

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang Penelitian**

Kemajuan teknologi dibidang komunikasi salah satunya ditandai dengan munculnya berbagai macam media sosial diantaranya adalah youtube. Youtube memiliki berbagai macam pengguna tanpa terkecuali pemerintah, akun youtube pemerintah sendiri memiliki fungsi sebagai salah satu media informasi bagi masyarakat untuk mengetahui berita terbaru. Manfaat yang bisa didapatkan dari penggunaan akun youtube pemerintah adalah dapat menganalisa sentimen masyarakat terhadap kebijakan-kebijakan yang dikeluarkan oleh pemerintah dari masyarakat pada komentar postingan akun pemerintah tersebut. Salah satunya adalah kebijakan yang berhubungan dengan pandemi yang sedang terjadi saat ini (Fitriyah dkk., 2022)

Pemerintah sebagai pembuat kebijakan sering kali menjadi sorotan utama publik, terutama saat menerapkan kebijakan yang berdampak langsung pada kehidupan sehari-hari, seperti subsidi energi, pajak, efisiensi anggaran, hingga bantuan sosial. Meskipun demikian, respon masyarakat terhadap kebijakan ini tidak selalu dapat terbaca secara cepat dan sistematis. Evaluasi kebijakan di Indonesia umumnya masih terfokus pada aspek administratif, seperti angka anggaran atau progres fisik proyek, sementara persepsi masyarakat yang tercermin di media sosial kurang mendapat perhatian yang memadai (Herawati, 2024).

Seiring dengan perkembangan *Natural Language Processing* (NLP) dan *machine learning*, analisis sentimen menjadi salah satu metode utama untuk mengkaji opini publik. NLP merupakan ide mendasar dibalik analisis sentimen yang bermanfaat untuk menentukan emosi/perasaan yang disampaikan dengan sebagian teks tertulis. Analisis sentimen memiliki sentimen positif, netral, negatif (Amrullah & Pane, 2023). Teknik ini telah banyak digunakan di bidang pemasaran dan bisnis, namun potensinya dalam evaluasi kebijakan publik masih belum sepenuhnya dimanfaatkan.

Beberapa penelitian sebelumnya telah membuktikan efektivitas analisis sentimen dalam mengukur opini masyarakat. Misalnya, penelitian yang menggunakan metode *Naïve Bayes* untuk mengklasifikasikan sentimen masyarakat terhadap kebijakan pembelajaran daring selama pandemi, penelitian dengan tingkat akurasi yang cukup tinggi.

Meskipun demikian, masih terdapat ruang eksplorasi, terutama pada penggunaan algoritma yang sederhana namun efisien seperti *K-Nearest Neighbor* (KNN). KNN dikenal sebagai algoritma klasifikasi berbasis kedekatan (*proximity*) yang tidak memerlukan proses pelatihan model secara eksplisit. Hal ini membuatnya sangat cocok untuk digunakan pada dataset berukuran kecil hingga sedang, serta mampu memberikan hasil yang kompetitif dalam konteks klasifikasi teks.

Dalam penelitian ini, data dikumpulkan dari komentar publik pada platform YouTube, mengingat karakteristiknya yang memungkinkan diskusi mendalam dan panjang terkait kebijakan pemerintah. Proses

pengambilan data dilakukan melalui YouTube API, dengan kata kunci yang disesuaikan dengan video kebijakan pemerintah yang sedang viral. Data yang diambil meliputi teks komentar, tanggal unggahan, serta metadata seperti jumlah like dan reply.

Dataset yang dikumpulkan kemudian melalui serangkaian tahapan pemrosesan, seperti pembersihan teks (*cleaning*), tokenisasi, penghapusan *stopwords*, dan *stemming*. Setelah itu, data diubah menjadi representasi numerik menggunakan teknik TF-IDF (*Term Frequency- Inverse Document Frequency*) sebelum diklasifikasikan menggunakan algoritma KNN.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, penelitian ini merumuskan permasalahan sebagai berikut:

1. Bagaimana proses preprocessing teks komentar YouTube untuk analisis sentimen?
2. Bagaimana penerapan algoritma KNN pada data komentar YouTube?
3. Bagaimana performa KNN diukur menggunakan akurasi, precision, recall, dan F1-score?
4. Bagaimana visualisasi hasil analisis sentimen pada aplikasi web berbasis desktop?

### 1.3 Batasan Masalah

Agar penelitian ini lebih terarah maka akan diberikan Batasan penelitian berdasarkan masalah yang telah didapatkan antara lain sebagai berikut :

1. Data yang digunakan berasal dari komentar publik di YouTube.
2. Sentimen yang diklasifikasikan terbagi menjadi tiga: positif, negatif, dan netral.
3. Metode klasifikasi yang digunakan hanya algoritma KNN.
4. Bahasa yang dianalisis adalah Bahasa Indonesia.
5. Aplikasi berbasis web hanya diakses melalui desktop, bukan mobile.

### 1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dirumuskan, penelitian ini bertujuan untuk:

1. Menerapkan algoritma KNN dalam analisis sentimen kebijakan pemerintah.
2. Membangun aplikasi berbasis desktop web untuk analisis sentimen.
3. Menyediakan visualisasi hasil analisis secara interaktif.

### 1.5 Kegunaan Penelitian

Penelitian ini diharapkan memberikan manfaat yang nyata bagi berbagai pihak, baik dalam dunia pendidikan, keluarga, maupun untuk pengembangan kemampuan penulis. Manfaat ini dirinci sebagai berikut:

## 1. Akademis

Menjadi referensi bagi penelitian sejenis terkait klasifikasi teks dan analisis sentimen.

## 2. Praktis

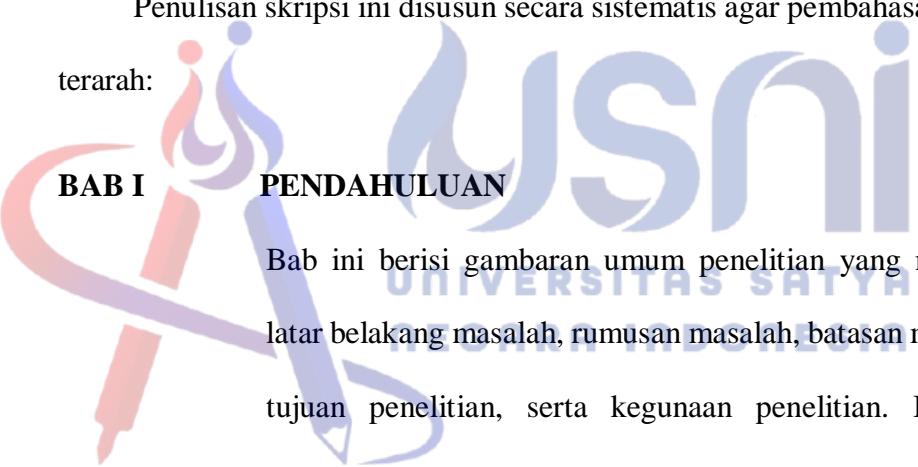
Memberikan alat bantu untuk mengetahui persepsi masyarakat terhadap kebijakan pemerintah.

## 3. Teknologis

Memberikan implementasi nyata dari *machine learning* berbasis web

## 1.6 Sistematika Penulisan

Penulisan skripsi ini disusun secara sistematis agar pembahasan lebih terarah:



### BAB I PENDAHULUAN

Bab ini berisi gambaran umum penelitian yang meliputi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, serta kegunaan penelitian. Bab ini memberikan dasar pemahaman awal mengenai pentingnya analisis sentimen masyarakat terhadap kebijakan pemerintah menggunakan algoritma K-Nearest Neighbor (KNN).

### BAB II

### TINJAUAN PUSTAKA

Bab ini memuat landasan teori dan penelitian terdahulu yang relevan dengan topik penelitian. Teori yang dibahas mencakup konsep K-Nearest Neighbor (KNN), analisis sentimen, tahapan preprocessing teks, serta pengembangan

aplikasi web berbasis desktop. Selain itu, bab ini juga menyajikan kajian terhadap penelitian-penelitian sebelumnya sebagai acuan dan pembeda penelitian ini.

### **BAB III**

### **METODE PENELITIAN**

Bab ini menjelaskan pendekatan dan metode yang digunakan dalam penelitian, meliputi waktu dan tempat penelitian, sumber dan teknik pengumpulan data, arsitektur sistem, alur proses penelitian, preprocessing data, metode ekstraksi fitur dengan TF-IDF, klasifikasi sentimen menggunakan algoritma KNN, evaluasi model, serta rancangan visualisasi dalam bentuk aplikasi berbasis desktop web.

### **BAB IV**

### **IMPLEMENTASI DAN PEMBAHASAN**

Bab ini menyajikan hasil implementasi penelitian mulai dari proses pengumpulan data komentar YouTube, preprocessing, ekstraksi fitur, pelatihan dan pengujian model KNN, evaluasi performa dengan akurasi, presisi, recall, dan F1-score, hingga visualisasi hasil dalam aplikasi web. Pada bagian ini juga dibahas interpretasi hasil analisis sentimen masyarakat terhadap kebijakan pemerintah serta kendala yang ditemui.

### **BAB V**

### **KESIMPULAN DAN SARAN**

Bab ini berisi kesimpulan yang diperoleh dari penelitian sesuai dengan tujuan yang telah ditetapkan, serta saran untuk pengembangan lebih lanjut, baik dari sisi teknis maupun cakupan penelitian, agar sistem dapat memberikan manfaat yang lebih luas dalam mendukung evaluasi kebijakan pemerintah.

