

ABSTRAK

Abstrak - Objek wisata adalah tempat atau lokasi yang menarik orang untuk mengunjungi dan menikmati tempat tersebut dengan keindahan alam dan budayanya. Saat ini teknologi informasi semakin berkembang dan media sosial pun semakin digunakan oleh banyak orang untuk berkomunikasi dengan banyak orang atau berbagi pengalaman yang telah dilaluinya dan termasuk dalam mencari tentang objek wisata. Sekarang media sosial bukan hanya digunakan oleh individu tetapi organisasi pariwisata untuk berbagi informasi. Salah satu yang memungkinkan pengunjung memberikan komentar atau ulasan terkait suatu objek wisata pada *Google My Business*.

Di *Google My Business*, fasilitas dari *Google* dan sudah terhubung dengan *Google Maps*. Layanan ini menyediakan informasi tentang suatu tempat, menyediakan rating dan menyediakan fasilitas untuk berkomentar atau mengulas tempat yang sudah dikunjungi. Komentar atau ulasan dari pengunjung sangat banyak, sehingga diperlukan waktu yang sangat lama untuk mengulas satu per satu komentar. Oleh karena itu, diperlukannya analisis sentimen untuk mengelompokkan komentar yang mengandung ulasan positif, netral dan negatif. Analisis sentiment adalah proses untuk menganalisis penilaian, sentiment, pandangan, perilaku yang berkaitan dengan produk, jasa, organisasi, objek aktivitas tertentu dan lain sebagainya.

Analisis sentimen membantu dalam memahami apakah ulasan atau komentar yang diberikan oleh pengunjung terkait suatu topik tertentu bersifat positif, netral dan negatif. Penelitian ini menggunakan analisis sentiment multidimensional pada tempat dan waktu terhadap objek wisata di Provinsi Jawa Timur, Jawa Tengah, Jawa Barat, Banten dan DKI Jakarta dengan rentang waktu januari, februari, dan maret 2023 menggunakan algoritma K-Nearest Neighbor (KNN). Pengambilan data dibantu dengan tools DataMiner, pada penelitian ini menggunakan suatu library yang dapat melakukan pelabelan secara otomatis. Data akan dilakukan pelabelan data dengan 2 library python untuk proses *Natural Language Processing* (NLP) yaitu pelabelan transformers dan pelabelan textblob serta dilakukan perbandingan pada rasio 70:30, 80:20 dan 90:10.

Implementasi analisis sentiment dengan klasifikasi algoritma K-Nearest Neighbor (KNN). Metode ini melakukan perhitungan dengan pencarian tetangga terdekat, Ketika menerima data baru yang ingin diklasifikasikan. Dengan mencari nilai K dan menghitung jarak *Euclidean* atau jarak *Manhattan*. Metode ini menggunakan *GridSearch* untuk mengkombinasi parameter yang baik dan menghasilkan performa model yang akurat dan optimal. Pengujian ini mendapatkan hasil akurasi, hasil akurasi yang paling tinggi yaitu transformers dengan rasio 80:20 (80% data training dan 20% data testing).

Dalam penelitian ini, untuk dapat mengevaluasi kinerja model dan mendapatkan pemahaman yang lebih detail hasil perbandingan dari klasifikasi yang dilakukan oleh library (prediksi) dengan hasil klasifikasi yang sebenarnya menggunakan confusion matrix. Tabel confusion matrix berguna untuk menganalisis performa model dalam mengklasifikasi data dengan membandingkan 4 nilai yaitu true positive, false positive, true negative, dan false negative.

Dalam penelitian ini, model yang dibuat memiliki akurasi 78% dengan rata-rata *precision* sebesar 65%, *recall* sebesar 49%, dan *f1-score* sebesar 51%. Selanjutnya, hasil prediksi divisualisasikan menggunakan Power BI. Power BI adalah alat yang digunakan untuk menggambarkan data secara visual, seperti menampilkan total data, menampilkan pembagian positif negative netral, pie chart, diagram batang, peta, dan tabel, kemudia menerapkannya dalam visualisasi yang menarik dan interatif. Informasi data yang ditampilkan dalam berbagai bentuk grafik, chart, tabel, *card*, *multi-row card*, *pie card*, *stacked bar card*, peta agar mempermudah pemahaman, pengambilan keputusan dan menjadi tampilan yang mudah dipahami.

Kata kunci - [Analisis sentimen, Analisis Sentimen Multidimensional, Transformers, TextBlob, K-Nearest Neighbor, Confusion Matrix, Power BI]