

BAB I

PENDAHULUAN

1.1. Latar Belakang

Teknologi informasi saat ini digunakan dalam beberapa bidang salah satunya adalah pariwisata. Teknologi informasi berperan penting dalam memajukan industri pariwisata, karena memberikan kemudahan akses informasi, kemudahan pelayanan bagi wisatawan, dan aspek lain yang bermanfaat bagi wisatawan[1]. Dalam bidang pariwisata, *google review* yang merupakan salah satu fitur dari platform *google maps* menjadi salah satu aspek penting dalam era big data untuk mengumpulkan jejak-jejak wisatawan ke suatu tempat[2]. Para wisatawan saat ini dapat dengan mudah untuk mencari referensi mengenai obyek wisata melalui platform ini.

Kota Sumenep merupakan salah satu daerah yang memiliki berbagai macam obyek wisata salah satunya adalah Pantai Sembilan. Pantai yang berada di pulau Giligenting ini sudah berada sejak tahun 2017 dan sudah menjadi icon wisata dari kota Sumenep. Setiap tahunnya ada 6000 pengunjung lebih yang mengunjungi obyek wisata ini[3]. Banyaknya pengunjung yang datang, tidak sedikit yang memberikan ulasan atau opini mengenai tempat wisata ini salah satunya pada aplikasi *google maps*. Ulasan ini menjadi hal yang sangat penting untuk diperhatikan, karena para calon pengunjung sebelum memutuskan untuk berkunjung ke sebuah tempat wisata, maka akan mencari tahu informasi terlebih dahulu tentang tempat wisata tersebut dari berbagai platform salah satunya di ulasan *google maps*[4]. Informasi yang diperoleh dari ulasan pengunjung yang mengunjungi obyek wisata ini juga dapat menjadi masukan bagi peningkatan layanan dan pengembangan obyek wisata[1]. Pengolahan informasi mengenai sentimen wisatawan ini dapat dilakukan dengan analisis sentimen.

Analisis sentimen atau yang sering disebut *opinion mining* merupakan studi komputasi untuk mengenali serta mengekspresikan sikap, opini, pendapat, emosi, subjektivitas, pandangan dan penilaian seseorang yang dilihat terhadap apa yang ditulis[5] atau juga bisa disebut sebagai proses analisis opini masyarakat yang diperoleh dari berbagai platform media sosial maupun internet[6].

Pada penelitian terdahulu yang dilakukan oleh mahasiswa Universitas Indonesia dengan membandingkan metode *Naïve Bayes* dan *KNN* dalam menganalisis sentimen kepuasan pelanggan *digital payment* di Indonesia, maka diperoleh hasil akurasi dengan metode *KNN* lebih tinggi dibandingkan dengan metode *Naïve Bayes*[7]. Pada penelitian yang lainnya juga yang membandingkan *KNN* dengan *Decision Tree C4.5*, didapatkan hasil akurasi metode *KNN* lebih tinggi dibandingkan *Algoritma C4.5*[8]. Hal ini menunjukkan bahwa metode *KNN* cukup baik dalam mengklasifikasikan data teks dibandingkan dengan *Algoritma C4.5* dan *Naïve Bayes*. Namun metode *KNN* mempunyai kekurangan jika bekerja dengan jumlah fitur yang cukup banyak[9].

Pada proses klasifikasi, banyaknya jumlah fitur dapat menyebabkan penurunan performa sehingga dibutuhkan proses seleksi fitur untuk mengurangi jumlah fitur. Beberapa metode yang dapat digunakan dalam seleksi fitur yaitu seleksi fitur *Information Gain*, seleksi fitur *Chi Square*, dan masih banyak lagi yang lainnya. Pada penelitian terdahulu yang membandingkan antara seleksi fitur *Information gain* dengan seleksi fitur *Chi square* dalam menganalisis sentimen masyarakat terkait bitcoin di twitter maka didapatkan seleksi fitur *Information Gain* lebih baik dalam meningkatkan akurasi dibandingkan dengan seleksi fitur *Chi square*[10].

Berdasarkan dari latar belakang tersebut, maka penulis mengajukan penelitian dengan judul “Analisis sentimen pada ulasan Pantai Sembilan Sumenep dengan menggunakan *K-Nearest Neighbor* dan seleksi fitur *Information Gain*”

1.2. Rumusan Masalah

1.2.1. Permasalahan

Sulitnya dalam melakukan klasifikasi opini masyarakat terhadap Pantai Sembilan secara manual dengan data dan jumlah fitur yang banyak.

1.2.2. Metode Usulan

Proses klasifikasi menggunakan metode *K-Nearest Neighbor* dengan seleksi fitur *Information Gain* sehingga menghasilkan akurasi yang terbaik sebagai acuan tingkat kepuasan masyarakat terhadap Pantai Sembilan.

1.2.3. Pertanyaan Penelitian

Berapa akurasi terbaik yang didapatkan metode *K-Nearest Neighbor* dengan seleksi fitur *Information Gain* dalam mengklasifikasikan sentimen masyarakat terhadap Pantai Sembilan.

1.3. Tujuan & Manfaat

1.3.1. Tujuan Penelitian

Mengetahui respon masyarakat terhadap Pantai Sembilan menggunakan *Information Gain* sebagai seleksi fitur dengan metode *K-Nearest Neighbor* sehingga dapat mengklasifikasikan ulasan ke dalam ulasan positif dan negatif.

1.3.2. Manfaat Penelitian

Adapun manfaat yang ingin dicapai dari penelitian ini adalah:

1. Penelitian ini memiliki manfaat untuk mengetahui ulasan masyarakat terhadap Pantai Sembilan di *google maps* berupa ulasan positif atau negatif agar dijadikan sebagai acuan untuk bahan evaluasi bagi pengelola wisata untuk pengembangan wisata dan sebagai referensi wisata bagi para calon pengunjung.
2. Penelitian ini memiliki manfaat untuk mengetahui seberapa besar pengaruh seleksi fitur *Information Gain* terhadap metode *KNN* dalam melakukan klasifikasi sentimen masyarakat terhadap Pantai Sembilan.

1.3.3. Batasan Masalah

1. Data ulasan diambil pada tanggal 3 November 2022 dari *google maps* berdasarkan kata kunci “Pantai Sembilan”
2. Data ulasan yang digunakan adalah ulasan yang berbahasa Indonesia.
3. Hasil klasifikasi ada 2 kelas yaitu positif dan negatif.
4. Nilai *threshold information gain* yang digunakan adalah 0,0008.

1.3.4. Sistematika Proposal

Sistematika penulisan proposal pada penelitian ini adalah sebagai berikut:

1. BAB I PENDAHULUAN

Pada bab ini berisi penjelasan tentang latar belakang, permasalahan, tujuan penelitian, batasan masalah, dan sistematika proposal

2. BAB II TINJAUAN PUSTAKA

Pada bab ini membahas tentang teori-teori yang berkaitan dengan penelitian ini

3. BAB III METODE USULAN

Pada bab ini menjelaskan proses atau langkah-langkah yang digunakan untuk menyelesaikan permasalahan menggunakan metode yang sudah ditentukan.