

## BAB I PENDAHULUAN

### 1.1 Latar Belakang

Program Makan Bergizi Gratis (MBG) merupakan salah satu kebijakan publik yang memicu perhatian luas dari masyarakat, bersifat viral, dan menimbulkan beragam respons di ruang publik digital. Media sosial, khususnya platform X (Twitter), menjadi sarana utama bagi masyarakat untuk menyampaikan opini, kritik, maupun dukungan terhadap program tersebut, sehingga menghasilkan data teks dengan volume yang sangat besar [1]. Opini yang disampaikan dalam media sosial umumnya berbentuk teks singkat yang bersifat spontan dan cenderung menggunakan bahasa yang tidak formal.

Dari sudut pandang linguistik, teks-teks opini di media sosial memiliki karakteristik yang berbeda dengan teks formal. Bahasa yang digunakan sering kali mengandung kata slang, singkatan, variasi ejaan, serta struktur kalimat yang tidak sepenuhnya mengikuti kaidah bahasa Indonesia baku [2]. Karakteristik kebahasaan ini menjadikan teks media sosial menarik untuk dikaji, sekaligus menimbulkan tantangan dalam proses pemrosesan bahasa alami (*Natural Language Processing/NLP*), khususnya ketika teks tersebut digunakan untuk analisis sentimen terhadap suatu kebijakan publik seperti Program Makan Bergizi Gratis.

Analisis sentimen merupakan salah satu cabang NLP yang bertujuan untuk mengidentifikasi dan mengklasifikasikan kecenderungan sikap atau opini penulis teks ke dalam kategori tertentu, seperti positif dan negatif [3]. Dalam konteks opini masyarakat terhadap Program Makan Bergizi Gratis, analisis sentimen dapat dimanfaatkan untuk memahami bagaimana respons publik terbentuk melalui penggunaan bahasa, pilihan kata, dan ekspresi kebahasaan yang muncul dalam teks tidak formal di media sosial, seperti menggunakan *Support Vector Machine (SVM)*, *K-Nearest Neighbor (KNN)*, dan *Random Forest* [4].

Performa algoritma klasifikasi sentimen sangat dipengaruhi oleh kualitas representasi fitur dan pemilihan parameter model. Untuk meningkatkan akurasi dan

pemilihan parameter yang tepat, diperlukan proses optimasi model seperti *Grid Search* [5]. Selain itu, representasi teks menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) banyak digunakan karena mampu mengubah data teks menjadi bentuk numerik yang dapat diproses oleh algoritma machine learning [6].

Penelitian ini bertujuan untuk mengoptimalkan kinerja algoritma SVM, KNN, dan *Random Forest* dalam analisis sentimen terhadap Program Makan Bergizi Gratis di media sosial Twitter (X) dengan menerapkan metode *Grid Search* dan ekstraksi fitur TF-IDF, sehingga dihasilkan model klasifikasi sentimen yang memiliki performa optimal dan akurat.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana tingkat akurasi metode klasifikasi *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), dan *Random Forest* dalam analisis sentimen terhadap program makan bergizi gratis di Twitter?
2. Bagaimana pengaruh penerapan optimasi *hyperparameter* menggunakan *Grid Search* terhadap peningkatan akurasi model klasifikasi dalam analisis sentimen tersebut?

## 1.3 Batasan Masalah

Beberapa batasan masalah ditetapkan sebagai berikut:

1. Data yang digunakan berupa *tweet* (cuitan) pengguna Twitter yang berkaitan dengan Program Makan Bergizi Gratis di Indonesia.
2. Analisis sentimen dibatasi pada dua kategori: positif dan negatif.
3. Metode klasifikasi yang digunakan terbatas pada SVM, KNN, dan *Random Forest*.
4. Metode ekstraksi fitur teks yang digunakan adalah *Term Frequency-Inverse Document Frequency* (TF-IDF).

5. Optimasi parameter pada masing-masing algoritma klasifikasi dilakukan menggunakan metode *Grid Search*.
6. Penelitian tidak membahas analisis mendalam terhadap isi, makna linguistik, maupun konteks sosial dan kebijakan secara kualitatif dari tweet yang dianalisis.

#### 1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh peneliti dalam penelitiannya adalah:

1. Menganalisis dan mengoptimasi kinerja algoritma klasifikasi *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), dan *Random Forest* dalam analisis sentimen terhadap Program Makan Bergizi Gratis menggunakan metode *Grid Search*.
2. Membandingkan performa hasil analisis sentimen sebelum dan sesudah dilakukan optimasi parameter menggunakan *Grid Search*.

#### 1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah:

1. Penelitian ini diharapkan dapat memberikan kontribusi keilmuan dalam bidang data mining dan analisis sentimen, khususnya terkait penerapan metode *Grid Search* untuk optimasi hiperparameter pada algoritma klasifikasi *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), dan *Random Forest*. Hasil penelitian ini dapat menjadi referensi bagi penelitian selanjutnya yang membahas optimasi model machine learning pada data teks.
2. Bagi pemerintah atau pihak terkait, penelitian ini dapat digunakan sebagai bahan evaluasi untuk memahami persepsi masyarakat terhadap Program Makan Bergizi Gratis berdasarkan data media sosial. Bagi peneliti dan praktisi teknologi informasi, penelitian ini dapat menjadi acuan dalam penerapan analisis sentimen dan optimasi parameter model klasifikasi.

## 1.6 Sistematika Penulisan

Sistematika penulisan dalam skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN, berisi latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA, membahas studi literatur yang berkaitan dengan penelitian sebelumnya serta dasar-dasar teori yang mendukung penelitian, meliputi konsep analisis sentimen, algoritma *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), *Random Forest*, metode TF-IDF, dan optimasi parameter menggunakan *Grid Search*.

BAB III METODE PENELITIAN, menjelaskan tahapan penelitian yang dilakukan, mulai dari pengumpulan data, *preprocessing* data, ekstraksi fitur, proses klasifikasi SVM, KNN, dan *Random Forest*, optimasi model menggunakan *Grid Search*, hingga metode evaluasi kinerja model.

BAB IV HASIL DAN PEMBAHASAN, memaparkan hasil eksperimen yang diperoleh dari penerapan masing-masing metode, perbandingan hasil antara model dengan dan tanpa optimasi.

BAB V PENUTUP, berisi kesimpulan yang diperoleh dari hasil penelitian serta saran yang dapat dijadikan acuan untuk pengembangan penelitian selanjutnya.