BAB I
PENDAHULUAN

1.1 Latar Belakang

Program Makan Bergizi Gratis (MBG) merupakan program nasional yang diluncurkan oleh Presiden Prabowo Subianto pada tahun 2025. Implementasi dari program MBG merupakan bagian dari pemenuhan janji-janji kampanye yang telah disampaikan oleh Prabowo-Gibran saat masih menjadi calon presiden. Sebagai bagian dari upaya strategis untuk meningkatkan sumber daya manusia (SDM), program makan bergizi gratis bertujuan untuk meningkatkan nutrisi anak sekolah dan ibu hamil (Pora et al., 2025). Program ini juga diharapkan mampu menekan angka *stunting* serta meningkatkan kualitas kesehatan. Secara konseptual, kebijakan MBG sejalan dengan upaya pemerintah dalam mewujudkan kesejahteraan sosial dan pemerataan pangan bergizi.

Namun, dalam proses implementasinya, Program Makan Bergizi Gratis tidak terlepas dari berbagai respons dan opini masyarakat. Perkembangan media sosial sebagai ruang diskusi terbuka membuat masyarakat dengan mudah menyampaikan pendapat, kritik, maupun dukungan terhadap kebijakan pemerintah. Media sosial X (sebelumnya Twitter) menjadi salah satu wadah paling berpengaruh dalam membentuk wacana publik. Platform ini populer karena kemampuannya menyebarkan informasi secara cepat melalui fitur seperti hashtag, yang memudahkan pencarian isu terkini (Aprianti et al., 2025). Media sosial X juga menjadi salah satu platform yang banyak digunakan untuk mengekspresikan opini secara *real-time* terkait isu-isu publik, termasuk pelaksanaan program MBG. Beragam pandangan yang muncul menunjukkan adanya sentimen positif, negatif, maupun netral dari masyarakat.

Perbedaan sentimen masyarakat terhadap Program Makan Bergizi Gratis menjadi hal yang penting untuk diperhatikan, karena persepsi masyarakat dapat memengaruhi tingkat kepercayaan masyarakat terhadap kebijakan pemerintah. Sentimen negatif yang tidak teridentifikasi dan dianalisis secara sistematis

berpotensi menimbulkan resistensi sosial, penyebaran informasi yang bias, serta menurunkan efektivitas program. Sebaliknya, sentimen positif dapat menjadi indikator penerimaan masyarakat serta keberhasilan awal kebijakan. Oleh karena itu, diperlukan suatu pendekatan yang mampu menganalisis opini masyarakat secara komprehensif dan objektif.

Analisis sentimen adalah cabang ilmu dari data mining yang berfungsi untuk menganalisis, memproses, dan mengekstrak data tekstual terkait entitas, seperti layanan, produk, individu, organisasi, peristiwa, atau topik tertentu. Analisis sentimen merupakan bidang baru dalam *Natural Language Processing* (NLP) yang bertujuan mengidentifikasi subjektivitas dalam teks serta melakukan klasifikasi sentimen pada opini (Wijaya et al., 2024). Metode ini memungkinkan pengelompokan opini ke dalam kategori sentimen positif, negatif, dan netral berdasarkan data teks yang tersedia. Analisis sentimen telah banyak diterapkan untuk mengkaji opini masyarakat sebagai cara memahami respons masyarakat terhadap suatu objek. Salah satu metode yang sering digunakan adalah Naïve Bayes karena kesederhanaan dan efektivitasnya dalam klasifikasi teks.

Metode Naïve Bayes adalah metode klasifikasi dalam penambangan teks yang digunakan dalam analisis sentimen (Samsir et al., 2021). Metode Naïve Bayes merupakan algoritma klasifikasi berbasis probabilistik yang dikenal memiliki kelebihan dalam kesederhanaan, efisiensi komputasi, serta kemampuan bekerja dengan baik pada data teks berukuran besar. Metode ini juga banyak digunakan dalam penelitian analisis sentimen karena mudah diimplementasikan dan memberikan hasil yang cukup baik, terutama pada data yang telah melalui tahap *pre-processing* yang tepat. Meskipun demikian, penerapan Naïve Bayes pada analisis sentimen kebijakan publik, khususnya Program Makan Bergizi Gratis, masih relatif terbatas.

Dalam penelitian yang dilakukan oleh Priyanto menunjukkan bahwa metode Naïve Bayes mampu mengklasifikasikan sentimen masyarakat terhadap pembangunan Ibu Kota Nusantara dengan menyimpulkan persepsi positif

sebesar 52,7%, negatif sebesar 27,6%, dan netral sebesar 19,7% serta tingkat akurasi sebesar 75,17% (Priyanto et al., 2025). Dalam penelitian lainnya yang dilakukan oleh Samsir menunjukkan metode Naïve Bayes digunakan untuk mengklasifikasikan sentimen masyarakat terhadap pembelajaran daring selama COVID-19 di media sosial Twitter dengan sentimen positif 30%, negatif 69%, dan netral 1% serta tingkat akurasi sebesar 97,15% (Samsir et al., 2021). Dalam penelitian lainnya yang dilakukan oleh Duei Putri metode Naïve Bayes mampu digunakan untuk mengklasifikasikan sentimen masyarakat terhadap kinerja Dewan Perwakilan Rakyat (DPR) dengan tingkat akurasi 80% (Duei Putri et al., 2022).

Penelitian serupa pernah dilakukan oleh Aprianti dalam menganalisis sentimen masyarakat terhadap Program MBG menggunakan metode Naïve Bayes dengan hasil akurasi, presisi, *recall*, dan F1-*score* secara berurutan: 86,78%, 86,86%, 86,78%, 86,58% (Aprianti et al., 2025). Penelitian serupa lainnya pernah dilakukan oleh Sitanggang dalam menganalisis sentimen masyarakat terhadap Program MBG menggunakan metode Naïve Bayes dengan hasil akurasi, presisi, *recall*, dan F1-*score* secara berurutan: 72,2%, 63,2%, 66,1%, 64% (Sitanggang et al., 2024).

Namun, sebagian besar penelitian analisis sentimen sebelumnya masih menggunakan metode pelabelan secara manual yang memerlukan waktu pelabelan yang lama atau menggunakan metode lexicon-based yang memiliki keterbatasan kosakata dan *slang* serta tidak mampu memahami konteks sentimen.

Seiring berkembangnya teknologi Artificial Intellegence, khususnya Large Language Model (LLM), proses pelabelan dapat dilakukan secara cepat dan otomatis dengan kemampuan memahami konteks lebih baik. Large Language Model (LLM) merupakan AI yang dirancang untuk memahami, memproses, dan menghasilkan teks bahasa secara alami. LLM dilatih menggunakan miliaran kata dari berbagai sumber untuk meniru bahasa alami manusia secara akurat (P. R. K. Nugroho et al., 2025).

Berdasarkan uraian tersebut, diperlukan penelitian yang menganalisis sentimen masyarakat terhadap Program Makan Bergizi Gratis menggunakan metode Naïve Bayes dengan memanfaatkan data dari media sosial X dan metode LLM dalam proses pelabelan data sentimen. Penelitian ini diharapkan dapat memberikan gambaran yang jelas mengenai kecenderungan opini publik terhadap program MBG.

1.2 Identifikasi Masalah

Berdasarkan uraian latar belakang yang telah disampaikan maka dapat diidentifikasikan beberapa permasalahan pada penelitian ini, di antaranya:

1. Program Makan Bergizi Gratis (MBG) menimbulkan berbagai opini pro dan kontra di media sosial X.
2. Opini masyarakat tersebar dalam jumlah besar dan tidak terstruktur di media sosial X.
3. Belum diketahui secara pasti kecenderungan sentimen masyarakat terhadap Program MBG berdasarkan data dari media sosial X.

1.3 Rumusan Masalah

Berdasarkan identifikasi masalah yang telah diuraikan, didapat rumusan masalah dalam penelitian ini sebagai berikut:

1. Bagaimana kecenderungan sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG) di media sosial X?
2. Bagaimana penerapan metode Naïve Bayes dalam mengklasifikasikan sentimen opini masyarakat terhadap Program Makan Bergizi Gratis (MBG).
3. Bagaimana penerapan metode pelabelan LLM dalam proses pelabelan pada dataset?
4. Seberapa baik kinerja metode Naïve Bayes dalam mengklasifikasikan sentimen opini masyarakat terhadap Program MBG di media sosial X?

1.4 Batasan Masalah

Agar penelitian tetap terfokus pada permasalahan yang telah diuraikan, adapun batasan masalah yang telah ditetapkan dalam penelitian ini adalah sebagai berikut:

1. Objek penelitian dibatasi hanya pada teks tweet di media sosial X yang membahas Program Makan Bergizi Gratis (MBG).
2. Data yang dianalisis hanya tweet dalam bahasa Indonesia.
3. Pengambilan data dilakukan pada periode bulan Juli tahun 2025 hingga bulan Januari tahun 2026 dengan jumlah data 6.992.
4. Klasifikasi sentimen dibatasi menjadi tiga kategori, yaitu positif, negatif, dan netral.
5. Metode klasifikasi yang digunakan adalah metode Naïve Bayes Classifier.
6. Tahapan *text mining* yang diterapkan meliputi *pre-processing* teks, pelabelan teks, pembobotan kata, dan proses klasifikasi.
7. Penelitian tidak mencakup konten non-teks seperti gambar, video, atau media lainnya.
8. Penelitian tidak menganalisis dampak kebijakan MBG secara langsung, melainkan hanya memetakan opini publik berdasarkan sentimen teks.

1.5 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan yang ingin dicapai dalam penelitian ini adalah:

1. Mengetahui kecenderungan sentimen masyarakat terhadap Program MBG berdasarkan opini yang disampaikan di media sosial X.
2. Menerapkan metode Naïve Bayes dalam proses klasifikasi sentimen opini masyarakat terhadap Program MBG.
3. Menerapkan metode pelabelan LLM dalam proses pelabelan dataset
4. Menganalisis kinerja metode Naïve Bayes dalam mengklasifikasikan sentimen opini masyarakat terhadap Program MBG di media sosial X.

1.6 Manfaat Penelitian

Adapun manfaat yang diharapkan dari hasil penelitian ini adalah sebagai berikut:

1.6.1 Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan manfaat secara teoritis, di antaranya:

1. Menambah ilmu pengetahuan di bidang data *mining* dan pemrosesan bahasa alami (*Natural Language Processing*), khususnya dalam penerapan analisis sentimen.
2. Memberikan kontribusi akademik terkait penerapan metode Naïve Bayes dalam mengklasifikasikan sentimen opini masyarakat di media sosial.
3. Menjadi referensi dan rujukan bagi penelitian selanjutnya yang mengkaji analisis sentimen terhadap kebijakan pemerintah dengan pendekatan metode klasifikasi teks.

1.6.2 Manfaat Praktis

Penelitian ini diharapkan dapat memberikan manfaat secara praktis, di antaranya:

1. Bagi Peneliti:
    a. Menjadi acuan implementatif dalam menerapkan metode Naïve Bayes untuk penelitian analisis sentimen berbasis data media sosial, khususnya pada media sosial X.
    b. Memberikan alternatif solusi analisis otomatis terhadap data opini masyarakat dalam jumlah besar.
    c. Menambah pengalaman peneliti dalam mengolah dan menganalisis data sentimen dari media sosial.
2. Bagi Satuan Pelayanan Pemenuhan Gizi (SPPG)
    a. Memberikan gambaran mengenai kecenderungan sentimen dan opini masyarakat terhadap Program MBG berdasarkan data yang diperoleh dari media sosial X.

b. Menjadi bahan pertimbangan berbasis opini publik bagi pemangku kebijakan dalam memahami respons masyarakat terhadap pelaksanaan Program MBG.
c. Mendukung upaya evaluasi dan penyempurnaan kebijakan program melalui pemanfaatan hasil analisis sentimen masyarakat di media sosial.

1.7 Pertanyaan Penelitian

Berdasarkan permasalahan yang telah dirumuskan dan fokus penelitian yang ditetapkan, penelitian ini dapat dirumuskan ke dalam beberapa pertanyaan penelitian sebagai berikut:

1. Bagaimana kecenderungan sentimen opini masyarakat di media sosial X terhadap Program MBG, apakah termasuk dalam kategori positif, negatif, atau netral?
2. Bagaimana penerapan metode Naïve Bayes dalam mengklasifikasikan sentimen opini masyarakat terhadap Program MBG di media sosial X?
3. Bagaimana penerapan metode pelabelan LLM dalam proses pemberian label pada dataset?
4. Bagaimana kinerja metode Naïve Bayes dalam mengklasifikasikan sentimen opini masyarakat terhadap Program MBG?

1.8 Metodologi Penelitian

Metodologi penelitian adalah serangkaian tahapan sistematis yang dilakukan peneliti untuk mengumpulkan, mengolah, serta menganalisis data untuk menyelesaikan permasalahan penelitian secara terarah dan terstruktur. Penelitian ini menggunakan metode klasifikasi Naïve Bayes. Data yang digunakan berupa tweet berbahasa Indonesia yang diperoleh dari media sosial X dengan kata kunci “makan bergizi gratis” dan “MBG” dengan periode waktu bulan Juli tahun 2025 hingga bulan Januari tahun 2026.

1.8.1 Teknik Pengumpulan Data

1. Observasi

Observasi dilakukan dengan mengamati secara langsung opini masyarakat terkait Program MBG pada media sosial X, khususnya pada unggahan tweet yang mengandung kata atau tagar MBG. Melalui observasi ini, peneliti dapat mengidentifikasi bahwa opini masyarakat terhadap Program MBG tersebar dalam jumlah besar dan tidak terstruktur, sehingga menyulitkan peneliti untuk mengetahui kecenderungan sentimen masyarakat terhadap Program MBG.

2. *Crawling Data*

   Tahapan penelitian dimulai dari pengumpulan data, yaitu proses pengambilan data sekunder berupa tweet yang berasal dari media sosial X melalui *crawling data* menggunakan API X dengan bantuan bahasa pemrograman Python pada platform Google Colab. Data tweet dikumpulkan berdasarkan kata kunci "makan bergizi gratis" dan "MBG" dengan periode waktu bulan Juli 2025 hingga Januari 2026.

3. Studi Pustaka

   Studi pustaka dilakukan dengan mengkaji dari berbagai sumber literatur seperti buku, jurnal ilmiah, *website* yang kredibel, dan penelitian terdahulu yang memiiliki kaitan dengan *Data Mining*, CRISP-DM, *Crawling data*, Naïve Bayes, hingga evaluasi kinerja metode. Studi pustaka ini juga digunakan untuk memperkuat kajian pustaka dan metodologis dalam penelitian analisis sentimen yang dilakukan

1.8.2 Metode Penyelesaian masalah

Metode penyelesaian masalah yang digunakan dalam penelitian ini adalah metode CRISP-DM. *Cross Industry Standard Process for Data Mining* (CRISP-DM) adalah model proses industri-independen yang biasa digunakan pada penelitian *data mining*. Metode ini terdiri dari 6 tahap, diantaranya *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment* (Kurniawan & Yasir, 2022). Metode CRISP-DM digunakan dalam penelitian ini karena metode ini memiliki kerangka kerja yang sistematis dan terstruktur, mulai dari pemahaman tujuan penelitian, pengolahan

data, pemodelan, dan evaluasi hasil. Metode ini sangat cocok dengan karakteristik penelitian analisis sentimen berbasis *text mining* yang menggunakan data mentah dan memerlukan tahapan *pre-processing* serta evaluasi model yang jelas.

1.9 Jadwal Kegiatan Penelitian

Jadwal pelaksanaan penelitian dibuat dengan tahapan yang jelas dalam bentuk tabel.

Tabel I.1 Jadwal Kegiatan Penelitian

Kegiatan	Bulan							
	Jan	**Feb**	**Mar**	**Apr**	**Mei**	**Jun**	**Jul**	**Agu**
1. Pengumpulan data	■							
1.1. Crawling data	■							
2. Penyusunan laporan	■							
2.1. Menyusun Bab 1	■							
2.2. Menyusun Bab 2	■	■						
2.3. Menyusun Bab 3		■	■	■				
2.4. Menyusun Bab 4				■	■	■		
2.5. Menyusun Bab 5						■	■	
3. Pre-processing	■							
3.1. Cleansing data	■							
3.1.1. Menghapus data yang tidak relevan	■	■						
3.1.2. Menghapus *mention* (@username)		■						
3.1.3. Menghapus URL		■						
3.1.4. Menghapus tagar		■						
3.1.5. Menghapus angka dan tanda baca			■					
3.1.6. Menghapus spasi yang berlebih			■					
3.2. Case folding				■				
3.3. Tokenizing				■				
3.4. Normalization				■				
3.5. Stopword removal				■				
3.6. Stemming				■				
4. Labeling menggunakan LLM		■						
4.1. Pelabelan batch 1		■						
4.2. Pelabelan batch 2		■	■					
4.3. Pelabelan batch 3			■					
5. Pembobotan TF-IDF				■				
5.1. Pembobotan TF				■				
5.2. Pembobotan IDF				■				
5.3. Menghitung nilai TF-IDF				■				
6. Pemodelan Naïve Bayes					■			

6.1. Menghitung probabilitas dataset menggunakan website SPK dengan metode Naïve Bayes								
7. Confusion Matrix								
7.1. Menghitung akurasi								
7.2. Menghitung presisi								
7.3. Menghitung recall								
7.4. Menghitung f1-score								