

Analisis Sentimen Komentar Youtube Terkait Vidio Demo DPR Menggunakan Algoritma Naïve Bayes

Zaehol Fatah¹, Khoiriyah Agustin Ulumiyah²
^{1,2}, Universitas Ibrahimy, Situbondo
Email: khoiriyahagustin684@gmail.com

Abstrak

Youtube telah berkembang menjadi situs media sosial yang populer bagi public untuk mengekspresikan pandangan dan reaksi mereka terhadap berbagai isu, termasuk peristiwa sosial dan politik seperti demonstrasi DPR. Dalam studi ini, metode Naïve Bayes digunakan untuk mengategorikan komentar teks berbahasa Indonesia pada video terkait dibagi menjadi tiga kategori: positif, negatif, dan netral. Hasilnya menunjukkan bahwa mayoritas komentar bersifat netral (88,1%), dengan 9,3% negatif dan hanya 2,4% positif, yang menunjukkan bahwa pengguna umumnya mengekspresikan reaksi non-emosional atau umum tanpa dukungan atau kritik eksplisit. Selain itu, efektivitas algoritma dalam mengklasifikasikan sentimen teks Bahasa Indonesia telah dibuktikan, terutama setelah Langkah-langkah pra proses yang tepat diterapkan. Studi ini berkontribusi pada literatur dengan mengkaji opini public terhadap isu-isu demonstrasi publik, menawarkan temuan yang dapat digunakan oleh pemangku kepentingan dalam memahami reaksi public yang cepat.

Kata Kunci: Analisis Sentimen, Naïve Bayes, Demo DPR, Youtube

Abstract

YouTube has grown in popularity as a social networking site for the public to express their views and reactions toward various issues, including social and political events such as the DPR demonstrations. In this study, the Naïve Bayes algorithm was used to classify Indonesian-language text comments on DPR-related videos divided positive, negative, and neutral. The findings indicate that the majority of comments are neutral (88.1%), with 9.3% negative and only 2.4% positive, indicating that users generally expressed non-emotional or general reactions without explicit support or criticism. Moreover, the effectiveness of the algorithm in classifying Indonesian text sentiment has been demonstrated, especially after proper preprocessing steps are applied. This research of literature looking into public opinions toward political demonstration issues, offering conclusions that might be used as a guide for stakeholders in understanding rapid public reactions

Keywords: Sentiment Analysis, Naïve Bayes, DPR Demo, YouTube.

PENDAHULUAN

Seiring dengan kemajuan teknologi internet dan platform media sosial, situs berbagi video seperti YouTube telah bertransformasi menjadi sarana utama bagi masyarakat untuk mengekspresikan pandangan dan reaksi mereka terhadap berbagai isu, termasuk peristiwa sosial dan politik seperti demonstrasi DPR. Komentar yang ada di YouTube menjadi sumber informasi yang kaya untuk menyelidiki

sentiment public mengenai suatu topik tertentu. Analisis sentimen merujuk pada metode yang digunakan untuk mengategorikan pandangan pengguna menjadi kategori positif, negative atau netral. Beragam studi telah menunjukkan bahwa algoritma Naïve Bayes sangat efektif dalam melakukan analisis sentiment berkat kesederhanaannya dan hasil yang cukup baik dalam mengklasifikasikan teks (Gudiato et al., 2024) (Aiswarya A S &

Rajeev, 2024)(Rambe et al., 2023). Pendekatan ini bisa membantu dalam mengidentifikasi tren sentiment masyarakat terkait video yang membahas demonstrasi DPR, sehingga mempermudah pemahaman terhadap pendapat publik yang terdapat di platform media sosial.

Dalam zaman digital modern, situs-situs web sosial seperti YouTube sudah bertransformasi mengubah saluran penting untuk orang-orang tentang mengekspresikan pandangan mereka, khususnya melalui ruang komentar yang terdapat di setiap video. Ulasan-ulasan ini mencerminkan berbagai pandangan dan reaksi publik mengenai masalah-masalah politik, termasuk video demonstrasi DPR yang sering kali memicu respon emosional dan pendapat yang bervariasi. Penting untuk melakukan analisis sentiment terhadap ulasan ini guna mendapatkan pemahaman yang lebih mendalam tentang persepsi masyarakat serta sebagai dasar pengambilan keputusan yang lebih akurat bagi para pembuat kebijakan dan pengelola situs media sosial(Andrew et al., 2024)(Firsttama et al., 2024).

Metode yang sering diterapkan dalam analisis sentiment adalah algoritma klasifikasi, salah satunya adalah Naïve Bayes. Karena efisiensinya dalam mengklasifikasikan teks dengan tingkat presisi yang tinggi dan proses komputasinya yang terbilang cepat. Naïve Bayes beroperasi berdasarkan prinsip probabilitas dan mengasumsikan bahwa fitur-fitur bersifat independen, sehingga sesuai untuk mengelompokkan komentar ke dalam kategori positif, negatif, atau netral. Berbagai penelitian sebelumnya menunjukkan bahwa Naïve Bayes mampu menghasilkan analisis yang cukup akurat terhadap komentar di platform media sosial, termasuk YouTube(Ardiyansyah et al., 2024)(Teknik et al., 2025).

Algoritma Naïve Bayes sebenarnya adalah sebuah model yang bersifat probabilistik, yang berlandaskan pada teorema Bayes dengan panduan bahwa antar fitur itu sendiri. Dalam dunia analisis

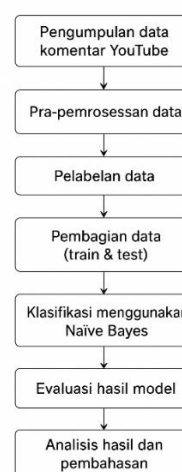
sentiment, algoritma tersebut menilai kemungkinan bahwa suatu komentar termasuk dalam kategori sentimen positif, negatif, atau netral. Naïve Bayes sering dijadikan pilihan karena kemampuannya dalam mengolah teks dengan cepat dan secara efisien, meskipun berhadapan dengan kumpulan data yang besar, serta cara penerapannya yang mudah(Angdresey et al., 2025)(Gudiato et al., 2024)(Aiswarya A S & Rajeev, 2024). Selain itu, berbagai studi yang mengkaji analisis sentimen di platform YouTube menunjukkan bahwa metode ini mampu mencapai tingkat akurasi yang cukup baik dalam mengklasifikasikan komentar yang bersifat subjektif dan beragam(Gudiato et al., 2024)(Armaeni et al., 2024).

Melalui studi ini, dicapai wawasan luas tentang pandangan masyarakat tema demonstrasi DPR, serta membuktikan seberapa efektif algoritma Naïve Bayes dalam menganalisis data teks dalam Bahasa Indonesia, terutama dalam lingkungan media sosial seperti YouTube.

METODE

Jenis studi kuantitatif ialah penggalian data dalam analisis sentimen. Sasaran dari penelitian ini adalah untuk mengelompokkan sentiment komentar pengguna YouTube mengenai video DPR kelompok, positif, negatif, dan netral, dengan memanfaatkan algoritma Naive Bayes.

METODE PENELITIAN



Gambar 1. Metode Penelitian

Pengumpulan Data

Data untuk studi dihimpun secara otomatis dengan menggunakan Youtube Data API v3. API berfungsi mengakses komentar video Youtube tema penelitian, yakni video yang berhubungan dengan Demo DPR. Pengambilan data dilaksanakan dengan memanfaatkan endpointcommentThreads pada Youtube Data API untuk mendapatkan komentar beserta data pendukung lainnya. Data terkumpul dengan sejumlah atribut penting.

Pra-Pemrosesan Data

Tahap Proses pra-pemrosesan dilakukan melalui beberapa langkah sebagai berikut:

1. Cleansing

Pada tahap ini, teks komentar dibersihkan dari elemen-elemen yang tidak relevan, seperti tanda baca, angka, url, emoji, dan simbol lain yang tidak memiliki maksna. Hasilnya adalah teks yang lebih bersih dan seragam untuk analisis lebih lanjut.

2. Case Folding

Setiap huruf di komentar dikonversi menjadi huruf yang lebih kecil untuk model tidak dapat membedakan huruf kapital dari huruf kecil, misalnya kata "Demo" dan "demo" dianggap identic.

3. Tokenizing

Proses ini memecah setiap kalimat komentar menjadi bagian-bagian kata. Sebagai contoh, kalimat "saya mendukung aksi demo" akan diubah menjadi daftar token ["saya", "mendukung", "aksi", "demo"].

4. Stopword Removal

Pada tahap ini, kata-kata biasa dan tidak signifikan terhadap evaluasi sentimen dihilangkan, seperti "yang", "di", "dan", "ke", dan sebagainya. Proses ini dilakukan menggunakan daftar stopwords Bahasa Indonesia dari library NLTK.

5. Stemming

Setiap kata diubah Kembali, contohnya kata "berjalan", "jalan"

dan "mendukung" menjadi "dukung". Tujuannya adalah agar kata-kata dengan makna yang sama dapat dikenali sebagai satu entitas.

Pelabelan Data

Setiap penilaian dikategorikan melalui proses manual yang mempertimbangkan arti kalimat.

Tabel 1. Pelabelan Data

Positif	Negatif	Netral
Akhirnya ada juga yang berani turun ke jalan buat suarain rakyat, salut sama demo hari ini.	Nggak jelas banget, ngomongnya soal rakyat tapi ujung-ujungnya rusuh, capek liatnya.	Tadi sempet liat demo di depan Gedung dpr, lumayan ramai tapi aman sih.

Pembagian Data

setelah selesai pelabelan, pembagian data dibagi menjadi dua segmen:

1. Data Latihan : 80% dari keseluruhan data.
2. Data pengujian : 20% dari keseluruhan data

Pembagian ini dilakukan secara acak dengan memanfaatkan fungsi train_test_split yang ada di library scikit-learn.

Klasifikasi Algoritma Naïve Bayes

Setelah proses pra-pemrosesan dan penandaan data komentar selesai, Langkah selanjutnya adalah melakukan klasifikasi sentimen dengan memanfaatkan algoritma Naive Bayes. Algoritma ini dipilih karena terbukti efektif dalam mengklasifikasikan teks dan memiliki efesiensi tinggi dalam menghitung probabilitas untuk masing-masing kategori sentimen. Rumus dasarnya adalah sebagai berikut:

$$P(X) = \frac{P(X).P(C)}{P(X)}$$

Dalam studi ini diterapkan varian Multinomial Naïve Bayes, yang sering dipakai untuk klasifikasi teks seperti analisis perasaan karena memperhatikan seberapa sering kata kata muncul.

Evaluasi Model

Penilaian kinerja model dilaksanakan melalui confusion matrix dengan menggunakan ukuran-ukuran berikut:

- Accuracy (akurasi)

- Precision (ketepatan)
- Recall (daya ingat)
- F1-Score (rata-rata konsisten precision dan recall)

Rumus:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

HASIL DAN PEMBAHASAN

Tahapan pra-pemrosesan dilakukan dengan langkah-langkah kunci: case folding, cleansing, tokenizing, stopword folding, dan stemming. Setelah metode pembersihan, data komentar yang sebelumnya mengandung huruf kapital, emoji, link, serta tanda baca telah disusun menjadi teks yang lebih teratur dan bersih.

Tabel 2. Hasil Pra-pemrosesan

Sebelum	Sesudah
Wah demo DPR gak jelas 😏👎!!#tolakRU U	Demo dpr tidak jelas tolak ruu

Tahapan ini sangat penting untuk menjamin bahwa data yang memasuki proses klasifikasi tidak terdistorsi oleh variasi bentuk kata yang berbeda namun bermakna sama.

Setiap komentar kemudian diberi label sentimen, dikelompokkan dalam kategori positif, negatif, dan netral. Pelabelan dilakukan berdasarkan interpretasi komentar terkait peristiwa demo dpr apakah komentar itu bersifat mendukung (positif), mengkritik (negatif), atau netral.

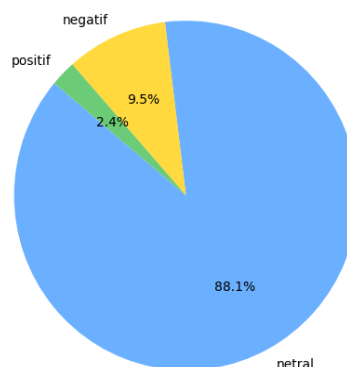
Tabel 3. Distribusi Sentimen Komentar.

Kategori	Jumlah	Presentase
Positif	45	2,4%
Negatif	175	9,3%
Netral	1658	88,1%
Jumlah	1882	100%

Table 3 ini mengindikasikan bahwa dari total 1882 komentar, Sebagian besar, yaitu 1658 komentar (88,1%), memiliki

sifat netral. Ada 175 komentar negative (9,3%), dan hanya 45 komentar yang mengandung nuansa positif (2,4%). Ini menunjukkan bahwa Sebagian besar pengguna cenderung memberikan tanggapan umum tanpa emosi yang jelas, sedangkan jumlah komentar negatif melebihi jumlah komentar positif.

Distribusi Sentimen Komentar YouTube



Gambar 2. Persentase Komentar Berdasarkan Sentimen

Gambar di atas memperlihatkan bagaimana sentiment komentar di youtube berkaitan dengan topik dibahas. Dari grafik tersebut, terlihat bahwa komentar yang bersifat netral menguasai dengan proporsi sekitar 88,1%, menandakan bahwa mayoritas pengguna memberikan tanggapan yang informatif atau tidak berpihak. Sementara itu, komentar negative mencakup 9,5%, menunjukkan bahwa hanya Sebagian kecil pengguna yang mengungkapkan kritik atau ketidakpuasan. Komentar positif hanya menunjukkan 2,4%, yang menunjukkan bahwa dukungan atau apresiasi dalam komentar tergolong rendah. Secara keseluruhan, temuan ini menunjukkan bahwa reaksi pengguna internet terhadap topik yang dibahas cenderung netral, dengan sedikit kecenderungan negatif.

Hasil dari gambaran menunjukkan bahwa terdapat lebih banyak sentiment netral, yang mengindikasikan bahwa sebagian besar pengguna internet memberikan tanggapan yang factual tanpa menunjukkan rasa dukungan atau penolakan yang jelas terhadap peristiwa demo dpr. Namun, komentar negatif juga cukup

signifikan, mengindikasikan adanya kritik atau ketidakpuasan terhadap jalannya aksi

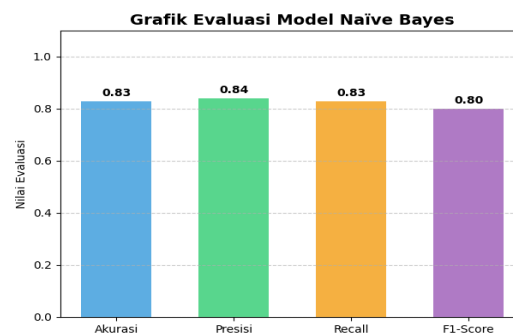
Setelah data dibersihkan dan dilabeli, proses klasifikasi dilakukan dengan memanfaatkan algoritma Multinomial Naïve Bayes, yang sering digunakan dalam analisis teks karena kemampuannya menangani data yang berbasis frekuensi kata (bag-of-words). Model dilatih menggunakan 80% dari data yang tersedia untuk pelatihan dan diuji dengan 20% data testing. Hasil evaluasi menunjukkan performa model yang cukup memuaskan dalam mengidentifikasi perbedaan komentar berdasarkan polaritasnya.

Tabel 4. Hasil Evaluasi Model Naïve Bayes

metrik	Nilai
Akurasi	0,83
Precision	0,84
Recall	0,83
F1-Score	0,80

Table 4 ini memperlihatkan kinerja model Naive Bayes dalam mengidentifikasi sentiment dari komentar YouTube. Tingkat akurasi mencapai 0,83, yang menunjukkan bahwa 83%. Nilai Precision sebesar 0,84 mengindikasikan bahwa model ini cukup akurat dalam menebak sentiment yang benar tanpa terlalu banyak kesalahan. Recall yang mencapai 0,83 mencerminkan kemampuan model dalam mendeteksi semua data yang relevan dengan konsisten. Sementara itu, F1-Score 0,80 menunjukkan adanya keseimbangan antara Precision dan Recall, meskipun sedikit lebih rendah, tetapi masih mencerminkan kinerja yang baik. Secara keseluruhan, tabel ini menunjukkan bahwa model Naïve Bayes cukup efektif dalam melakukan analisis sentimen.

Hasil klasifikasi kemudian divisualisasikan melalui grafik batang untuk memperlihatkan penyebaran jumlah komentar dalam setiap kategori.



Gambar 3. Grafik Evaluasi Model Naïve Bayes

Memperlihatkan berdasarkan empat matrik penilaian. Hasil yang diperoleh menunjukkan bahwa model ini berfungsi dengan baik, dengan akurasi mencapai 0,83 yang berarti 83% dari prediksinya akurat. Precision yang sebesar 0,84 menunjukkan bahwa model ini cukup efektif dalam mengenali sentiment yang tepat, sementara nilai Recall 0,83 menjelaskan seberapa baik model ini dapat mengenali semua data sentimen dengan konsisten. Nilai F1-Score yang tercatat adalah 0,80, yang mencerminkan keseimbangan antara precision dan recall, meskipun sedikit lebih rendah dibandingkan dengan metrik lainnya. Secara keseluruhan, grafik ini mengindikasikan bahwa model beroperasi dengan baik dan efisien dalam mengklasifikasikan sentiment pada komentar youtube.

SIMPULAN

Hasil analisis mengungkapkan bahwa mayoritas komentar di YouTube mengenai video Demo DPR bersifat netral (88,1%), yang menandakan bahwa masyarakat cenderung memberikan tanggapan yang informatif, deskriptif, atau hanya sekedar menyampaikan pendapat tanpa menunjukkan bahwa diskusi di kolom komentar lebih condong pada penyampaian pandangan umum mengenai kejadian tersebut. Di samping itu, efektivitas algoritma Naïve Bayes telah terbukti dalam mengklasifikasikan sentiment teks berbahasa Indonesia setelah melewati proses preprocessing yang memadai.

DAFTAR PUSTAKA

- Aiswarya A S, & Rajeev, H. (2024). Youtube Comment Sentimental Analysis. *Indian Journal of Data Mining*, 4(1), 5–8. <https://doi.org/10.54105/ijdm.a1633.04010524>
- Andrew, C., Budiyantara, A., & Lewenusa, I. (2024). Analisis Sentimen Komentar Pengguna YouTube Untuk Ponsel Di Indonesua Berbasis Web. *Scientica: Jurnal Ilmiah Sains Dan Teknologi*, 3 No. 2, 175–188. <https://jurnal.kolibri.org/index.php/scientica/article/view/4212>
- Angdresey, A., Sitanayah, L., & Tangka, I. L. H. (2025). Sentiment Analysis for Political Debates on YouTube Comments using BERT Labeling, Random Oversampling, and Multinomial Naïve Bayes. *Journal of Computing Theories and Applications*, 2(3), 342–354. <https://doi.org/10.62411/jcta.11668>
- Ardiyansyah, R., Makmur, I., & Phai, S. P. (2024). SEMINAR NASIONAL CORISINDO Sentimen Komentar Youtube Dengan Sentiment Intensity Analyzer Dari NLTK. 31–36.
- Armaeni, P. P., Wiguna, I. K. A. G., & Parwita, W. G. S. (2024). Sentiment Analysis of YouTube Comments on the Closure of TikTok Shop Using Naïve Bayes and Decision Tree Method Comparison. *Jurnal Galaksi*, 1(2), 70–80. <https://doi.org/10.70103/galaksi.v1i2.15>
- Firsttama, R. A., Arifiyanti, A. A., & Kartika, D. S. Y. (2024). Analisis Sentimen Komentar Youtube Konferensi Tingkat Tinggi G20 Menggunakan Metode Naive Bayes. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 6(2), 282–285. <https://doi.org/10.47233/jteksis.v6i2.1263>
- Gudiato, C., Cahyaningtyas, C., & P., N. (2024). Analisis Sentimen Komentar Youtube Terhadap Tayangan #Terbaru! Temuan dan Masalah Ahlak Di Ponpes Al-zaytun Menggunakan Metode Naïve Bayes. *G-Tech : Jurnal Teknologi Terapan*, 8(1), 186–195. <https://ejournal.uniramalang.ac.id/index.php/g-tech/article/view/1823/1229>
- Rambe, T. S., Hasibuan, M. N. S., & Dar, M. H. (2023). Sentiment Analysis of Beauty Product Applications using the Naïve Bayes Method. *Sinkron*, 8(2), 980–989. <https://doi.org/10.33395/sinkron.v8i2.12303>
- Teknik, J., Informasi, T., Andhini, A., Handayani, F. N., Diasih, I., Ji, A., No, B., Serengan, K., Surakarta, K., & Tengah, J. (2025). Analisis Sentimen Opini Publik pada Channel Youtube Mata Najwa Menggunakan Metode SVM Universitas Duta Bangsa Surakarta , Indonesia Perkembangan pesat media sosial dan platform video seperti YouTube telah menciptakan ruang baru bagi masyarakat untuk menge. 5.
- Zaehol Fatah, M. (2025). In *TIK dan Masyarakat* (pp. 144-147). Yogyakarta: PT. Penamuda Media.
- Fauzan, R. A., & Mufti, M. (2023). Analisis Sentimen Komentar Youtube Program Kampus Merdeka Berbasis Web Menggunakan Algoritma Multinomial Naïve Bayes. *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 2(2), 864–871. <https://senafti.budiluhur.ac.id/index.php/senafti/article/view/929/563>
- Harahap, S. M., & Kurniawan, R. (2024). Analisis Sentimen Komentar Youtube Terhadap Food Vlogger Dengan Menggunakan Metode Naïve Bayes. *MEANS (Media Informasi Analisa Dan Sistem)*, 9(1), 87–96. <https://doi.org/10.54367/means.v9i1.3912>
- Huwaida, S. F., Kusumawati, R., & Isnaini, B. (2024). Analisis Sentimen Komentar YouTube terhadap Pindahan Ibu Kota Negara Menggunakan Metode Naïve Bayes.

Jambura Journal of Informatics, 6(1),
26–39.

<https://doi.org/10.37905/jji.v6i1.24718>

Mulyati, S., Syawali, F., Slameto, I. N.,
Nugraha, R. S., & Ratama, N. (2022).
*Pengembangan Aplikasi Web
Penerimaan Beasiswa Menggunakan
Model Waterfall*. 5(4), 226–235.
<https://doi.org/10.32493/jtsi.v5i4.21125>

Rambe, T. S., Hasibuan, M. N. S., & Dar,
M. H. (2023). Sentiment Analysis of
Beauty Product Applications using the
Naïve Bayes Method. *Sinkron*, 8(2),
980–989.
<https://doi.org/10.33395/sinkron.v8i2.12303>