleh karena itu, saat melakukan analisis data menggunakan teknik penambangan data, perhatian khusus harus diberikan pada perlindungan data pribadi dan data sensitif.

## 3. Metode Naïve Bayes Classifier

Metode Naïve Bayes classifier merupakan salah satu metode paling popular di gunakan dengan tingkat keakuratan yang baik. Banyak penelitian tentang pengklasifikasian yang telah di lakukan dengan menggunakan algoritma ini. Berbeda dengan metode pengklasifikasian dengan logisticregression ordinal maupun nominal, pada algoritma naïve bayes pengklasifikasian tidak membutuhkan adanya pemodelan maupun uji statistic. Deangan melakukan analisa suatu topik yang memiliki sifat Positif, Negatif, Netral. Metode Naïve Bayes Classifier sendiri mempunyai kemampuan untuk klarifikasi data serupa dengan decession tree dan neural network. Metode Naïve Bayes classifier telah terbukti mempumyai akurasi kecepatan tinggi yang dapat di aplikasikan ke dalam database dengan data yang besar.

Naïve bayes merupakan sebuah konsep dasar dari Teorema Bayes dan merupakan metode yang terdapat dalam data mining yang digunakan untuk mengklasfikasikan data. Adapun proses pengelompokkan atau klasifikasi yang dilakukan dibagi menjadi dua fase, yaitu learning/training dan testing/classify (Riadi et al., 2019). Metode bayes digunakan untuk menghitung probabilitas dalam suatu peristiwa berdasarkan pengamatan yang dilakukan. Metode naïve bayes berbasis probabilitas sederhana yang melakukan prediksi dengan mengandalkan nilai asumsi pada keanggotaan sebuah kelas.

## 4. Flowchart

Flowchart merupakan gambaran tentang langkah-langkah dan prosedur suatu aplikasi. Flowchart membantu analis dan pemrogram membuat suatu gambaran dari aplikasi yang dibuat dan memecahnya 13 menjadi segmen yang lebih kecil untuk analisis yang lebih mudah. Flowchart digunakan dalam desain aplikasi untuk menggambarkan logika aplikasi (Londjo, 2021)

Tabel 1. Daftar Simbol Flowchart Diagram (Jogiyanto, 2001)

Simbol	Nama	Keterangan
	Terminator Awal / Akhir Program	Simbol untuk memulai dan mengakhiri suatu program
SAN LESTIN	MUH4 DokumenS	Menunjukkan dokumen berupa dokumen input dan output pada proses manual dan proses berbasis computer
	Proses Manual	Menunjukkan kegiatan proses yang dilakukan secara manual
1000	Proses Komputer	Menunjukkan kegiatan proses yang dilakukan secara komputerisasi
	Arah Aliran Data	Menunjukkan arah aliran dokumen antar bagian yang terkait pada suatu system

Penyimpanan Manual	Menunjukkan media penyimpanan data / infomasi secara manual
Data	Simbol input/output digunakan untuk mewakili data input/output

## **B.** Penelitian Terkait

Penelitian M. Qorib, dkk (2023) pada penelitiannya (Qorib et al., 2023) menjelaskan tentang analisis sentimen dengan dataset pada twitter mengenai keraguan vaksin covid-19. Penelitian ini menggunakan metode Random Forest, Logistics Regression, Decision Tree, LinearSVC dan Naïve Bayes dengan kombinasi Doc2Vec, CountVectorizer dan TF-IDF. Hasil terbaik yang dihasilkan pada penelitian ini yaitu menggunakan eksperimen kombinasi library TextBlob, feature weighting TF-IDF dan metode klasifikasi LinearSVC dengan akurasi sebesar 0.96752.

Kemudian A. M. U. D. Khanday, dkk (2022) pada penelitiannya (Khanday et al., 2022) yang menggunakan fitur ekstraksi mengungkapkan bahwa TF-IDF juga dapat mengoptimalkan proses klasifikasi. Penelitian tersebut mengambil topik ujaran kebencian dengan dataset pada twitter dengan akurasi sebesar 97%.

Penelitian lainnya dilakukan oleh Muzaki dan Witanti yang melihat sentimen masyarakat terhadap pemerintah mengenai pelaksanaan pilkada pada tahun 2020 dengan adanya wabah covid – 19. Data yang didapatkan oleh Muzaki didapatkan pada media sosial twitter dengan hasil evaluasi dari penggunaan

algoritme naïve bayes untuk data training adalah 92,3%, validation 50,79%, data testing 54,1%.

V. A. Fitri, dkk (2019) pada penelitiannya (Fitri et al., 2019) membahas tentang analisis sentimen dengan isu anti LGBT menggunakan metode naïve bayes, decision tree dan random forest. Penelitian tersebut mengungkapkan bahwa metode klasifikasi yang memiliki akurasi terbaik adalah naïve bayes yaitu 86,43%.

Penelitian yang dilakukan oleh Yulita dan kawan – kawan pada analisis sentimen opini masyarakat terhadap vaksin covid – 19 dengan menggunakan algoritme naïve bayes yang menghasilkan opini masyarakat lebih cenderung ber sentimen positif dengan opini positif memiliki persentase sebesar 60,3%, untuk negatif 5.4%, dan netral sebesar 34,4%. Kemudian hasil dari penggunaan algoritme ini menghasilkan akurasi sebesar 0,93 atau 93%.

Penelitian lainnya yang dilakukan Winda dan kawan – kawan adalah melakukan analisa sentimen pada opini masyarakat terhadap vaksinasi covid – 19 dengan algoritme naïve bayes dengan data yang digunakna berasal dari twitter yang didapatkan total jumlah data sebanyak 3.780 tweet. Hasil evaluasi dari penelitian ini berupa opini masyarakat yang dominan terhadap sentimen positif sebesar 2.278 (60,3%), netral 1.299 (34,4%), dan negatif 203 (5,4%). Kemudian nilai evaluasi yang didapatkan setelah menggunakan algoritme naïve bayes adalah 0,93 (93%).

Penelitian selanjutnya dilakukan oleh Santoso dan Nugroho mengenai analisis sentimen yang dilakukan terhadap calon presiden Indonesia pada tahun 2019 yang menggunakan sumber data dari media sosial facebook. Pengumpulan data pada media sosial tersebut dilakukan dengan menggunakan syarat dari tanggal 17 April 2019 hingga 22 May 2019. Hasil dari penelitian ini membuktikan bahwa Joko Widodo memiliki polaritas sentimen yang lebih unggul daripada Prabowo Subianto.

#### **BAB III**

#### METODE PENELITIAN

## A. Tempat dan Waktu Penelitian

Penelitian ini akan di lakukan dengan cara Pengambilan data di media social *Twitter* yang terkait dengan Pilpres2024,waktu yang di gunakan oleh peniliti akan di lakukan maksimal 3 bulan. Bertempat di Ruang Baca Lt 3 Menara IqraUnsimuh Makassar.

#### B. Bahan dan Alat

Adapun alat yang di gunakan dalam penelitian ini yaitu

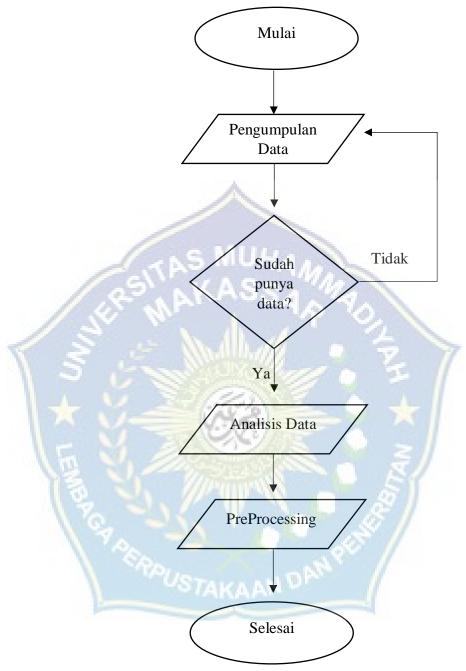
- 1. Kebutuhan Hardware (Perangkat Keras)
  - a. Laptop Asus X441M
  - b. Print
- 2. Kebutuhan Software (Perangkat Lunak)
  - a. Windows 10
  - b. Phyton
  - c. Microsoft Excel 2010
  - d. Goggle Colab

## C. Perancangan Sistem

Untuk mempermudah melakukan analisis sentiment, peneliti merancang Flochart, sehingga analisis sentiment dapat di lakukan secra terstruktur.

#### 1. Flowchart

Flowchart atau bagan alur adalah diagram yang menampilkan Langkahlangkah dan keputusan untuk melakukan sebuah proses dari suatu program. Setiap Langkah di gambarkan dalam bentuk diagram dan di hubungkan dengan garis atau arah panah



Gambar 2. Flowchart System

## D. Teknik Pengujian sistem

Teknik pengujian sistem yang akan di gunakan pada pengujian ini adalah menggunakan Skenario dengan pemisahan data training dan data testing.

Pengujian Data Training dan Testing: Bagian pertama dari pengujian sistem adalah membagi dataset Anda menjadi dua subset, yaitu data training dan data testing. Data training digunakan untuk melatih model Naive Bayes, sedangkan data testing digunakan untuk menguji kinerja model. Pastikan bahwa data testing mencakup beragam sentimen yang ada dalam dataset dan mewakili data yang belum pernah dilihat oleh model.

Salah satu metode pengujian sistem yang di gunakan adalah cross validation. Teknik ini dapat digunakan untuk menguji stabilitas dan konsistensi model Naive Bayes. Dalam pengujian cross-validation, dataset Anda dibagi menjadi beberapa subset yang saling bergantian sebagai data training dan testing. Misalnya, dengan menggunakan metode "k-fold cross-validation", dataset Anda dibagi menjadi k subset, di mana setiap subset digunakan sebagai data testing satu kali dan subset yang lain digunakan sebagai data training. Hasil pengujian di setiap iterasi dapat digabungkan untuk mendapatkan pengukuran kinerja yang lebih robust.

Pengujian Confusion Matrix: Teknik ini digunakan untuk memvisualisasikan dan menghitung jumlah klasifikasi yang benar dan salah untuk setiap kelas sentimen. Confusion matrix akan memberikan gambaran yang lebih lengkap tentang kinerja model dalam mengklasifikasikan sentimen. Dari confusion matrix, Anda dapat menghitung akurasi, presisi, recall, dan F1-score untuk masingmasing kelas sentimen.

Pengujian Akurasi: Teknik ini bertujuan untuk mengukur sejauh mana model dapat mengklasifikasikan sentimen dengan benar. Hitung akurasi model dengan membandingkan hasil klasifikasi sentimen dari model dengan label sentimen yang sebenarnya pada data testing. Akurasi dapat dihitung dengan menggunakan persamaan: Berikut merupakan rumus perhitungan akurasi confusion matrix (Rahmad et al., 2020).

#### E. Teknik Analisis Data

Teknik analisis data adalah proses mencari data, Menyusun secara sistematis data yang di peroleh dari hasil wawancara,catatan lapangan, dan dokumentasi, dengan cara mengorganisasikan data ke dalam kategori, menjabarkan ke dalam unit unit, melakukan sintesis, Menyusun ke dalam pola memilih mana yang penting yang akan di pelajari dan membuat kesimpulan hingga mudah di pahami oleh diri sendiri maupun orang lain.

Adapun Langkah – Langkah untuk menganilisis data dalam penelitian ini adalah sebagai berikut

## 1. Pengumpulan Data

Pengumpulan data dari media socia di kumpulkan menjadi satu untuk di evakuasi dan di bentuk agar menjadi sebuah penelitian, pengambilan data ini atau biasa di sebut proses crawling data twitter ini menggunakan *API Key* Twitter. *Crawling dara di Twitter* dapat di lakukan menggunakan by keyword yaitu pencarian menggunakan penggalan kata kuci dengan total tweet yang di unduh dalam sekali proses maksimun 100 Tweet. Sedangkan pencarian dengan by user yitu pencarian berdasarkan nama akun user di twitter dengan total tweet yang di unduh dalam sekali proses maksimum 200 tweet.

### 2. Preprocessing

Preprocessing merupakan suatu tahapan yang bertujuan untuk memudahkan dalam proses pengolahan data untuk diolah pada tahapan selanjutnya. Proses preprocessing dilakukan untuk menyeleksi data serta mengubahnya menjadi data yang lebih terstruktur. Proses preprocessing membuat data yang banyak dan tidak terpakai akan tereliminasi terlebih dahulu sebelum dataset diproses

#### 3. Analisis Data

Pada tahapan analisis, hal yang di lakukan adalah menganalisis hasil dari perhitungan dan kode program menggunakan analisis sentiment. Analisis sentiment merupakan bagian dari Tekt mining yang sangat sering di lakukan, Analisis sentiment adalah bidang ilmu yang menganalisis opini orang — orang, sentimen, evaluasi, dan emosi terhadap produk, layanan, individu, organisasi, masalah, topik, peristiwa tertentu.

Analisis sentiment di lakukan untuk mengetahui arah polaritas kalimat opini, sehingga dapat di temukan anggota penyusun dari kelompok kalimat positif, netral maupun negative, terdapat dua fungsi dalam klasifikasi sentiment yang di gunakan untuk mengklasifikasi atau memprediksi tweet, yaitu Polarityy dan subjective. Subjectivity digunakan untuk melihat value dari tweet itu adalah opini atau faktual. Yang artinya semakin tinggi subjectivity maka tweet tersebut bisa didefinisikan sebagai opini. Sedangkan Polarity digunakan untuk melihat seberapa positif atau negatif sebuah teks. Sehingga semakin tinggi polarity, maka dapat dikatakan feel/emotion dari sebuah tweet semakin positif juga. Setelah semua tweet sudah diklasifikasi, polarity dari semua tweet akan ditampilkan dalam bentuk rata-rata dan label 'Positif', 'Netral' atau 'Negatif' sesuai dengan mean dari polarity semua tweet.

### 4. Evaluasi

Evaluasi terhadap suatu model analisis umumnya dilakukan dengan membandingkan nilai positif, netral dan negatif agar dapat digunakan untuk menarik kesimpulan dari hasil klasifikasinya. Data yang digunakan untuk menghitung nilai adalah data uji yang sudah diketahui sentimennya.

Nilai Positif (%) = 
$$\frac{Tweet\ Positif}{Total\ Hasil\ Sentimen} x\ 100\ \%$$

Nilai Netral (%) = 
$$\frac{Tweet\ Netral}{Total\ Hasil\ Sentimen} x\ 100\ \%$$

Nilai Negatif (%) = 
$$\frac{Tweet\ Negatif}{Total\ Hasil\ Sentimen} x\ 100\ \%$$

Selanjutnya hasil data akan divisualisasi menjadi sebuah grafik untuk mempermudah dalam membaca hasil analisis sentimen. Hasil evaluasi ini akan menjadi gambaran opini masyarakat dalam menyambut Pemilihan Umum Indonesia tahun 2024 yang akan mendatang



#### **BAB IV**

#### HASIL DAN PEMBAHASAN

## A. Deskripsi Data

Data diakses dari media sosial Twitter berupa Tweet dalam bentuk file berekstensi.csv dengan memasukkan kata kunci yang diinginkan. Kata kunci yang digunakan dalam penelitian ini adalah yang berhubungan dengan tokoh berpotensi dalam Pilpres 2024 di Indonesia, seperti "Anies Presiden 2024", "Ganjar Presiden 2024", dan "Prabowo Presiden 2024". Data didapatkan menggunakan pustaka bahasa pemrograman python, yaitu SNScrape. SNScrape memungkinkan pengguna melakukan pengambilan data lebih dari 7 hari dan memerlukan akun Developer Twitter terlebih dahulu untuk menggunakannya. Data yang berhasil didapatkan oleh Pustaka pemrograman python ini yaitu 30.000 data. Data yang berhasil didapat tersebut terbagi kedalam 10.000 data untuk masing-masing kandidat (Anies Baswedan, Ganjar Pranowo, dan Prabowo Subianto). Berikut ringkasan data mentah yang berhasil di ambil oleh peneliti.

Tabel 2 Tabel Ringkasan Data Mentah

No	Text
1	b'@Mi9aWempy Kolaborasi pemenang pemilu 2024'
2	b'Minggu,30 April 2023\nRakor PPS, SEKERTARIAT PPS, PKD, dan Kepala Desa terkait persiapan pleno DPSHP pada Pemilu tahun 2024. https://t.co/6mSJh9gcc1'
3	b'@sociotalker Wow warbiasah \n\nBagi mereka, tidaklah cukup Anda mencoblos Jokowi 3x saat pemilu. Anda harus menyembah, tegak lurus nderek Jokowi, tutup mata thdp kebobrokan kebijakan rezimnya, tetap memuja senantiasa. \n\nJokowi pemimpin sempurna jaya jaya \xf0\x9f\xa4\xaa\xf0\x9f\x98\x86\xf0\x9f\xa4\xaa'
4	b'Bentrok TNI-Polri, Pengamat Intelijen Ingatkan Jangan Ada Pihak yang Bermain Jelang Pemilu 2024\nhttps://t.co/La12r58Quk'
5	b'Yosafat juga menyampaikan syarat administrasi apa saja yang perlu disiapkan partai politik saat mengajukan nama- nama bakal calon anggota legislatifnya. Juga demikian halnya bagi peserta pemilu perseorangan (DPD) saat mendaftarkan dirinya.'
6	b'@rmol_id Banyak orang justru meyakini prabowo kalah, pemilu 2024 sbagai pelampiasan dendam pendukung prabowo yg dulu dikhianati prabowo, mereka akan mencoblos asal bukan prabowo, jumlah mereka nyaris separo republik ini'
7	b'Gantung Mati Pelaku Kecurangan Pemilu sebab kesalahannya amat berat terhadap konstitusi dan rakyat jadi bila ada yg mau melakukan kecurangan maka layak dipidanakan dengan dakwaan Pengkhianat Konstitusi shg layak digantung mati.\n#TolakKetuaKPUCurang \n#TolakKetuaKPUCurang https://t.co/tYj61n0lrQ'

## **B.** Data Preprocessing

preprocessing adalah tahap yang krusial dalam penelitian ini. proses preprocessing data ini bertujuan untuk memastikan bahwa data yang digunakan dalam analisis sentimen dengan metode Naive Bayes adalah bersih, konsisten, dan siap digunakan. Proses ini dimulai dengan pengumpulan data dari data yang sudah dapatkan.

Tabel 3. Tabel DataSet Bersih

No	Text
1	Semoga Pak Anies Baswedan Makin Banyak Yang Mendukung Dan Yg Mendoakan Beliau Hingga Presiden Ri Biar Menghanguskan Rencana Jahat Mereka Hempaskan Orang Yg Penuh Halusinasi Biar Stres Sendiri Si Nganu Yg Tidak Bisa Kerja
2	Tni Mau Minta Bantuan Banser Atau Apaan Kek Rakyat Akan Kawal Anies Agar Ada Restorasi Indonesia Yang Adil Sejahtera Indonesia Baru Keadilan Ditegakkan Rakyat Disejahterakan Kawal For Presiden No Other No Else
3	Makar Alloh Swt Lebih Bagus Dan Berlaku In Sya Alloh Pak Anies Baswedan The Next Presiden Ri Aamiinx Yra
4	Sebelumnya Ketua Majelis Syura Partai Ummat Amien Rais Menyatakan Sepakat Bakal Mendukung Prabowo Subianto Untuk Maju Sebagai Calon Presiden Capres Jika Anies Baswedan Gagal Maju Atau Tidak Dapat Tiket Capres
5	Yang Mengusung Anies Baswedan Sebagai Calon Presiden Di Pilpres Prabowo Subianto
6	Semoga Kalau Anies Jadi Presiden Orang Ini Dicopot Dari Jabatannya
7	Gerakan Muda Terus Bergerak Merangkul Merajut Tenun Kebangsaan Indonesia Perubahan Indonesia Sejahtera Anies Presiden

## C. Pelabelan

Tahapan pelabelan dilakukan untuk memberikan pembagian kelas positif negatif dan netral terhadap data Tweet. Banyak teknik yang dapat digunakan untuk pelabelan bentimen, diantaranya Transformer, TextBlob, dan VADER. Namun, pelabelan pada sentimen juga dapat dilakukan secara manual . Karena pelabelan secara manual menunjukkan tingkat akurasi lebih tinggi dan akurat.

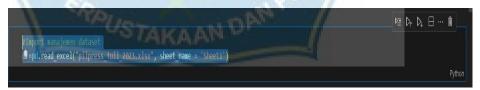
Tabel 4. Tabel dataset yang sudah di label

No	Text	Label
1	Anies Baswedan Itu Mengayomi Seluruh Lapisan Masyarakat Nkri Banget Tanpa Harus Teriak Saya Nkri Sangat Cocok Untuk Indonesia Yg Majemuk	1
2	Anies Mendadak Batal Jadi Calon Presiden Tahun Oalah Padahal Dana Akselerasi Ke Beberapa Negara Sudah Habis Banyak	2
3	Doaku Semoga Pak Ganjar Pranowo Presiden	3
4	Dengan Melepas Kaus Bertuliskan Ganjar Presiden Kelompok Relawan Gp Mania Resmi Membatalkan Dukungannya Kepada Ganjar Pranowo	4
5	Dukungan Terhadap Prabowo Subianto Sebagai Calon Presiden Capres Periode Terus Mengalir Kali Ini Dukungan Datang Dari Komunitas Para Ibu Atau Emakemak Di Jember Jawa Timur	5
6	Elektabilitas Calon Presiden Capres Untukprabowo Subiantodi Provinsi Jawa Barat Menjadi Yang Terrendah	6
7	Bentrok TNI-Polri, Pengamat Intelijen Ingatkan Jangan Ada Pihak Yang Bermain Jelang Pemilu 2024	7

## D. Proses Penerapan Metode Naive Bayes

Proses menggunakan Metode Naive Bayes melibatkan serangkaian langkah untuk membuat, melatih, dan mengevaluasi model Naive Bayes. Berikut adalah penjelasan umum tentang tahapan proses tersebut:

Tahap pertama yaitu proses Import Data set dari Excel ke dalam variabel data frame DF pada gambar 4.2 dan selanjutnya akan di lakukan pembacaan dan menampilkan tampilan data frame df seperti pada gambar 4.3.

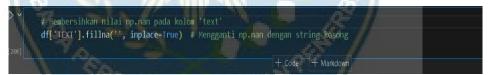


Gambar 3. import data set



Gambar 4. tampillan data

 Selanjutnya adalah proses Cleaning data untuk data dari kolom label yang berisi data error atau NaN (String Kosong)data error atau NaN (String Kosong). Proses cleaning data ini merupakan langkah kritis dalam analisis data yang bertujuan untuk meningkatkan kualitas data dan menghilangkan potensi gangguan yang dapat mempengaruhi hasil analisis.



Gambar 5. proses cleaning

Pembagian data bframe untuk variabel x sebagai inputan text (dari kolom TEXT) dan y sebagai label atau target (dari kolom LABEL). Dimana variabel x dideskripsikan sebagai wadah untuk menampung teks atau informasi yang akan di jadikan input dalam analisis teks. Kolom TEXT dapat berisi teks atau dokumen yang akan di analisis, dalam konteks ini adalah analisis sentimen. Sedangkan variabel y di tetapkan sebagai wadah untuk menampung label iatau target yang ingin kita prediksi atau analisis,

kolom LABEL dapat berisi kategori atau label yang akan dihubungkan dengan atau di prediksi berdasarkan analisis text pada variabel x. Deskripsi variabel x dan y dari gambar 7 dari gambar bisa di lihat bahwa jumlah data sesuai input text dan label sebanyak 20825



Gambar 6. pembagian data frame kolom



Gambar 7. jumlah data sesuai dengan pengimputan

Selanjutnya adalah proses spliting, Setelah mempersiapkan variabel x sebagai input teks dan y sebagai label atau target, langkah berikutnya adalah membagi data menjadi dua bagian utama: data training dan data testing. Pembagian ini dilakukan untuk melatih model pada sebagian data dan menguji kinerja model pada bagian lainnya. Pembagian ini mengikuti proporsi 90% untuk data training dan 10% untuk data testing. Berikut merupakan keterangan

x\_train adalah variabel input untuk data training (Latih)

x\_test adalah variabel input untuk data test (uji)

y\_train adalah variabel target untuk data training (Latih)

y\_test adalah variabel output untuk data test (uji)

```
from sklearn.model_selection import-train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y,random_state=0,test_size=0.1)
```

Gambar 8. pembagian data training dan testing

• Selanjutnya adalah tahap tokenisasi dan vektorisasi data train dengan tfidfvectorizer(). Proses tokenisasi dan vektorisasi bertujuan untuk mengubah teks menjadi representasi numerik sehingga dapat digunakan oleh model machine learning.sedangkan TF-IDF membantu dalam mengekstraksi informasi penting dari setiap teks, memberikan bobot berdasarkan frekuensi dan signifikansi kata dalam korpus dokumen.

```
vectorizer = TfidfVectorizer()
x_train_vec = vectorizer.fit_transform(x_train)
x_test_vec = vectorizer.transform(x_test)
```

Gambar 9. proses tokenisasi, vektorisasi menggunakan tf-idf

• Setelah melakukan proses tokenisasi dan vektorisasi pada data training, langkah selanjutnya adalah menguji data testing (x\_test) ke dalam model Naive Bayes yang telah dilatih sebelumnya. Proses ini bertujuan untuk mengevaluasi sejauh mana model dapat melakukan prediksi dengan akurat terhadap data yang belum pernah dilihatnya sebelumnya. Dari gambar di bawa ini terlihat akurasi mencapai 64-65 //

```
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score

#accuracy = accuracy_score(y_test, y_pred)
#print("Akurasi: {:.2f}%".format(accuracy * 100))

# Hitung metrik evaluasi
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='macro')
recall = recall_score(y_test, y_pred, average='macro')

# fampilkan hasil evaluasi
print(f'Accuracy: {accuracy}')
print(f'Accuracy: {accuracy}')
print(f'Recuracy: {accuracy}')
print(f'Recuracy: {accuracy}')
print(f'Reall: {recall}')
print(f'Reall: {recall}')

# Accuracy: 8.6476236197791647
Precision: 0.6561488116619648
Recall: 0.6496209209513613
F1 Score: 0.6502678263682623
```

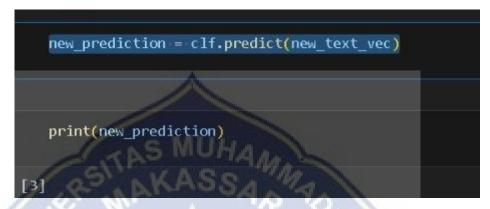
Gambar 10. proses uji coba

Tahap Pengujian dengan data baru yang coba dimasukkan, "presiden jokowi masuk ke area masjid sheikh zayed seusai salat jumat terlihat presiden jokowi keluar melalu pintu vip dengan gubernur jawa tengah ganjar pranowo" Untuk text di atas maka label sebenarnya adalah 3 atau Positif Ganjar Pranowo



Gambar 11. proses uji coa kata secara manual

 Selanjutnya dilakukan pengujian dengan model naive bayes dan ternyata prediksi yang dilakukan benar yakni 3 atau Ganjar Pranowo Positif



Gambar 12. proses uji coba naive bayes

selanjutnya menampilkan data hasil prediksi perbandingan antara True
 Label (Label Asli) dan Predicted Label (Label Prediksi)



Gambar 13. hasil prediksi

Index: 0

Text: nah gini kan bagus postingan kau lae adem gak usah nyinyirin anies lagi gak usah bikin gaduh lagi biarlah mereka bersaing dng ide dan gagasannya siapapun yg jadi presiden udh kehendak tuhan

True Label: 1
Predicted Label: 2
Index: 1 STAS MIUHA
Text: sir dimana ada gula disitu ada semut dimana ada pak anies yg manie
disitu ada masyarakat yang bahagia menyambutnya anies presiden
indonesia sejahtera
True Label: 1
Predicted Label: 1
Index: 2

Text: hasil survei prabowo sebagai calon presiden dengan elektabilitas paling tinggi bisa menjadi penyemangat masyarakat capres dekade mendingprabowo terusmajubersamaprabowo terbukti kerja prabowo mending prabowo

True Label: 5
Predicted Label: 5

Index: 3
Text: Sepertinya Pemilu 2024 Golput aja untuk pertama kalinya
True Label: 7
Predicted Label: 7
Index: 4
Text: ngaku presiden tapi koq ngomong nya kayak broker impor cuman
orang goblok yang masih percaya penipu berkedok pemimpin yang
merakyat apanya yang berhasil ngaku aja bersama rakyat nyatanya
ninggalin rakyat tenggelam hukumnya wajib tenggelamkan prabowo di
True Label: 6
Predicted Label: 2
Index: 5
Text: waras bismillah pilih anies baswedan presiden
True Label: 1
Predicted Label: 1
Index: 6
Text: survei media survei nasional median mencatat elektabilitas ketua
umum partai gerindra prabowo subianto berada di posisi pertama sebagai

calon presiden di pilpres setyoko

True Label: 6
Predicted Label: 6
Index: 7
Text: komunitas nelayan pesisir ganjar pranowo sulsel kabupaten sinjai
festival nelayan amp deklarasi presiden ganjar menang nelayan sejahterah
presiden
True Label: 3
Predicted Label: 3
Index: 8
Text: ujung dr ikhtiar kita utk memenangkan mas adalah dgn mengetuk
pintu langit spt yg dilakukan tmnx sekber jateng aamiin kan setiap doa yg
terbaik utk mas anies agar terpilih jd presiden ri di aamiin sir relawan
True Label: 2
Predicted Label: 1
STAKAANDA
Index: 9
Text: Relawan sedulur saklawase ajak seluruh lapisan masyarakat kawal
pemilu aman dan damai di pesta demokrasi 2024
True Label: 7
Predicted Label: 7

Index: 10
Text: sy malah gak pilih pdip tapi saya tetap pilih pak jokowi jd presider
klo ganjar capres sy akan pilih tapi tidak pdip
True Label: 4
Predicted Label: 2
SAS MUHAM
Index: 11 S KASS
Text: sekarang saya merasa bersyukur dan saya merasa bangga telah
bergabung dengan presiden joko widodo imbuh prabowo subianto capres
pemilu dekade terusmajubersamaprabowo mendingprabowo saatnya
prabowo lihat dekade
00
True Label: 5
Predicted Label: 5
Index: 12
STAKAAND
Text: laksono sekarang presiden kita masih megawati ntar mau puan atau
ganjar juga presiden kita tetep megawati
True Label: 3
Predicted Label: 4

Index: 13 Text: berita menhan jokowi mania dukung prabowo subianto subianto presiden tribun timur True Label: 6 Predicted Label: 6 Index: 14 Text: asmara luuuaaar biaaaasaaaa sip demokrat ahy wapresanies presiden True Label: 2 Predicted Label: 2 Index: 15 Text: Tahapan & amp True Label: 7 Predicted Label: 7

Index: 16

Text: madali sukses selalu untuk nasdem mengusung anies baswedan calon presiden ri dgn slogan perubahan segalanya perlu perubahan demi rakyat indonesia yg lebih sejahtera menuju negara yg berwibawa

True Label: 1

Index: 17
Text: Mari Jaga Netralitas ASN Dalam Pemilu 2024
True Label: 7
Predicted Label: 7
Index: 18 S MUHA
Text: Coba jelaskan dg ilmiah kematian ratusan petugas kpps saat pemilu
2019 lalu bisa tidak
True Label: 7
Predicted Label: 7
Index: 19
Text: dengan menggunakan baju putih presiden jokowi masuk ke area
masjid sheikh zayed seusai salat jumat terlihat presiden jokowi kelua
melalu pintu vip dengan gubernur jawa tengah ganjar pranowo
True Label: 3
Predicted Label: 4
<del></del>
Index: 20
Text: asalamuallaikum mudahan tahun hok prahowo subianto terpilil

menjadi presiden indonesia amin yra

True Label: 6

Predicted Label: 6

Index: 21

Text: cebong pada kepanasan anies jadi presiden berbagai cara fitnah di buatnya ya jangan heran sih kemana tuh ngalirnya kalo gak buat bayar orang seperti ini keturuan anda apa kodok atau babiatau anjing ni liat faktanya

True Label: 2

Predicted Label: 2

Index: 22

Text: letjen tni purn bapak prabowo subianto dilantik sebagai menhan oleh presiden republik indonesia pada tanggal oktober dalam kabinet indonesia maju masa jabatan

True Label: 6

Predicted Label: 6 STAKAAN DAN

Index: 23

Text: kegiatan pelatihan pengelolaan air bersih dan doa bersama untuk ganjar pranowo presiden sekaligus buka puasa bersama pondok pesantren asy syuro desa galang kec pinyuh kabupaten mempawah kalbar

True Label: 3

Predicted Label: 4
Index: 24
Text: terlihat rakyat begitu mencintainya anies presiden ri
True Label: 1
Predicted Label: 1
S MUHAMA
Index: 25
Text: tokoh agama dki jakarta bersyukur tiga partai dukung anies sebagai
capres masyarakat jakarta sangat antusias dan berharap serta berdoa anies
menjadi presiden di tahun insaallah suara anies dapat mencapai persen di
jakarta
True Label: 1
Predicted Label: 1
Index: 26
Text: ganjar cocok jadi presiden pranowo
True Label: 4
Predicted Label: 3

Index: 27

Text: Mas Bargaining power kecil itu secara partai tapi secara dukungan rakyat nyatanya sangat besarnPTS 20% pemilu 2024 barubisa mengusung capres di 2029 Udah keburu ancur negara ininLagian emang yakin pemilu bisa jurdil

True Label: 7

Predicted Label: 7

Index: 28

Text: saewad yg penting jokowi dah tamat besok yg jadi presiden mau anies atau prabowo rakyat tetep happy masa jahiliyah masa suram dah lewat

True Label: 6

Predicted Label: 2

Index: 29

Text: pak jokowi jgn sampai pak ganjar pranowo terganjal elitx politikbagaimanapun caranya pak jokowi harus menjadikan pak ganjar presiden penerus bapakkarena pak ganjar itu simbol nkripancasilabhinekatoleranpribumibapak harus berusaha agar pak ganjar presiden dari pribumi

True Label: 3

Index: 30
Text: siap kita ganti presiden
True Label: 5
Predicted Label: 1
Index: 31 S MUHA
Text: dan calon wakil presiden yang akan maju pada pemilihan presiden
prabowo subianto
True Label: 5
Predicted Label: 6
Index: 32
Text: warga laporkan relawan ganjar bagikan kaos dan dompet ganjar
presiden
True Label: 3
Predicted Label: 3
Index: 33
Text: kami sangat yakin pak prabowo akan jadi presiden di ujar anggota
dor ri asal sumbar itu prabowo subianto

True Label: 6

Predicted Label: 6

-----

Index: 34

Text: najeges sesungguhnya presiden sudah ada presiden yang terpilih namun kita tidak tahu siapa dia bapak prabowo subianto bapak ganjar pranowo atau bapak anis baswedan saya pilih bapak anis baswedan manusia hanya bisa berusaha namun allah yang menentukan semoga qobul aamiin

True Label: 6

Predicted Label: 1

\_\_\_\_\_

Index: 35

Text: dan hanya diikuti tiga capres dengan data hasil survei pasangan caprescawapres yang dipilih responden pada pemilu ucapnya prabowo subianto presiden

True Label: 6

Predicted Label: 6

-----

Index: 36

Text: tekad yang kuat semoga aja pak ganjar pranowo bisa maju di pilpres

dan menjadi presiden ri ke

True Label: 3

Index: 37
Text: masyarakat adat papua disebut mendukung prabowo subianto sebagai
calon presiden capres capres pemilu dekade terusmajubersamaprabowo
mendingprabowo prabowo tak gentar mending pasti prabowo
True Label: 5
Predicted Label: 5
CS KASS
Index: 38
Text: kedua tokoh hebat indonesia ini direstui sama pak presiden jokowi
untuk memimpin indonesia pada tahun mendatang
True Label: 5
Predicted Label: 3
Index: 39
Text: kesira gerindra tak kenal lelah melakukan hal baik prabowo presiden
yesss
True Label: 6
Predicted Label: 6

Index: 40

Text: surya paloh pegang kartu as jokowi dinilai tak akan berani geser menterimenteri nasdem jokowi firaun kades umur minions kevin kls lisa shines in riyadh

True Label: 1 Predicted Label: 1 Index: 41 Text: Gw auto teringat pemilu 9297 dimana hanya sihijau dan simerah lah yang konsisten jadi partai politik dijaman Orbaditengah tekanan yg dasyat dari pemerintah saat itunMEREKA ADALAH PARTAI LEGEND True Label: 7 Predicted Label: 7

Index: 42

Text: lembaga survei indonesia lsi merilis tingkat elektabilitas sosok potensial calon presiden capres pada pemilu prabowo subianto

True Label: 5

Predicted Label: 5

Index: 43

Text: Mari Jaga Netralitas ASN Dalam Pemilu 2024

True Label: 7

	Predicted Label: 7
	Index: 44
	Text: kegiatan canvasing door to door di jl pelabuhan talang duku rt desa
	muara kumpeh kec kumpeh ulu kab muaro jambi prov jambi
	True Label: 4
	Predicted Label: 4
	KASSA MA
d	Index: 45
	Text: Pentingnya netralitas ASN dalam pemilu
	True Label: 7
	Predicted Label: 7
	Index: 46
	Text: Saya percaya & amp
	True Label: 7
	Predicted Label: 1
	Index: 47
	Text: siti nurizka anak muda benteng masa depan prabowo presiden
	True Label: 6

Predicted Label: 6
Index: 48
Text: Selamat Memperingati Hari Keterbukaan Informasi Nasional 30
April 2023 Mari bersama saling bersinergi untuk mewujudkan keterbukaan
informasi dan Pemilu 2024 yang jujur dan
terbuka # Ayo Awasi Bersaman # Pemilu Serentak 2024n # Keterbuka an Inform
asiPublik AS MUHA
True Label: 7
Predicted Label: 7
Index: 49
Text: asalamualaikum pak ganjar kami dari bandung pak semoga bapak
ganjar menjadi presiden penerus pak presiden jokowi pada pemilu tahun
True Label: 3
Predicted Label: 3
POUSTAKAAN DAN
Index: 50
Text: samosir semoga pasangan calon presiden amp wakil presiden anies
rasyid baswedan amp agus harimurti yudhoyono menang di pemilu
True Label: 2

Index: 51 Text: selalu berinovasi itulah gamjar pranowo sang penerus jokowi jadi presiden jateng jadi nomor satu indonesia pun akan jadi nomor satu di bawah kepemimpinan ganjar pranowo jangan ada yg iri ya ganjar memang oke True Label: 3 Predicted Label: 3 Index: 52 Text: KPU Kabupaten Sigi melakukan pelayanan kepada Perwakilan Partai Politik Peserta Pemilu Tahun 2024 di kantor KPU Kabupaten Sigi Sabtu 2942024 True Label: 7 Predicted Label: 7 Index: 53 Text: ini bukti keseriusan anies baswedan dan partai koalisi pendukungnya mewujudkan harapan para relawan untuk bersungguhsungguh menjadi kandidat calon presiden di pilpres nanti True Label: 1

Index: 54 Text: embe tetep dukung ganjar jadi presiden palestina True Label: 3 Predicted Label: 3 Index: 55 Text: musyawarah rakyat musra kexiii jawa timur di surabaya menjaring namanama calon presiden yg dinilai layak maju pemilu presiden prabowo subianto jodoh civic arsenal pssi kuat maruf tulus utamakan rakyat cak nun north london ratu tisha True Label: 5 Predicted Label: 5 Index: 56 Text: masyaallah tabarrakallahu pak anies kunjungan yg bagus warga natar adalah warga yg ramah warga yg siap berubah lebih maju menuju keadilan sosial bersama capres anies yg insyaallah dpt dimenangkan menjadi presiden ri pd pilpres aamiin

True Label: 1

Index: 57

Text: presiden jokowi menyampaikan sosok seperti prabowo subianto erick thohir sandi uno hingga mahfud md merupakan tokoh potensial di pilpres mendatang

True Label: 6

Predicted Label: 6

-----

Index: 58

Text: aliansi nasional indonesia sejahtera kabupaten sintang terus bergerak silahturahmi sinegritas memenangkan menjadi presiden bersama tokoh masyarakat ketua parpol kolisi perubahan nasdem pks demokrat simpul relawan anies

True Label: 2

Predicted Label: 1

Index: 59

Text: syaa juga mua pak bapak saya doakan jadi presiden

True Label: 6

Predicted Label: 2

\_\_\_\_\_

Index: 60

Text: Adapun Daerah Pemilihan dan Alokasi Kursi Anggota DPRD Kabupaten Enrekang untuk Pemilu Tahun 2024 tertuang dalam PKPU

Nomor 6 Tahun 2023

True Label: 7

Predicted Label: 7

-----

Index: 61

Text: prabowo subianto mengatakan bakal melanjutkan kebijakan jokowi bila dia terpilih menjadi presiden dalam pilpres capres pemilu dekade mendingprabowo

True Label: 5

Predicted Label: 5

Index: 62

Text: salut dan bangga dengan berbagai aksi positif yang dilakukan oleh relawan usbat semoga relawan lain juga turut melakukan berbagai aksi yang bagus ini demi perjalanan pak ganjar menuju pemilihan presiden tahun

True Label: 3

Predicted Label: 3

-----

Index: 63
Text: gibran lempar pujian untuk pak prabowo
True Label: 5
Predicted Label: 5
Index: 64
Text: anies baswedan presiden indonesia
True Label: 6
Predicted Label: 1
5 5 5 5 6 5 6 5 6 6 6 6 6 6 6 6 6 6 6 6
Index: 65
Text: alhamdulillah rabbal aalamiin capres yang gagah semoga allah swt
mengangkat derajat bang anies rasyid baswedan menjadi presiden republik
indonesia tahun dan seterusnya agar keadilan dan kesejahteraan dapat
dinikmati oleh seluruh rakyat indonesia
True Label: 1
Predicted Label: 1
Index: 66

Text: deklarasikan ganjar suai gerakan rakyat riau ini dukung ganjar

pranowo jadi presiden ri

True Label: 3

Predicted Label: 4

-----

Index: 67

Text: Siapa yg disini waktu pemilu serentak 2019 ngak milih MPRDPRRI DPRD dan DPD hanya milih presiden dan wapres Kalo saya hanya pilih presiden dan wapres saja yg lain pass karena ngak kenal n males nyoblos

partainya

True Label: 7

Predicted Label: 7

\_\_\_\_

Index: 68

Text: ganjar pranowo presiden aamin amanda begal windah moana

demokrat doge morning

True Label: 3

Predicted Label: 3

Index: 69

Text: partai demokrat menyatakan kader partai gerindra sandiaga uno tak memenuhi kriteria calon wakil presiden cawapres pendamping anies baswedan di pilpres

True Label: 1

-----

Index: 70

Text: sudah saatnya untuk pak ganjar pranowo maju di pilpres dan menjadi presiden ri ke karena semua kalangan masyarakat mengidamidamkan beliau untuk menjadi presiden ri

True Label: 4

Predicted Label: 3

• Hasilnya kemudian di eksport ke dalam file Excel "hasil prediksi"



Gambar 14. ekpor data hasil prediksi

Tabel 5. Dataset Hasil Prediksi

NO	TEXT	LABEL	HASIL PREDIKSI
1.	Dimana Ada Gula Disitu Ada Semut Dimana Ada Pak Anies Yg Manies Disitu Ada Masyarakat Yang Bahagia Menyambutnya Anies Presiden Indonesia Sejahtera	1	1
2.	Hasil Survei Prabowo Sebagai Calon Presiden Dengan Elektabilitas Paling Tinggi Bisa Menjadi Penyemangat Masyarakat Capres Dekade Mendingprabowo Terusmajubersamaprabowo Terbukti Kerja Prabowo Mending Prabowo	5	5
3.	Sepertinya Pemilu 2024 Golput Aja Untuk Pertama Kalinya	7	7
4.	Ngaku Presiden Tapi Koq Ngomong Nya Kayak Broker Impor Cuman Orang Goblok Yang Masih Percaya Penipu Berkedok Pemimpin Yang Merakyat Apanya Yang Berhasil Ngaku Aja Bersama Rakyat Nyatanya Ninggalin Rakyat Tenggelam Hukumnya Wajib Tenggelamkan Prabowo	6	2
5.	Waras Bismillah Pilih Anies Baswedan Presiden	511	2
100		§"//	
2084.	Kenapa Pemilu Saling Serang Bukannya Tunjukin Aja Lu Yang Paling Baik	7	7

## E. Teknik Pengujian Sistem

Teknik pengujian sistem memainkan peran krusial untuk memastikan kualitas dan kehandalan sistem. Proses ini melibatkan serangkaian lagkah langkah yang di rancung untuk mengevaluasi berbagai aspek kinerja dan fungsionalitas system.

Pengujian Data Training dan Testing: Bagian pertama dari pengujian sistem adalah membagi dataset menjadi dua subset, yaitu data training dan data testing. Data training digunakan untuk melatih model Naive Bayes, sedangkan data testing digunakan untuk menguji kinerja model. data testing mencakup beragam sentimen yang ada dalam dataset dan mewakili data yang belum pernah dilihat oleh model.

Pengujian Confusion Matrix: Teknik ini digunakan untuk memvisualisasikan dan menghitung jumlah klasifikasi yang benar dan salah untuk setiap kelas sentimen. Confusion matrix akan memberikan gambaran yang lebih lengkap tentang kinerja model dalam mengklasifikasikan sentimen. Dari confusion matrix, Anda dapat menghitung akurasi, presisi, recall, dan F1-score untuk masing-masing kelas sentimen. Adapun pengujian yang di hasilkan. Accuracy: 64 %, Precission: 65 %, Recall: 65 %, F1 Score: 65 %.

#### **BAB V**

#### **PENUTUP**

## A. Kesimpulan

Berdasarkan penelitian yang telah dilaksanakan maka dapat disimpulkan sebagai berikut:

- Data calon Presiden didapatkan melalui platform Twitter sejumlah 30.000 data untuk dianalisis.yang dimana data tersebut terdapat 10.000 di masingn masing kandidat.
- 2. Tahapan preprocessing dilakukan untuk membersihkan data agar mudah diterapkan pada metode. Data yang digunakan untuk dianalisis berjumlah 20.800 setelah melalui tahapan preprocessing.
- 3. Rata-rata akurasi untuk berdasarkan pengujian pada metode Naïve Bayes yaitu 64-65 %, Precission :65 %, Recall:65 %, F1 Score :65 %.
- 4. Interpretasi penelitian sebagai pembuktian bahwa metode naive bayes mampu melakukan perhitungan akurasi yang akurat dan menjadi proyeksi Pilpres tahun 2024 seperti sentimen dan sebaran suara.

#### B. Saran

Saran yang dapat diberikan terkait penelitian ini diharapkan mampu menghilangkanakun atau Tweet yang berasal dari pengguna berbayar (bot) dan pengguna berbayar (buzzer).dan lebih meningkatkan prosesing data yang lebih akurat. Dikarenakan pembahasan terkait politik, tidak menutup kemungkinan hal tersebut banyak digunakan untuk menggiring opini publik di media sosial. Hasil data yang bersih perlu dimiliki untuk selanjutnya dianalisis, salah satunya *noise* yang ada pada data yaitu ketersediaan akun bot dan buzzer yang perlu dihilangkan.

#### **DAFTAR PUSTAKA**

- Fais Sya' bani, Muhammad Raihan, Ultach Enri, and Tesa Nur Padilah. 2022. "Analisis Sentimen Terhadap Bakal Calon Presiden 2024 Dengan Algoritme Naïve Bayes." *JURIKOM (Jurnal Riset Komputer)* 9 (2): 265. https://doi.org/10.30865/jurikom.v9i2.3989.
- Febryanti, Dian Rousta, Zaki Khudzaifi Mahmud, and Santika Vania Putri. 2022. "Tipologi Hate Speech Di Twitter Terkait Kebijakan Pemerintah Selama Pandemi COVID-19" 11 (2).
- Muzaki, Akhmad, and Arita Witanti. 2021. "Sentiment Analysis of the Community in the Twitter To the 2020 Election in Pandemic Covid-19 By Method Naive Bayes Classifier Sentimen Analisis Masyarakat Di Twitter Terhadap Pilkada 2020 Ditengah Pandemic Covid-19 Dengan Metode Na I"Ve Bayes Classifier" 2 (2): 101–7.
- Nardilasari, Aprilia Putri, April Lia Hananto, Shofa Shofia Hilabi, and Bayu Priyatna. 2024. "Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM" 7 (1): 11–18.
- Perwira, Aditya, Joan Dwitama, and Kata Kunci. 2021. "Deteksi Ujaran Kebencian Pada Twitter Bahasa Indonesia Menggunakan Machine Learning: Reviu Literatur." *Jurnal SNATi* 1 (1): 31–39.
- Qorib, Miftahul, Timothy Oladunni, Max Denis, Esther Ososanya, and Paul Cotae. 2023. "Analisis Sentimen Dengan Dataset Pada Twitter Mengenai Keraguan Vaksin Covid-19 Menggunakan Metode Random Forest." *Expert Systems with Applications* 212 (August 2022): 118715. https://doi.org/10.1016/j.eswa.2022.118715.
- Rahayu, A. 2019. "ANALISIS SENTIMEN DEBAT PILPRES 2019 MENGGUNAKAN METODE NAIVE BAYES." *Http://E-*

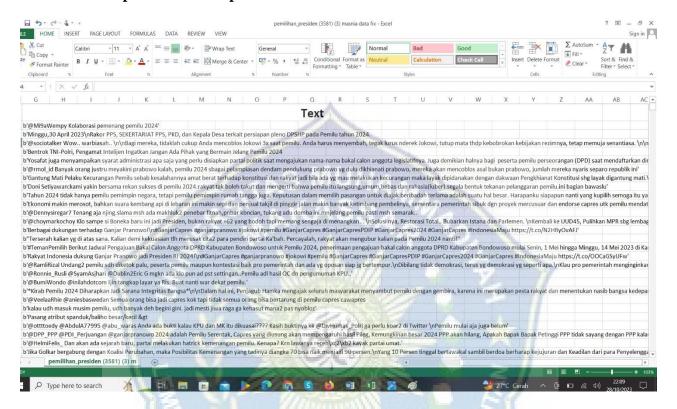
- *Journal.Uajy.Ac.Id*/7244/4/3TF03686.Pdf, no. 2010: 15–48. http://e-journal.uajy.ac.id/7244/4/3TF03686.pdf.
- Rasyadi, Muhammad Hadiyan. 2017. "Analisis Sentimen Pada Twitter Menggunakan Metode Naïve Bayes (Studi Kasus Pemilihan Gubernur Dki Jakarta 2017)," 1–17.
- Santoso, Eko Budi, and Aryo Nugroho. 2019. "Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook." *Eksplora Informatika* 9 (1): 60–69. https://doi.org/10.30864/eksplora.v9i1.254.
- Syarli, and Asrul Ashari Muin. 2016. "Metode Naive Bayes Untuk Prediksi Kelulusan." *Jurnal Ilmiah Ilmu Komputer* 2 (1): 22–26. https://media.neliti.com/media/publications/283828-metode-naive-bayes-untuk-pre
- diksi-kelulu-139fcfea.pdf.
- Vindua, Raditia, and Achmad Udin Zailani. 2024. "Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python" 10 (2): 479–87. https://doi.org/10.30865/jurikom.v10i2.5945.
- Yulita, Winda. 2021. "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier." *Jurnal Data Mining Dan Sistem Informasi* 2 (2): 1. https://doi.org/10.33365/jdmsi.v2i2.1344.
- Amirullah, Fuad, Syariful Alam, and M. Imam Sulistyo S. 2024. "Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini Twitter Menggunakan Naïve Bayes." 2(3):69–76.
- Cetak, Issn, and Issn Online. 2024. "DECODE: Jurnal Pendidikan Teknologi Informasi." 3(2):236–45.
- Kurniawan, Wahidin Deni. 2023. "PERJUANGAN PADA TWITTER MENGGUNAKAN METODE NAÏVE ANALYSIS OF PUBLIC SENTIMENT ON PDI PERJUANGAN ON TWITTER USING THE NAÏVE BAYES METHOD." 2(September):1127–34.

Pendahuluan, I. 2024. "Penerapan Algoritma Naive Bayes Dan PSO Pada Analisis Sentimen Kandidat Calon Presiden 2024." 7:1367–80.

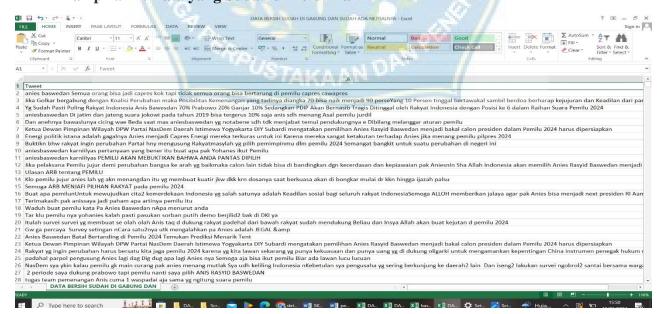


#### **LAMPIRAN**

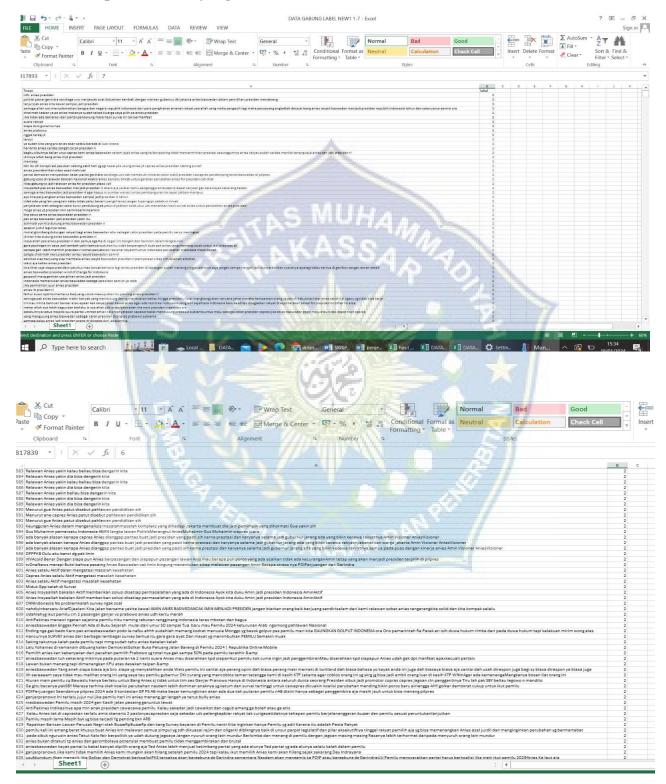
### Lampiran 1. Deskripsi Data

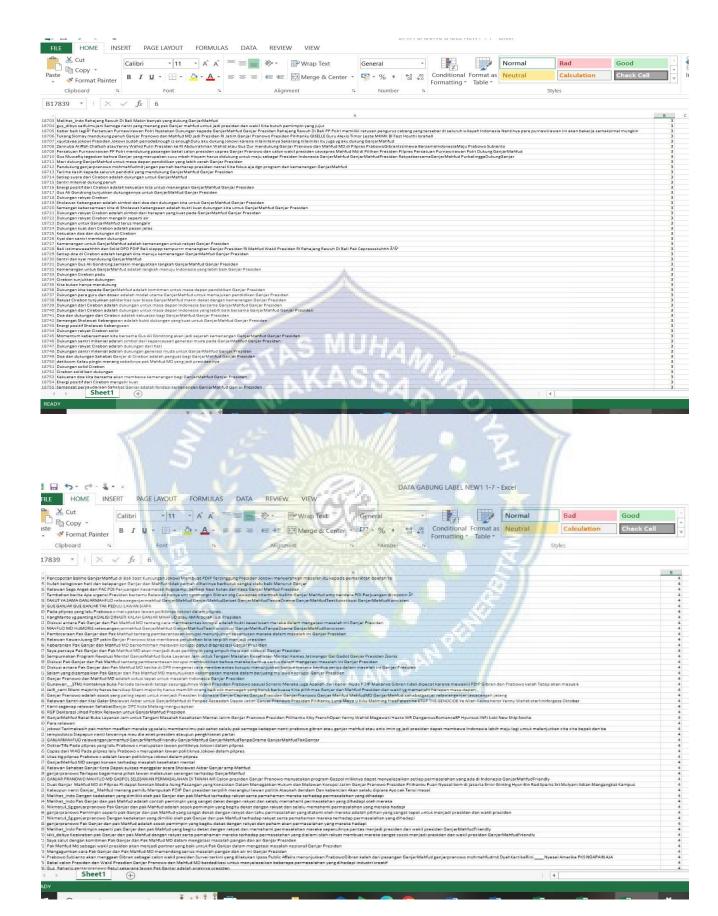


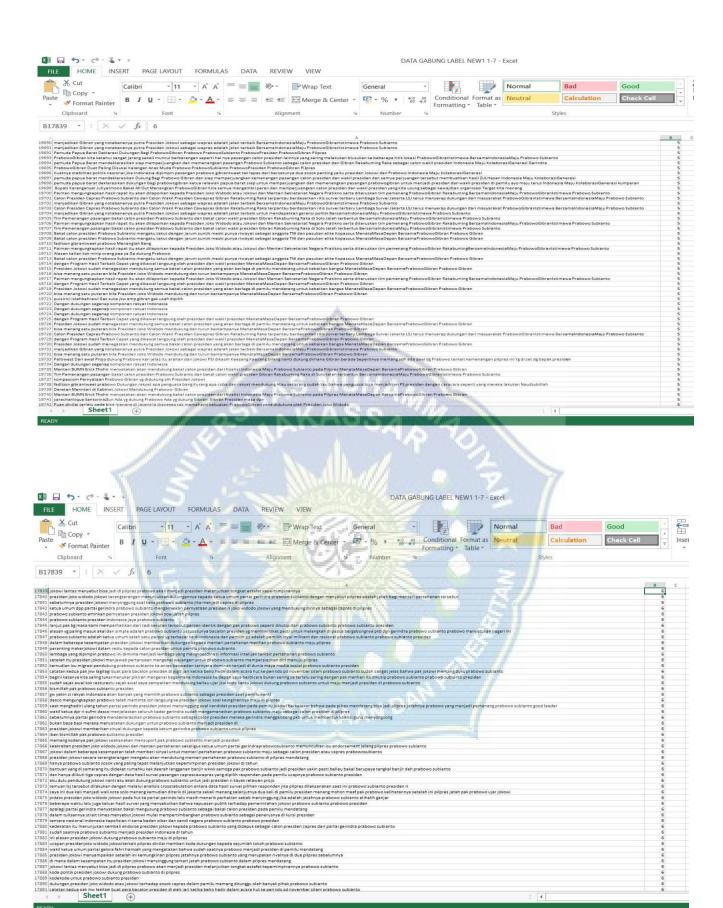
## Lampiran 2. Data yang Sudah di Bersihkan

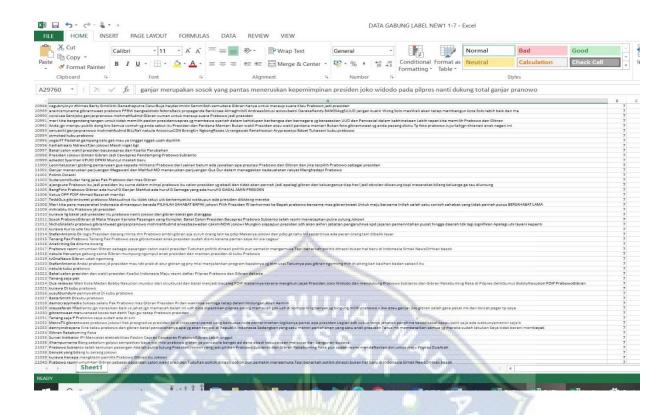


# Lampiran 3. Data yang sudah di Label









#### Lampiran 4. Tampilan Data



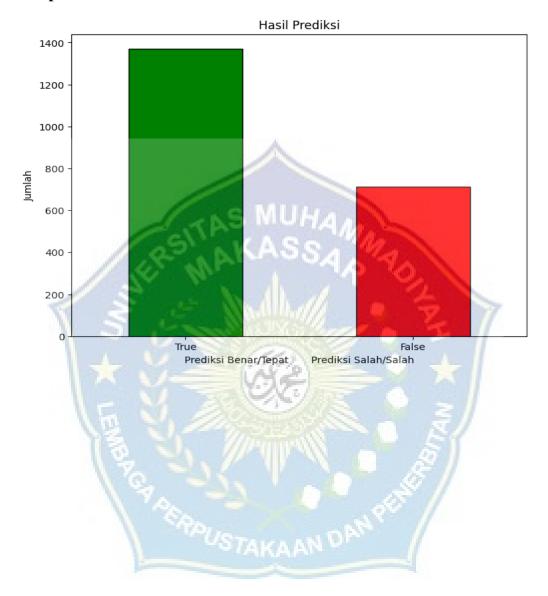
#### Lampiran 5. Source Code

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
import sklearn
warnings.simplefilter("ignore")
#Import dataset
df=pd.read excel("pilpress full 2023
df
df.describe()
x = df['TEXT']
y = df['LABEL']
x.describe()
from sklearn.model selection import train test split
x_train, x_test, y_train, y_test =
train test split(x,y,random state=0,test size=0.1)
from sklearn.naive bayes import MultinomialNB
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics import accuracy score
import nltk
from nltk.tokenize import word tokenize
vectorizer = CountVectorizer()
x train vec = vectorizer.fit transform(x train)
x test vec = vectorizer.transform(x test)
print(x_test)
non_string_values = df[df['TEXT'].apply(lambda x: not isinstance(x,
str))]['TEXT']
print("Non-string values in 'TEXT' column:")
print(non_string_values)
clf = MultinomialNB()
clf.fit(x train vec, y train)
y_pred = clf.predict(x_test_vec)
v pred
```

```
sklearn.metrics
from
                           import
                                     accuracy score,
                                                        precision score,
recall_score, f1_score, confusion_matrix, classification_report
# Prediksi menggunakan model yang telah dilatih
y_pred = clf.predict(x_test_vec)
# Menghitung beberapa metrik evaluasi
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='weighted')
recall = recall_score(y_test, y_pred, average='weighted')
f1 = f1_score(y_test, y_pred, average='weighted')
conf_matrix = confusion_matrix(y_test, y_pred)
class_report = classification_report(y_test, y_pred)
# Menampilkan hasil
print(fAccuracy: {accuracy:.2f}')
print(f'Precision: { precision:.2f}')
print(f'Recall: {recall:.2f}')
print(f'F1 Score: {f1:.2f}')
# Data baru dalam bentuk string
new_text = ["Ganjar presiden"]
new_text_vec = vectorizer.transform(new_text)
print(new text vec)
new_prediction = clf.predict(new_text_vec)
print(new prediction)
for i, (text, true_label, predicted_label) in enumerate(zip(x_test, y_test,
y_pred)):
  print(f"Index: {i}")
  print(f"Text: {text}")
  print(f"True Label: {true_label}")
  print(f"Predicted Label: {predicted_label}")
  print("-----
```

```
# Membuat DataFrame dari X_test, y_test, dan hasil prediksi
df_results = pd.DataFrame({'Text': x_test, 'True Label': y_test, 'Predicted
Label': y_pred})
# Menyimpan DataFrame ke file Excel
df_results.to_excel('hasil_prediksi.xlsx', index=False)
import matplotlib.pyplot as plt
# Membuat DataFrame untuk menghitung jumlah prediksi yang benar dan
salah
                                         df results['True Label']
df_results['Correct
df_results['Predicted Label']
result_counts = df_results['Correct Prediction'].value_counts()
# Membuat diagram batang
plt.figure(figsize=(8, 6))
result_counts.plot(kind='bar', color=['green', 'red'], edgecolor='black')
plt.title('Hasil Prediksi')
                                    Prediksi Salah/Salah')
plt.xlabel('Prediksi Benar/Tepat
plt.ylabel('Jumlah')
plt.xticks(rotation=0)
plt.show()
```

## Lampiran 6 Grafik



Adapun Hasil presentasi dari hasil data yang telah di uji berdasarkan data uji yaitu :

Anies Positif = 14%

Anis Negatif = 13%

Prabowo Positif = 14%

Prabowo Negatif = 14 %

Ganjar Positif =14%

Ganjar Negatif =16%

Netral = 14%





#### MAJELIS PENDIDIKAN TINGGI PIMPINAN PUSAT MUHAMMADIYAH UNIVERSITAS MUHAMMADIYAH MAKASSAR UPT PERPUSTAKAAN DAN PENERBITAN

Alamat kantor: Jl.Sultan Alauddin NO.259 Makassar 90221 Ttp.(0411) 866972,881593, Fax (0411) 865588



#### SURAT KETERANGAN BEBAS PLAGIAT

UPT Perpustakaan dan Penerbitan Universitas Muhammadiyah Makassar, Menerangkan bahwa mahasiswa yang tersebut namanya di bawah ini:

Nama

: Masnia

Nim

: 105841100918

Program Studi: Teknik Informatika

Dengan nilai:

No	Bab	Nilai	Ambang Batas
1	Bab I	9%	10 %
2	Bab 2	23 %	25 %
3	Bab 3	8%	10 %
4 .	Bab 4	0%	10 %
5	Bab 5	5 %	5%

Dinyatakan telah lulus cek plagiat yang diadakan oleh UPT- Perpustakaan dan Penerbitan Universitas Muhammadiyah Makassar Menggunakan Aplikasi Turnitin.

Demikian surat keterangan ini diberikan kepada yang bersangkutan untuk dipergunakan seperlunya.

> Makassar, 10 Januari 2024 Mengetahui,

Kepala UP7 n dan Pernerbitan,

Jl. Sultan Alauddin no 259 makassar 90222 Telepon (0411)866972,881 593,fax (0411)865 588 Website: www.library.unismuh.ac.ld E-mail: perpustakaan@unismuh.ac.id



Submission date: 10-Jan-2024 05:30AM (UTC+0700)

Submission ID: 2268584255

File name: TURNITIN\_BAB\_I.docx (29.83K)

Word count: 1135 Character count: 7473

### BAB I Masnia 105841100918





Submission date: 10-Jan-2024 05:31AM (UTC+0700)

Submission ID: 2268584522

File name: TURNITIN\_BAB\_II.docx (140.76K)

Word count: 1427 Character count: 9283

## BAB II Masnia 105841100918

23% SIMILARITY INDEX	19% INTERNET SOURCES	8% PUBLICATIONS	20% STUDENT PAPERS
PRIMARY SOURCES			
repositor Internet Sour	ori.uin-alauddin.	ac.id	6%
2 Submitt Student Pape	ed to Universita	s Islam Indone	esia 5 <sub>%</sub>
3 WWW.Sli Internet Sour	deshare.net		3 / 4%
4 Ip2m.un Internet Sour	na.ac.id		2%
reposito	ory.pelitabangsa	.ac.id	2%
6 Submitt Student Pape	ed to Universita	s Brawijaya	2%
7 Submitt Student Pape	ed to Sriwijaya l	University	2%
Exclude quotes Exclude bibliography	Off	Exclude matches	< 2%

# BAB III Masnia 105841100918

by TutupTahap



Submission date: 10-Jan-2024 05:32AM (UTC+0700)

Submission ID: 2268584763

File name: TURNITIN\_BAB\_III.docx (44.84K)

Word count: 945 Character count: 5950

## BAB III Masnia 105841100918

ORIGINALITY REPORT			
8% SIMILARITY INDEX	4% INTERNET SOURCES	4% PUBLICATIONS	6% STUDENT PAPERS
PRIMARY SOURCES			
Cuaca k	ti Dwiyanti, Cahy Kota Jakarta Mer n Forest", Jurnal	nggunakan Me	etode
2 eprints.	upj.ac.id	OAP A	2,
Perpust	ed to Forum Ko akaan Pergurua sia (FKPPTKI)		2 <sub>9</sub>
4 Submitt Student Pape	ed to Universita	s Brawijaya	£ 29
Exclude quotes Exclude bibliography	Off Off	AN DAN	< 2%



Submission date: 10-Jan-2024 05:42AM (UTC+0700)

Submission ID: 2268588033

File name: TURNITIN\_BAB\_IV.docx (1.46M)

Word count: 1981 Character count: 12093

### BAB IV Masnia 105841100918





Submission date: 10-Jan-2024 05:32AM (UTC+0700)

Submission ID: 2268585018

File name: TURNITIN\_BAB\_V.docx (26.58K)

Word count: 175 Character count: 1126

## BAB V Masnia 105841100918

