

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital saat ini telah mengubah cara manusia berbagi informasi dan pengalaman dalam berbagai aspek kehidupan, termasuk pada sektor pariwisata. Dengan perkembangan internet mempermudah wisatawan untuk membagikan pengalaman melalui ulasan di berbagai platform seperti Google Review, Tripadvisor, maupun media sosial. Salah satu sumber informasi bagi calon pengunjung sebuah objek wisata adalah ulasan pengunjung terdahulu [1]. Ulasan tersebut tidak hanya memberikan deskripsi perjalanan, tetapi juga opini yang mencerminkan kepuasan atau ketidakpuasan pengunjung terhadap objek wisata tersebut.

Salah satu destinasi yang sering menjadi pilihan wisatawan adalah Kawasan Gunung Merapi. Keindahan bentang alam yang megah dan daya tarik sejarah yang kuat menjadikan kawasan ini menjadi tujuan populer dari berbagai daerah. Tingginya minat kunjungan ke kawasan wisata Gunung Merapi menjadikan jumlah ulasan wisatawan terus meningkat di setiap tahunnya. Calon pengunjung sering membaca ulasan yang dibagikan di media sosial oleh pengunjung terdahulu sebagai bahan informasi dari sebuah objek wisata.

Namun, permasalahan muncul ketika jumlah ulasan yang diberikan pengunjung sangat banyak sehingga menjadi tantangan untuk menganalisis secara manual ulasan tersebut. Padahal, ulasan tersebut mengandung informasi penting mengenai persepsi positif atau negatif dari wisatawan. Oleh karena itu diperlukan metode otomatis untuk menganalisis data ulasan dalam jumlah besar secara efektif.

Analisis sentimen merupakan salah satu metode pendekatan yang dapat digunakan untuk mengidentifikasi opini dalam teks. Metode ini mampu memberikan informasi mengenai sentimen pengunjung tentang kelebihan dan kekurangan sebuah objek wisata [2]. Penelitian analisis sentimen menghadapi berbagai tantangan, seperti keberagaman kata, penggunaan sarkasme, penggunaan

kalimat majemuk, serta bentuk kalimat yang tidak struktur. Analisis sentimen mengkaji berbagai macam jenis database, teknik seleksi fitur, algoritma klasifikasi, evaluasi akurasi guna memperoleh metode yang paling efektif dalam text mining.

Pada umumnya analisis sentimen dapat dilakukan dengan menggunakan berbagai macam algoritma machine learning seperti Decision Tree, Naive Bayes, Random Forest, K-Nearest Neighbor, dan Support Vector Machine. Performa setiap algoritma klasifikasi berbeda tergantung pada data yang diolah. Dari berbagai algoritma tersebut, Support Vector Machine (SVM) dikenal memiliki performa yang baik dalam menangani data teks karena kemampuannya dalam memprediksi label sentimen [3]. Peneliti [4] menegaskan bahwa kinerja dari Support Vector Machine dengan pendekatan TF-IDF memiliki kinerja yang baik dalam melakukan analisis sentimen pada data review objek wisata di Ubud. Selain bidang pariwisata, peneliti [5] mendapatkan akurasi rata-rata sebesar 88,18% untuk algoritma SVM dimana hasil tersebut menunjukkan bahwa SVM secara konsisten memberikan akurasi lebih tinggi dibandingkan algoritma Decision Tree dan Logistic Regression dalam analisis sentimen terhadap ulasan pengguna Netflix.

Selain tantangan dalam klasifikasi sentimen itu sendiri, besarnya jumlah ulasan wisata yang dihasilkan menimbulkan kebutuhan untuk merangkum informasi secara otomatis agar dapat dipahami isi ulasan tanpa membaca semua isi dokumen. Salah satu pendekatan *auto text summarization* adalah Latent Semantic Analysis (LSA). LSA bekerja dengan memetakan hubungan semantik antar kata dan dokumen melalui dekomposisi matriks, sehingga mampu menangkap struktur makna laten di dalam kumpulan teks. Peneliti [6] menunjukkan bahwa LSA mampu menghasilkan ringkasan yang efektif dan unggul dibandingkan dengan metode lain seperti LexRank dalam konteks bahasa Indonesia.

Berdasarkan hal tersebut, penelitian ini berfokus pada analisis sentimen ulasan pengunjung Kawasan Gunung Merapi dengan menggunakan metode Support Vector Machine dan memanfaatkan metode *text summarization* berbasis Latent Semantic Analysis untuk membantu merangkum informasi dari kumpulan ulasan. Penelitian ini diharapkan dapat memberikan wawasan dan bahan referensi

bagi penelitian yang akan datang dalam mengembangkan analisis sentimen yang lebih baik khususnya dalam bidang pariwisata.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka permasalahan yang akan dibahas dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana penerapan metode Support Vector Machine untuk mengklasifikasikan sentimen positif dan negatif pada ulasan pengunjung objek wisata Kawasan Gunung Merapi yang diperoleh dari TripAdvisor?
2. Berapa tingkat *accuracy*, *precision*, *recall*, dan *F1-Score* yang dapat dicapai oleh model SVM sebelum dan sesudah LSA pada dataset ulasan pengunjung wisata Kawasan Gunung Merapi?

1.3 Batasan Masalah

Batasan dari penelitian ini adalah:

1. Bahasa Pemrograman yang digunakan adalah Python.
2. Algoritma yang digunakan adalah Support Vector Machine.
3. Klasifikasi sentimen dibagi menjadi dua yaitu, positif dan negatif.
4. Peneliti hanya menggunakan data ulasan berbahasa Indonesia yang bersifat publik dari TripAdvisor.
5. Labeling data menggunakan pendekatan *lexicon based*.
6. Kernel trick menggunakan *linear kernel*.
7. Latent Semantic Analysis menggunakan pendekatan TruncatedSVD

1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh penelitian ini adalah:

1. Mengetahui klasifikasi sentimen pada ulasan pengunjung objek wisata Kawasan Gunung Merapi yang diperoleh dari TripAdvisor..
2. Mengetahui tingkat *accuracy*, *precision*, *recall*, dan *F1-Score* yang dapat dicapai oleh model SVM sebelum dan sesudah LSA pada dataset ulasan pengunjung wisata Kawasan Gunung Merapi.

1.5 Manfaat Penelitian

Manfaat penelitian sebagai berikut :

1. Mengetahui tingkat akurasi metode Support Vector Machine dalam mengklasifikasikan ulasan pengunjung objek wisata Kawasan Gunung Merapi.
2. Menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan klasifikasi teks, pengolahan bahasa alami, dan penerapan algoritma SVM pada data ulasan berbahasa Indonesia di sektor pariwisata.

1.6 Sistematika Penulisan

Laporan penelitian ini disusun dalam sistematika penulisan agar tertata rapi dan mudah dipahami, berikut adalah struktur penulisannya :

BAB I PENDAHULUAN

Berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Berisi studi literatur dan dasar-dasar teori yang digunakan pada penelitian sebagai landasan dalam proses penelitian.

BAB III METODE PENELITIAN

Berisi tinjauan umum tentang objek penelitian, alur penelitian, serta alat dan bahan yang digunakan dalam penelitian.

BAB IV HASIL DAN PEMBAHASAN

Berisi tahapan pembahasan dan hasil dari penelitian mengenai analisis sentimen yang telah dilakukan.

BAB V PENUTUP

Berisi kesimpulan dan saran dari penelitian yang telah dirangkum selama proses penelitian.