

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Kualitas sumber daya manusia merupakan salah satu faktor utama yang menentukan keberhasilan pembangunan suatu negara. Salah satu aspek yang berpengaruh terhadap kualitas tersebut adalah terpenuhinya kebutuhan gizi masyarakat. Asupan gizi yang memadai tidak hanya mendukung pertumbuhan fisik, tetapi juga berperan dalam perkembangan kemampuan kognitif, kesehatan, dan produktivitas individu. Meskipun berbagai upaya perbaikan gizi telah dilakukan, Indonesia masih menghadapi sejumlah permasalahan di bidang tersebut. Berdasarkan hasil Survei Status Gizi Indonesia (SSGI) tahun 2023, prevalensi stunting nasional tercatat sebesar 21,5%, yang menunjukkan bahwa masalah kekurangan gizi masih menjadi perhatian penting dalam pembangunan kesehatan masyarakat [1]. Kondisi ini mendorong pemerintah untuk terus mengembangkan berbagai kebijakan dan program yang berorientasi pada peningkatan status gizi masyarakat.

Salah satu program yang diluncurkan pemerintah untuk mendukung upaya tersebut adalah Program Makan Bergizi Gratis (MBG). Program yang mulai diimplementasikan secara nasional pada 6 Januari 2025 ini menjadi bagian dari strategi pemerintah dalam meningkatkan kualitas gizi kelompok sasaran, terutama peserta didik, balita, ibu hamil, dan ibu menyusui. Pelaksanaan MBG berada di bawah koordinasi Badan Gizi Nasional (BGN) melalui Satuan Pelayanan Pemenuhan Gizi (SPPG) yang tersebar di berbagai daerah [2]. Melalui penyediaan makanan bergizi secara rutin, program ini diharapkan dapat membantu menekan angka stunting, meningkatkan kualitas kesehatan masyarakat, serta mempersiapkan generasi yang lebih sehat dan produktif dalam menyongsong visi Indonesia Emas 2045.

Sebagai program yang menjangkau jutaan penerima manfaat dan didukung oleh anggaran yang besar, MBG menarik perhatian luas dari masyarakat. Berbagai pihak memberikan tanggapan yang beragam terhadap pelaksanaan program tersebut. Sebagian masyarakat menilai MBG sebagai langkah positif karena dinilai mampu membantu pemenuhan kebutuhan gizi kelompok rentan. Namun, di sisi lain, sejumlah persoalan yang muncul selama implementasi turut memunculkan

kritik dari publik. Sepanjang tahun 2025, beberapa kasus dugaan keracunan makanan yang melibatkan ribuan siswa dilaporkan terjadi di sejumlah wilayah Indonesia. Data Badan Gizi Nasional menunjukkan bahwa ribuan siswa terdampak hingga September 2025, sementara lembaga pemantau independen melaporkan jumlah yang lebih besar [3]. Selain isu keamanan pangan, masyarakat juga menyoroti berbagai aspek lain, seperti kualitas makanan yang disajikan, kesiapan sistem distribusi, efektivitas pelaksanaan, serta transparansi pengelolaan anggaran. Perbedaan pandangan tersebut menunjukkan bahwa respons publik terhadap Program Makan Bergizi Gratis bersifat beragam dan dinamis.

Perkembangan teknologi informasi dan media sosial memberikan kesempatan yang lebih luas bagi masyarakat untuk menyampaikan pendapat mengenai berbagai isu publik. Salah satu platform yang banyak digunakan untuk berdiskusi adalah YouTube. Video yang membahas Program Makan Bergizi Gratis sering kali memperoleh respons dalam jumlah besar melalui kolom komentar. Komentar-komentar tersebut memuat berbagai bentuk opini, mulai dari dukungan, kritik, harapan, hingga tanggapan netral terhadap program yang sedang berjalan. Oleh karena itu, komentar YouTube dapat dimanfaatkan sebagai sumber data yang merepresentasikan persepsi masyarakat terhadap implementasi MBG.

Meskipun menyediakan informasi yang berlimpah, data komentar di media sosial memiliki karakteristik yang tidak terstruktur sehingga sulit dianalisis secara manual, terutama ketika jumlahnya mencapai ribuan. Selain itu, komentar pengguna umumnya mengandung bahasa tidak baku, singkatan, kesalahan penulisan, penggunaan emoji, serta variasi ekspresi yang beragam. Kondisi tersebut menyebabkan proses identifikasi opini secara manual menjadi kurang efisien. Untuk mengatasi permasalahan tersebut, diperlukan pendekatan yang mampu melakukan pengolahan dan klasifikasi data teks secara otomatis. Salah satu pendekatan yang banyak diterapkan adalah analisis sentimen, yaitu teknik yang digunakan untuk mengidentifikasi kecenderungan opini dalam suatu teks dan mengelompokkannya ke dalam kategori sentimen tertentu, seperti positif, negatif, atau netral [4].

Dalam penelitian ini, proses klasifikasi sentimen dilakukan menggunakan algoritma Naive Bayes. Algoritma tersebut merupakan metode klasifikasi probabilistik yang bekerja berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen satu sama lain [5]. Naive Bayes dipilih karena memiliki proses komputasi yang relatif sederhana, mampu menangani data teks berdimensi tinggi, serta tetap menunjukkan performa yang baik pada berbagai permasalahan

klasifikasi teks [6, 7]. Selain itu, algoritma ini dapat dikombinasikan dengan representasi fitur TF-IDF sehingga sesuai untuk digunakan dalam analisis sentimen berbasis data komentar media sosial [8].

Penelitian mengenai sentimen masyarakat terhadap Program Makan Bergizi Gratis telah dilakukan menggunakan berbagai pendekatan, seperti *Long Short-Term Memory* (LSTM), *Support Vector Machine* (SVM), maupun *Random Forest* [9, 10]. Namun demikian, kajian yang secara khusus memanfaatkan algoritma Naive Bayes untuk menganalisis opini publik Indonesia terhadap Program Makan Bergizi Gratis berdasarkan komentar YouTube masih relatif terbatas. Oleh sebab itu, penelitian ini dilakukan untuk mengimplementasikan algoritma Naive Bayes dalam mengklasifikasikan sentimen masyarakat terhadap Program Makan Bergizi Gratis serta mengevaluasi performa model yang dihasilkan berdasarkan berbagai metrik evaluasi klasifikasi.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, permasalahan yang menjadi fokus penelitian ini adalah sebagai berikut.

1. Bagaimana implementasi algoritma Naive Bayes dalam melakukan analisis sentimen opini publik Indonesia terhadap program Makan Bergizi Gratis pada komentar YouTube?
2. Bagaimana tingkat akurasi algoritma Naive Bayes dalam melakukan analisis sentimen opini publik Indonesia terhadap program Makan Bergizi Gratis pada komentar YouTube berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*?

1.3 Batasan Permasalahan

Penelitian ini dibatasi pada hal-hal berikut.

1. Sumber data berasal dari 27 video YouTube yang dipublikasikan oleh media nasional terverifikasi, meliputi KOMPAS TV, Metro TV, TV One, iNews, CNN Indonesia, Beritasatu, dan Indosiar. Seluruh video membahas program Makan Bergizi Gratis (MBG) dari berbagai sudut pemberitaan.
2. Pengambilan data dilakukan menggunakan YouTube Data API versi 3. Periode komentar yang berhasil dikumpulkan mencakup rentang Januari 2025

hingga Juni 2026, sedangkan proses pengambilan data dilaksanakan pada Juni 2026.

3. Jumlah data yang digunakan dalam penelitian ini sebanyak 2.974 komentar yang telah melalui proses *cleaning* minimal, yaitu penghapusan duplikat, komentar berbahasa asing, komentar yang hanya berisi simbol atau emoji, komentar yang terlalu pendek (kurang dari tiga kata bermakna), dan komentar mengandung indikasi spam.
4. Pelabelan sentimen dilakukan secara manual oleh peneliti berdasarkan kriteria yang telah ditetapkan. Sentimen positif diberikan pada komentar yang menunjukkan dukungan, apresiasi, atau harapan positif terhadap program MBG. Sentimen negatif diberikan pada komentar yang mengandung kritik, kekhawatiran, atau penilaian negatif terhadap implementasi program. Sentimen netral diberikan pada komentar yang bersifat informatif atau deskriptif tanpa kecenderungan evaluatif tertentu.
5. Klasifikasi sentimen dibatasi pada tiga kategori, yaitu positif, negatif, dan netral.
6. Metode klasifikasi naive yang digunakan dalam penelitian ini adalah Multinomial Naive Bayes dan Complement Naive Bayes dengan pembobotan fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Untuk menangani masalah ketidakseimbangan distribusi kelas dalam data, diterapkan teknik *Synthetic Minority Over-sampling Technique* (SMOTE-Tomek) pada data latih.
7. Penelitian ini tidak membahas program MBG dari sudut pandang kebijakan publik, ekonomi, maupun kesehatan secara mendalam. Fokus utama penelitian berada pada aspek pengolahan teks, klasifikasi sentimen, dan evaluasi performa model.

1.4 Tujuan Penelitian

Sesuai dengan rumusan masalah yang telah ditetapkan, penelitian ini bertujuan untuk:

1. Mengimplementasikan algoritma Naive Bayes dalam melakukan analisis sentimen opini publik Indonesia terhadap program Makan Bergizi Gratis pada komentar YouTube; dan

2. Mengukur tingkat akurasi algoritma Naive Bayes dalam melakukan analisis sentimen opini publik Indonesia terhadap program Makan Bergizi Gratis pada komentar YouTube berdasarkan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut.

1. Dari sisi akademik, penelitian ini dapat menjadi referensi penerapan algoritma Naive Bayes pada analisis sentimen komentar YouTube berbahasa Indonesia, khususnya untuk topik program kebijakan pemerintah. Penelitian ini juga dapat memperkaya kajian mengenai penggunaan TF-IDF, SMOTE-Tomek, dan Naive Bayes pada data teks yang tidak seimbang.
2. Dari sisi praktis, hasil analisis sentimen dapat memberikan gambaran mengenai kecenderungan opini publik Indonesia terhadap program MBG, yang berpotensi menjadi masukan bagi pemangku kepentingan dalam mengevaluasi penerimaan masyarakat terhadap program tersebut.
3. Bagi penelitian lanjutan, penelitian ini dapat dijadikan acuan untuk pengembangan metode serupa dengan cakupan data yang lebih luas, jumlah kelas sentimen yang lebih rinci, maupun perbandingan dengan algoritma lain seperti *Support Vector Machine*, *Random Forest*, atau model berbasis *deep learning*.

1.6 Sistematika Penulisan

Sistematika penulisan dalam laporan ini disusun dalam lima bab dengan rincian sebagai berikut.

1. **Bab 1 Pendahuluan** memuat latar belakang masalah, rumusan masalah, batasan permasalahan, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.
2. **Bab 2 Landasan Teori** memuat teori dan konsep yang mendasari penelitian, meliputi analisis sentimen, *text mining*, *natural language processing*, preprocessing teks, pelabelan data, TF-IDF, Naive Bayes, penanganan

ketidakseimbangan data, evaluasi model, YouTube sebagai sumber data opini publik, dan penelitian terdahulu.

3. **Bab 3 Metodologi** memuat cara dan metode penelitian yang digunakan, tahapan penelitian, pengumpulan data, *cleaning* minimal, pelabelan data manual, preprocessing teks, pembagian data, penyeimbangan data dengan SMOTE-Tomek, ekstraksi fitur TF-IDF, klasifikasi dalam Naive Bayes, dan evaluasi dari model.
4. **Bab 4 Hasil dan Diskusi** memuat deskripsi dataset, hasil preprocessing, distribusi sentimen, hasil klasifikasi, dan evaluasi performa model.
5. **Bab 5 Simpulan dan Saran** memuat kesimpulan dari penelitian yang telah dilakukan dan saran yang diberikan untuk pengembangan penelitian selanjutnya.

