

BAB I PENDAHULUAN

I.1 Latar Belakang

Indonesia mendapatkan sumber devisa yang tinggi salah satunya karena adanya pariwisata (Soputan, Kumenaung, and Kawung 2022a). Objek Wisata yang menjadi tujuan destinasi para wisatawan berpengaruh terhadap banyak tidaknya wisatawan yang datang ke tempat wisata. Setiap objek wisata mempunyai daya tarik sendiri dan sangat beragam baik di darat maupun di laut, sehingga pengelolaan terhadap daerah wisata harus menjadi prioritas, terutama dalam sumber daya alam. Hal tersebut menjadikan daya tarik bagi pengelola wisata untuk memanfaatkan sebagai objek wisata yang diminati masyarakat.

Wisata yang berkunjung ke Indonesia dibagi menjadi 2, yaitu wisatawan mancanegara dan wisatawan domestik. Di mana wisatawan mancanegara merupakan wisatawan yang melakukan perjalanan wisata ke luar negrinya sendiri, sedangkan wisatawan domestik merupakan wisatawan yang melakukan perjalanan wisata dalam negaranya sendiri (Soputan, Kumenaung, and Kawung 2022b). Data statistika wisatawan nusantara berdasarkan Badan Pusat Statistika (BPS) mengurutkan Provinsi mana saja yang paling banyak dikunjungi oleh wisatawan nusantara berdasarkan data tersebut terdapat 5 urutan provinsi teratas di Indonesia pada tahun 2021, yaitu Jawa Timur, Jawa Tengah, Jawa Barat, Banten dan DKI Jakarta (BPS 2022).



Gambar I-1 Provinsi Tujuan Wisata Nusantara Terbanyak Tahun 2021

Teknologi informasi berperan penting dalam kemajuan sektor pariwisata terutama untuk mengakses informasi layanan bagi para wisatawan dan aspek-aspek penunjang lainnya. Wisatawan dapat mencari referensi mengenai objek wisata ataupun menuliskan komentar melalui platform media social salah satunya *Google My Business*. Merupakan sebuah *tools* yang disediakan oleh *Google* untuk mempermudah wisatawan mengetahui informasi wisata (Yuliani and Novita 2019). Berdasarkan pengalaman wisatawan yang sudah berkunjung ke tempat wisata, unggahan dari wisatawan dapat mempengaruhi minat wisatawan lain untuk berkunjung pada wisata tersebut.

Dari sekian banyak pengunjung yang datang, tidak sedikit dari mereka meninggalkan kesan berupa komentar atau ulasan terhadap tempat yang telah dikunjunginya. Ulasan ini bisa menjadi informasi yang sangat penting bagi pengelola wisata maupun wisatawan lainnya (Sibyan and Hasanah 2022a). informasi tersebut sebagai bahan evaluasi bagi pengelola wisata sedangkan bagi wisatawan bisa menjadi bahan pertimbangan atau referensi sebelum berkunjung ke tempat wisata. Sehingga diperlukannya analisis sentimen untuk mengatasi banyaknya komentar tersebut. Wisatawan dapat mengetahui apakah objek wisata tersebut termasuk sentiment positif, negatif atau netral. Multidimensional

digunakan untuk menggantikan data pada dimensi daerah dan waktu berguna untuk memahami variabilitas yang ada dalam sentimen.

Metode yang digunakan untuk mengambil *sampel* adalah *simple random sampling*. *Sampel* data terdiri dari komentar-komentar yang diberikan oleh wisatawan tentang objek wisata di 5 provinsi. Kelima provinsi ini dipilih karena provinsi tersebut memiliki jumlah kunjungan wisatawan terbanyak selama tahun 2021, yang mencakup sekitar 78% dari total kunjungan. Dari kelima provinsi ini, kami akan memilih sepuluh objek wisata unggulan yang mewakili berbagai jenis objek wisata, termasuk wisata alam, wisata buatan, wisata religi, dan jenis-jenis wisata lainnya. Pengumpulan data akan dilakukan selama tiga bulan, yaitu pada Januari, Februari, dan Maret 2023. Dengan menggunakan analisis sentimen dilakukan dengan menggunakan suatu metode algoritma yang mempermudah melakukan penelitian ini.

Beberapa algoritma yang umum digunakan untuk melakukan klasifikasi pada data teks, salah satunya adalah metode K-Nearest Neighbor (K-NN). K-NN adalah salah satu algoritma dalam machine learning yang digunakan untuk melakukan klasifikasi berdasarkan kedekatan atau jarak antara data. Dalam algoritma ini, data diklasifikasikan berdasarkan seberapa dekatnya dengan data lainnya. Algoritma K-NN melakukan klasifikasi dengan mencari sejumlah k tetangga terdekat dari data yang akan diklasifikasikan, kemudian mengambil kelas mayoritas dari tetangga-tetangga tersebut. Salah satu keunggulan dari algoritma K-NN adalah performanya yang baik dan hasil klasifikasinya yang akurat pada data teks (Purnamawati, Winarto, and Mailasari 2023).

Berdasarkan penelitian terdahulu oleh Sibyan dan Hasan, telah melakukan penelitian, yaitu mengidentifikasi dan mengklasifikasikan ulasan pengunjung wisata dieng dengan menggunakan algoritma K-Nearest Neighbor (K-NN) akurasi sebesar 86% (Sibyan and Hasanah 2022b). selain itu, Cherlina Helean P Panjaitan, Lisda Julian Pangaribuan dan Catra Indra Cahyadi melakukan penelitian menggunakan algoritma K-Nearest Neighbor (K-NN) dan aplikasi Rapidminer dengan nilai akurasi 83,33%, Hasil analisis membuktikan bahwa

metode algoritma K-Nearest Neighbor (K-NN) yang digunakan telah mampu sebagai acuan untuk merekomendasi tempat wisata kepada wisatawan di Labuan Bajo (Panjaitan, Pangaribuan, and Cahyadi 2022).

Dengan dasar dari ulasan literatur yang disajikan, peneliti memutuskan untuk menerapkan algoritma K-Nearest Neighbor (K-NN) dalam analisis sentimen melalui platform Google My Business. Berdasarkan konteks yang telah dijelaskan sebelumnya, fokus dari tugas akhir ini adalah untuk mendeteksi sentimen yang terkait dengan objek wisata di 10 lokasi wisata yang terdapat dalam 5 provinsi berbeda. Penilaian sentimen ini didasarkan pada ulasan-ulasan yang diberikan oleh pengguna pada platform Google My Business. Metode yang akan digunakan dalam analisis ini adalah K-NN, yang merupakan algoritma klasifikasi yang memanfaatkan konsep jarak antara data.

Data komentar yang ada pada *Google My Business* digunakan untuk memperoleh hasil representasi umum wisatawan terhadap objek wisata yang ada pada masing-masing provinsi. Hasil dari representasi tersebut digunakan untuk klasifikasi sentimen pada penilaian pengelola wisata sebagai evaluasi dari metode algoritma K-Nearest Neighbor (K-NN) serta harapannya hasil tersebut dapat meningkatkan pelayanan pada objek wisata di Provinsi Jawa Timur, Jawa Tengah, Jawa Barat, Banten dan DKI Jakarta sedangkan bagi wisatawan sebagai referensi tempat wisata.

I.2 Perumusan Masalah

Berdasarkan latar belakang diatas, maka rumusan permasalahan untuk tugas akhir ini adalah sebagai berikut:

1. Bagaimana cara mengimplementasikan analisis sentimen dari sisi multidimensional pada platform *Google My Business* terhadap objek wisata di Provinsi Jawa Timur, Jawa Tengah, Jawa Barat, Banten dan DKI Jakarta dengan menggunakan metode Algoritma K-Nearest Neighbor (K-NN)?

2. Bagaimana perbandingan sentimen pada setiap dimensi dalam multidimensional analisis berdasarkan metode algoritma K-Nearest Neighbor (K-NN)?
3. Bagaimana hasil performa metode algoritma K-Nearest Neighbor (K-NN) pada analisis sentimen berdasarkan komentar wisatawan terhadap objek wisata di *Google My Business*?

I.3 Tujuan Tugas Akhir

Tugas akhir ini bertujuan untuk:

1. Mengimplementasikan analisis sentimen dari sisi multidimensional pada *platform Google My Business* terhadap objek wisata di Provinsi Jawa Timur, Jawa Tengah, Jawa Barat, Banten dan DKI Jakarta dengan menggunakan metode algoritma K-Nearest Neighbor (K-NN).
2. Membandingkan sentimen pada setiap dimensi dalam *multidimensional* analisis berdasarkan metode algoritma K-Nearest Neighbor (K-NN).
3. Mengetahui hasil performa metode algoritma K-Nearest Neighbor (K-NN) pada analisis sentimen berdasarkan komentar wisatawan terhadap objek wisata di *Google My Business*.

I.4 Batasan Tugas Akhir

Batasan tugas akhir ini adalah sebagai berikut:

1. Data yang digunakan hanya data komentar dari wisatawan pada *Google My Business* dengan rentang waktu bulan Januari 2023 sampai dengan Maret 2023 pada 5 besar Provinsi di Indonesia dengan jumlah perjalanan wisatawan paling banyak.
2. Analisis sentimen ini hanya menggunakan *platform Google My Business* dan metode algoritma K-Nearest Neighbor (K-NN).
3. Analisis *multidimensional* menggunakan dimensi daerah yaitu Jawa Timur, Jawa Tengah, Jawa Barat, Banten dan DKI Jakarta serta dimensi waktu yang digunakan dalam satuan bulanan.

I.5 Manfaat Tugas Akhir

Manfaat tugas akhir ini:

1. Bagi Universitas Telkom, penelitian ini dapat berguna dalam memajukan efektivitas proses bisnis yang ada sehingga hubungan antara mahasiswa dan dosen menjadi lebih produktif dalam bidang akademis.
2. Bagi peneliti, dapat mengubah data mentah menjadi informasi menggunakan analisis sentimen untuk mendapatkan akurasi dari metode algoritma K-Nearest Neighbor (K-NN)
3. Bagi Dinas Pariwisata, penelitian ini dapat berguna dalam mengembangkan strategi pemasaran objek wisata serta meningkatkan pelayanan.
4. Bagi peneliti selanjutnya, penelitian ini berguna untuk menambah wawasan baru serta dapat menjadi bahan acuan penelitian selanjutnya yang berguna dengan analisis sentimen objek wisata dalam *multidimensional* analisis.

I.6 Sistematika Penulisan

Tugas akhir ini diuraikan dengan sistematika penulisan sebagai berikut:

Bab I Pendahuluan

Pada bab ini berisi uraian mengenai konteks permasalahan, latar belakang permasalahan, perumusan masalah yang diambil dari latar belakang permasalahan, tujuan penelitian, manfaat yang didapatkan dari penelitian, dan batasan masalah dari ruang lingkup penelitian.

Bab II Tinjauan Pustaka

Bab ini berisi literatur yang relevan dengan permasalahan yang diambil dan dibahas pula hasil-hasil referensi buku/ penelitian/ referensi lainnya yang terdiri atas konsep analisis sentimen, pengambilan data, algoritma yang digunakan hingga pengukuran dan evaluasi berupa *confusion matrix*.

Bab III Metodologi Penelitian

Bab ini menjelaskan metode / konsep / kerangka kerja yang dipilih untuk penyelesaian masalah mulai dari proses pengumpulan data, tahap implementasi hingga proses evaluasi hasil analisis yang telah dilakukan.

Bab IV Analisis dan Perancangan

Bab ini menjelaskan proses perancangan mulai dari proses pengambilan data, pengolahan data, analisis data, hingga visualisasi data.

Bab V Hasil dan Evaluasi

Pada bab ini, disajikan hasil dari implementasi algoritma dan analisis sentimen dengan data berupa teks yang berasal dari *Google My Business*. Hasil implementasi selanjutnya dievaluasi dengan *confusion matrix*. Selanjutnya akan ditampilkan hasil visualisasi menggunakan bantuan *tools Power BI*.

Bab VI Kesimpulan dan Saran

Pada bab ini dijelaskan kesimpulan dari penyelesaian masalah yang dilakukan serta jawaban dari rumusan permasalahan yang ada pada bagian pendahuluan. Saran dari solusi sebagai bahan perbaikan untuk penelitian selanjutnya.