

# Analisis Sentimen Masyarakat X Terhadap Kepemimpinan Presiden Prabowo Subianto 2025 Menggunakan Algoritma *Naïve Bayes*

<sup>1</sup>Albert Julio Tampubolon, <sup>2</sup>Wita Manik, <sup>3</sup>Michael Ginting, <sup>4</sup>Sardo Pardington Sipayung  
Universitas Katolik Santo Thomas, Medan, Indonesia  
[tampubolonmurini@gmail.com](mailto:tampubolonmurini@gmail.com), [witasihotang14@gmail.com](mailto:witasihotang14@gmail.com), [michaelgginting@gmail.com](mailto:michaelgginting@gmail.com),  
[pinsarsiphom@gmail.com](mailto:pinsarsiphom@gmail.com)

Submit : 22 Des 2025 | Diterima : 26 Jan 2026 | Terbit : 28 Jan 2026

## ABSTRAK

Penelitian ini dirancang untuk mengevaluasi pandangan publik terhadap gaya kepemimpinan Presiden Prabowo di tahun 2025 melalui penerapan metode Naive Bayes. Sumber data utama berasal dari platform Twitter yang mengandung ragam perspektif dan respon masyarakat terkait strategi pemerintahan, inisiatif pembangunan, dan karakteristik kepemimpinan yang ditunjukkan. Metodologi penelitian mencakup beberapa fase utama: koleksi data, proses pembersihan dataset, pengolahan teks menggunakan pendekatan Natural Language Processing (NLP), serta kategorisasi sentimen ke dalam dua klasifikasi utama yaitu sentimen positif dan negatif. Implementasi algoritma Naive Bayes digunakan untuk menganalisis dataset yang telah melalui proses pelabelan manual. Hasil evaluasi menunjukkan model mampu mengklasifikasikan sentimen dengan akurasi 72,22%, dengan kinerja yang lebih baik pada identifikasi sentimen positif dibanding negatif. Meskipun akurasi masih perlu ditingkatkan, hasil ini memberikan indikasi awal yang cukup memadai untuk memahami kecenderungan opini publik terhadap kepemimpinan Presiden Prabowo. Temuan ini dapat menjadi fondasi untuk penelitian lanjutan dengan optimalisasi model dan penambahan fitur untuk meningkatkan performa klasifikasi sentimen pada analisis opini digital.

**Kata Kunci:** Analisis Sentimen, Naive Bayes, Presiden Prabowo, Natural Language Processing, Data Twitter

## PENDAHULUAN

Sebagai negara yang berada di kawasan Asia Tenggara, Indonesia menerapkan sistem pemerintahan presidensial dengan presiden yang berperan ganda sebagai kepala negara dan kepala pemerintahan. Kedudukan presiden dalam tata kelola negara Indonesia dilandasi oleh Undang-Undang Dasar 1945 yang menempatkan presiden sebagai pemegang otoritas tertinggi dalam hierarki pemerintahan. Pasal 7 UUD 1945 menetapkan bahwa periode jabatan Presiden dan Wakil Presiden adalah lima tahun dengan kemungkinan terpilih kembali untuk satu periode tambahan. Ketentuan ini membatasi masa kepemimpinan presiden maksimal dua periode atau selama sepuluh tahun.

Berakhirnya dua periode kepemimpinan Presiden Joko Widodo menandai dimulainya era baru dalam dinamika politik Indonesia. Hasil pemilihan umum 2024 menetapkan pasangan Prabowo Subianto dan Gibran Rakabuming Raka sebagai presiden dan wakil presiden yang terpilih. Pasca pengumuman hasil tersebut, reaksi publik terhadap kepemimpinan Prabowo mulai bermunculan di berbagai platform media sosial, khususnya Twitter, yang sering dijadikan sarana bagi masyarakat untuk menyampaikan perspektif, pendapat, dan evaluasi terhadap perkembangan isu politik. Ekspresi yang disampaikan umumnya menunjukkan polarisasi antara sentimen positif berupa dukungan dan sentimen negatif berupa kritik atau penolakan (Al Muqsith Prasetyo & Hermawan, 2025).

Perkembangan platform digital telah menghadirkan transformasi signifikan terhadap cara berkomunikasi dan distribusi informasi selama beberapa dekade terakhir. Twitter merupakan salah

satu wadah yang populer digunakan masyarakat untuk mengekspresikan opini dan ide-ide mereka.. Dengan ciri khas pesan singkat, Twitter memfasilitasi pengguna untuk berbagi ide, menyebarkan berita, dan berinteraksi dengan masyarakat luas secara langsung. Hal ini menjadikan Twitter sebagai salah satu media efektif untuk memantau opini masyarakat terhadap berbagai isu, termasuk kepemimpinan tokoh publik seperti Presiden Prabowo Subianto. Twitter memberi ruang bagi pengguna untuk mengikuti perkembangan terkini, berpartisipasi dalam diskusi, serta menyuarakan pandangan mereka mengenai kepemimpinan dan kebijakan pemerintah. Di Indonesia, Twitter digunakan secara luas untuk berbagi informasi, mengorganisasi kegiatan sosial, dan menyampaikan aspirasi atau kritik terhadap pemimpin nasional. Dalam konteks ini, Twitter menjadi wadah penting dalam memahami bagaimana sentimen masyarakat terbentuk dan berkembang terhadap kepemimpinan Presiden Prabowo Subianto (Diana Dwi Rahayu et al., 2024).

Analisis sentimen merupakan cabang ilmu yang mempelajari bagaimana seseorang menyampaikan opini, sikap, penilaian, evaluasi, hingga emosinya terhadap sebuah topik, layanan, produk, individu, organisasi, atau aktivitas tertentu (August et al., 2024). Analisis sentimen bertujuan untuk mengidentifikasi apakah sebuah opini atau komentar terkait suatu permasalahan cenderung bersifat positif atau negatif. Hasil dari analisis ini dapat digunakan sebagai dasar untuk memperbaiki layanan atau meningkatkan mutu produk yang ditawarkan (Nurjanah et al., 2017).

Penelitian ini memiliki beberapa tujuan utama. Pertama, untuk mengungkap pola dan kecenderungan opini publik terkait kepemimpinan Presiden Prabowo tahun 2025 berdasarkan analisis teks dari unggahan dan komentar di platform media sosial X (sebelumnya Twitter). Kedua, untuk mengukur performa akurasi metode Naive Bayes dalam proses klasifikasi sentimen terhadap dataset yang digunakan. Ketiga, untuk mengidentifikasi sentimen positif maupun negatif yang secara langsung berkaitan dengan Presiden Prabowo. Dengan pendekatan Naive Bayes, penelitian ini bertujuan menggali dan memahami pola-pola sentimen yang dominan muncul, sehingga dapat memberikan gambaran menyeluruh mengenai pandangan masyarakat terhadap kepemimpinan Presiden Prabowo (Amanda et al., 2024).

Berbagai penelitian terkait analisis sentimen terhadap tokoh politik telah dilakukan menggunakan algoritma yang beragam. Algoritma Support Vector Machine (SVM) digunakan oleh (Sari et al., 2023) untuk menganalisis sentimen publik terhadap kebijakan pemerintah dan menghasilkan tingkat akurasi sebesar 78,5%. Selanjutnya, (Wijaya & Kusuma, 2024) menerapkan algoritma Random Forest dalam konteks analisis sentimen pemilihan presiden dan mencatat akurasi sebesar 82,3%. Penelitian lainnya oleh (Rahman et al., 2024) menggunakan algoritma K-Nearest Neighbor (KNN) dan memperoleh akurasi 75,6% dalam klasifikasi sentimen politik.

Pendekatan berbasis deep learning juga menunjukkan hasil yang menjanjikan. Studi yang dilakukan oleh (Pratama & Sari, 2024) menerapkan metode Long Short-Term Memory (LSTM) dan memperoleh nilai akurasi maksimal mencapai 85,2%, meskipun membutuhkan sumber daya komputasi yang besar dan dataset yang kompleks. Sementara itu, (Handayani et al., 2023) menggunakan algoritma Decision Tree dengan akurasi 73,8%, namun algoritma ini diketahui rentan terhadap overfitting.

Dalam konteks analisis sentimen berbahasa Indonesia, algoritma Naive Bayes menunjukkan performa yang konsisten dan efisien. (Nugroho & Wijaya, 2024) melaporkan bahwa Naive Bayes mampu memberikan hasil stabil dengan akurasi 81,7%, berkat kemampuannya dalam menangani data berdimensi tinggi serta performa yang baik pada dataset berukuran kecil hingga menengah. Keunggulan tersebut menjadi salah satu alasan pemilihan algoritma ini dalam penelitian yang dilakukan.

## TINJAUAN PUSTAKA

### Analisis Sentimen

Analisis sentimen adalah metode kategorisasi teks yang diterapkan untuk mengklasifikasikan konten teks berdasarkan kecenderungan atau arah sentimen yang terkandung di dalamnya, apakah bernada positif atau negatif. Analisis ini termasuk dalam metode *data mining* yang digunakan untuk mengolah opini atau pendapat seseorang dengan bantuan *Natural Language Processing* (NLP), komputasi linguistik, dan analisis teks. Dalam penelitian ini, penentuan sentimen dilakukan secara manual dengan membaca setiap opini dan memberikan label sentimen

positif atau negatif sesuai makna yang terkandung dalam teks. Tujuan analisis ini adalah untuk mengidentifikasi polaritas sentimen masyarakat terhadap kepemimpinan Presiden Prabowo Subianto pada tahun 2025 (Permana et al., 2025).

## X

X merupakan layanan media sosial microblogging yang dikembangkan untuk mendistribusikan informasi dengan format ringkas, efektif, dan waktu nyata melalui batasan karakter di bawah 140. Layanan ini difungsikan untuk menyebarkan informasi kepada seluruh pengguna, baik yang memiliki hubungan personal maupun yang tidak dikenal sebelumnya. X memfasilitasi penggunaannya dalam menulis dan membagikan aktivitas atau pendapat mereka, serta menjadi salah satu sumber utama data mentah dalam analisis sentimen. Dengan jumlah pengguna aktif di Indonesia yang sangat besar, X memungkinkan pemetaan distribusi dan kekuatan sentimen melalui fitur retweet, like, dan hashtag. Selain itu, X mempermudah pemantauan respons publik terhadap peristiwa besar secara cepat karena pesan yang diunggah bersifat real-time dan singkat (Satriyo, 2022).

## Klasifikasi Naïve Bayes

Naïve Bayes dikenal sebagai salah satu metode klasifikasi yang sering dimanfaatkan dalam analisis sentimen karena keunggulannya dalam hal kesederhanaan, kecepatan, serta tingkat akurasi yang baik dalam mengelompokkan data. Dalam salah satu studi, metode ini berhasil digunakan untuk menganalisis data dari platform X dengan fokus pada pencarian merek mobil terlaris menurut data GAIKINDO, dan menghasilkan akurasi hingga 93% dengan precision positif dan negatif sebesar 90% serta sentimen netral mencapai 100%. Hal ini menunjukkan kemampuan Naïve Bayes dalam memberikan hasil analisis dengan akurasi yang tinggi.

Metode ini juga sering digunakan untuk memprediksi probabilitas, karena memiliki karakteristik statistik yang menyerupai pola pikir seorang pakar dan didukung landasan matematika yang kuat. Dalam pengujiannya pada diagnosa penyakit dan hama tanaman jagung, metode ini menunjukkan akurasi sebesar 90%.

Naïve Bayes Classifier memberikan proses klasifikasi yang efisien dan mudah dipahami algoritmanya. Pada berbagai penelitian, metode ini terbukti memiliki kinerja yang baik untuk klasifikasi data berkat kesederhanaan dalam perhitungannya. Salah satu penerapannya adalah dalam mendukung proses seleksi penerima beasiswa, di mana penggunaan Naïve Bayes mampu mempercepat proses penentuan calon penerima dan meminimalkan kesalahan, dengan akurasi mencapai 80%. Selain itu, algoritma ini juga mampu menangani data berukuran besar, sehingga sangat sesuai untuk diterapkan dalam analisis sentimen pada tweet terkait Prabowo-Gibran (Satriyo, 2022).

## Algoritma Naïve Bayes

Algoritma *Naive Bayes* adalah salah satu metode machine learning yang termasuk dalam kategori *supervised learning* dan memiliki popularitas tinggi, yang mendasarkan kerjanya pada konsep *Teorema Bayes* dengan asumsi bahwa setiap fitur bersifat independen. Metode ini beroperasi dengan menggunakan dataset yang sudah memiliki label untuk mempelajari pola-pola dalam proses klasifikasi. Istilah *Naive* merujuk pada anggapan bahwa setiap fitur atau karakteristik dalam data dianggap saling bebas satu sama lain, sehingga masing-masing fitur memberikan kontribusi secara independen terhadap hasil klasifikasi. Sementara itu, istilah *Bayes* mengacu pada penggunaan Teorema Bayes sebagai dasar perhitungannya. Salah satu keunggulan utama dari algoritma ini adalah kemampuannya melakukan klasifikasi meskipun hanya menggunakan sejumlah kecil data pelatihan. Secara umum, *Naive Bayes* berfungsi untuk memprediksi probabilitas dari setiap kelas yang mungkin, sehingga dapat digunakan untuk mengklasifikasikan data secara efisien, bahkan dalam kasus-kasus dengan kompleksitas tinggi di dunia nyata (Sofyan & Iqbal, 2025).

## METODE PENELITIAN

### Jenis Penelitian

Penelitian ini dikategorikan sebagai studi kuantitatif dengan menggunakan metode deskriptif. Studi ini memanfaatkan data tekstual dari tweet pengguna yang diolah untuk mengidentifikasi orientasi sentimen masyarakat terhadap kepemimpinan Presiden Prabowo. Melalui proses pengolahan data secara sistematis dan penerapan algoritma *Naive Bayes*, penelitian ini bertujuan mengklasifikasikan opini publik ke dalam kategori sentimen positif, negatif, sehingga diperoleh gambaran umum tentang persepsi masyarakat di media sosial.

### Pengumpulan Data

Pada penelitian ini, data sentimen diambil dari konten tweet berbahasa Indonesia yang diunggah secara langsung oleh para pengguna di platform media sosial Twitter. Twitter dipilih karena merupakan platform terbuka yang memungkinkan siapa pun menyampaikan opini, memberikan tanggapan, menyukai, atau membagikan ulang suatu unggahan. Platform ini sering dimanfaatkan masyarakat untuk menyuarakan pandangan terhadap isu-isu aktual, termasuk kepemimpinan nasional. Untuk mendukung fokus penelitian terhadap persepsi publik terhadap Presiden Prabowo pada tahun 2025, data dikumpulkan menggunakan kata kunci seperti “kepemimpinan Prabowo”, “Presiden Prabowo”, dan istilah relevan lainnya yang berkaitan dengan topik tersebut (Al Muqsith Prasetyo & Hermawan, 2025).

Spesifikasi Pengumpulan Data :

1. Kata Kunci : “Kepemimpinan Prabowo”
2. Rentang Waktu : 1 January 2025 – 20 Juni 2025
3. Bahasa Indonesia (lang:id)
4. Jumlah data awal: 1.200 tweet
5. Sumber : LATEST tab
6. Platform: Google Collab + Node.js
7. Link Notebook: [https://colab.research.google.com/drive/1RyMMM0I4e-3XniX\\_fILfIoqnytBuB-4m](https://colab.research.google.com/drive/1RyMMM0I4e-3XniX_fILfIoqnytBuB-4m)

### Pre-processing

Tahap *pre-processing* merupakan langkah lanjutan setelah proses pelabelan data secara manual. Pada tahap ini, data mentah yang telah diperoleh melalui metode *crawling* akan diproses dan disiapkan agar dapat digunakan dalam sistem analisis. Proses *pre-processing* sangat penting dalam pengolahan data berbasis teks, karena setiap karakter akan diberi bobot yang akan digunakan dalam perhitungan klasifikasi. Tahapan ini juga berfungsi untuk meningkatkan kualitas data dan mengoptimalkan kinerja algoritma *Naive Bayes* saat diterapkan. Tahap preprocessing dilakukan untuk memastikan kualitas dataset melalui lima tahapan sistematis filtering:(Al Muqsith Prasetyo & Hermawan, 2025), yaitu:

### Filtering

Tahapan selanjutnya dalam proses pra-pemrosesan data pada penelitian ini adalah stopword removal, yaitu proses penyaringan kata-kata dalam teks untuk mengurangi jumlah kata yang tidak memberikan makna signifikan terhadap analisis. Stopword merupakan kumpulan kata yang sering muncul dalam kalimat namun tidak memiliki bobot penting dalam penentuan sentimen, seperti kata ganti, kata penghubung, kata tanya, atau kata seruan. Dengan menghapus kata-kata tersebut, teks menjadi lebih ringkas dan fokus pada kata-kata yang memiliki nilai informasi lebih tinggi dalam konteks analisis sentimen terhadap kepemimpinan Presiden Prabowo tahun 2025. Proses penghapusan stopwords ini dilakukan menggunakan library Python bernama Sastrawi, yang mendukung bahasa Indonesia. Hasilnya, teks yang dianalisis menjadi lebih bersih dan efisien untuk diproses oleh algoritma *Naive Bayes* (Steven et al., 2025).

### Pelabelan Data Manual

Pelabelan data secara manual diperlukan karena algoritma *Naive Bayes* termasuk dalam metode *supervised learning*, yang mengharuskan setiap data mempunyai label atau kategori target

yang jelas. Proses pelabelan ini tidak dapat dilakukan secara sembarangan, karena dibutuhkan pemahaman yang memadai untuk dapat mengidentifikasi dan menentukan jenis sentimen dengan tepat. Oleh karena itu, tingkat literasi tertentu sangat diperlukan agar sentimen yang terkandung dalam data dapat diklasifikasikan secara akurat. Tabel 1 menyajikan contoh data hasil *crawling* yang telah melalui proses pelabelan sentimen.

Tabel 1. Sampel Data yang Digunakan

Data Ke	Sentimen	Deskripsi
1	Papua tetap bagian Indonesia makin mantap karena komunitas internasional juga dukung via kepemimpinan Presiden Prabowo	Positif
2	Di masa kepemimpinan Presiden Prabowo Subianto Indonesia benar-benar berdiri di atas kaki sendiri soal pangan	Negatif

Tabel 1 menyajikan contoh data hasil *crawling* dari media sosial Twitter yang telah melalui proses pelabelan sentimen secara manual. Setiap entri merupakan opini masyarakat yang secara langsung menyinggung kepemimpinan Presiden Prabowo. Sentimen dari masing-masing pernyataan dikategorikan ke dalam kelas positif dan negatif berdasarkan makna yang terkandung dalam teks. Pelabelan ini bertujuan untuk membedakan tanggapan masyarakat yang mendukung ataupun mengkritik kepemimpinan Presiden Prabowo, sebagai bagian dari proses pelatihan model *Naive Bayes* dalam analisis sentimen.

### Pembagian Dataset

Dataset yang telah dipreprocessing dibagi menjadi 2 bagian :

1. Data training : 80% dari total dataset (492 tweet)
2. Data testing : 20% dari total dataset (122 tweet)

### Klasifikasi

Dalam konteks pembelajaran mesin, klasifikasi merupakan proses memprediksi kategori atau label dari suatu data berdasarkan fitur-fitur yang dimilikinya. Pada penelitian ini, klasifikasi dilakukan dengan mengubah data tweet menjadi sentimen, yang dikelompokkan ke dalam kategori positif dan negatif. Proses klasifikasi tersebut dilakukan menggunakan pendekatan *Naive Bayes Classification* untuk menentukan kecenderungan opini masyarakat terhadap topik yang dianalisis (Sofyan & Iqbal, 2025).

### Evaluasi Model

Penilaian kinerja model dilakukan dengan menggunakan metode *confusion matrix*, yang meliputi beberapa metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*. Evaluasi ini bertujuan untuk mengukur seberapa efektif model *Naive Bayes* dalam mengelompokkan sentimen masyarakat secara tepat sesuai dengan kategori yang sebenarnya (No et al., 2025).

### Visualisasi Dan Analisis Data

Hasil klasifikasi sentimen terhadap kepemimpinan Presiden Prabowo akan divisualisasikan menggunakan diagram batang dan diagram lingkaran untuk menggambarkan proporsi masing-masing kategori sentimen, yaitu positif dan negatif. Selain itu, *word cloud* juga digunakan untuk menampilkan kata-kata yang paling sering muncul dalam setiap kategori sentimen. Analisis lanjutan dilakukan guna mengidentifikasi isu-isu utama yang banyak dibicarakan masyarakat, seperti penilaian terhadap kebijakan pemerintah, kekhawatiran terhadap transparansi, serta pandangan mengenai peran teknologi dalam kepemimpinan Prabowo (No et al., 2025).

## HASIL DAN PEMBAHASAN

### Hasil Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan dengan metode *crawling* menggunakan API dari Twitter Developer yang telah terdaftar sebelumnya. Proses ini dijalankan melalui platform Google Colab, dengan bantuan bahasa pemrograman Python serta library Tweepy untuk mengakses dan mengambil data teks dari Twitter. Awalnya diperoleh sebanyak 1.200 tweet, namun setelah

melalui proses penyaringan dan seleksi data yang tidak relevan, jumlah data yang digunakan dalam analisis akhir berkurang menjadi 614 tweet. Dari total tersebut, 434 dikategorikan sebagai sentimen positif dan 180 sebagai sentimen negatif, sebagaimana ditampilkan pada Tabel 2.

Tabel 2. Hasil *Crawling* Data

No	Text
1	@genidaa41665 Pekerja punya tameng ekstra lewat satgas PHK di bawah kepemimpinan Prabowo
2	@KotaNusantara Harapan besar buat sektor pertanian Seluma di bawah kepemimpinan Presiden Prabowo dan Brigade Pangan
.....	.....
1200	@reaareaa_ Presiden Prabowo tunjukkan kepemimpinan sejati dengan dukung penuh upaya pemerintah basmi premanisme

Hasil *crawling* data menggunakan Tweepy menghasilkan satu atribut utama, yaitu teks. Atribut ini berisi isi tweet yang menjadi fokus utama dalam analisis sentimen. Melalui data teks inilah proses pelabelan sentimen dilakukan secara manual, untuk kemudian dianalisis menggunakan metode klasifikasi *Naive Bayes* dalam mengidentifikasi persepsi masyarakat terhadap kepemimpinan Presiden Prabowo (Al Muqstith Prasetyo & Hermawan, 2025).

### **Pre-processing data**

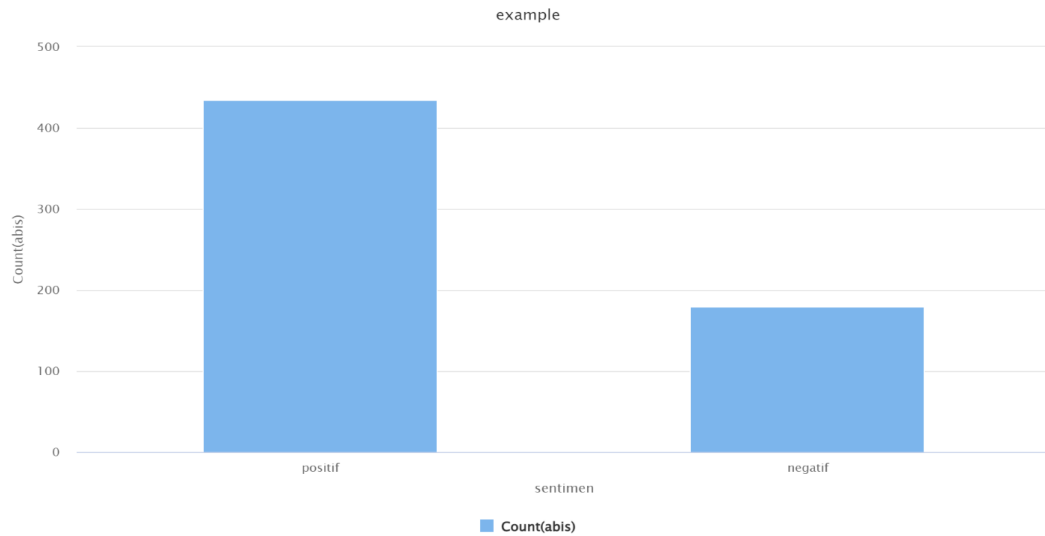
Data yang telah dikumpulkan dan diberi label secara manual selanjutnya diproses dalam tahap *pre-processing*. Pada tahap ini, data mentah diolah dan dikonversi menjadi data yang siap digunakan dalam analisis. *Pre-processing* terdiri dari enam langkah utama, di mana setiap langkah memiliki peran dan fungsi tersendiri dalam mempersiapkan data. Hasil dari masing-masing tahapan *pre-processing* tersebut dapat dilihat pada Tabel 3.

Tabel 3. Contoh Hasil *pre-processing*

No	Tahapan Preprocessing	Deksripsi	Jumlah Tweet Tersisa
1	Data awal	Data hasil <i>crawling</i> dengan <i>search</i> ="Kepemimpinan Prabowo"	1.300
2	Menghapus tweet retweet	Menggunakan <i>parameters</i> = "RT"	1.293
3	Menghapus Duplikasi	Menghapus tweet yang sama	868
4	Menghapus tweet karakter aneh dan link	Menggunakan <i>parameters</i> = "@.*?" dan "https.*"	866
5	Menghapus tweet yang berisi link	Menggunakan <i>parameters</i> = "@.*?" dan "https.*"	614

### **Pelabelan data**

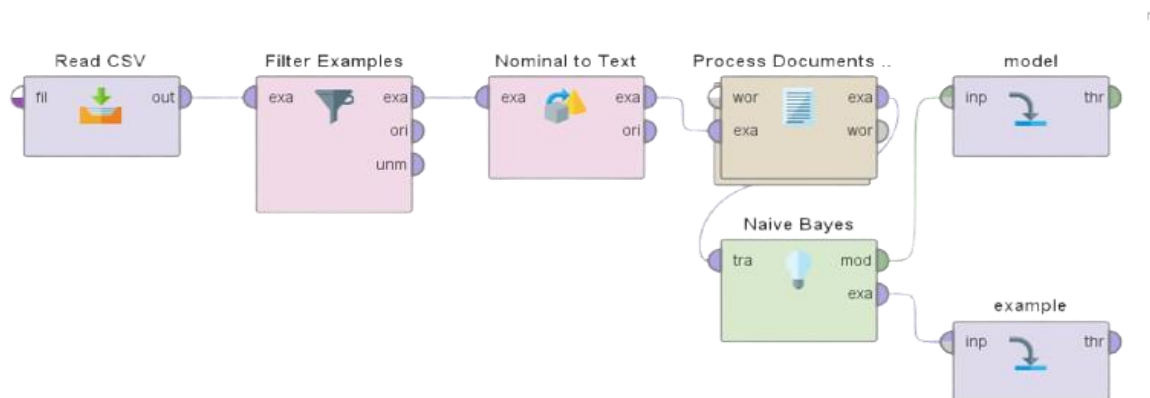
Proses pelabelan sentimen dilakukan secara manual dengan bantuan perangkat lunak Microsoft Excel, untuk menentukan apakah suatu tweet mengandung sentimen positif atau negatif, berdasarkan klasifikasi yang telah ditentukan sebelumnya. Pelabelan secara otomatis masih sulit dilakukan, terutama karena keterbatasan mesin dalam memahami konteks bahasa Indonesia secara mendalam. Oleh karena itu, pelabelan manual dipilih untuk memastikan akurasi data. Gambar 2 menunjukkan sebaran data hasil *crawling* yang telah melalui proses pembersihan, dengan total 614 tweet setelah menghapus data duplikat dan yang tidak relevan.



Gambar 1. Sebaran data

### Pelatihan Model Menggunakan Data Testing

Mengacu pada diagram alur penelitian, proses pelatihan model diawali dengan membaca dataset dalam format CSV yang telah melalui tahap penyaringan awal. Data tersebut kemudian diproses melalui tahap *Filter Examples* untuk menghapus entri yang tidak memenuhi kriteria. Selanjutnya, dilakukan konversi dari format nominal ke teks menggunakan *Nominal to Text* guna mempersiapkan data untuk tahap analisis teks. Proses *pre-processing* lanjutan dilakukan melalui *Process Documents ..*, yang mencakup tahapan seperti *case folding*, *tokenization*, *stopword removal*, dan *stemming* menggunakan algoritma Sastrawi yang sesuai untuk Bahasa Indonesia. Setelah itu, dataset dibagi menjadi dua bagian, yaitu data latih dan data uji dengan rasio 80:20. Pada gambar 2, Pelabelan data dilakukan secara manual dengan mengkategorikan tweet ke dalam sentimen positif dan negatif sesuai dengan konteks isi tweet tersebut.



Gambar 2. Melatih model data testing

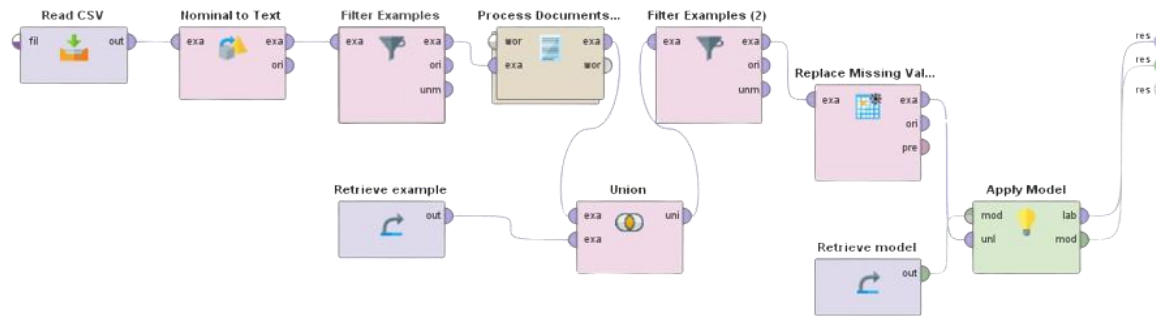
### Pengujian Model Menggunakan Data Training

Setelah proses pelatihan model selesai dilakukan, langkah berikutnya adalah menguji model menggunakan data uji pada gambar 3. Berdasarkan diagram alur pengujian, proses dimulai dengan mengambil dataset yang telah diproses sebelumnya melalui Retrieve Model-dataset. Data uji tersebut kemudian digabungkan dengan model yang telah dilatih menggunakan operator Union. Tahap Filter Examples diterapkan untuk memastikan format data uji selaras dengan data latih, sehingga tidak terjadi ketidaksesuaian dalam struktur data.

Selanjutnya, operator Replace Missing Values digunakan untuk mengatasi kemungkinan

adanya data kosong atau nilai hilang dalam data uji, guna menjamin kelancaran proses klasifikasi. Model Naive Bayes yang telah dilatih sebelumnya kemudian diambil melalui Retrieve Model-Senti, dan diterapkan pada data uji menggunakan Apply Model untuk menghasilkan prediksi sentimen, baik positif maupun negatif.

Langkah terakhir adalah mengevaluasi hasil prediksi menggunakan metrik akurasi, presisi, dan recall guna menilai sejauh mana model mampu mengklasifikasikan sentimen dalam tweet secara akurat dan objektif.



Gambar 3. Menguji model data training

### Evaluasi model

Evaluasi model dilakukan untuk mengukur sejauh mana kinerja algoritma Naive Bayes dalam mengklasifikasikan sentimen masyarakat secara akurat. Pengujian ini memanfaatkan data uji (test data) yang sebelumnya telah dipisahkan dari data latih (training data) dengan proporsi 80:20. Hasil evaluasi disajikan dalam bentuk confusion matrix, yang mencakup beberapa metrik utama seperti akurasi, presisi, dan recall.

Dengan melakukan evaluasi ini, peneliti dapat memahami efektivitas model dalam mengenali sentimen positif dan negatif, serta mengidentifikasi kelemahan yang mungkin ada agar dapat ditingkatkan pada pengujian berikutnya.

#### a. Accuracy

accuracy: 72.22%

	true negatif	true positif	class precision
pred. negatif	0	0	0.00%
pred. positif	50	130	72.22%
class recall	0.00%	100.00%	

Gambar 4. Hasil accuracy

#### b. Precision

precision: unknown (positive class: negatif)

	true positif	true negatif	class precision
pred. positif	130	50	72.22%
pred. negatif	0	0	0.00%
class recall	100.00%	0.00%	

Gambar 5. Hasil Precision

c. *Reccal*

recall: 0.00% (positive class: negatif)

	true positif	true negatif	class precision
pred. positif	130	50	72.22%
pred. negatif	0	0	0.00%
class recall	100.00%	0.00%	

Gambar 6. Hasil *reccal*

### KESIMPULAN

Berdasarkan hasil penelitian yang berjudul *Analisis Sentimen Masyarakat Terhadap Kepemimpinan Presiden Prabowo 2025 Menggunakan Algoritma Naive Bayes*, dapat disimpulkan hal-hal berikut:

Dataset yang digunakan dalam penelitian ini terdiri dari 1.200 tweet yang membahas kepemimpinan Presiden Prabowo. Setelah melalui proses pelabelan manual ke dalam dua kategori sentimen (positif dan negatif), serta penyaringan untuk menghapus duplikasi dan entri tidak relevan, diperoleh 614 tweet yang layak dianalisis. Data ini kemudian dibagi menjadi dua bagian: 80% digunakan sebagai data latih dan 20% sebagai data uji. Model klasifikasi Naive Bayes dilatih menggunakan data latih, dan performanya diuji terhadap data uji yang belum dikenali sebelumnya oleh model.

Hasil evaluasi menunjukkan bahwa model Naive Bayes memiliki tingkat akurasi sebesar 72,22%, yang termasuk dalam kategori cukup baik namun belum optimal. Model menunjukkan kinerja yang memadai dalam mengenali sentimen positif, dengan precision dan recall masing-masing sebesar 72,22%, yang berarti sebagian besar data dengan opini mendukung berhasil diklasifikasikan secara benar.

Namun demikian, performa model dalam mendeteksi sentimen negatif masih rendah, dengan precision dan recall hanya sebesar 27,78%, yang mengindikasikan ketidakmampuan model dalam mengenali pola kalimat bernada kritik atau penolakan secara akurat.

Secara keseluruhan, model lebih efektif dalam mengklasifikasikan opini positif dibandingkan dengan negatif. Hasil ini menjadi catatan penting bahwa meskipun Naive Bayes mampu memberikan prediksi dasar yang cukup akurat, diperlukan pengembangan lebih lanjut untuk meningkatkan performa, terutama dalam mengenali sentimen negatif, agar hasil analisis lebih seimbang dan representatif terhadap keseluruhan opini publik.

### REFERENSI

- Al Muqstih Prasetyo, P., & Hermawan, A. (2025). Analisis sentimen twitter terhadap pemilihan presiden menggunakan algoritma naive bayes. *INFOTECH: Jurnal Informatika & Teknologi*, 4(2), 224–233. <https://doi.org/10.37373/infotech.v4i2.863>
- Amanda, R., Yuandhika, A., Sabrina, N., Aulia, J., Melati, C. E., & Sholihatin, E. (2024). *Postingan Dan Komentar Di Platform X Dalam Pemilihan Presiden 2024 Di Indonesia Melalui Pemrograman Python Dan Metode Naive Bayes*. 9(April), 85–92.
- August, M. R., February, R., February, A., & April, P. (2024). *FOR 2024 ON TWITTER USING THE MULTINOMIAL*. 2(1), 31–40.
- Diana Dwi Rahayu, Muhammad Fatchan, & Alfonsus Ligouri. (2024). Analisis Sentimen Twitter Terpilihnya Prabowo - Gibran Menggunakan Metode Neural Network. *Tematik*, 11(1), 85–91. <https://doi.org/10.38204/tematik.v11i1.1943>
- No, V., Juni, A., Siregar, K. A., Nasution, S., & Nabawy, P. (2025). *Analisis Sentimen Netizen Indonesia Terhadap Kampanye Penggunaan Kecerdasan Buatan Oleh Pemerintah Menggunakan Algoritma Naive Bayes*. 1(4), 112–121.

- Nurjanah, W. E., Setya Perdana, R., & Fauzi, M. A. (2017). Analisis sentimen terhadap tayangan televisi berdasarkan opini masyarakat pada media sosial twitter menggunakan metode k-kearest neighbor dan pembobotan jumlah retweet. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(12), 1750–1757.
- Permana, H., Herry Chrisnanto, Y., & Ashaury, H. (2025). Analisis Sentimen Masyarakat X Terhadap Presiden Prabowo Dengan Algoritma Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(5), 3257–3264. <https://doi.org/10.36040/jati.v7i5.7309>
- Satriyo, M. A. (2022). *Tinjauan Pustaka Dan Dasar Teori Statistik*. 151, 10–17. [https://repository.unikom.ac.id/60158/1/STAT 03 %28Penyajian Data%29.pdf](https://repository.unikom.ac.id/60158/1/STAT%2003%28Penyajian%20Data%29.pdf)
- Sofyan, S. N., & Iqbal, M. (2025). *Analisis Sentimen Terhadap Dampak Inflasi Menggunakan Naive Bayes*. 6(1), 4–11. <https://doi.org/10.47065/bit.v5i2.1796>
- Steven, A., Prasetya, D., & Aji, A. S. (2025). *Public Sentiment Analysis Towards Mayor Teddy Indra Wijaya with Logistic Regression Approach Analisis Sentimen Publik Terhadap Mayor Teddy Indra Wijaya dengan Pendekatan Logistic Regression*. 5(January), 222–231.