

BAB I

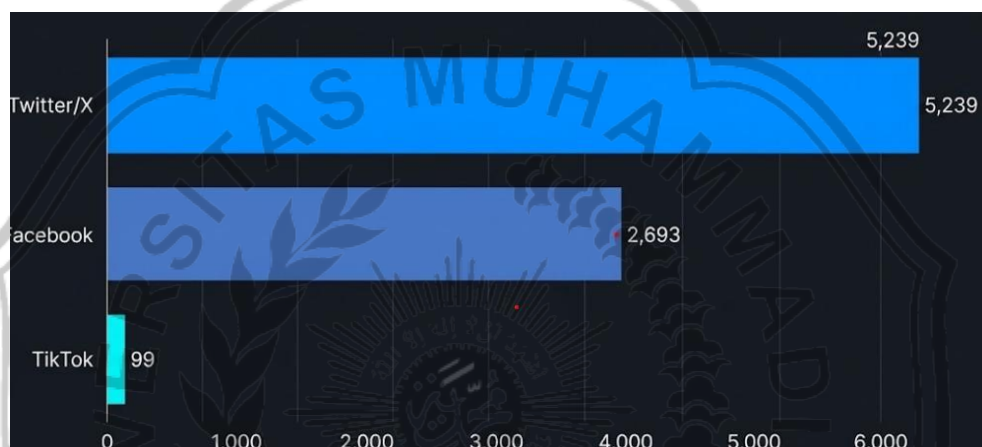
PENDAHULUAN

1.1 Latar Belakang

Indonesia rutin meluncurkan program sosial untuk mendukung kesejahteraan warganya. Perbaikan gizi anak menjadi salah satu fokus yang paling banyak disorot dalam beberapa tahun terakhir. Presiden Prabowo Subianto secara resmi meluncurkan Program Makan Bergizi Gratis (MBG) pada 6 Januari 2025. Badan Gizi Nasional (BGN) mengelola program ini dengan anggaran Rp71 triliun dalam APBN 2025, mencakup 19,47 juta penerima manfaat mulai dari jenjang PAUD hingga SMA, balita, ibu hamil, dan ibu menyusui (Kementerian Keuangan Republik Indonesia, 2025). Jumlah penerima manfaat ini direncanakan bertambah secara bertahap hingga mencapai 82 juta orang (Sekretariat Kabinet Republik Indonesia, 2025). Skala besar dan anggaran yang signifikan menjadikan MBG salah satu kebijakan yang banyak mendapat perhatian publik sepanjang 2025. Pelaksanaannya tidak lepas dari kontroversi, mulai dari kasus keracunan massal hingga kritik terhadap tata kelola distribusi. Berbagai persoalan ini menjadikan MBG topik yang relevan untuk dikaji dari sudut pandang persepsi publik.

Platform X kerap dipilih sebagai sumber data opini karena statusnya sebagai platform mikroblog Berdasarkan analisis pasar digital yang dilakukan oleh Hidayat dan Pratama (2026), Facebook dan TikTok diakui sebagai platform dengan basis pengguna terbesar di Indonesia karena dominasi jangkauan pengguna yang melampaui 110 juta pengguna aktif. Data tersebut menunjukkan bahwa kedua platform memiliki basis pengguna yang sangat besar di Indonesia Kedua platform tersebut lebih menonjolkan konten visual berupa gambar dan video pendek. Interaksi pada kedua platform ini lebih banyak berlangsung melalui kolom komentar yang menyertai konten visual, sehingga proses pengumpulan dan pengelompokan data teks menjadi lebih menantang dibandingkan platform X. Platform X berbeda karena mengutamakan format teks singkat, sehingga pengguna dapat menyampaikan sikap secara langsung melalui fitur *tweet*, *reply*, *quote tweet*, dan *hashtag*. Karakteristik tersebut menjadikan platform X

ruang diskusi yang aktif untuk membahas isu sosial, politik, maupun kebijakan negara secara *real time*, dan memudahkan pengamatan terhadap perkembangan opini publik. platform X juga menyediakan data yang relatif lebih mudah diakses melalui proses *crawling*, yang cocok digunakan untuk studi *text mining* berbasis *machine learning*. Gambar 1.1 memperlihatkan perbandingan jumlah data hasil *crawling* dari platform X, Facebook, dan TikTok. Perbandingan tersebut menjadi salah satu pertimbangan dalam memilih platform X sebagai sumber data penelitian.



Gambar 1. 1 Perbandingan Jumlah Data Hasil Crawling di Sosial Media

Berdasarkan Gambar 1.1, proses *crawling* data yang dilakukan pada tahun 2025 dengan kata kunci "MBG" di platform X berhasil mengumpulkan sebanyak 5.239 *tweet*. Jumlah tersebut lebih besar dibandingkan hasil *crawling* dari grup Facebook "*MBG Makan Bergizi Gratis BGN Badan Gizi Nasional SPPG Forum KDMP UMKM*", yang menghasilkan 2.693 postingan. Sementara itu, proses *crawling* pada kolom komentar akun TikTok @sppg_panti01_jember menghasilkan 99 komentar. Berdasarkan proses *crawling* yang dilakukan, platform X menghasilkan jumlah data yang lebih besar dibandingkan sumber data Facebook dan TikTok yang digunakan dalam penelitian ini. Kondisi tersebut menjadi salah satu pertimbangan dalam memilih platform X sebagai sumber data utama penelitian.

Pilihan ini berpijak pada fakta bahwa diskusi di platform X didominasi oleh teks tertulis yang lugas, memiliki pembahasan yang terpusat, serta pengelompokan opini yang jelas melalui tagar. Karakteristik ini mempermudah proses pengambilan

data (*crawling*). Kondisi tersebut berbeda dengan Facebook, yang interaksinya lebih menyebar dalam komunitas dan bercampur dengan media visual. Sementara itu, opini di TikTok cenderung terpecah dalam video pendek dan menggunakan bahasa tidak baku, sehingga mempersulit proses klasifikasi data. Peneliti menetapkan platform X sebagai instrumen pengumpul data utama untuk memperoleh gambaran akurat mengenai pandangan publik atas Program Makan Bergizi Gratis (MBG).

Sebagai program yang menjangkau jutaan penerima manfaat, MBG memunculkan beragam tanggapan dari masyarakat. Opini publik terhadap MBG di platform X dikelompokkan menjadi tiga kategori sentimen, yaitu positif, negatif, dan netral. Mukarom dkk, (2025) menemukan sentimen positif mendominasi dengan persentase 67,9%, diikuti sentimen netral sebesar 23,2% dan sentimen negatif sebesar 8,9%. Kualitas makanan merupakan aspek yang paling banyak disorot, sedangkan pelaksanaan program paling banyak menerima kritik. Wacana program makan bergizi gratis yang berkembang di media sosial juga telah memunculkan berbagai tanggapan, baik berupa dukungan maupun kritik.

Berbagai penelitian telah menerapkan algoritma *Naïve Bayes* dan pembobotan kata TF-IDF untuk menganalisis sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG) pada data platform X. Seperti penelitian Wibowo dkk, (2026), Singarimbun dan Iqbal (2026) melalui pendekatan *K-Fold*, serta Ihdahayuningsih dkk, (2026). Meskipun demikian, analisis sentimen masih menghadapi tantangan berupa ketidakseimbangan kelas (*class imbalance*), yaitu kondisi ketika distribusi data pada setiap kelas sentimen tidak proporsional sehingga model cenderung memiliki performa yang lebih baik pada kelas mayoritas. Salah satu teknik yang digunakan untuk mengatasi permasalahan tersebut adalah *Synthetic Minority Oversampling Technique* (SMOTE), yang menghasilkan sampel sintesis pada kelas minoritas agar distribusi data menjadi lebih seimbang. Penelitian Najib dkk, (2026) menunjukkan bahwa penerapan kombinasi *Naïve Bayes*, TF-IDF, dan SMOTE pada data komentar YouTube terkait MBG meningkatkan *F1-score macro* dari 27,60% menjadi 36,24%, yang mengindikasikan peningkatan kemampuan model dalam mengklasifikasikan kelas sentimen minoritas.

Meskipun demikian, penelitian Najib dkk, (2026) menggunakan data komentar YouTube, sedangkan penelitian yang memanfaatkan platfrom X sebagai sumber data umumnya belum menerapkan *Synthetic Minority Oversampling Technique* (SMOTE) atau teknik resampling sejenis untuk menangani ketidakseimbangan kelas. Perbedaan karakteristik data platfrom X dan YouTube, baik dari segi panjang teks, gaya bahasa, maupun struktur percakapan, menunjukkan bahwa kinerja kombinasi *Naïve Bayes Classifier*, TF-IDF, dan SMOTE pada data platfrom X masih perlu dikaji. *Naïve Bayes Classifier* dipilih karena sederhana, efisien, dan sesuai untuk klasifikasi data teks. Namun, tanpa penanganan ketidakseimbangan kelas, model cenderung memberikan kinerja yang lebih baik pada kelas mayoritas dibandingkan kelas minoritas. Oleh karena itu, penerapan SMOTE dipandang relevan untuk meningkatkan keseimbangan distribusi data sehingga kemampuan model dalam mengklasifikasikan seluruh kelas sentimen dapat ditingkatkan. Kesenjangan tersebut menjadi salah satu dasar pelaksanaan penelitian ini.

Berdasarkan fenomena dan kendala yang telah dipaparkan, masalah utama dalam penelitian ini berkaitan dengan klasifikasi sentimen opini masyarakat secara akurat menggunakan data platfrom X. Tantangan utamanya terletak pada ketidakseimbangan jumlah data pada tiap label sentimen, yang berisiko menurunkan keandalan hasil prediksi model. Penelitian ini memadukan tiga metode untuk menangani masalah tersebut, yaitu TF-IDF untuk menghitung bobot kata, SMOTE untuk mengatasi ketidakseimbangan kelas, dan *Naïve Bayes Classifier* untuk mengklasifikasikan sentimen. Performa model diukur menggunakan empat metrik, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Pendekatan ini diharapkan dapat menghasilkan performa klasifikasi yang lebih baik pada sentimen positif, negatif, dan netral, terutama dalam mengenali kelas minoritas. Hasil akhir ini membuat penelitian berjudul ANALISIS SENTIMEN OPINI MASYARAKAT TENTANG PROGRAM MAKAN BERGIZI GRATIS (MBG) DI TWITTER MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER sangat penting bagi pemerintah maupun publik untuk memahami sudut pandang masyarakat. Penelitian ini bertujuan untuk menghasilkan model klasifikasi sentimen yang optimal dan menyajikan informasi yang valid mengenai evaluasi publik terhadap Program

Makan Bergizi Gratis (MBG).

1.2 Rumusan Masalah

Penelitian ini bertujuan menganalisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG) di X menggunakan *Naïve Bayes Classifier* dan SMOTE. Rumusan masalah penelitian ini mencakup:

1. Bagaimana hasil analisis sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG) di X dikategorikan ke dalam sentimen positif, negatif, dan netral menggunakan *Naïve Bayes Classifier* ?
2. Bagaimana kinerja *Naïve Bayes Classifier* dalam mengklasifikasikan sentimen masyarakat terhadap Program MBG sebelum dan sesudah penerapan *Synthetic Minority Oversampling Technique* (SMOTE) ?

1.3 Batasan Penelitian

Agar penelitian memiliki ruang lingkup yang jelas dan terarah serta pembahasan tidak melebar dari tujuan penelitian, maka diperlukan batasan penelitian. Batasan ini ditetapkan untuk memfokuskan proses pengumpulan, pengolahan, dan analisis data sehingga hasil penelitian dapat diperoleh secara lebih sistematis. Adapun batasan dalam penelitian ini adalah sebagai berikut:

1. Pengumpulan data pada penelitian ini menggunakan *tweet* yang diperoleh dari media sosial X (Twitter) dengan kata kunci dan *hashtag* yang berkaitan dengan Program Makan Bergizi Gratis, yaitu "Makan Bergizi Gratis", "MBG", #MakanBergiziGratis, dan #MBG, dengan rentang waktu pengambilan data mulai 1 Juni 2025 hingga 31 Desember 2025
2. Data yang diambil pada sebanyak 5.239 *tweet* dan setelah melakukan *preprocessing* data menghasilkan 5.232 *tweet*
3. Data set hanya menggunakan bahasa Indonesia.
4. Menggunakan bahasa pemrograman *Python* dan *software* yang digunakan berupa *Google colab*.
5. Pelabelan data dilakukan validator bahasa yaitu Bapak Ardiyansah M.Pd dan dibantu oleh ibu Fatdhillahtullah S.Pd

6. Pelabelan sentimen pada penelitian ini hanya menggunakan tiga kategori, yaitu positif, negatif, dan netral. Hasil pelabelan menghasilkan 2.788 data sentimen negatif, 1.320 data sentimen positif, dan 1.123 data sentimen netral, yang selanjutnya digunakan sebagai data acuan dalam proses klasifikasi.
7. Dataset dibagi menggunakan rasio 80:20, yang menghasilkan 4.184 data *training* dan 1.047 data *testing*. Rasio ini memberikan proporsi yang ideal antara proses pembelajaran model dan evaluasi performa pada data baru.

1.4 Tujuan Penelitian

Beberapa tujuan utama yang ingin dicapai melalui penelitian ini mencakup:

1. Untuk menganalisis sentimen publik terhadap Program Makan Bergizi Gratis (MBG) di X
2. Untuk mengevaluasi kinerja *Naïve Bayes Classifier* dalam analisis sentimen tersebut sebelum dan sesudah penerapan *Synthetic Minority Oversampling Technique* (SMOTE).

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan kontribusi sebagai berikut :

1. Manfaat Teoretis

Penelitian ini diharapkan dapat memberi kontribusi pada bidang penambangan data teks (*text mining*), khususnya pada analisis sentimen terhadap kebijakan publik. Selain itu, penelitian ini diharapkan dapat memperkaya kajian mengenai penerapan *Multinomial Naïve Bayes*, *Term Frequency-Inverse Document Frequency* (TF-IDF), dan *Synthetic Minority Oversampling Technique* (SMOTE) dalam mengatasi ketidakseimbangan data pada klasifikasi teks.

2. Manfaat Praktis

Penelitian ini disusun untuk memberi manfaat praktis bagi tiga pihak. Bagi pemerintah, hasil analisis dapat membantu evaluasi dan pengambilan keputusan dalam pelaksanaan Program Makan Bergizi Gratis (MBG). Bagi akademisi, penelitian ini berkontribusi pada kajian analisis sentimen dan machine learning. Bagi peneliti selanjutnya, metodologi dalam penelitian ini dapat di replikasi atau

dikembangkan lebih lanjut, khususnya pada teknik optimasi algoritma klasifikasi dalam evaluasi kebijakan publik.

