

BAB I PENDAHULUAN

1.1 Latar Belakang

Analisis sentimen merupakan bagian dari pemrosesan bahasa alami yang mengolah kata untuk melacak persepsi publik tentang suatu produk atau topik berdasarkan opini individu dalam kategori positif, netral, atau negatif [1]. Seiring berkembangnya media sosial, analisis ini semakin penting untuk memahami pola opini publik secara cepat. Hasil analisis ini berperan penting dalam pengambilan keputusan, seperti peningkatan fitur, perbaikan kualitas respons, dan peningkatan kepuasan pengguna.

Dalam penelitian pada opini publik terhadap Bank BSI, algoritma Naïve Bayes mencapai akurasi sebesar 76% dalam melakukan klasifikasi data sentimen, menunjukkan efektivitasnya sebagai metode berbasis probabilitas untuk mengolah data teks yang kompleks [2].

Pada penelitian yang sama dengan menggunakan beberapa algoritma menghasilkan Support Vector Machine (SVM) sebagai algoritma dengan akurasi tertinggi sebesar 88% dengan berbagai teknik pemisahan data [2]. Ini menunjukkan efektivitas SVM dalam mengklasifikasikan data sentimen, berkat pendekatan margin maksimum yang membedakan kelas-kelas. Akurasi yang tinggi membuktikan kemampuannya dalam menangani data kompleks.

Random Forest adalah algoritma ensemble learning yang menggabungkan banyak pohon keputusan untuk meningkatkan akurasi klasifikasi dan regresi. Dalam penelitian Reza tentang analisis sentimen ulasan aplikasi Peduli Lindungi, metode ini mencapai akurasi 72% dengan 500 pohon, meningkat menjadi 89% ketika mempertimbangkan dua sentimen (positif dan negatif) [3]. Sementara itu, penelitian Audrey menunjukkan bahwa Random Forest mencapai akurasi 93% pada klasifikasi sentimen tweet dengan parameter optimal untuk tiga kelas [4]. Algoritma ini unggul dalam menangani kompleksitas data dan meningkatkan akurasi melalui penggabungan hasil.

ChatGPT adalah model bahasa besar dari OpenAI yang menggunakan pembelajaran mesin untuk menghasilkan teks mirip manusia. Dilatih pada data teks yang luas, model ini mendukung berbagai tugas seperti pembuatan teks, penyelesaian, dan percakapan [5]. Meski bermanfaat, ketergantungan berlebihan pada ChatGPT berpotensi melemahkan kreativitas dan kemampuan berpikir kritis pengguna. Oleh sebab itu, penulis mencoba mengimplementasikan sentimen ChatGPT dalam penelitian ini.

Penelitian ini bertujuan untuk membandingkan kinerja Naive Bayes, SVM, dan Random Forest dalam analisis sentimen terhadap ChatGPT. Dengan memahami keunggulan dan kelemahan masing-masing algoritma, diharapkan penelitian ini dapat memberikan panduan bagi peneliti dan praktisi dalam memilih metode yang paling sesuai untuk analisis sentimen dalam konteks tertentu.

1.2 Rumusan Masalah

Penelitian ini memiliki batasan-batasan sebagai berikut:

Bagaimana perbandingan kinerja algoritma Naive Bayes, Support Vector Machine (SVM), dan Random Forest berdasarkan akurasi, recall, presisi, skor F1, dan waktu pemrosesan?

1.3 Batasan Masalah

Penelitian ini memiliki batasan-batasan sebagai berikut:

1. Hanya membahas tiga algoritma yaitu Naive Bayes, SVM, dan Random Forest.
2. Bahasa pemrograman yang digunakan hanya bahasa pemrograman python
3. Data yang digunakan hanya akan diklasifikasikan dalam tiga kategori sentimen: positif, negatif, dan netral.
4. Evaluasi performa algoritma dibatasi pada metrik akurasi, precision, recall, dan F1-score.
5. Penelitian ini tidak membahas optimisasi parameter untuk setiap algoritma secara mendalam.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah :

1. Membandingkan akurasi, recall, dan F1-score dari algoritma Naive Bayes, SVM, dan Random Forest dalam klasifikasi sentimen pada teks yang dihasilkan dari ChatGPT.
2. Memberikan rekomendasi algoritma yang paling sesuai untuk digunakan dalam aplikasi analisis sentimen untuk ChatGPT.

1.5 Manfaat Penelitian

Penelitian ini memiliki 2 manfaat dari segi teoritis dan praktis.

1. Manfaat Teoritis

- a. Kontribusi pada pengembangan ilmu pembelajaran mesin dan analisis sentimen.
- b. Pemahaman kinerja algoritma Naive Bayes, SVM, dan Random Forest dalam klasifikasi sentimen pada ChatGPT.
- c. Referensi bagi penelitian lanjutan tentang kelebihan dan kekurangan algoritma tersebut.

2. Manfaat Praktis

- a. Berguna bagi organisasi seperti perusahaan, pemerintah, dan pengembang yang membutuhkan analisis sentimen ChatGPT.
- b. Rekomendasi algoritma yang efisien dan akurat untuk mendukung pengambilan keputusan.
- c. Membantu pemilihan algoritma optimal sesuai kebutuhan dan sumber daya untuk analisis yang lebih cepat.

1.6 Sistematika Penulisan

Sistematika didasarkan untuk mempermudah penyusunan skripsi. Berikut sistematika penulisan yang digunakan untuk menyusun skripsi :

BAB I PENDAHULUAN

Berisi penjelasan Latar Belakang, Rumusan Masalah, Batasan Masalah, Tujuan Penelitian, Manfaat Penelitian, dan Sistematika Penulisan.

BAB II TINJAUAN PUSTAKA

Pada BAB II menguraikan dasar-dasar teori yang digunakan dalam penulisan laporan skripsi dan menjelaskan definisi secara detail dari referensi yang dijadikan rujukan untuk penulisan laporan skripsi ini.

BAB III METODE PENELITIAN

Dalam BAB III berisi penjelasan mengenai Objek Penelitian, Alur Penelitian, dan tools yang digunakan dalam penelitian.

BAB IV HASIL DAN PEMBAHASAN

BAB IV menguraikan hasil implementasi tiap algoritma pada analisis sentimen untuk ChatGPT dan perbandingan performa pada tiap algoritma.

BAB V PENUTUP

BAB V berisi Kesimpulan dan Saran.

