

BAB I

PENDAHULUAN

1.1 Latar Belakang

Media sosial telah menjadi *platform* utama bagi masyarakat untuk berbagi opini dan tanggapan terhadap isu-isu sosial. X (sebelumnya *Twitter*) merupakan salah satu *platform* media sosial yang berfungsi sebagai tempat untuk menyampaikan opini, apresiasi, atau kritik terhadap kebijakan pemerintah. Menurut data yang diperoleh dari statista, Indonesia merupakan salah satu negara dengan jumlah pengguna X terbesar di dunia. Pada April 2024, tercatat bahwa 24,85 juta pengguna X berasal dari Indonesia, angka inilah yang menempatkan Indonesia berada di posisi keempat. Karena bersifat *real-time*, transparan, dan interaktif, media sosial berperan sebagai sarana yang efektif untuk menelusuri opini masyarakat mengenai kebijakan pemerintahan (Kusuma et al., 2024).

Salah satu kebijakan pemerintahan yang kini menarik perhatian masyarakat adalah Program Makan Bergizi Gratis (MBG). Program ini merupakan bagian dari upaya pemerintah untuk meningkatkan kualitas sumber daya manusia di Indonesia dengan memperbaiki status gizi, khususnya bagi anak sekolah dan masyarakat berpenghasilan rendah (Kiftiyah et al., 2023). Menurut laporan Sekretariat Kabinet Republik Indonesia (2025), program MBG bertujuan untuk menargetkan 82,9 juta penerima manfaat secara bertahap, dengan alokasi anggaran dana sebesar Rp71 triliun dalam RUU APBN 2025. Hingga pertengahan Juli 2025, program ini telah menjangkau sekitar 6,2 juta penerima manfaat di 26 provinsi (Amrin, 2025).

Meskipun bertujuan untuk meningkatkan gizi dan ketahanan pangan, program MBG menerima tanggapan yang beragam. Sebagian orang memandangnya sebagai langkah positif, sementara yang lain mengkritik kualitas, distribusi, dan transparansinya. Perbedaan persepsi ini terlihat jelas di media sosial, khususnya di X sebagai tempat untuk menyampaikan opini mereka dengan bebas.

Secara sistematis, untuk memahami persepsi publik diperlukan metode analisis data otomatis yang mampu memproses opini dalam skala besar. Metode ini dikenal sebagai analisis sentimen, yang merupakan cabang dari *text mining* dan bertujuan untuk mengklasifikasikan teks atau opini ke dalam kategori positif, negatif, atau netral. Hal ini memungkinkan kita untuk mengidentifikasi tren opini publik terkait isu, kebijakan, atau program tertentu (Cui et al., 2023).

Analisis sentimen terhadap program Makanan Bergizi Gratis (MBG) sangat penting karena memberikan wawasan mendalam tentang persepsi publik terhadap kebijakan ini di dunia digital. Analisis ini memungkinkan pemerintah untuk menilai tingkat penerimaan publik, mengidentifikasi dukungan dan kritik yang muncul, serta mengidentifikasi area yang perlu ditingkatkan dalam implementasi program. Selain itu, analisis sentimen sangat penting untuk menilai dan mengembangkan strategi komunikasi yang lebih efektif, dan dapat meningkatkan kepercayaan dan partisipasi publik terhadap keberlanjutan program MBG.

Berbagai macam algoritma *machine learning* telah digunakan untuk analisis sentimen, diantaranya adalah *Naïve Bayes* (NB) dan *Logistic Regression* (LR). *Naïve Bayes* beroperasi berdasarkan prinsip teorema Bayes, asumsi independensi antar fitur, sehingga metode ini cepat, stabil, dan ideal untuk data yang berdimensi tinggi sehingga banyak digunakan dalam analisis teks (Jessica et al., 2025). Sementara itu, *Logistic Regression* adalah algoritma klasifikasi diskriminatif yang memprediksi probabilitas data yang termasuk dalam kelas tertentu menggunakan fungsi logistik (*sigmoid*). Model ini efisien dan akurat karena secara langsung mempelajari hubungan antara fitur dan label tanpa mengasumsikan distribusi data, dan cocok untuk kumpulan data kompleks seperti teks atau opini publik (J. Jansen, 2024).

Sejumlah penelitian menunjukkan bahwa tingkat akurasi metode metode analisis sentimen dapat bervariasi, hal ini bergantung pada karakteristik data dan studinya. Sebagaimana penelitian yang dilakukan oleh (Siddiqui et al., 2025), menunjukkan bahwa algoritma *Logistic Regression* (LR) mencapai tingkat akurasi sebesar 91.2%, dan *Naïve Bayes* (NB) memperoleh akurasi 89,7% dalam menganalisis sentimen teks menggunakan *tweet* media sosial. Sementara itu,

penelitian yang dilakukan oleh (Mantika et al., 2024), dalam analisis sentimen pemilihan presiden 2024, hasil penelitian menunjukkan bahwa algoritma *Naïve Bayes* (NB) menunjukkan hasil yang lebih baik dari hasil *Logistic Regression* (LR) dalam mengklasifikasikan data sentimen terhadap kandidat presiden 2024. Akurasi yang dicapai adalah 63% untuk Anies Baswedan, 77% untuk Ganjar Pranowo, dan 44% untuk Prabowo Subianto. Penelitian ini hanya menggunakan 1.573 *tweet* yang dikumpulkan. Penelitian lain yang dilakukan oleh (Dahlian & Sitanggang, 2023), hasil penelitian menunjukkan bahwa algoritma *Support Vector Machine* (SVM) memberikan kinerja terbaik dalam klasifikasi sentimen, melampaui *Logistic Regression* (LR) dan *Naïve Bayes Multinomial* (NB). Akurasi SVM adalah 94%, LR 90%, dan NB 88%. Temuan ini menunjukkan bahwa SVM lebih unggul dalam mengkategorikan sentimen publik terkait migrasi televisi digital di Indonesia, dan hasil LR lebih tinggi dibandingkan NB.

Adapun hasil dari berbagai penelitian sebelumnya menunjukkan bahwa kinerja algoritma *Naïve Bayes* dan *Logistic Regression* tidak sepenuhnya konsisten. Kedua metode dapat menghasilkan tingkat akurasi yang bervariasi, bergantung pada karakteristik data, konteks topik yang dianalisis, dan kualitas pra-pemrosesan yang diterapkan. Hal ini menyoroti pentingnya penelitian lebih lanjut untuk mengkaji efektivitas kedua algoritma dalam analisis sentimen terhadap isu-isu kebijakan sosial di Indonesia, khususnya Program Makan Bergizi Gratis (MBG) yang saat ini sedang menjadi sorotan publik.

Berdasarkan latar belakang tersebut, penulis melakukan penelitian berjudul “Perbandingan Metode *Naïve Bayes* dan *Logistic Regression* dalam Analisis Sentimen Pengguna X terhadap Program Makan Bergizi Gratis (MBG)”. Penelitian ini bertujuan untuk menemukan metode klasifikasi yang paling efektif dalam mengidentifikasi opini publik di media sosial dan memberikan rekomendasi yang dapat mendukung pemerintah dalam meningkatkan kualitas implementasi dan keberhasilan program Makan Bergizi Gratis.

1.2 Rumusan Masalah

Adapun rumusan masalah yang akan dibahas pada penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan algoritma *Naïve Bayes* dan *Logistic Regression* dalam analisis sentimen opini publik mengenai program Makanan Bergizi Gratis (MBG) berdasarkan data teks dari media sosial X?
2. Bagaimana perbandingan performa *Naïve Bayes* dan *Logistic Regression* dalam analisis sentimen masyarakat terhadap program MBG di media sosial X?

1.3 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dari penelitian ini antara lain adalah sebagai berikut:

1. Mengidentifikasi dan menjelaskan penerapan algoritma *Naïve Bayes* dan *Logistic Regression* dalam menganalisis sentimen publik terhadap program Makan Bergizi Gratis (MBG) menggunakan data teks media sosial X, meliputi tahapan pra-pemrosesan teks, ekstraksi fitur, hingga proses klasifikasi sentimen.
2. Membandingkan kinerja kedua algoritma dalam mengklasifikasikan sentimen publik menggunakan metrik evaluasi standar seperti akurasi, presisi, *recall*, dan *F1-Score*, untuk menentukan metode yang paling efektif dan memberikan rekomendasi terbaik untuk penerapan analisis sentimen pada kebijakan publik di media sosial.

1.4 Manfaat Penelitian

Adapun manfaat yang dihasilkan dari penelitian ini adalah sebagai berikut:

1. Menambah wawasan lebih dalam tentang penerapan dan perbandingan algoritma *Naïve Bayes* dan *Logistic Regression* dalam analisis sentimen dari *tweet* pada media sosial X.
2. Menjadi referensi bagi peneliti selanjutnya dalam menganalisis opini masyarakat berbahasa Indonesia di media sosial.

3. Memberikan gambaran yang lebih nyata tentang persepsi publik terhadap program Makan Bergizi Gratis, yang dapat digunakan sebagai dasar untuk meningkatkan kebijakan dan layanan agar lebih sesuai dengan kebutuhan masyarakat.

1.5 Ruang Lingkup dan Batasan Masalah

Adapun ruang lingkup dan batasan masalah dari penelitian ini diantara lain adalah sebagai berikut:

1. Penelitian ini hanya memfokuskan pada persepsi publik terhadap program MBG menggunakan *tweet* yang berbahasa Indonesia dari platform X.
2. Tahapan yang digunakan meliputi pengolahan teks dan pengklasifikasian sentimen menggunakan metode *Naïve Bayes* (NB) dan *Logistic Regression* (LR).
3. Pada metode *Naïve Bayes*, model *Multinomial Naïve Bayes* (MNB) dipilih karena data yang dikelola berupa teks berdasarkan frekuensi kata.
4. Penelitian ini membatasi klasifikasi sentimen menjadi tiga kelas utama, yaitu positif, netral dan negatif sebagai bentuk respon masyarakat terhadap program Makan Bergizi Gratis (MBG).
5. Rentang data yang digunakan dalam penelitian ini sebanyak 1.500 data dan dibatasi pada periode Januari 2025 hingga Januari 2026, dengan kata kunci “Makanan Bergizi Gratis (MBG) dan Badan Gizi Nasional (BGN)” sesuai dengan kebutuhan penelitian, untuk memastikan bahwa analisis sentimen mencerminkan kondisi terkini.
6. Hasil pengujian kedua algoritma dibandingkan menggunakan *confusion matrix* dengan parameter *accuracy*, *precision*, *recall*, dan *F1-Score*.
7. Sistem yang dibangun merupakan aplikasi berbasis *website* yang dirancang untuk satu pengguna (*admin*) dalam analisis sentimen terkait program MBG pada sosial media X.