

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Program Makan Bergizi Gratis (MBG) yang diluncurkan pada tanggal 6 Januari 2025 menjadi salah satu langkah strategis dalam memenuhi gizi pada anak dan upaya dalam meningkatkan kualitas sumber daya manusia di Indonesia (Basyari, 2025). Program Makan Bergizi Gratis ini ditujukan pada anak-anak sekolah, pesantren dan ibu hamil untuk memastikan pemenuhan gizi yang merata. Diharapkan juga program ini dapat mendorong pertumbuhan ekonomi lokal dengan membantu Usaha Mikro, Kecil, dan Menengah (UMKM) dalam menyediakan bahan makanan lokal (Report, 2024).

Seiring munculnya rencana program ini hingga pada pengimplementasiannya sekarang menghadapi masalah yang memicu kekhawatiran publik sehingga menuai berbagai tanggapan khususnya di media sosial, baik dalam bentuk dukungan atau kritikan. Diantaranya yaitu terkait perkiraan anggaran yang mencapai Rp 450 triliun per tahun. Perhatian masyarakat juga tertuju pada rencana penggunaan dana dari APBN pendidikan dan dana BOS. Hal ini memicu kekhawatiran masyarakat mengenai dampak pada biaya pendidikan, gaji guru dan stabilitas keuangan negara (Triningsih et al., 2025). Selain itu, menurut laporan BBC News Indonesia, telah terjadi sejumlah kasus keracunan makanan di berbagai daerah sejak awal tahun 2025 yang diduga berkaitan dengan kualitas bahan makanan,

proses pengolahan, maupun penyimpanan yang kurang sesuai standar. Kondisi tersebut menimbulkan kekhawatiran masyarakat terhadap aspek keamanan dan kualitas pelaksanaan Program Makan Bergizi Gratis (Ibrahim et al., 2025). Laporan muncul dari wilayah SPPG di Bandung diperkirakan bahan mentah yang digunakan akan tersedia dalam 10 hari ke depan, dalam menyikapi hal ini Badan Gizi Nasional menerjunkan tim investigasi gabungan untuk melakukan uji laboratorium terhadap sampel makanan dan bahan makanan yang digunakan, ini menjadi salah satu faktor terjadinya keracunan tersebut (Badan Gizi Nasional, 2025).

Permasalahan yang muncul selama Program Makan Bergizi Gratis menjadi perhatian luas dari masyarakat terutama di media sosial salah satunya TikTok. Masyarakat menyampaikan dukungan, kritikan, keluhan maupun kekhawatiran secara terbuka melalui komentar. Kondisi ini menjadi pertanda bahwa kebijakan program tidak hanya dinilai dari regulasi dan mekanismenya, tetapi juga pengalaman nyata yang dialami masyarakat sebagai penerima program di lapangan. TikTok dipilih sebagai sumber data penelitian karena merupakan salah satu platform media sosial yang banyak digunakan masyarakat untuk menyampaikan opini dan berinteraksi melalui kolom komentar. Selain itu, berdasarkan Survei APJII tahun 2025, TikTok merupakan salah satu platform media sosial yang paling banyak diakses oleh masyarakat Indonesia dengan persentase penggunaan sebesar 35,17%. Selain itu, penelitian Mustofa et al., (2026) menunjukkan bahwa integrasi IndoBERT-SVM memperoleh performa paling optimal pada data *platform*

TikTok dengan akurasi 95,73% dibandingkan *platform* YouTube dan X dengan akurasi 92,53%. Temuan tersebut memberikan dukungan empiris bahwa data komentar TikTok dapat menjadi sumber data yang relevan untuk analisis sentimen menggunakan pendekatan berbasis IndoBERT-SVM. Oleh karena itu, penelitian ini menggunakan TikTok sebagai sumber data untuk menganalisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG). Meskipun opini publik terhadap Program Makan Bergizi Gratis tersebar luas diberbagai platform media sosial, informasi tersebut masih berupa data tidak terstruktur sehingga sulit dianalisis secara manual dalam jumlah besar. Selain itu, bahasa informal, singkatan, serta ekspresi sarkastik banyak ditemukan pada komentar publik di media sosial.

Menjadi salah satu program prioritas nasional dengan cakupan penerima yang luas dan alokasi anggaran yang besar, persepsi publik terhadap Program Makan Bergizi Gratis menjadi faktor penting dalam keberhasilan implementasi kebijakan. Opini yang berkembang di masyarakat dapat memengaruhi tingkat kepercayaan publik, penerimaan program, serta efektivitas pelaksanaannya di lapangan. Oleh karena itu, diperlukan suatu pendekatan yang mampu mengidentifikasi dan mengukur kecenderungan sentimen masyarakat secara sistematis berdasarkan data yang tersedia di media sosial.

Analisis sentimen berbasis kecerdasan buatan menjadi pendekatan yang tepat dalam mengolah opini publik dalam jumlah besar dan dengan variasi bahasa yang kompleks, termasuk penggunaan bahasa gaul, singkatan, dan

struktur kalimat yang beragam. Analisis sentimen mencakup identifikasi kata-kata dan ungkapan yang bersifat positif, negatif, atau netral, serta menganalisis pola penggunaan kata-kata tersebut. Tujuannya untuk mengetahui opini, sikap, atau emosi publik dalam mengomentari topik tertentu. Oleh karena itu, analisis sentimen relevan untuk menilai sudut pandang publik terhadap Program Makan Bergizi Gratis secara sistematis dan objektif (M. S. Dewi & Hasugian, 2025).

Penelitian terdahulu terkait analisis sentimen program Makan Bergizi Gratis yang dilakukan oleh Saleh & Imanda, (2025) dengan membandingkan metode *Naïve Bayes* dan *Support Vector Machine* (SVM) menghasilkan akurasi metode yang paling besar yaitu SVM dengan akurasi 87% dan *Naïve Bayes* 68%, namun pendekatan pelabelan berbasis NLTK dan VADER memiliki keterbatasan dalam memahami konteks bahasa Indonesia, bahasa informal, serta variasi ekspresi yang umum digunakan pada media sosial sehingga berpotensi memengaruhi kualitas pelabelan sentimen. Sementara itu, penelitian lainnya yang dilakukan oleh Destryanto et al., (2025) menunjukkan efektivitas IndoBERT dalam menganalisis sentimen respon publik terhadap Program Makan Bergizi Gratis dengan hasil akurasi 99,89% pada pelatihan dan 93,95% pada pengujian. Selain itu, penelitian Asshiddiq & Witanti, (2025) terkait analisis sentimen penundaan pengangkatan CPNS 2025 menunjukkan model IndoBERT mampu mengklasifikasikan sentimen dengan andal dengan akurasi 84,27%. Di sisi lain, model *deep learning*

berbasis *transformer* seperti IndoBERT menunjukkan keunggulannya dalam memahami konteks kalimat berbahasa Indonesia (Manoppo et al., 2025).

Berdasarkan penelitian sebelumnya, terdapat keterbatasan dalam memahami konteks kalimat berbahasa Indonesia, karena keberagaman variasi kata dan terus berkembang. Masyarakat juga sering menggunakan kalimat sarkasme yaitu dengan konotasi negatif dalam mengeksikan suatu yang positif. Oleh karena itu, diperlukannya suatu metode yang handal dalam memahami konteks kalimat. Penelitian sebelumnya terkait Program Makan Bergizi Gratis umumnya menggunakan algoritma klasifikasi konvensional atau memanfaatkan model *transformer* secara mandiri. Namun, penelitian yang mengintegrasikan kemampuan representasi kontekstual IndoBERT dengan kemampuan klasifikasi *Support Vector Machine* (SVM) pada data komentar TikTok terkait Program Makan Bergizi Gratis masih terbatas. Kombinasi kedua metode tersebut berpotensi menghasilkan representasi fitur yang lebih baik sekaligus meningkatkan performa klasifikasi sentimen. Oleh karena itu, penelitian ini menggabungkan dua algoritma yaitu IndoBERT sebagai representasi bahasa dan *Support Vector Machine* (SVM) sebagai model klasifikasi data.

Selain tantangan dalam memahami konteks bahasa Indonesia, penelitian analisis sentimen pada media sosial juga dihadapkan pada permasalahan distribusi data yang tidak seimbang. Pada umumnya komentar publik terhadap suatu isu didominasi oleh satu kelas sentimen tertentu sehingga model mengklasifikasikan cenderung lebih baik dalam mengenali

kelas mayoritas dibandingkan kelas minoritas (Ritonga et al., 2025). Kondisi ini berpotensi menurunkan kemampuan model dan memengaruhi evaluasi klasifikasi. Oleh karena itu, penelitian ini juga menganalisis pengaruh distribusi data terhadap performa melalui penanganan ketidakseimbangan data.

Penelitian ini berfokus pada analisis sentimen publik terhadap Program Makan Bergizi Gratis yang diluncurkan pemerintah berdasarkan komentar pada platform TikTok. Tujuan penelitian ini yaitu untuk menganalisis distribusi sentimen publik, mengevaluasi performa model dalam mengklasifikasikan sentimen, dan menganalisis pengaruh distribusi data terhadap hasil klasifikasi untuk mengetahui efektivitas model pada kondisi data yang tidak seimbang. Penelitian ini diharapkan dapat memberikan informasi mengenai persepsi publik terhadap Program Makan Bergizi Gratis sebagai bahan evaluasi bagi pemerintah. Selain itu, penelitian ini diharapkan dapat berkontribusi pada pengembangan penelitian di bidang *Natural Language Processing* (NLP), khususnya dalam penerapan kombinasi metode IndoBERT dan *Support Vector Machine* untuk analisis sentimen berbahasa Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan yang telah dijelaskan, maka dapat dirumuskan permasalahan penelitian ini sebagai berikut:

- a. Bagaimana distribusi sentimen pengguna *platform* TikTok terhadap Program Makan Bergizi Gratis berdasarkan hasil klasifikasi menggunakan metode IndoBERT – *Support Vector Machine* (SVM)?
- b. Bagaimana performa penerapan metode IndoBERT – *Support Vector Machine* (SVM) dalam menganalisis sentimen masyarakat terhadap Program Makan Bergizi Gratis berdasarkan data komentar TikTok?
- c. Bagaimana pengaruh distribusi data terhadap performa metode IndoBERT – *Support Vector Machine* (SVM)?

1.3 Batasan Masalah

Terdapat beberapa batasan masalah dalam penelitian ini yaitu sebagai berikut:

- a. Data yang digunakan adalah komentar pada beberapa video TikTok yang membahas Program Makan Bergizi Gratis (MBG). Pengambilan data dilakukan berdasarkan kata kunci #ProgramMakanBergiziGratis pada akun @BadanKomunikasiPemerintahRI, @bbcnewsindonesia dua video, dan @badangizinasional.ri dua video, untuk tagar #Permasalahanmbg yaitu satu video pada akun @KabarTVOne, @Ferryirwandy, @Dennis-FoodTechnologist @MetroTV, dan @bbcnewsindonesia.
- b. Jumlah data yang digunakan dalam penelitian ini sebanyak 3.590 data komentar, yang terdiri atas 2.872 data latih (80%) dan 718 data uji (20%).
- c. Analisis sentimen pada penelitian ini dibatasi pada tiga kategori sentimen, yaitu positif, negatif, dan netral.

- d. Pelabelan sentimen dilakukan menggunakan pendekatan *hybrid labelling*, yaitu kombinasi antara pelabelan otomatis berbasis *lexicon* dan anotasi manual pada data yang bersifat ambigu.
- e. Penelitian ini menggabungkan 2 metode yaitu IndoBERT dengan menghasilkan representasi numerik dari setiap komentar dan *Support Vector Machine* (SVM) untuk klasifikasi sentimen.
- f. Penelitian ini menganalisis pengaruh distribusi data terhadap performa model melalui penerapan *class weight* pada proses klasifikasi.
- g. Proses pengolahan dan analisis data dilakukan menggunakan bahasa pemrograman *Python* dengan bantuan platform *Google Colaboratory* (*Google Colab*).

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah sebelumnya maka tujuan penelitian ini antara lain:

- a. Mengetahui distribusi sentimen pengguna *platform* TikTok terhadap Program Makan Bergizi Gratis menggunakan metode IndoBERT – *Support Vector Machine* (SVM).
- b. Mengevaluasi performa penerapan metode IndoBERT – *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen masyarakat terhadap Program Makan Bergizi Gratis berdasarkan data komentar TikTok.

- c. Mengetahui pengaruh distribusi data terhadap performa metode IndoBERT – *Support Vector Machine* (SVM) dalam proses klasifikasi sentimen.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan kontribusi baik dalam hal teoritis maupun praktis. Manfaat tersebut dijelaskan sebagai berikut:

1.5.1 Manfaat Teoritis

- a. Penelitian ini diharapkan dapat menambah pemahaman terkait kombinasi model IndoBERT – *Support Vector Machine* (SVM) dalam analisis sentimen pada media sosial, khususnya bahasa Indonesia.
- b. Penelitian ini memberikan bukti empiris terkait tingkat akurasi dan performa model IndoBERT – *Support Vector Machine* (SVM) dalam melakukan klasifikasi sentimen. Selain itu, dapat memperluas pemahaman mengenai efektivitas model *hybrid* yang menggabungkan model berbasis *deep learning* dengan algoritma klasifikasi tradisional.
- c. Penelitian ini menambah pemahaman mengenai pengaruh distribusi data terhadap performa model analisis sentimen. Hal ini penting mengenai dampak distribusi data pada proses pelatihan model dalam bidang *Natural Language Processing*.

1.5.2 Manfaat Praktis

- a. Hasil analisis sentimen dapat digunakan sebagai masukan bagi pemerintah atau pihak terkait dalam memahami persepsi masyarakat terhadap program makan bergizi gratis.
- b. Penelitian ini menyajikan gambaran mengenai kombinasi model IndoBERT – *Support Vector Machine* (SVM) sehingga dapat menjadi acuan dalam pengembangan dan pengimplementasian sistem analisis sentimen yang lebih akurat dan efisien.
- c. Penelitian ini dapat dimanfaatkan sebagai acuan atau pembandingan bagi penelitian selanjutnya terkait analisis sentimen, baik isu sosial lainnya atau menggunakan model dan teknik yang berbeda.

