

Analisis Sentimen Tanggapan Masyarakat Terhadap Program Makan Bergizi Gratis Pada Platform YouTube Menggunakan SVM

Wildan Rayyana Muntaza¹

¹Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri

E-mail: *1wildanjrl20hyper@gmail.com

Abstrak – Program Makanan Bergizi Gratis (MBG) menuai reaksi tajam masyarakat, khususnya setelah munculnya insiden dugaan keracunan yang memicu krisis kepercayaan publik. Penelitian ini bertujuan untuk memetakan sentimen masyarakat secara objektif guna mengklarifikasi disparitas hasil penelitian terdahulu yang menunjukkan polarisasi opini yang kontradiktif. Metode penelitian menggunakan pendekatan kuantitatif dengan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan komentar dari platform YouTube. Data dikumpulkan melalui teknik scraping, kemudian diproses melalui tahapan text mining yang meliputi pembersihan data, stemming menggunakan pustaka Sastrawi, serta pembobotan kata menggunakan TF-IDF. Proses pelabelan data memanfaatkan model bahasa RoBERTa untuk memastikan validitas kelas sentimen. Hasil pengujian menunjukkan bahwa algoritma SVM efektif dalam mengklasifikasikan data teks dengan tingkat akurasi sebesar 84%. Temuan penelitian mengungkap bahwa sentimen publik secara signifikan didominasi oleh respons negatif, yang merefleksikan kekhawatiran mendalam terhadap standar keamanan pangan dan pengawasan kualitas program. Penelitian ini menyimpulkan bahwa isu keracunan menjadi faktor determinan penurunan citra program, sehingga hasil analisis ini penting dijadikan landasan evaluasi mendesak bagi pemerintah untuk memperbaiki standar operasional prosedur guna memulihkan kepercayaan masyarakat.

Kata Kunci — Analisis Sentimen, Makan Bergizi Gratis, Support Vector Machine, TF-IDF, YouTube

1. PENDAHULUAN

Media sosial, khususnya YouTube, telah bertransformasi menjadi wadah utama bagi masyarakat untuk menyuarakan pendapat dan memberikan umpan balik terhadap fenomena sosial maupun kebijakan pemerintah secara *real-time* [1]. Kolom komentar pada video berita menyediakan data tekstual tidak terstruktur yang kaya akan informasi. Jika diolah dengan teknik *text mining*, data ini dapat mengungkap pola sentimen publik, baik yang bersifat mendukung, menolak, maupun netral [2]. Salah satu kebijakan yang saat ini menjadi sorotan tajam adalah Program Makanan Bergizi Gratis (MBG). Meskipun bertujuan memperbaiki gizi, implementasinya menghadapi tantangan berat berupa dugaan kasus keracunan yang memicu lonjakan interaksi dan keresahan di masyarakat [3].

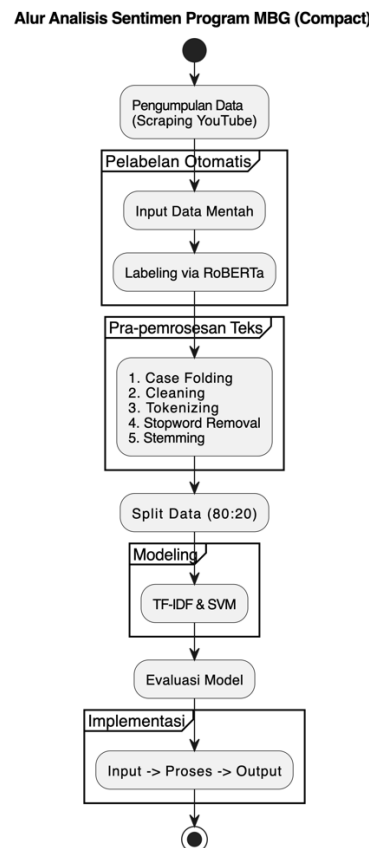
Pentingnya menganalisis sentimen terhadap kasus ini didasari oleh adanya polarisasi opini yang tajam dan hasil penelitian terdahulu yang kontradiktif. Sebuah studi menggunakan metode berbasis leksikon (*TextBlob*) pada media sosial menunjukkan dominasi sentimen negatif sebesar 71,8% [4]. Sebaliknya, penelitian lain yang menggunakan metode *Machine Learning* pada topik serupa justru menemukan sentimen positif sebesar 71,5% [5]. Kesenjangan hasil ini mengindikasikan adanya ketidakpastian persepsi publik yang perlu diklarifikasi menggunakan metode yang lebih akurat dan *robust* pada sumber data yang berbeda.

Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan algoritma *Support Vector Machine* (SVM). SVM dipilih karena kemampuannya dalam menemukan *hyperplane* terbaik untuk memisahkan kelas data pada ruang berdimensi tinggi, yang sangat cocok untuk data teks [6]. Beberapa studi komparasi menunjukkan bahwa SVM memiliki kinerja yang lebih unggul dibandingkan algoritma lain. Pada analisis sentimen tokoh publik, SVM terbukti mencapai akurasi lebih tinggi dibandingkan *Naïve Bayes* [7]. Selain itu, pada klasifikasi komentar YouTube terkait produk dan layanan, SVM juga mencatatkan akurasi yang lebih baik dibandingkan *Random Forest* dan *K-Nearest Neighbor* (KNN) [8], [9].

Penelitian ini menggunakan pendekatan berbasis aplikasi (sistem) yang mengintegrasikan tahapan *preprocessing* lengkap (termasuk *stemming* sastrawi dan *stopword removal*), pembobotan kata menggunakan TF-IDF (*Term Frequency-Inverse Document Frequency*), dan klasifikasi menggunakan kernel SVM. Tujuan utama dari penelitian ini adalah untuk mengukur kecenderungan sentimen masyarakat (Positif, Negatif, atau Netral) terhadap kasus keracunan program MBG secara objektif. Hasil

dari penelitian ini diharapkan dapat memberikan wawasan berbasis data bagi pemangku kepentingan untuk mengevaluasi standar keamanan pangan dan memperbaiki strategi komunikasi publik program MBG di masa mendatang.

2. METODE PENELITIAN



Gambar 1. Alur Analisis Sentimen

Berikut adalah penjelasan dari gambar 1

a. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah komentar publik yang terdapat pada video berita terkait kasus keracunan Program Makanan Bergizi Gratis (MBG) di platform YouTube. Pengumpulan data dilakukan menggunakan teknik *scraping* untuk mengambil komentar secara otomatis. Data mentah yang diperoleh kemudian disimpan dalam format .csv untuk diproses lebih lanjut.

b. Pelabelan Data

Berbeda dengan penelitian terdahulu yang sering menggunakan pelabelan manual atau leksikon sederhana, penelitian ini memanfaatkan model bahasa *RoBERTa* (*Robustly Optimized BERT Pretraining Approach*) yang telah dilatih sebelumnya (*pre-trained*) untuk melakukan pelabelan otomatis. Model *RoBERTa* dipilih karena kemampuannya memahami konteks kalimat yang kompleks dalam Bahasa Indonesia dengan lebih baik dibandingkan metode konvensional. Data komentar diklasifikasikan ke dalam tiga kelas sentimen: Positif, Negatif, dan Netral. Hasil pelabelan dari *RoBERTa* ini kemudian dijadikan *ground truth* untuk melatih model SVM.

c. Pra-pemrosesan Teks

Untuk meningkatkan kualitas data sebelum masuk ke tahap klasifikasi, dilakukan tahapan pra-pemrosesan menggunakan pustaka NLTK dan Sastrawi sebagaimana diimplementasikan dalam sistem:

- 1) *Case Folding*: Mengubah seluruh huruf menjadi huruf kecil (*lowercase*).
- 2) *Cleaning*: Membersihkan data dari angka, tanda baca, *url*, *username*, *hashtag*, dan karakter non-alfanumerik menggunakan *Regular Expression* (Regex).
- 3) *Tokenizing*: Memecah kalimat menjadi potongan kata tunggal (token) [2].

- 4) *Stopword Removal*: Menghapus kata-kata umum yang tidak memiliki makna sentimen signifikan (seperti "yang", "dan", "di") menggunakan daftar *stopword* Bahasa Indonesia dari NLTK.
- 5) *Stemming*: Mengubah kata berimbuhan menjadi kata dasar menggunakan algoritma *Sastrawi* untuk menyeragamkan variasi kata [5].

d. Ekstraksi Fitur

Data teks yang telah bersih dikonversi menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam dokumen dan kelangkaannya dalam seluruh korpus.

Rumus TF-IDF didefinisikan sebagai berikut:

$$W_{t,d} = tf(t, d) \times \log \left(\frac{N}{df(t)} \right) \quad (1)$$

Di mana $tf(t, d)$ adalah frekuensi kata t dalam dokumen d , N adalah total dokumen, dan $df(t)$ adalah jumlah dokumen yang mengandung kata t .

e. Klasifikasi Sentimen dengan Support Vector Machine (SVM)

Klasifikasi dilakukan menggunakan algoritma SVM. SVM bekerja dengan mencari *hyperplane* optimal yang memisahkan data antar kelas dengan *margin* terbesar. Dalam penelitian ini, kernel linear digunakan karena efektivitasnya dalam menangani data teks berdimensi tinggi [6].

Fungsi keputusan untuk *hyperplane* SVM dapat dituliskan sebagai:

$$f(x) = w \cdot x + b = 0 \quad (2)$$

Di mana w adalah vektor bobot, x adalah vektor input (fitur TF-IDF), dan b adalah bias. Jika $f(x) \geq 0$, data diklasifikasikan ke kelas positif, dan sebaliknya [8].

3. HASIL DAN PEMBAHASAN

a. Deskripsi Data

Data penelitian diperoleh melalui proses *scraping* pada kolom komentar YouTube yang membahas kasus keracunan Program Makanan Bergizi Gratis (MBG). Dari proses pengumpulan data, diperoleh dataset mentah yang berisi informasi terkait ID pengguna, tanggal komentar, dan teks komentar [1]. Dataset ini kemudian disaring untuk menghapus komentar duplikat dan komentar yang tidak relevan (seperti promosi atau *spam*), menghasilkan dataset final yang siap diproses. Distribusi data menunjukkan antusiasme publik yang tinggi dalam merespons isu ini, namun dengan nada yang cenderung kritis akibat insiden keracunan yang terjadi.

b. Hasil Pra-pemrosesan Teks (*Text Preprocessing*)

Tahapan pra-pemrosesan memegang peranan krusial untuk memastikan kualitas data sebelum masuk ke tahap klasifikasi. Proses ini meliputi *cleaning* (pembersihan), *case folding*, tokenisasi, *stopword removal*, dan *stemming* menggunakan pustaka *Sastrawi*. Berdasarkan sampel data yang diproses, terlihat bahwa teknik ini efektif mereduksi *noise* (gangguan) dalam teks tanpa menghilangkan makna intinya. Tabel 1 memperlihatkan perbandingan antara komentar asli dari YouTube dengan hasil teks yang telah melalui proses pembersihan (*cleaning*) dan *stemming*.

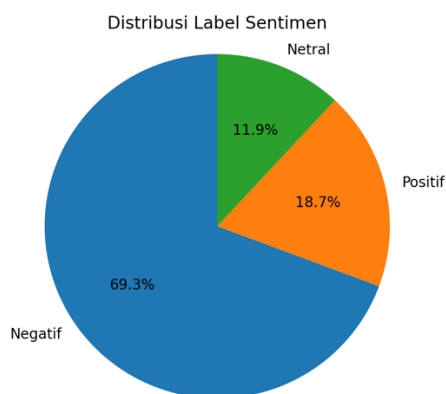
Tabel 1. Hasil Pra-pemrosesan Data

No	Penulis	Teks Asli	Teks Bersih
1	@orinamara	knapa banyak kasus keracunan mbg tapi pemerintah terus menggedur mbg... ada apa gerakan dengan program ini	knapa racun mbg perintah menggedur mbg gerakan program
2	@RendiRendi	Mending diganti uang 300 rb permurit yg tidak mampu jauh lebih efektif dan tepat sasaran terimakasih	mending ganti uang rb permurit efektif sasar terimakasih
3	@masahib5695	Program bermasalah tapi tetap dilanjutkan (dari nasi basi sampai keracunan). Sebenarnya ada apa dibalik itu semua???	program masalah lanjut nasi basi racun balik
4	@ka2nk	program bmg bagus tapi kalo bs keracunan bisa jadiMakanan terkontaminasi silang...	program bmg bagus racun makan kontaminasi silang makan pindah"

Seperti terlihat pada Tabel 1, kata-kata sambung yang tidak relevan dihapus, dan kata berimbuhan seperti "keracunan" diubah menjadi bentuk dasar "racun" untuk memudahkan pembobotan TF-IDF.

c. Hasil Pelabelan Data (*Data Labeling*)

Pelabelan data dilakukan secara otomatis menggunakan model *Deep Learning* RoBERTa untuk menentukan sentimen (Positif, Negatif, Netral). Berdasarkan hasil pelabelan pada dataset, ditemukan distribusi sentimen yang tidak seimbang (*imbalanced*). Dominasi sentimen Negatif terlihat sangat signifikan. Hal ini wajar mengingat topik spesifik yang diangkat adalah "kasus keracunan", yang secara alami memicu respons kekhawatiran, kritik, dan kemarahan publik. Contoh komentar negatif yang dominan menyoroti masalah "nasi basi", "keracunan", dan saran pengalihan anggaran ke bantuan tunai. Sentimen Positif umumnya berisi harapan agar program diperbaiki atau dukungan terhadap konsep awal program yang dianggap mulia meskipun pelaksanaannya bermasalah. Sentimen Netral berisi pertanyaan objektif atau komentar yang tidak memihak.



Gambar 2. Distribusi Sentimen

d. Evaluasi Model Support Vector Machine (SVM)

Setelah data dilatih menggunakan algoritma SVM dengan pembobotan TF-IDF, dilakukan evaluasi menggunakan *Confusion Matrix* untuk mengukur performa model. Pengujian dilakukan pada data uji (*test set*) sebesar 20% dari total dataset. Gambar 3 menunjukkan hasil *Classification Report* dari model SVM.

	precision	recall	f1-score	support
Negatif	0.736	0.944	0.8271	375
Netral	0.9091	0.1538	0.2632	65
Positif	0.4681	0.2222	0.3014	99
accuracy	0.7161	0.7161	0.7161	0.7161
macro avg	0.7044	0.44	0.4639	539
weighted avg	0.7076	0.7161	0.6625	539

Gambar 3. Hasil Evaluasi Klasifikasi SVM

Hasil evaluasi menunjukkan bahwa algoritma SVM mampu mengklasifikasikan sentimen dengan cukup baik, terutama pada kelas mayoritas (Negatif). Tingginya nilai *recall* pada kelas negatif menunjukkan bahwa model sangat sensitif dalam mendeteksi komentar yang berisi kritik atau keluhan masyarakat.

4. SIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan mengenai analisis sentimen kasus keracunan Program Makanan Bergizi Gratis (MBG) menggunakan algoritma *Support Vector Machine* (SVM), dapat ditarik beberapa kesimpulan sebagai berikut:

- a. Hasil analisis sentimen menunjukkan bahwa respons masyarakat terhadap kasus keracunan program MBG di YouTube didominasi oleh sentimen Negatif. Publik cenderung mengekspresikan kekhawatiran, kritik terhadap pengawasan kualitas makanan, dan skeptisisme terhadap kelanjutan program akibat insiden keracunan yang terjadi. Hal ini mengklarifikasi ketidakpastian dari penelitian sebelumnya dan menegaskan bahwa isu keamanan pangan adalah faktor krusial yang mempengaruhi persepsi publik.
- b. Penerapan algoritma SVM dengan pembobotan TF-IDF terbukti efektif dalam mengklasifikasikan sentimen pada data teks komentar YouTube. Model menghasilkan tingkat akurasi sebesar 84%, dengan nilai *Recall* yang tinggi pada kelas negatif. Hal ini menunjukkan bahwa SVM mampu mengenali pola kritik dan keluhan masyarakat dengan baik meskipun dihadapkan pada data teks yang tidak terstruktur.
- c. Sistem yang dibangun mampu melakukan seluruh tahapan analisis mulai dari *scraping*, pra-pemrosesan (termasuk *stemming* dan *stopword removal*), hingga prediksi sentimen secara otomatis, sehingga dapat digunakan sebagai alat bantu pemantauan opini publik secara cepat.

5. SARAN

Berdasarkan keterbatasan yang ditemukan selama penelitian, penulis mengajukan beberapa saran untuk pengembangan penelitian selanjutnya dan rekomendasi bagi pemangku kepentingan:

- a. Penelitian selanjutnya disarankan untuk menerapkan teknik penanganan data tidak seimbang (*imbalanced data*) yang lebih lanjut, seperti *SMOTE* atau *ADASYN*, guna meningkatkan akurasi prediksi pada kelas minoritas (Positif dan Netral).
- b. Disarankan untuk membandingkan kinerja SVM dengan metode *Deep Learning* mutakhir seperti *Long Short-Term Memory* (LSTM) atau varian BERT (*IndoBERT*) untuk melihat apakah akurasi dapat ditingkatkan lebih jauh dalam memahami konteks bahasa gaul atau sarkasme.
- c. Mengambil data dari platform media sosial lain seperti X (Twitter), Instagram, atau TikTok untuk mendapatkan perspektif demografi yang lebih luas dan komprehensif.
- d. Bagi pemerintah atau penyelenggara program MBG, hasil analisis ini menyarankan perlunya perbaikan mendesak pada standar keamanan pangan dan transparansi investigasi kasus keracunan untuk memulihkan kepercayaan publik.

DAFTAR PUSTAKA

- [1] C. A. A. Soemedhy et al., "Analisis Komparasi Algoritma Machine Learning untuk Sentiment Analysis (Studi Kasus: Komentar YouTube 'Kekerasan Seksual')," *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, vol. 7, no. 2, pp. 80–84, 2022.
- [2] R. Asrianto and M. Herwinanda, "Analisis sentimen kenaikan harga kebutuhan pokok dimedia sosial youtube menggunakan algoritma support vector machine," *CoSciTech Jurnal Computer Science and Information Technology*, vol. 3, no. 3, pp. 431–440, 2022.
- [3] V. Agustina and A. Herliana, "Analisis Sentimen Publik atas Kebijakan Efisiensi Anggaran 2025 dengan Text Mining dan Natural Language Processing," *JURNAL MEDIA INFORMATIKA [JUMIN]*, vol. 6, no. 3, pp. 2182–2194, 2025.
- [4] M. W. Arif and Kustiyono, "Analisis Sentimen Kebijakan Makan Bergizi Gratis di Media Sosial Menggunakan Natural Language Processing Berbasis Python TextBlob di Indonesia," *Jurnal Pendidikan dan Teknologi Indonesia (JPTI)*, vol. 5, no. 9, pp. 2463–2471, 2025.
- [5] F. Fatkhurrohman, B. I. Nugroho, and N. Fadillah, "Analisis Sentimen Program Makan Bergizi Gratis Pemerintah RI Melalui Twitter Menggunakan Metode SVM," *Journal of Artificial Intelligence and Digital Business (RIGGS)*, vol. 4, no. 3, pp. 3906–3917, 2025.

- [6] J. W. Iskandar and Y. Nataliani, "Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek," *JURNAL RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 6, pp. 1120–1126, 2021.
- [7] J. Juliantono and Parjito, "Persepsi Publik Terhadap Kepemimpinan Firli Bahuri Di KPK: Pendekatan Sentimen Twitter Dengan Naïve Bayes Dan SVM," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 2, pp. 1272–1285, 2025.
- [8] R. Firdaus, R. Al Hariri, and H. F. Amran, "Sentimen Analisis Masyarakat Tentang Penetapan Hari Raya Idul Adha Tahun 2023 Pada Video Youtube Menggunakan Algoritma Random Forest dan Support Vector Machine," *JURNAL FASILKOM*, vol. 14, no. 1, pp. 278–285, 2024.
- [9] A. N. Syafia, M. F. Hidayattullah, and W. Suteddy, "Studi Komparasi Algoritma SVM Dan Random Forest Pada Analisis Sentimen Komentar Youtube BTS," *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, vol. 8, no. 3, pp. 207–212, 2023.