



Analisis Sentimen Komentar Netizen Terhadap 17+8 Tuntutan Rakyat Pada X Menggunakan Naive Bayes Classifier

Francisco Lucky Halawa^{1*}, Rudi Heriansyah², Indah Permatasri³

^{1,2,3} Prodi Teknik Informatika, Universitas Indo Global Mandiri Palembang

Korespondensi penulis: fransiscoholawa14@gmail.com

Abstract. This study analyzes netizen sentiment concerning the 17+8 public aspirations circulating the digital platform X spanning the period from August 18 through October 31, 2025. 1,837 comments obtained through scraping method. Classification Research stages include data preprocessing, sentiment weighting based on lexicon, and feature extraction using TF-IDF. Data 80% used for learning purposes and the remaining 20% utilized for validation. The findings reveal that the majority of comments, amounting to 81.14%, contained negative sentiment, while the remaining 18.86% were positive. The outcomes demonstrate that community reactions toward the 17+8 People's Demands were dominated by unresponsive views. From a theoretical standpoint this scholarly work offers to enriching knowledge concerning public opinion classification on political issues through a computational approach, while also serving as a reference for future research focused on improving the accuracy of sentiment analysis related to political dynamics and the behavior of state institutions.

Keywords: Sentiment Analysis; 17+8 People's Demands; Naïve Bayes Classifier; X

Abstrak. Studi ini menganalisis sentimen netizen terhadap 17+8 Tuntutan Rakyat beredar di platform media sosial X dalam rentang waktu 18 Agustus hingga 31 Oktober 2025. 1.837 komentar yang diperoleh melalui metode scraping. Klasifikasi Tahapan penelitian meliputi pra-pemrosesan data, pemberian bobot sentimen berbasis leksikon, serta ekstraksi fitur menggunakan TF-IDF. Data 80% untuk latih 20% untuk pengujian. Temuan penelitian mengungkapkan bahwa mayoritas komentar, yakni sebesar 81,14%, mengandung sentimen negatif, sedangkan 18,86% sisanya bersifat positif. Hasil ini mengindikasikan bahwa respons masyarakat terhadap 17+8 Tuntutan Rakyat didominasi oleh pandangan yang tidak mendukung. Secara teoritis, penelitian ini berkontribusi dalam memperkaya pemahaman tentang klasifikasi opini publik pada isu-isu politik melalui pendekatan komputasional, sekaligus menjadi referensi bagi penelitian lanjutan yang berfokus pada peningkatan akurasi analisis sentimen terkait dinamika isu-isu lain. **Kata kunci:** Analisis Sentimen; 17+8 Tuntutan Rakyat; Naïve Bayes Classifier; X

1. LATAR BELAKANG

Evolusi *platform* telah mentransformasi pola perilaku dan kehidupan manusia dalam mengkomunikasikan pemikiran dan kritik mereka tentang masalah sosial melalui saluran daring. Melalui komentar, unggahan, dan percakapan di *platform* seperti Twitter/X, persepsi masyarakat terhadap kebijakan pemerintah dapat diamati secara langsung dan *real time*. Dalam kerangka ini, analisis sentimen berfungsi sebagai teknik yang ampuh untuk memetakan kecenderungan opini publik mengenai suatu peristiwa atau kebijakan tertentu (Siregar Alfarizi et al., 2025).

Sejumlah kajian ilmiah sebelumnya telah menunjukkan bahwa teknik *Naïve Bayes Classifier* dapat secara efektif mengkategorikan sentimen dalam teks Bahasa Indonesia dengan presisi tinggi. Penelitian mengenai sentimen terhadap kinerja DPR (Putri Duei, Dianati, Forda, Gigih, Sulistiano Eko, 2022). Penelitian terdahulu menjelaskan bahwa

penelitian tekstual, proses pra-pemrosesan dan pelabelan menjadi kunci keberhasilan klasifikasi. Studi komparatif analisis ulasan aplikasi “KitaLulus” menggunakan *Naive Bayes* melaporkan bahwa metode tersebut efisien dan menghasilkan akurasi tinggi (91 %), setelah dilakukan pembersihan data, tokenisasi, penghapusan kata umum (*stopwords*), dan *stemming* (Mandasari et al., 2022).

Menurut Amini & Setiawan (2024), penelitian tentang penerapan *Naive Bayes* dalam opini politik dan kebijakan juga telah dilakukan di ranah internasional. Dalam kajian ilmiah “Analisis Sentimen Terhadap Calon Wakil Presiden Indonesia 2024 menggunakan *Naive Bayes*” berhasil menghasilkan akurasi sekitar 82,19 % pada data twitter (Amini & Setiawan, 2024). Selain itu, aplikasi *Naive Bayes* untuk *sentiment* publik terhadap kebijakan pandemi (PPKM) mengungguli beberapa algoritma, sehingga hasil-hasil ini mendukung bahwa *Naive Bayes* relevan dan kompetitif dalam klasifikasi sentimen (Fransiscus & Girsang, 2022).

Dalam penelitian “Analisis Sentimen Publik atas Kebijakan Efisiensi Anggaran 2025 menggunakan *Naive Bayes* dan *NLP*” menemukan bahwa masyarakat sangat merespons instruksi presiden terkait penghematan anggaran dengan kritis, terutama karena kekhawatiran bahwa pemangkasan dapat berdampak pada layanan publik (Agustina & Herliana, 2025). Menurut Fitrianti & Yudhistira 2025, dalam penelitian berjudul “Analisis Sentimen Media Sosial Terhadap Calon Pilkada 2024 dengan metode *Naive Bayes*”, Pendekatan *Naive Bayes* telah menunjukkan efektivitas kajian yang berkelanjutan dalam analisis emosi tulisan berbahasa Indonesia, khususnya dalam kumpulan data media sosial antara lain twitter atau *platform* “X” (Fitrianti & Yudhistira, 2025). Penelitian lain analisis sentimen terhadap “seleksi CPNS tahun 2024 berbasis media sosial X, menggunakan *naive bayes*” memperoleh akurasi 82,02 %, yang menegaskan bahwa metode ini cukup andal untuk mendeteksi opini publik dalam konteks kebijakan atau tindakan negara yang sensitif (Yudistira & Isnain Rahman, 2025).

Dalam beberapa waktu terakhir, ruang publik Indonesia ramai oleh munculnya fenomena 17+8 Tuntutan Rakyat, yaitu gelombang aspirasi masyarakat yang menyuarakan tekanan kepada kondisi sosial, ekonomi, serta politik nasional. Gerakan ini berkembang pesat di media sosial dan menjadi pusat perhatian publik setelah aksi demonstrasi yang terjadi di berbagai kawasan pada Agustus 2025. Aspirasi yang disampaikan mencakup 17 butir jangka pendek dan 8 butir jangka panjang terkait

transparansi anggaran, reformasi DPR, perlindungan hak sipil, hingga evaluasi kebijakan ekonomi. Gerakan ini turut diperkuat oleh dukungan figur publik dan organisasi masyarakat sipil, sehingga respons masyarakat terhadap isu ini semakin luas dan beragam (Setyaningrum, 2025).

Beberapa artikel dan pemberitaan media yang membahas tentang isi 17+8 Tuntutan Rakyat seperti penelitian (Salsabila et al., 2025), menyatakan isi 17+8 Tuntutan Rakyat tidak sekedar daftar teknis, melainkan dukungan agar masyarakat mendapatkan keadilan terhadap cara negara mengelola demokrasi dan menyerap aspirasi masyarakat, dan menurut penelitian (Khusein et al., 2025) menyatakan isi Tuntutan 17+8 Rakyat Ini merupakan bentuk penolakan terhadap sejumlah kebijakan legislatif yang dianggap tidak pro-rakyat. Serta pemberitaan kenal berita (Ranawati, 2025) memberitakan bahwa beberapa tokoh terkemuka, termasuk aktivis Andhyta Firselly Utami (Afu) dan berbagai influencer, menyoroti bahwa tuntutan yang terdapat dalam Gerakan Rakyat 17+8 tidak eksklusif untuk satu faksi saja, melainkan mencerminkan opini kolektif masyarakat Indonesia. Selain itu, unsur-unsur Gerakan Rakyat 17+8 mencakup seruan dari 211 kelompok masyarakat sipil, yang rinciannya dapat dilihat disitus web YLBHI.

Oleh karena itu, penelitian ini bertujuan untuk mengevaluasi teori awal mengenai perasaan masyarakat terhadap aspek-aspek 17+8 Tuntutan Rakyat, menilai apakah reaksi tersebut menguntungkan atau tidak menguntungkan, dengan menggunakan analisis sentimen. Nilai dari penelitian ini terletak pada penyampaian analisis sentimen terhadap 17+8 Tuntutan Rakyat yang belum pernah dikaji sebelumnya secara kuantitatif menggunakan pendekatan klasifikasi teks.

Berdasarkan gap penelitian terdahulu di atas, dengan mempertimbangkan kekurangan pada penelitian-penelitian sebelumnya yang disebutkan di atas, penelitian ini dilakukan untuk menguji penerapan Naive Bayes Classifier terhadap 17+8 tuntutan rakyat yang muncul sebagai bentuk protes dan aspirasi masyarakat terhadap berbagai kebijakan dan kinerja pemerintah serta DPR RI. Penelaahan ini diharapkan mampu memberikan gambaran yang tidak memihak mengenai pola pendapat warga terhadap seruan 17+8 aspirasi rakyat.

2. KAJIAN TEORITIS

Analisis Sentimen

Analisis sentimen menurut Bhatara & Suryono (2024), merupakan suatu tahapan klasifikasi yang bertujuan untuk mengelompokkan kalimat berdasarkan kecenderungan opini yang terkandung di dalamnya, Metode melakukan ekstraksi informasi dan pengambilan data tekstual mulai dari memasukkan kalimat ke file csv, pra-pemrosesan teks, hingga pengklasifikasian memakai algoritma machine learning untuk merangkum pendapat publik secara rinci pada tiap aspek (Bhatara & Suryono, 2024).

Netizen

Merujuk pada kajian berbasis korpus yang menganalisis kolokasi kosakata warganet dalam pemberitaan Indonesia tahun 2020, penulis menjelaskan bahwa netizen dipahami sebagai pengguna Internet yang aktif terlibat dalam aktivitas komunikasi, diskusi, dan partisipasi daring, terutama melalui platform media sosial dan ruang publik digital. Netizen tidak hanya menjadi konsumen informasi, tetapi juga berperan sebagai pihak yang memberikan komentar, kritik, pendapat, dan respons terhadap isu-isu sosial melalui media *online* (Sumiyati et al., 2020).

Media Sosial X (Twitter)

Menurut Jannah (2023), Jejaring sosial adalah sarana berbasis digital yang memberi kesempatan kepada penggunanya untuk saling berkomunikasi, bertukar informasi, serta menjalin hubungan antar sesama secara virtual sendiri dan tetap terhubung dengan teman-teman untuk mengobrol dan berbagi berbagai hal. Studi yang berjudul "Analisis Wacana Opini Publik di Twitter oleh Van Dijk" ini meneliti topik tersebut, menyatakan bahwa penggunaan media sosial seperti X/ twitter menjadi sarana penting bagi netizen untuk mengutarakan opini mereka. Media sosial memungkinkan setiap individu untuk berpartisipasi dalam membentuk opini publik (Jannah et al., 2023).

Komentar

Komentar adalah Respons atau reaksi yang diberikan melalui ucapan, seperti pujian, keberatan, kritik, kesan, atau pernyataan apa pun yang harus diungkapkan dengan cara yang tepat. Komentar dapat dipahami sebagai respons tertulis atau lisan yang diberikan oleh seseorang setelah melihat atau membaca suatu konten, gagasan, atau peristiwa, yang berfungsi sebagai ekspresi opini, evaluasi, atau tanggapan terhadap objek tersebut. Dalam konteks media sosial, komentar adalah bagian dari interaksi digital dimana pengguna menyampaikan pemikiran mereka secara terbuka, memungkinkan adanya dialog dua arah antara pengunggah konten dan audiens (Ketut & Ida, 2023).

Google Colab

Colaboratory, atau *Colab* menurut Irfon & Soen (2022), *Google Colaboratory* merupakan layanan yang dikembangkan oleh *Google Research* yang menyediakan lingkungan berbasis *cloud* guna membuat dan mengoperasikan kode *Python* secara *real-time* melalui peramban apa pun menggunakan peramban *web*, dan sangat bagus untuk hal-hal seperti pembelajaran mesin, analisis data, dan pengajaran. (Irfon & Soen, 2022).

Python

Python merupakan bahasa pemrograman yang populer berkat struktur penulisannya yang ringkas dan mudah dimengerti. Bahasa ini digagas dan dibangun oleh Guido van Rossum, dengan peluncuran perdananya pada tahun 1991. Dalam perkembangannya, *Python* mengalami peningkatan popularitas serta signifikan dan pada era sekarang banyak dimanfaatkan dalam berbagai bidang, khususnya ilmu data dan pembelajaran mesin. *Deep learning* karena fleksibel, mudah digunakan, serta memiliki ekosistem pustaka yang sangat kaya seperti *NumPy*, *Pandas*, *Scikit-Learn*, dan *TensorFlow* (Alfariz sirfatullahi et al., 2023).

Scraping

Dapat dipahami sebagai suatu cara yang digunakan untuk memperoleh data secara otomatis dari halaman *web* dengan mengekstraksi elemen-elemen yang terdapat pada struktur situs tersebut, proses ini bekerja dengan membaca dan memproses komponen halaman *web* seperti *HTML*, *DOM*, *CSS selector*, atau *XPath* untuk menemukan bagian informasi yang dibutuhkan (Dwicahyo, 2021).

Text Mining

Metode analitik yang diterapkan untuk menjalankan proses pengelompokan terhadap data berbentuk teks dan dapat dipandang sebagai pengembangan dari konsep data *mining*. Pendekatan ini bertujuan untuk mengidentifikasi pola-pola tertentu dari kumpulan teks dalam jumlah besar. Dalam penerapannya, *text mining* mengadaptasi metode-metode pada data *mining* untuk mengekstraksi pola dan informasi tersembunyi, sehingga menghasilkan temuan yang relevan dan bermanfaat sesuai dengan kebutuhan analisis (Munthe et al., 2022).

Naïve Bayes Classifier

Pendekatan kategorisasi berbasis perhitungan statistik yang diterapkan guna memperkirakan probabilitas sebuah kumpulan informasi termasuk 23 kategori tertentu.

Pendekatan ini berlandaskan pada *Teorema Bayes* dengan memanfaatkan prinsip probabilitas dan statistika dalam proses perhitungannya. Melalui mekanisme tersebut, metode ini mampu memperkirakan kemungkinan suatu kejadian berdasarkan pola atau data yang telah diamati sebelumnya (Simanjuntak et al., 2022).

Klasifikasi

Klasifikasi adalah cara melihat item data dan memasukkannya ke dalam salah satu kelompok yang sudah ditetapkan. Klasifikasi dilakukan dengan membangun representasi yang dikonstruksi dengan informasi telah tersedia, lalu memanfaatkan pola yang telah terbentuk tersebut dalam rangka mengelompokkan informasi ke dalam kategori tertentu. Secara umum, klasifikasi dapat dipahami sebagai proses pembentukan fungsi yang mengarahkan beragam ciri atau karakteristik data menuju salah satu label kelas yang sudah ditetapkan lebih awal (Utomo & mesran, 2020).

Data Preprocessing

Data *preprocessing* menurut Jon (2023), suatu proses pengolahan informasi tekstual yang tidak beraturan menjadi susunan yang lebih teratur, dengan tujuan mengurangi kompleksitas data dan meningkatkan efisiensi komputasi. Langkah ini berperan penting sebagai persiapan sebelum analisis lebih lanjut dilakukan pada data *preprocessing* membantu menghilangkan elemen - elemen yang tidak relevan sehingga informasi utama dalam dokumen dapat lebih mudah diidentifikasi dan dianalisis, yang terisi dalam tahapan *cleaning*, *case folding*, *tokenization*, normalisasi, *stemming*, *stopwords removal* (Jon, 2023).

TF-IDF

Salah satu pendekatan dalam pengolahan teks mengonversi dokumen ke representasi numerik berbentuk vektor. Metode mengukur nilai kepentingan tiap kata berdasarkan dari dua elemen pokok, yaitu seberapa sering kata muncul pada dokumen dan tingkat kelangkaannya pada seluruh dokumen dalam korpus. Melalui cara kerja ini, diksi yang kerap muncul pada suatu berkas hampir tidak dijumpai pada berkas teks lainnya diberi nilai kepentingan lebih tinggi, oleh karena itu dinilai lebih mampu mewakili dalam mencerminkan kandungan dokumen (Putra Tri et al., 2023).

Confusion Matrix

Dapat dipahami sebagai sarana evaluasi berstruktur tabel matriks difungsikan dalam rangka menilai tingkat keberhasilan sistem pemilahan data dalam mengenali data

dari berbagai kelas. Dalam konteks analisis data dan *machine learning*, metode ini menyajikan ringkasan hasil prediksi model melalui perbandingan antara kelas sebenarnya serta kelas yang diprediksi. Kerangka tersebut memuat keterangan terkait banyaknya hasil dugaan akurat serta yang keliru pada masing-masing kelas, dengan demikian memberikan penjelasan yang terperinci mengenai tingkat akurasi serta kesalahan klasifikasi yang dihasilkan oleh model (Hidayah & Dodiman, 2024).

3. METODE PENELITIAN

Pengumpulan Data

Dalam fase ini, penulis mengadopsi teknik penghimpunan data primer melalui teknik *web scraping* untuk memperoleh data komentar netizen dari media sosial X (Twitter). Proses *scraping* dilakukan dengan mengumpulkan komentar pada postingan yang relevan dengan topik 17+8 Tuntutan Rakyat. Pengambilan data ini menggunakan pustaka *snsrape* melalui *platform Google Colab*, sehingga peneliti dapat mengekstraksi teks komentar secara otomatis dan efisien.

Data Preprocessing

Tahap *preprocessing* data mencakup sejumlah proses transformasi, seperti *cleaning*, *case folding*, *tokenisasi*, *stemming*, penghapusan *stopword*. Pada tahap pertama ini atau tahap *cleaning* dilakukan eliminasi karakter yang tidak diperlukan, seperti simbol yang tidak ada arti, serta penghapusan URL atau tautan, hashtag, yang terdapat pada komentar.

Case Folding Pada tahap ini dilakukan proses konversi seluruh huruf kapital menjadi huruf kecil. *Case folding* merupakan salah satu tahapan dalam *text preprocessing* bertujuan untuk menyeragamkan bentuk karakter pada data teks. Proses ini dilakukan dengan mengubah seluruh huruf dari rentang “A–Z” menjadi “a–z”, sehingga tidak terdapat perbedaan penulisan akibat penggunaan huruf besar dan kecil.

Normalisasi merupakan proses mengganti kata-kata dalam teks menjadi bentuk yang baku atau standar. Tahap ini bertujuan untuk mengatasi variasi pemakaian diksi dimana menyimpang secara standar linguistik, misalnya mengubah “gk” menjadi “tidak”, “nggak” menjadi “tidak”, atau “tdk” menjadi “tidak”. Melalui proses normalisasi, data menjadi lebih konsisten dan mudah diproses, sehingga sistem dapat memahami makna teks secara akurat tanpa terpengaruh oleh perbedaan penulisan.

Tokenize fase ini, masing-masing diksi yang termuat dalam kalimat dalam berkas, setiap kata dalam suatu kalimat pada dokumen dipisahkan satu per satu. Pemisahan kata umumnya dilakukan dengan menggunakan spasi. Meskipun metode penulisan dapat bervariasi, tujuan utama dari proses ini adalah memecah kalimat menjadi kata-kata penyusunnya.

Pembuangan kata penghubung Dalam fase ini diterapkan prosedur penyaringan terhadap kata-kata yang memiliki frekuensi kemunculan sangat tinggi maupun sangat rendah, yang umumnya dikenal sebagai *stopword*. Tahapan ini disebut sebagai Pembuangan kata penghubung, yaitu proses penghapusan kosakata yang tidak berperan berarti bagi makna teks. Dengan menghilangkan kata-kata yang kurang informatif tersebut, performa model klasifikasi dapat 35 meningkat, sekaligus mengurangi tingkat sparsitas data dan memperkecil dimensi ruang fitur.

Stemming dapat mengubah diksi yang ada dalam berkas teks beralih menjadi bentuk dasar, tahap reduksi kata di dalam teks berbahasa Indonesia tergolong cukup kompleks karena memerlukan penghapusan seluruh mubahan yang melekat pada kata-kata tersebut.

Table 1. Hasil *Preprocessing*

| Sebelum <i>preprocessing</i> | Setelah <i>preprocessing</i> |
|--|--|
| Rakyatnya keren bisa bikin tuntutan semua warga dalam 3 jam, langsung bikin websitenya buat memantau, tapi pemerintahnya kayak fir'aun sombong2 banget najiss... | rakyat, keren, buat, tuntutan, warga, 3, jam, langsung, buat, situs, pantau, pemerintah, firaun, sombong, jijik] |

Labeling Data

Proses pelabelan dengan *InSet Lexicon* mencakup perhitungan polaritas sentimen teks melalui penjumlahan nilai polaritas yang terkait dengan setiap kata. Skor polaritas dialokasikan untuk setiap kata sesuai dengan 36 kemunculannya dalam sentimen *lexicon* yang telah ditentukan sebelumnya (Kurniawan & Iriananda, 2025). Nilai sentimen total diperoleh dengan menjumlahkan seluruh skor polaritas kata dalam teks, yang dirumuskan sebagai berikut:

$$\text{Sentiment Score} = \sum_{i=1}^n \text{Polarity}(w_i) \quad (1)$$

Pada rumus tersebut, "n" melambangkan banyaknya kata secara keseluruhan yang terdapat, adapun (w_i) mencerminkan nilai sentimen diasosiasikan diksi ke-i. Skor sentimen yang melebihi 0 menunjukkan polaritas positif teks. Sebaliknya, jika skor kurang dari 0, teks tersebut diklasifikasikan sebagai negatif.

Ekstraksi Fitur

Setelah data melalui tahapan *preprocessing* dan *labelling*, tahap selanjutnya adalah melakukan ekstraksi atribut guna memperlancar tahapan pengelompokan. penggalian atribut dilaksanakan untuk katateknik pemberian bobot pada istilah, mengukur sejauh mana suatu term dianggap relevansi sebuah istilah pada suatu berkas teks mengacu kepada seberapa sering istilah tersebut muncul serta distribusinya pada seluruh dokumen, yang dirumuskan sebagai berikut:

$$TF = \frac{n_{t,d}}{\text{(Total number of term in document)}} \quad (2)$$

$$IDF = \log \frac{D}{DF} \quad (3)$$

$$TF - IDF = Tf \times IDF \quad (4)$$

Pembagian Data

Setelah proses pembobotan selesai, tahap berikutnya lanjutan pembagian dataset menjadi data latih dan data uji sebagai bagian dari prosedur validasi model. Data latih dimanfaatkan untuk membangun dan melatih model agar mampu mempelajari pola yang terdapat dalam dataset, sedangkan data uji 42 digunakan untuk mengevaluasi kemampuan model dalam melakukan prediksi terhadap data yang belum dikenali. Pada kajian ini, proporsi pemisahan data ditetapkan dengan komposisi 80% untuk proses pembelajaran dan 20% diperuntukkan bagi proses evaluasi.

Klasifikasi Metode Naïve Bayes

Dalam alur penerapannya, kumpulan informasi yang telah melalui serangkaian proses pembersihan pengolahan awal berikutnya dibagi ke dalam himpunan himpunan latihan dan himpunan informasi evaluasi. Model kemudian menghitung probabilitas masing-masing kelas sentimen untuk menentukan kategori berdasarkan fitur yang tersedia. Tahap akhir dari proses ini adalah pengukuran kemampuan Teknik

memanfaatkan ukuran kemampuan yakni tingkat keakuratan, ketepatan, kelengkapan, dan indeks F1.

Pada tahap awal perhitungan metode *Naïve Bayes*, peneliti menghitung probabilitas kelas untuk setiap kelas sentimen berdasarkan distribusi label pada data latih (*training data*). Yang dihitung menggunakan rumus dibawah ini:

$$P(c) = \frac{\text{jumlah komentar pada kelas}}{\text{total komentar}} \quad (5)$$

Hasil perhitungan menggunakan seluruh dataset 1837 kumpulan komentar yang berikutnya dipisahkan ke dalam himpunan data latihan berjumlah 1.469 himpunan data evaluasi 277 komentar menunjukkan hasil yang ditampilkan pada Tabel 2.

Table 2 Distribusi Kelas Sentimen pada Dataset

| Kelas Sentimen | Jumlah Data | Probabilitas | Presentase |
|----------------|-------------|--------------|------------|
| Negatif | 1506 | 0.8114 | 81,14% |
| Positif | 331 | 0,1856 | 18,86% |

Menghitung Probabilitas *Kondisional* $P(x|c)$ Pada tahap perhitungan probabilitas kondisional $P(x|c)$, yaitu peluang kemunculan suatu kata pada kelas sentimen tertentu. Jika pada tahap sebelumnya metode sudah mengetahui peluang awal kelas positif dan negatif, maka pada tahap ini metode mulai belajar dari istilah-istilah yang hadir dalam data latih. Melalui diksi lain, di sinilah proses *training* benar-benar terjadi pada level fitur teks. Hasil dari Menghitung *Kondisional* $P(x|c)$ ditampilkan pada Gambar 1.

$$P(x|c) = \prod_{t=1}^n P(x_t | c_i) \quad (6)$$

| stemming_str | P_x_given_Positive | P_x_given_Negative |
|--|------------------------|------------------------|
| jokes doang nder petisi kayak mah banget kemarin tuntutan 17 8 gubris angkat bilang not worth and turns out they were right seru biar stres hehe | 7.931434335429016e-86 | 2.2034332183949823e-83 |
| asli tuntutan 17 8 receh akan demo kemarin buang waktu | 2.23018986506198e-28 | 4.6695124595956187e-26 |
| wih endingnya tuntutan rakyat 17 8 lupa | 9.24667732828782e-17 | 1.1862660577458861e-15 |
| 17 8 tuntutan rakyat revisi tambah | 2.634589723453847e-12 | 2.023306609527065e-11 |
| iya usul 17 8 tuntutan rakyat masuk dalam | 1.0848663271358412e-19 | 7.047740102208009e-18 |

Gambar 1. Hasil Perhitungan Probabilitas *Kondisional*

Menghitung Probabilitas *Posterior* $P(c|x)$ Pada tahap perhitungan sebuah komentar termasuk ke dalam kelas sentimen tertentu setelah mempertimbangkan kata-

kata yang muncul di dalam komentar tersebut. Jika pada tahap sebelumnya sistem sudah mengetahui probabilitas awal kelas $P(c)$ dan probabilitas setiap kata terhadap kelas $P(x|c)$, maka pada tahap ini kedua nilai tersebut digabungkan untuk menentukan keputusan akhir sentimen, yang dihitung menggunakan rumus dibawah ini:

$$P(c|x) = \frac{P(x|c) \cdot P(c)}{P(x)} \quad (7)$$

Hasil dari menghitung probabilitas posterior ditampilkan pada Gambar 2.

| stemming_str | Posterior_Positive | Posterior_Negative | Predicted_Sentiment |
|--|------------------------|------------------------|---------------------|
| jokes doang nder petisi kayak mah banget kemarin tuntutan 17 8 gubris angkat bilang not worth and turns out they were right seru biar stres hehe | 2.2019768704783863e-86 | 1.5917016226511387e-83 | Negatif |
| asli tuntutan 17 8 receh akan demo kemarin buang waktu | 6.191599516503048e-29 | 3.3731317549719035e-26 | Negatif |
| wih endingnya tuntutan rakyat 17 8 lupa | 2.5671232647941147e-17 | 8.569270868964568e-16 | Negatif |
| 17 8 tuntutan rakyat revisi tambah | 7.314320952430384e-13 | 1.4615829454776346e-11 | Negatif |
| ija usul 17 8 tuntutan rakyat masuk dalam | 3.011877119429934e-20 | 5.0911002262547784e-18 | Negatif |

Gambar 2. Hasil Perhitungan Probabilitas Posterior

Setelah proses pelatihan metode selesai, dilakukan evaluasi terhadap performa metode untuk mengetahui kemampuan metode dalam mengklasifikasikan data training, yang ditampilkan pada Gambar 3. Dan menggunakan rumus yang dijelaskan dibawah ini:

a. Akurasi

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$\text{Akurasi} = \frac{206 + 1180}{206 + 1180 + 12 + 71} = \frac{1386}{1469} = 0.9435 \text{ atau } 94.35\%$$

b. Presisi

$$\text{Presisi(positif)} = \frac{TP}{TP + FP} \quad (10)$$

$$\text{Presisi(negatif)} = \frac{TN}{TN + FN}$$

$$\text{Presisi(positif)} = \frac{206}{206 + 12} = \frac{206}{218} = 0.9449 \text{ atau } 94.49\%$$

$$\text{Presisi(negatif)} = \frac{1180}{1180 + 71} = \frac{1180}{1251} = 0.9432 \text{ atau } 94.32\%$$

c. Recall

$$\text{Recall(positif)} = \frac{TP}{TP + FN} \quad (11)$$

$$\text{Recall(negatif)} = \frac{TN}{TN + FP}$$

$$Recall(\text{positif}) = \frac{206}{206 + 71} = \frac{206}{277} = 0.7436 \text{ atau } 74.36\%$$

$$Recall(\text{negatif}) = \frac{1180}{1180 + 12} = \frac{1180}{1192} = 0.9899 \text{ atau } 98.99\%$$

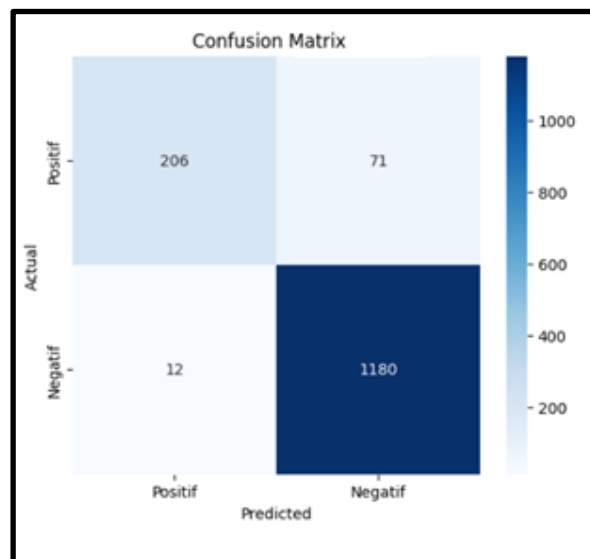
d. *F1-Score*

$$F1(\text{positif}) = 2 \times \frac{Precision \cdot Recall}{Precision + Recall} \quad (12)$$

$$F1(\text{negatif}) = 2 \times \frac{Precision \cdot Recall}{Precision + Recall}$$

$$F_1(\text{positif}) = 2 \times \frac{0.9449 \times 0.7436}{0.9449 + 0.7436} = 0,8322 = 83,22\%$$

$$F_1(\text{negatif}) = 2 \times \frac{0.9432 \times 0.9899}{0.9432 + 0.9899} = 0,9659 = 96,59\%$$



Gambar 3. Hasil *Confusion Matrix*

4. HASIL DAN PEMBAHASAN

Berdasarkan hasil pengujian menggunakan *Confusion Matrix*, memperoleh skor keakuratan 94.35%. Nilai kecermatan tersebut dihitung dari proporsi banyaknya komentar yang berhasil dikelompokkan secara tepat Positif Benar dan Negatif Benar terhadap keseluruhan data pengujian. Jika ditinjau berdasarkan masing-masing kelas, diperoleh nilai presisi sebesar 94.49% untuk kelas positif dan 94.32% untuk kelas negatif. Sementara itu, skor kelengkapan pada kategori positif mencapai 74.36% sedangkan pada golongan negatif senilai 98.99%. Nilai F1-score yang dihasilkan

adalah 83,22% untuk kelas positif dan 96,59% untuk kelas negatif. Hasil tersebut secara langsung menjawab rumusan masalah penelitian, yaitu bagaimana hasil penelaahan opini ulasan warganet terkait aspirasi 17+8 poin rakyat melalui penerapan pendekatan *Naive Bayes Classifier*. Berdasarkan hasil analisis sentimen yang dapat dilihat dari distribusi jumlah data pada kelas *sentiment* yang menunjukkan bahwa komentar negatif lebih banyak dibanding komentar positif. Hasil analisis menunjukkan terdapat 81,14% komentar negatif, dan 18,86% komentar positif. Dengan demikian menunjukkan bahwa hasil analisis sentimen berdasarkan 1837 komentar netizen pada media sosial X terkait 17+8 Tuntutan Rakyat, peneliti menyimpulkan bahwa komentar netizen cenderung bersentimen negatif. Selain itu, hasil penelitian ini juga telah mencapai tujuan penelitian, yaitu mengetahui kecenderungan sentimen komentar netizen apakah bersifat positif atau negatif. Melalui proses klasifikasi teks melalui penerapan pendekatan algoritma *Naive Bayes* kajian ini sukses mengidentifikasi bahwa opini publik pada periode 18 Agustus hingga 31 Oktober 2025 lebih banyak mengarah pada sentimen negatif. Dari sisi manfaat penelitian, dimanfaatkan guna menyajikan deskripsi berbasis angka terkait sudut pandang warga terkait aspirasi 17+8 poin rakyat di jejaring sosial X. Temuan dari kajian ini difungsikan dapat berfungsi sebagai landasan bagi kajian berikutnya dimana ingin melakukan analisis sentimen terhadap kebijakan publik atau isu sosial lainnya dengan menggunakan pendekatan klasifikasi teks.

5. KESIMPULAN DAN SARAN

Mengacu pada temuan studi tersebut, kini sudah diselesaikan kemudian ditarik konklusi bahwa penelaahan opini kepada ulasan warganet terkait 17+8 Tuntutan Rakyat di jejaring sosial X melalui penggunaan pendekatan *Naive Bayes Classifier* menunjukkan bahwa komentar netizen cenderung bersentimen negatif, hasil tersebut menggambarkan kecenderungan opini masyarakat yang terekam melalui data komentar pada periode pengambilan data penelitian. Luaran penilaian menggunakan *Confusion Matrix* mengungkapkan pendekatan *Naive Bayes* memperoleh nilai ketepatan 94.35%, didapatkan melalui proporsi banyaknya komentar tersebut berhasil dikategorikan secara akurat kepada semua komentar. Nilai presisi pada kategori positif mencapai 94.49% sementara kategori negatif mencapai 94.32, sedangkan nilai recall untuk kelas positif sebesar 74.36% dan kelas negatif sebesar 98.99%. Adapun nilai F1-score yang diperoleh adalah sebesar 83,22% untuk kelas positif dan 96,59% untuk kelas negatif.

Hasil analisis menunjukkan terdapat 81,14% komentar negatif, dan 18,86% komentar positif. Dengan demikian menunjukkan bahwa hasil analisis sentimen berdasarkan 1837 ulasan warganet di jejaring sosial X perihal aspirasi 17+8 poin rakyat, penulis menyatakan bahwa ulasan warganet lebih banyak bernada negatif. Hasil kajian ini ditargetkan berfungsi sebagai landasan bagi kajian berikutnya di ranah analisis sentimen terhadap isu sosial atau kebijakan publik menggunakan pendekatan klasifikasi teks.

Bertolak dari temuan kajian yang sudah lebih dulu diselesaikan serta keterbatasan dimana masih terdapat dalam penelitian ini, terdapat saran sebagai bahan pertimbangan untuk pengembangan penelitian selanjutnya yaitu, disarankan untuk menerapkan atau membandingkan metode klasifikasi, seperti Support Vector Machine, Random Forest, atau cara kerja Deep aringan saraf tiruan mendalam dan metode pemilahan data lainnya, guna 81 meningkatkan kinerja metode dalam memberikan hasil pada penelitian yang dilakukan.

DAFTAR REFERENSI

- Agustina, V., & Herliana, A. (2025). Analisis Sentimen Publik atas Kebijakan Efisiensi Anggaran 2025 dengan Text Mining dan Natural Language Processing. *Jurnal MediaInformatika*, 6(3), 2182–2194.
<https://ejournal.sisfokomtek.org/index.php/jumin/article/view/6301>
- Alfariz sirfatullahi, M. R., Al-farish Zidan, M., Taufiqurrahman, Muhamad Ardiansah, G., & Elgar, M. (2023). Penggunaan Python Sebagai Bahasa Pemrograman Untuk Machine Learning Dan Deep Learning. 2, 1–6.
- Amini, A. T., & Setiawan, K. (2024). Application of the Naive Bayes Algorithm in Twitter Sentiment Analysis of 2024 Vice Presidential Candidate Gibran Rakabuming Raka using Rapidminer. *International Journal Software Engineering and Computer Science (IJSECS)*, 4(1), 234–246.
<https://doi.org/10.35870/ijsecs.v4i1.2236>
- Bhatara, D. W., & Suryono, R. R. (2024). Analisis Sentimen Aplikasi Bca Mobile Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine. 9(4), 1907–1917.
- Dwicahyo, K. (2021). *Perbandingan Metode Web Scraping Dalam Pengambilan Data : Kajian Literatur*.
- Fitrianti, S., & Yudhistira, A. (2025). Analisis Sentimen Media Sosial Terhadap Calon Pilkada 2024 Dengan Metode Naïve Bayes Social Media Sentiment Analysis Towards 2024 Regional Election Candidates Using the Naïve Bayes Method. 5(1), 167–176.
- Fransiscus, & Girsang, A. S. (2022). Sentiment Analysis of COVID-19 Public Activity Restriction (PPKM) Impact using BERT Method. *International Journal of Engineering Trends and Technology*, 70(12), 281–288.
<https://doi.org/10.14445/22315381/IJETT-V70I12P226>

- Hidayah, N., & Dodiman. (2024). Implementasi Algoritma Multinomial Naive Bayes , TF-IDF dan Confusion Matrix dalam Pengklasifikasian Saran Monitoring dan Evaluasi Mahasiswa Terhadap Dosen Teknik Informatika Universitas Dayanu Ikhsanuddin. 10, 8–15.
- Irfon, G., & Soen, E. (2022). *Implementasi Cloud Computing dengan Google Colaboratory Pada Aplikasi Pengolah Data Zoom Participants*. 6(1), 24–30.
- Jannah, N. F., Fatonah, R. A., & Turnip, M. (2023). Analisis Wacana Van Dijk Terhadap Opini Publik Di Media Sosial Twitter Program Studi Penerbitan , Politeknik Negeri Jakarta * Corresponding Author PENDAHULUAN Pada Tanggal 27 Oktober 2023 , Masyarakat Dikejutkan Dengan Kasus Pembunuhan Berencana Yang Dila. 05(November), 179–195.
- Jon, A. M. (2023). *Analisis Sentimen Pada Media Sosial Instagram Klub Persija Jakarta Menggunakan Metode Naive Bayes*.
- Ketut, N., & Ida, A. (2023). *Variasi dan Karakteristik Komentar Netizen di Media Sosial*. November, 7–9.
- Khusein, A., Nilamsari, R. D., Selvia, F., & Ulfah, S. M. (2025). *Pengaruh Paparan Media Sosial Instagram Terhadap Persepsi Mahasiswa Ilmu Administrasi Negara Angkatan 2023 Universitas Palangka Raya Terkait Tuntutan 17 + 8 Di Indonesia Tahun 2025*. 12(2), 905–914.
- Kurniawan, R. P., & Iriananda, S. W. (2025). Analisis Sentimen Ulasan Aplikasi LinkedIn Berbasis Lexicon Dan Long Short-Term Memory (Lstm). 9(2), 2315–2324.
- Mandasari, S., Hayadi, B. H., & Gunawan, R. (2022). Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD Analisis Sentimen Pengguna Transportasi Online Terhadap Layanan Grab Indonesia Menggunakan Multinomial Naive Bayes Classifier Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD. 5, 118–126.
- Munthe, C. J. E., Hasibuan, N. A., & Hutabarat, H. (2022). Penerapan Algoritma Text Mining Dan TF-RF Dalam Menentukan Promo Produk Pada Marketplace. 2(3), 110–115.
- Putra Tri, K., Hariyadi Amin, M., & Crysdian, C. (2023). Perbandingan Feature Extraction Tf-Idf Dan Bow Untuk. 1449–1463.
- Putri Duei, Dianati. Forda, Gigih. Sulistiano Eko, W. (2022). Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier. *Jurnal Informatika Dan Teknik Elektro Terapan*, 10(1), 34–40. <https://doi.org/10.23960/jitet.v10i1.2262>
- Ranawati, N. K. (2025). Apa Itu 17+8 Tuntutan Rakyat yang Ramai di Medsos dan Maknanya. 02 Sep 2025. <https://www.detik.com/jabar/berita/d-8092490/apa-itu-17-8-tuntutan-rakyat-yang-ramai-di-medsos-dan-maknanya>
- Salsabila, K. A., Anastasya, N., Sartiza, Z., Simanjuntak Christina, A., Nafis, M., Situmorang R, K., Manurung, R., & Annisa, S. (2025). Dpr Pada Era Presiden Prabowo (17 + 8 People ' S Demands : Crisis Of Trust And Reform In The Dpr During President Prabowo ' S Era). September, 16686–16695.
- Setyaningrum, P. (2025). Apa Itu 17+8 Tuntutan Rakyat yang Ramai di Media Sosial? Ini Latar Belakang, Arti, dan Isinya. Kompas.Com. <https://www.kompas.com/jawa-tengah/read/2025/09/02/060338388/apa-itu-178-tuntutan-rakyat-yang-ramai-di-media-sosial-ini-latar?page=all#page2>
- Simanjuntak, A. Y., Septian, I., & Simatupang, S. (2022). Implementasi Data Mining

- Menggunakan Metode Naïve Bayes Classifier Untuk Data Kenaikan Pangkat Dinas. 4307(1), 85–91.
- Siregar Alfarizi, K., Nasution, S., & Nabway, P. (2025). Analisis Sentimen Netizen Indonesia Terhadap Kampanye Penggunaan Kecerdasan Buatan Oleh Pemerintah Menggunakan Algoritma Naïve Bayes. *Jurnal Ilmu Komputer Dan Informatika*, 1(4), 9.
- Sumiyati, N., Fauziah, E., & Basid, A. (2020). Corpus-Based Analysis : Collocation Of Netizen S ' Words In Indonesian News In 2020. 236–251.
- Utomo, D. P., & mesran. (2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung. 4(April), 437–444. <https://doi.org/10.30865/mib.v4i2.2080>
- Yudistira, F., & Isnain Rahman, A. (2025). Analisis Sentimen Terhadap Seleksi CPNS Tahun 2024 Berbasis Media Sosial X Menggunakan Algoritma Naïve Bayes Sentiment Analysis Of The 2024 Cpns Selection Based On Social Media X Using The Naïve Bayes Algorithm. *Jurnal Pendidikan Dan Teknologi Indonesia (Jpti)*, 5(3), 887–897.