

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan dalam bidang teknologi informasi dan komunikasi menghadirkan transformasi signifikan dalam cara masyarakat berinteraksi serta menyampaikan pandangan terhadap berbagai isu sosial dan kebijakan pemerintah. Media sosial menjadi arena diskusi publik yang aktif, di mana masyarakat dapat dengan bebas mengekspresikan opini, kritik, maupun dukungan terhadap program-program nasional, maupun dukungan terhadap program-program nasional. Kondisi ini menjadikan media sosial sebagai salah satu sumber informasi yang mencerminkan persepsi publik secara luas, sekaligus memberi gambaran mengenai respons masyarakat terhadap kebijakan yang sedang dijalankan pemerintah.

Salah satu kebijakan pemerintah yang banyak mendapat perhatian publik dalam beberapa tahun terakhir adalah Program Makan Bergizi Gratis (MBG). Program ini merupakan inisiatif strategis dari pemerintah dalam meningkatkan status gizi dan kesehatan peserta didik di seluruh Indonesia. Berdasarkan penelitian [1], kebijakan MBG dipandang sebagai upaya nyata untuk mewujudkan pemerataan kesejahteraan sosial dan keadilan pangan. Namun demikian, pelaksanaannya masih menghadapi sejumlah tantangan, terutama terkait keterbatasan anggaran dan kesiapan infrastruktur di berbagai daerah.

Dalam konteks komunikasi publik, berbagai tanggapan masyarakat terhadap pelaksanaan Program MBG banyak disampaikan melalui *platform X* (sebelumnya Twitter). *Platform* ini memungkinkan masyarakat untuk menyalurkan opini secara langsung dan terbuka, baik dalam bentuk dukungan, kritik, maupun saran. Banyaknya opini publik yang muncul di media sosial menggambarkan beragam persepsi masyarakat terhadap kebijakan tersebut,

sehingga analisis terhadap data ini menjadi penting untuk memahami sejauh mana tingkat penerimaan dan respons publik terhadap Program MBG.

Beberapa penelitian terdahulu telah memanfaatkan media sosial sebagai sumber data untuk menjalankan analisis sentimen terhadap berbagai kebijakan pemerintah menggunakan metode *machine learning*. [2] dalam penelitiannya yang berjudul “Analisis Sentimen Program Makan Gratis pada *platform X* Menggunakan Algoritma *Naïve Bayes*” berhasil mencapai tingkat akurasi sebesar 86% dengan pembobotan TF-IDF. Sementara itu, [3] menggunakan metode *Support Vector Machine* (SVM) dalam penelitian berjudul “Analisis Sentimen Program Makan Bergizi Gratis” Pemerintah RI dan memperoleh akurasi sebesar 82% dengan kecenderungan opini publik bersifat netral. Selanjutnya, [4] dalam penelitiannya yang berjudul Analisis Sentimen pada Ulasan Pengguna Aplikasi Bibit dan Bareksa dengan Algoritma KNN berhasil mencapai tingkat akurasi sebesar 85,14% untuk aplikasi Bibit dan 81,70% untuk aplikasi Bareksa.

Ketiga penelitian tersebut menunjukkan efektivitas algoritma *machine learning* dalam analisis sentimen berbasis teks. Namun, hingga saat ini belum ada penelitian yang secara spesifik menerapkan algoritma *K-Nearest Neighbor* (KNN) untuk menganalisis sentimen publik terhadap kebijakan MBG. Selain itu, sebagian besar penelitian terdahulu cenderung berfokus pada penggunaan satu algoritma tanpa menjalankan perbandingan performa antar metode dalam konteks kebijakan tersebut. Padahal, perbandingan semacam ini penting untuk mengetahui sejauh mana setiap algoritma mampu beradaptasi dengan karakteristik bahasa Indonesia yang kompleks dan tidak terstruktur.

Dengan demikian, penelitian ini bertujuan menjembatani celah tersebut dengan membandingkan dua algoritma populer, yaitu *Naïve Bayes* (NB) dan *K-Nearest Neighbor* (KNN). Algoritma *Naïve Bayes* dikenal efisien dalam mengklasifikasikan teks berdimensi tinggi melalui pendekatan probabilistik, sedangkan KNN bekerja berdasarkan prinsip kedekatan antar data dengan menghitung jarak antar fitur. Melalui pendekatan komparatif ini, studi

diharapkan dapat mengidentifikasi metode yang paling efektif dalam menganalisis sentimen publik terhadap Program MBG.

Perbandingan kinerja kedua algoritma dilakukan dengan mengukur performanya memakai metrik evaluasi akurasi, presisi, *Naïve Bayes*, dan *F1-score*. Data yang dipakai diperoleh dari platform *X* dengan kata kunci yang relevan dengan Program MBG selama periode dua bulan terakhir. Proses pengolahan data mencakup tahapan *pre-processing* (*case folding*, *tokenizing*, *stopword removal*, dan *stemming*) serta pembobotan fitur menggunakan metode TF-IDF (*Term Frequency–Inverse Document Frequency*). Hasil klasifikasi dibagi ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral.

Melalui hasil penelitian ini, diharapkan dapat diperoleh gambaran empiris mengenai persepsi publik terhadap Program Makan Bergizi Gratis, sekaligus menjadi bahan pertimbangan bagi pemerintah dalam menyusun strategi komunikasi publik dan pengambilan kebijakan berbasis data yang lebih efektif.

1.2 Rumusan Masalah

Berdasarkan paparan latar belakang di atas, maka permasalahan yang dirumuskan dalam studi ini adalah “Bagaimana perbandingan kinerja algoritma *Naïve Bayes* dan *K-Nearest Neighbor* (KNN) dalam menjalankan Klasifikasi sentimen publik terhadap Program Makan Bergizi Gratis (MBG) berdasarkan data yang diperoleh dari media sosial *X* (Twitter) memakai pembobotan TF-IDF?”

1.3 Tujuan Penelitian

Tujuan dari studi ini adalah untuk membandingkan kinerja algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dalam menganalisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG) berdasarkan data dari media sosial *X* (Twitter).

1.4 Batasan Masalah

Penelitian ini berfokus pada hal-hal berikut:

1. Data yang digunakan berupa *tweet* dari *platform* media sosial *X* (sebelumnya Twitter) yang hanya mencakup unggahan dalam bahasa Indonesia.
2. Pengambilan data dilakukan menggunakan beberapa kata kunci yang relevan dengan Program Makan Bergizi Gratis (MBG), seperti “*MakanBergiziGratis*,” dan “*mbg*.”
3. Periode pengumpulan data ditetapkan selama dua bulan terakhir, terhitung dari 1 Agustus 2025 hingga 30 September 2025.
4. Klasifikasi sentimen dalam penelitian ini dibatasi pada tiga kategori, yaitu positif, negatif, dan netral.
5. Analisis perbandingan kinerja algoritma difokuskan pada dua metode *machine learning*, yaitu *Naïve Bayes* (NB) dan *K-Nearest Neighbor* (KNN), dengan teknik pembobotan fitur TF-IDF (*Term Frequency–Inverse Document Frequency*).
6. Perbandingan hasil kedua algoritma dilakukan berdasarkan empat metrik evaluasi utama, yaitu:
 - Akurasi, untuk mengukur tingkat ketepatan klasifikasi model.
 - Presisi, untuk mengetahui seberapa tepat model dalam mengklasifikasikan data positif secara benar.
 - *Recall*, untuk menilai kemampuan model dalam mendeteksi seluruh data yang relevan pada tiap kelas.
 - *F1-score*, sebagai ukuran keseimbangan antara presisi dan *Naïve Bayes* dalam satu nilai komposit.

1.5 Manfaat Penelitian

1. Penelitian ini memberi gambaran perbandingan kinerja algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dalam analisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG), sehingga dapat diketahui metode yang paling efektif untuk pengolahan dan klasifikasi opini publik di media sosial.
2. Hasil penelitian ini dapat menjadi masukan bagi pemerintah dalam memantau persepsi dan respons masyarakat terhadap pelaksanaan Program MBG, mengidentifikasi isu yang paling sering dikeluhkan maupun diapresiasi, serta mendukung proses evaluasi dan penyempurnaan kebijakan secara lebih tepat sasaran berbasis data.

1.6 Sistematika Penulisan

Sistematika penulisan dalam studi ini disusun sebagai berikut:

BAB I PENDAHULUAN

Bab ini merupakan bagian awal yang menjelaskan gambaran umum penelitian, meliputi latar belakang masalah, rumusan masalah, tujuan penelitian, batasan, serta manfaat penelitian. Pada bagian ini juga dijelaskan konteks kebijakan Program Makan Bergizi Gratis (MBG) dan pentingnya analisis sentimen publik terhadap implementasinya memakai pendekatan machine learning.

BAB II TINJAUAN PUSTAKA

Bab ini menguraikan teori-teori yang menjadi dasar penelitian, termasuk konsep analisis sentimen, algoritma *Naïve Bayes* dan *K-Nearest Neighbor* (KNN), serta studi-studi terdahulu yang relevan. Bagian ini berfungsi memberikan landasan ilmiah dan memperkuat posisi penelitian dalam konteks pengembangan kajian serupa.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode studi yang meliputi pendekatan, pengumpulan data, serta pengolahan data memakai *Google Colab*. Data diambil dari platform *X* (Twitter) melalui proses *Confusion Matrix* dengan kata kunci terkait Program Makan Bergizi Gratis (MBG), kemudian diberi label sentimen positif, negatif, dan netral secara manual.

Tahap pengolahan data meliputi *cleaning* awal untuk menghapus, emoji, dan tautan, dilanjutkan dengan *pre-processing* berupa *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Data kemudian divisualisasikan melalui diagram lingkaran dan *word cloud* untuk melihat distribusi serta kata yang sering muncul.

Untuk klasifikasi, data dibagi menjadi data latih dan data uji, lalu dibobotkan memakai metode TF-IDF. Dua algoritma, *Naïve Bayes* (NB) dan *K-Nearest Neighbor* (KNN), dipakai untuk menganalisis sentimen. Hasilnya dievaluasi menggunakan metrik akurasi, *presisi*, *recall*, dan *F1-score* guna menentukan metode dengan performa terbaik dalam mengKlasifikasikan opini publik terhadap Program MBG.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi hasil pengolahan data yang diperoleh melalui penerapan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* (KNN) pada dataset yang dikumpulkan dari media sosial *X* (Twitter). Pada bagian ini dilakukan analisis terhadap performa kedua algoritma serta pembahasan mengenai kecenderungan sentimen publik terhadap Program MBG. Hasil analisis kemudian diinterpretasikan berdasarkan teori dan temuan dari penelitian terdahulu.

BAB V PENUTUP

Bab ini memuat kesimpulan dari hasil penelitian yang telah dilakukan serta saran untuk pengembangan penelitian selanjutnya. Kesimpulan disusun

berdasarkan hasil analisis dan pembahasan, sedangkan saran ditujukan bagi pihak-pihak terkait agar dapat memanfaatkan hasil penelitian dalam evaluasi maupun pengambilan kebijakan yang lebih tepat sasaran.