ANALISIS SENTIMEN TEMPAT WISATA DI KABUPATEN BANDUNG JAWA BARAT

1st Ridwan Ramadhan
School of Electrical Engineering
Telkom University
Bandung, Indonesia
ridwanr@student.telkomuniversity.ac.id

2nd Rita Purnamasari
School of Electrical Engineering
Telkom University
Bandung, Indonesia
ritapurnamasari@telkomuniversity.ac.id

3rd Efri Suhartono
School of Electrical Engineering
Telkom University
Bandung, Indonesia
esuhartono@telkomuniversity.ac.id

***Abstrak* — Kabupaten Bandung memiliki sektor pariwisata yang menjadi salah satu pilar utama ekonomi daerah, dengan kekayaan objek wisata alam dan budaya yang menarik minat wisatawan domestik maupun mancanegara. Di sisi lain, wisatawan sangat bergantung pada ulasan di Google Maps sebagai referensi utama, namun informasi yang tersedia sering kali tidak terstruktur dan kualitasnya beragam, sehingga menyulitkan pengelola pariwisata dalam pengambilan keputusan. Untuk mengatasinya penelitian ini mengembangkan sistem analisis sentimen berbasis machine learning yang secara otomatis mengumpulkan, memproses, dan menganalisis ulasan wisatawan dari Google Maps. Sistem ini menerapkan tiga algoritma klasifikasi (Naive Bayes, SVM, dan K-Nearest Neighbors) dengan serangkaian pre-processing teks mencakup pembersihan teks, tokenisasi, penghapusan *stopword*, dan *stemming*. Fitur teks diekstraksi menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF), kemudian ulasan diklasifikasikan ke dalam kategori sentimen positif, negatif, atau netral. Hasil evaluasi menunjukkan model SVM memberikan kinerja terbaik dengan akurasi 88%, diikuti oleh Naive Bayes (85%) dan K-NN (67%). Penelitian ini membuktikan bahwa analisis sentimen dapat memberikan wawasan bagi pengelola destinasi wisata dalam memahami persepsi pengunjung dan mendukung pengambilan kebijakan yang baik untuk meningkatkan kualitas layanan pariwisata di Kabupaten Bandung.**

***Kata kunci* — analisis sentimen, google maps, klasifikasi teks, *machine learning*, pariwisata**

I. PENDAHULUAN

Pariwisata Kabupaten Bandung di Jawa Barat berperan sebagai salah satu sektor utama penopang perekonomian daerah. Dengan kekayaan objek wisata alam, budaya, dan kuliner, wilayah ini berpotensi menarik wisatawan domestik maupun mancanegara [1]. Peningkatan jumlah kunjungan wisata pasca pandemi tercatat signifikan pada tahun 2023 kunjungan wisatawan diperkirakan mencapai 7,04 juta orang, naik dari 6,55 juta di tahun 2022 (data Dinas Kebudayaan dan Pariwisata). Tren pertumbuhan positif tersebut dibarengi berbagai tantangan yang memengaruhi kualitas pengalaman wisata, seperti keterbatasan infrastruktur, aksesibilitas yang kurang memadai, kerusakan lingkungan, hingga adanya ulasan negatif di media digital yang dapat merusak citra destinasi. Meskipun potensi wisatanya besar, para pemangku kepentingan (pemerintah daerah dan pengelola destinasi) masih kesulitan mengelola sektor pariwisata secara efektif, terutama dalam memahami persepsi wisatawan secara menyeluruh. Wisatawan sendiri kerap mengalami kesulitan menentukan destinasi yang sesuai minat dan kurang memahami sentimen pengunjung lain terhadap tempat wisata yang ada, karena minimnya informasi yang lengkap, terpercaya, dan terstruktur mengenai destinasi tersebut. Akibatnya, upaya peningkatan kualitas layanan dan pengembangan infrastruktur sering tidak optimal. Selain itu, belum ada standar evaluasi objektif terhadap kualitas destinasi, sehingga pengelola beranggapan telah memberikan yang terbaik tanpa acuan evaluasi yang jelas [2].

Di era digital saat ini, wisatawan sangat mengandalkan ulasan dan penilaian di platform online seperti Google Maps sebagai referensi utama sebelum mengunjungi suatu destinasi. Ironisnya, informasi berharga dari ulasan pengguna tersebut kerap tidak terorganisir, bervariasi kualitasnya, dan tidak mudah diakses secara *real-time*. Data opini wisatawan tersebar di berbagai platform digital tanpa ada integrasi, menyulitkan pemerintah daerah dan pengelola wisata untuk memanfaatkannya secara cepat sebagai dasar kebijakan atau perbaikan layanan [3]. Hingga kini belum tersedia sistem terintegrasi yang mampu menghimpun, mengolah, dan memvisualisasikan ulasan-ulasan tersebut, sehingga banyak wawasan penting tentang pengalaman dan kepuasan pengunjung tidak terserap ke dalam proses pengambilan keputusan. Kondisi ini menyebabkan prioritas perbaikan destinasi lebih sering ditentukan berdasarkan pertimbangan administratif daripada umpan balik langsung wisatawan. Dengan kata lain, peluang untuk memahami dan menanggapi pandangan wisatawan secara optimal masih belum dimanfaatkan sepenuhnya. Analisis sentimen telah terbukti efektif untuk merangkum opini pengguna skala besar dan membantu pengambil kebijakan meningkatkan kualitas layanan. Studi terkini pada ulasan aplikasi publik menunjukkan bahwa pendekatan pembelajaran mesin sederhana namun terukur, seperti *Naive Bayes Multinomial* dengan fitur TF-IDF, mampu mencapai kinerja tinggi pada data ulasan yang di *scraping* dari platform resmi pada kasus ulasan Google Play Store [4].

Analisis sentimen berbasis data ulasan online menawarkan pendekatan untuk mengatasi masalah di atas. Teknik ini dapat mengungkap pola persepsi publik terhadap destinasi wisata secara lebih jelas, bahkan untuk aspek-aspek yang sulit diidentifikasi dengan metode konvensional. Penelitian sebelumnya telah membuktikan potensi analisis sentimen dalam konteks pariwisata. Misalnya, algoritma *Support Vector Machine* (SVM) telah diterapkan untuk klasifikasi sentimen ulasan destinasi wisata di Google Maps, sementara *Naive Bayes* berhasil digunakan untuk menganalisis sentimen ulasan objek wisata pantai [5]. Hasil-hasil tersebut menunjukkan bahwa teknik *machine*

learning mampu mengklasifikasikan sentimen wisatawan dengan akurasi tinggi, sehingga dapat membantu mengidentifikasi kekuatan dan kelemahan suatu destinasi. Analisis sentimen dapat menjadi dasar perumusan strategi pengembangan pariwisata yang lebih efektif, misalnya dengan meningkatkan daya tarik destinasi, memperbaiki kepuasan pengunjung, dan mendorong pertumbuhan ekonomi lokal melalui prinsip pariwisata berkelanjutan [5].

Penelitian ini bertujuan untuk mengembangkan sebuah sistem analisis sentimen otomatis yang mampu menangkap nuansa sentimen wisatawan secara akurat dan relevan dalam lingkup pariwisata Kabupaten Bandung. Sistem yang diusulkan dirancang untuk mengumpulkan ulasan wisatawan dari Google Maps, melakukan prapemrosesan data teks, melakukan ekstraksi fitur, dan mengklasifikasikan sentimen (positif, negatif, netral) menggunakan algoritma *machine learning* [2]. Dengan adanya sistem ini, informasi tak terstruktur dari ulasan wisatawan dapat diubah menjadi insight yang bernilai bagi para pemangku kepentingan pariwisata. Hasil analisis sentimen diharapkan dapat membantu pemerintah daerah dan pengelola destinasi dalam mengambil keputusan berbasis data untuk meningkatkan kualitas layanan wisata sesuai kebutuhan dan ekspektasi pengunjung.

II. KAJIAN TEORI

A. *Data Collection*

Google Maps dipilih sebagai sumber utama karena menyediakan ulasan berbasis teks yang mudah dikategorikan dan divalidasi untuk analisis sentimen. Pengambilan data dilakukan otomatis melalui web scraping menggunakan ekstensi *Instant Data Scraper* di Chrome dan software *scraping* lainnya menggunakan *stack* C# (.NET) berbasis CEF dengan pustaka EPPlus, HtmlAgilityPack, ExcavatorSharp, CEFSharp, RestSharp, Newtonsoft.JSON, dan log4net. Sehingga ekstraksi ulasan dan penyimpanan ke CSV/Excel jauh lebih efisien dibanding cara manual. Dari sisi cakupan dan kualitas, Google Maps unggul atas Tripadvisor maupun media sosial. Destinasi populer umumnya memiliki ribuan ulasan, sedangkan platform pembanding hanya ratusan atau bahkan puluhan komentar relevan setelah penyaringan. Karena ditulis langsung oleh pengunjung di lokasi, ulasannya cenderung autentik dan kontekstual, sehingga *noise* lebih rendah [7]. Untuk menjaga kemutakhiran sekaligus menangkap pola musiman, rentang data dibatasi satu tahun terakhir dengan kriteria inklusi berbasis lokasi (Kabupaten Bandung) serta ambang minimal jumlah ulasan per destinasi. Kombinasi horizon waktu, seleksi destinasi, dan *pipeline* pembersihan data menghasilkan korpus yang representatif untuk klasifikasi sentimen [8].

B. *Data Preprocessing*

Pra-pemrosesan data teks merupakan tahap krusial untuk menjamin keberhasilan analisis sentimen. Tahap ini mencakup serangkaian langkah sistematis sebelum data ulasan dianalisis. Pembersihan teks (text cleaning) dilakukan dengan menghapus karakter khusus, tanda baca, angka, serta mengubah semua huruf menjadi huruf kecil (*case folding*) agar format data tetap konsisten [3]. Selanjutnya, tokenisasi memecah kalimat menjadi unit-unit kata (token) berdasarkan delimiter seperti spasi atau tanda baca. Setelah itu dilakukan penghapusan *stopword*, yakni eliminasi kata-kata umum (misalnya “yang”, “dan”, “di”) yang tidak memiliki makna signifikan terhadap sentimen. Terakhir, *stemming* atau *lemmatisasi* mengubah kata berimbuhan ke bentuk dasar (kata root) untuk menyatukan variasi kata yang sebenarnya memiliki arti sama. Kombinasi langkah-langkah preprocessing tersebut memastikan data ulasan dalam kondisi bersih, rapi dan 9iap diproses lebih lanjut. misalnya ekstraksi fitur berbasis TF-IDF sebagai representasi numerik sebelum klasifikasi [9].

C. *Feature Extraction*

Term Frequency–Inverse Document Frequency (TF-IDF) adalah metode feature extraction yang digunakan untuk mengubah data teks menjadi representasi numerik berdasarkan tingkat kepentingan setiap kata dalam korpus dokumen. Term Frequency–Inverse Document Frequency memberikan bobot lebih tinggi pada kata yang sering muncul dalam sebuah dokumen (term frequency tinggi) namun jarang muncul di keseluruhan koleksi dokumen (inverse document frequency tinggi). Intuitifnya, TF-IDF menurunkan bobot kata-kata umum yang muncul di banyak dokumen, dan menekankan kata-kata khas yang relatif unik pada dokumen tertentu. Rumusan TF-IDF secara umum adalah perkalian antara frekuensi term pada dokumen dengan log kebalikan frekuensi term tersebut di seluruh dokumen. Hasilnya adalah matriks fitur berupa vektor nilai TF-IDF untuk setiap ulasan, yang siap digunakan sebagai input algoritma klasifikasi [9]. Dengan TF-IDF, sistem dapat mengenali kata-kata kunci yang paling memengaruhi sentimen dalam ulasan wisata.

D. *Data Labelling*

Bidirectional Encoder Representations from Transformers (BERT) telah membawa perubahan dalam bidang pemrosesan bahasa alami dengan kemampuan pemahaman konteks bidireksional yang superior. Model BERT menggunakan arsitektur transformer dengan mekanisme attention untuk menghasilkan representasi kontekstual yang kaya makna dari teks input [10]. Model transformer memungkinkan pemrosesan data sekuensial secara paralel, yang secara signifikan meningkatkan retensi konteks dan kemampuan menangani kompleksitas tugas seperti terjemahan serta ringkasan teks dapat ditangani lebih efektif. Selain itu, *Large Language Models* (LLMs) berbasis transformer menunjukkan kemampuan luar biasa dalam memahami kompleksitas dan nuansa bahasa manusia, meskipun masih menghadapi isu bias, tantangan interpretabilitas, dan tingginya biaya komputasi.

E. *Naive Bayes*

Naive Bayes merupakan algoritma klasifikasi yang didasarkan pada Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen satu sama lain. Pada konteks analisis sentimen teks, biasanya digunakan varian Multinomial Naive Bayes yang cocok untuk data diskrit seperti frekuensi kata. Algoritma ini menghitung probabilitas sebuah ulasan termasuk ke kelas sentimen tertentu (misal positif atau negatif) berdasarkan probabilitas prior kelas dan probabilitas likelihood dari kata-kata dalam ulasan tersebut. Praktik tersebut banyak digunakan pada korpus ulasan aplikasi dan menjadi *baseline* kuat sebelum membandingkan dengan SVM atau K-NN [4]. Meskipun asumsi independensi fitur jarang terpenuhi sempurna, pendekatan Naive Bayes sering memberikan hasil cukup baik karena kesederhanaan model mampu menghindari overfitting pada data teks berukuran besar. Naive Bayes tergolong efisien dan telah banyak digunakan dalam klasifikasi teks. Dalam kasus analisis sentimen wisata, algoritma ini belajar dari pola kemunculan

kata pada ulasan berlabel untuk memperkirakan sentimen ulasan baru secara cepat [11].

F. *Support Vector Machine* (SVM)

SVM adalah algoritma *machine learning* supervisi yang sangat populer untuk tugas klasifikasi teks. SVM berusaha mencari *hyperplane* optimal yang memisahkan data ke dalam dua kelas dengan margin terlebar, sehingga pemisahan antar kelas menjadi sejelas mungkin. Keunggulan SVM terletak pada kemampuannya menangani data berdimensi tinggi dan menghasilkan performa yang stabil. Penelitian terdahulu menunjukkan SVM mampu mencapai akurasi tinggi dalam tugas analisis sentimen, mengungguli metode tradisional seperti Naive Bayes [12]. Dalam tugas klasifikasi sentimen wisata, SVM memetakan ulasan (yang telah diubah menjadi vektor fitur TF-IDF) ke ruang berdimensi tinggi, lalu memisahkannya berdasarkan sentimen dengan memaksimalkan margin antar kelas. Pemilihan parameter seperti jenis *kernel* (linear, RBF, dll.), nilai regularisasi (*C*), dan parameter lain memengaruhi kinerja SVM dan biasanya ditentukan melalui validasi. SVM dikenal unggul untuk data teks berukuran besar karena kemampuannya mencari *decision boundary* yang optimal [12].

G. *K-Nearest Neighbor* (K-NN)

K-Nearest Neighbors merupakan algoritma klasifikasi non-parametrik berbasis instance. Metode ini menentukan kelas suatu data baru dengan melihat mayoritas kelas dari *k* tetangga terdekatnya dalam ruang fitur. Dalam analisis sentimen, setiap ulasan direpresentasikan sebagai vektor TF-IDF, lalu jarak (misalnya Euclidean) antara vektor ulasan baru dan vektor ulasan-ulasan dalam data latih dihitung. Kelas sentimen yang paling dominan di antara *k* ulasan terdekat tersebut akan dipilih sebagai prediksi kelas ulasan baru [13]. Parameter penting K-NN adalah nilai *k* (jumlah tetangga) dan metrik jarak yang digunakan. K-NN relatif mudah diimplementasikan namun kurang efektif untuk data berdimensi tinggi seperti teks, karena perhitungannya sangat bergantung pada metrik jarak [14]. Dalam konteks ulasan wisata yang fitur TF-IDF-nya berukuran besar, kinerja K-NN dapat menurun akibat *curse of dimensionality*. Meskipun beberapa studi melaporkan K-NN mampu mencapai akurasi di atas 80% dengan pemilihan parameter yang tepat, secara umum performanya masih berada di bawah SVM untuk tugas klasifikasi sentimen [13]. Oleh sebab itu, K-NN sering dijadikan pembanding *baseline* dalam penelitian ini.

H. *K-Means Clustering*

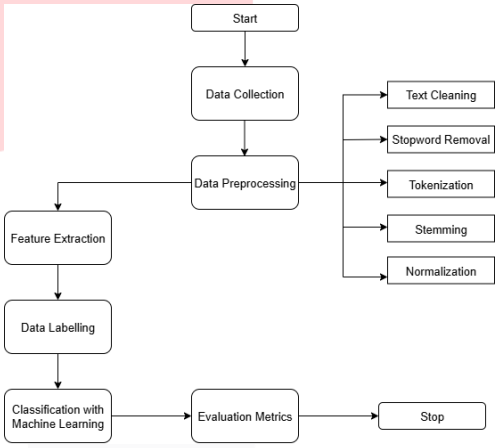
K-Means clustering merupakan metode pengelompokan data non-hierarkis yang efektif untuk membagi data ke dalam kelompok-kelompok berdasarkan karakteristik tertentu. Dalam analisis sentimen pariwisata, *K-Means* terbukti berguna untuk mengelompokkan ulasan berdasarkan pola sentimen dan mengidentifikasi segmen wisatawan dengan preferensi yang berbeda [15]. Penelitian terbaru menunjukkan bahwa *K-Means clustering* dapat menghasilkan wawasan berharga tentang preferensi wisatawan terhadap atribut destinasi, seperti proporsi nilai unik (*unique selling proposition*) dan citra merek (*brand image*) [15]. Penerapan *K-Means clustering* dalam analisis data pariwisata juga memungkinkan identifikasi pola perilaku dan preferensi wisatawan yang berkaitan dengan demografi, pola perjalanan, dan pilihan destinasi.

I. *Evaluation Performation*

Evaluasi model klasifikasi umumnya mengandalkan confusion matrix untuk membandingkan prediksi dengan label aktual; dari matriks ini dihitung akurasi, precision, recall, dan F1-score. *Precision* mengukur proporsi prediksi positif yang benar, *recall* menilai kemampuan model menangkap seluruh kasus positif yang relevan, sedangkan F1 menyeimbangkan keduanya [16]. Pada data tidak seimbang, evaluasi perlu dibarengi penanganan seperti resampling (oversampling/undersampling) atau pemberian bobot kelas agar model tak bias pada kelas mayoritas; dalam praktiknya, SMOTE kerap dipakai untuk memperbanyak contoh dari kelas minoritas [17].

III. METODE

Tahapan yang dilakukan pada penelitian ini meliputi *data collection*, *data preprocessing*, *feature extraction*, *data labelling*, pelatihan model *Machine Learning*, dan *Evaluation Performation*, sebagaimana digambarkan pada rancangan sistem pada Gambar 1.



Gambar 1. Desain Sistem

A. *Data Collection*

Dataset yang digunakan dalam penelitian ini dikumpulkan dari ulasan pengguna Google Maps terhadap objek wisata berfokus pada tempat wisata yang berlokasi di Kabupaten Bandung, Jawa Barat. Data ulasan dikumpulkan dari platform Google Maps dengan rentang waktu satu tahun terakhir agar mencerminkan kondisi terkini dan mengakomodasi variasi musiman. Proses pengumpulan data dilakukan secara otomatis menggunakan teknik *web scraping*. Alat bantu ekstensi Chrome *Instant Data Scraper* digunakan untuk mengekstraksi konten ulasan dari halaman Google Maps. Data yang diperoleh mencakup teks ulasan, nama objek wisata, dan metadata waktu ulasan. Data tersebut berisi ulasan berbentuk teks yang ditulis oleh pengunjung mengenai berbagai destinasi seperti taman alam, museum, dan tempat rekreasi. Semua data ulasan kemudian disimpan dalam bentuk terstruktur (format CSV) untuk keperluan analisis ke tahap lanjutan. Setiap ulasan diklasifikasikan secara manual ke dalam sektor tertentu berdasarkan konten ulasan untuk mendukung analisis sentimen berbasis aspek.

B. *Data Preprocessing*

Sebelum masuk ke tahap klasifikasi, dataset ulasan berlabel tersebut melalui tahap pre-processing teks seperti yang telah dijelaskan pada bagian kajian teori.

- *Text cleaning*: Proses mengubah seluruh teks menjadi huruf kecil (lowercase) dan setiap ulasan dibersihkan dari karakter aneh (emoji, simbol, URL). Langkah ini bertujuan untuk menormalkan variasi bentuk kata yang

disebabkan oleh perbedaan penggunaan huruf kapital, baik di awal, tengah, maupun akhir kata.

- Penghapusan Stopword (*Stopword Removal*): *Stopword* adalah kata-kata umum yang dianggap. (contoh: *yang*, *di*, *ke*, *dari*, *untuk*, dll.)
- *Tokenization*: Komentar pengguna yang telah melalui proses cleaning dan case folding dipecah menjadi kata-kata tunggal melalui proses tokenisasi. Proses ini membantu mengekstrak setiap kata dari sebuah kalimat untuk dianalisis lebih lanjut. Sebagai contoh, kalimat ulasan "Pelayanannya kurang memuaskan dan tempatnya kotor" akan diubah menjadi token kata dasar seperti ["layan", "kurang", "puas", "tempat", "kotor"].
- *Stemming*: *Stemming* adalah proses mengubah kata menjadi bentuk dasarnya dengan menghapus imbuhan.
- *Normalization*: Proses ini melibatkan perbaikan bahasa informal atau penghapusan huruf berulang, seperti mengubah “bangettt” menjadi “banget”

Proses ini mengurangi keragaman fitur kata tanpa menghilangkan esensi informasi, sehingga mempermudah model dalam mempelajari pola.

C. *Feature Extraction*

Setiap ulasan direpresentasikan sebagai vektor fitur berukuran *N* ((jumlah kata unik yang dipilih pada korpus setelah prapemrosesan). Nilai setiap dimensi vektor adalah bobot TF-IDF dari kata tertentu dalam ulasan tersebut. Untuk sebuah kata *t* pada dokumen *d*, nilai TF-IDF dihitung menggunakan persamaan berikut secara umum:

$$TF - IDF(t,d) = tf(t,d)\log(df(t)N) \quad (1)$$

Dengan keterangan:

• tf(t,d) adalah frekuensi kemunculan term *t*

• df(t) adalah jumlah dokumen yang mengandung term *t*.

• N adalah jumlah total dokumen.

Dengan transformasi TF-IDF ini, teks ulasan yang semula unstructured berubah menjadi data numerik terstruktur siap untuk modeling.

D. *Data Labelling*

Karena data ulasan yang diperoleh tidak memiliki label sentimen (tidak ada keterangan langsung apakah sebuah ulasan bernada positif, netral, atau negatif), maka dilakukan proses pelabelan data secara otomatis dengan pendekatan *lexicon based*. Penelitian ini memanfaatkan kamus lexicon sentiment yang berisi daftar kata-kata positif dan negatif beserta skor polaritasnya. Setiap kalimat ulasan dianalisis dengan menghitung skor total polaritas, Kata positif menambah skor, sedangkan kata negatif mengurangi skor. Jika hasil penjumlahan skor suatu ulasan bernilai positif (>=0), maka ulasan tersebut diberi label positif, jika bernilai negatif (<0) maka diberi label negatif.

Ulasan yang memiliki skor netral atau tidak mengandung kata berpolaritas signifikan dianggap netral. Dengan metode ini, seluruh dataset ulasan berhasil dilabeli ke dalam tiga kelas sentimen (positif, netral, negatif) secara otomatis. pendekatan berbasis lexicon memiliki keterbatasan akurasi, namun digunakan sebagai langkah awal untuk menghasilkan *silver standard* dataset yang kemudian dilatih ulang menggunakan algoritma *machine learning*.

E. *Classification with Naïve Bayes*

Dalam penelitian ini, *Naive Bayes* (NB) digunakan sebagai *baseline* karena ringan, cepat, dan efektif pada data teks berdimensi tinggi. Varian yang dipakai adalah Multinomial NB dengan *smoothing* (Laplace) agar peluang kata langka tetap terdefinisi. Algoritma ini memanfaatkan probabilitas bersyarat berdasarkan asumsi independensi antar fitur, sehingga cocok untuk pengolahan data teks dengan jumlah fitur yang besar. Penyesuaian prior kelas mengikuti proporsi data membantu mengurangi bias ketika distribusi sentimen tidak seimbang (positif biasanya dominan). Pada korpus ulasan, model *Naive Bayes* diimplementasikan dengan menggunakan bobot fitur hasil ekstraksi TF-IDF untuk menentukan kelas sentimen pada setiap ulasan. Persamaan dasar yang digunakan dalam *Naive Bayes* adalah sebagai berikut:

$$\log P(y|x) \propto \log P(y) + \sum_{i=1}^{n} x_i \log P(x_i|y) \quad (2)$$

Keterangan:

- P(y|x): probabilitas suatu kelas y diberikan data fitur x.
- P(y): probabilitas awal (*prior probability*) kelas y.
- $P(x_i|y)$: probabilitas fitur x_i muncul dalam kelas y (*likelihood*).
- x_i: nilai fitur ke-i
- n : jumlah fitur dalam dataset.

Dengan konfigurasi tersebut, NB memberikan *baseline* yang ringan, cepat, dan cukup akurat untuk teks ulasan, serta menjadi pembanding yang baik terhadap SVM dan K-NN pada studi ini.

F. *Classification with* SVM

Pada tugas klasifikasi sentimen, *Support Vector Machine* (SVM) dipilih karena kinerjanya yang stabil pada ruang fitur berdimensi tinggi dan sparce seperti TF-IDF, serta ketahanannya terhadap *overfitting* melalui kontrol margin dan regularisasi. Implementasi menggunakan kernel *linear* (LinearSVC) yang secara empiris paling sesuai untuk teks, dikombinasikan dengan skema *one-vs-rest* (OvR) untuk tiga kelas sentimen. SVM bekerja dengan cara mencari *hyperplane* optimal yang memisahkan titik data dari kelas yang berbeda dengan margin maksimum. Dalam formulasi standarnya (*hard-margin*), permasalahan optimasi didefinisikan sebagai berikut:

$$f(x) = sign(w \cdot x + b) \quad (3)$$

Dengan keterangan:

- w adalah *weight vector* (vektor bobot)
- x adalah *feature vector* (vektor fitur)
- b adalah *bias*

G. *Classification with* K-NN

Dalam penelitian ini, algoritma K-Nearest Neighbor (KNN) juga digunakan sebagai metode klasifikasi sentimen. KNN bekerja dengan mengklasifikasikan data uji berdasarkan kedekatannya dengan data latih terdekat dalam ruang fitur. Jarak antar dokumen dihitung menggunakan

Euclidean, dengan vektor telah dinormalisasi (L2) agar perbandingan antar ulasan adil meskipun panjang kalimat berbeda, sehingga data uji akan diberi label sesuai mayoritas kelas dari tetangga terdekatnya. K-NN mudah diimplementasikan, namun pada ruang fitur yang sangat tinggi dan sparce ia rawan terkna efek *curse of dimensionality* sehingga jarak menjadi kurang informatif dan kinerjanya kurang baik. Jarak Euclidean antara dua titik data x dan y dengan n fitur didefinisikan sebagai:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(xi - yi)2} \qquad (4)$$

Dimana:

- x_i dan y_i masing-masing adalah nilai fitur ke-i dari titik data x dan y.
- n adalah jumlah fitur.

Pada penelitian ini, nilai k ditentukan berdasarkan eksperimen untuk mendapatkan performa klasifikasi terbaik.

H. *Evaluation Metrics*

Setelah model dilatih, dilakukan evaluasi kinerja menggunakan data uji. Metrik evaluasi yang dianalisis mencakup akurasi, precision, recall, dan F1-score untuk setiap algoritma. Evaluasi dilakukan menggunakan metrik sebagai berikut:

- *Accuracy*: proporsi prediksi yang benar terhadap seluruh data. (mengukur persentase prediksi sentimen yang tepat)
- *Precision*: tingkat ketepatan prediksi untuk setiap label.
- *Recall*: kemampuan model dalam menangkap data yang sebenarnya untuk setiap label.
- *F1-Score*: rata-rata antara presisi dan *recall*. Sebagai indikator keseimbangan kinerja model terutama pada data yang kelasnya tidak seimbang.
- *Confusion Matrix*: digunakan untuk mengamati kesalahan klasifikasi antar label atau menilai distribusi prediksi model (benar atau salah per kelas)

IV. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penelitian analisis sentimen beserta interpretasinya terkait berbagai aspek layanan pariwisata. Analisis dilakukan pada kumpulan ulasan pengguna yang diperoleh dari sumber-sumber terkait pariwisata. Proses pelatihan (*training*) dilakukan dengan memasukkan data ulasan yang telah dipreproses dan difeaturisasi (TF-IDF) beserta label sentimennya ke algoritma masing-masing. Proporsi pembagian data latih dan uji ditetapkan sehingga model dapat dievaluasi performanya pada data yang belum pernah dilihat. Dalam hal ini, sekitar 80% data digunakan sebagai training set, sedangkan 20% sisanya sebagai testing set. Dengan total dataset yang mencapai puluhan ribu ulasan, data uji berjumlah cukup besar (sekitar 39 ribu ulasan) untuk memperoleh estimasi performa yang reliabel. Ulasan-ulasan tersebut dikategorikan ke dalam beberapa sektor yang telah ditentukan sebelumnya. Sentimen dari setiap ulasan diklasifikasikan ke dalam tiga kategori yaitu negatif, netral, dan positif.

Selama pelatihan, parameter model dioptimalkan menggunakan pengaturan default yang umum digunakan. Pada model *Naive Bayes Multinomial*, digunakan nilai *smoothing alpha* = 1 (*Laplace smoothing*) untuk menghindari probabilitas nol. Pada model SVM (LinearSVC), parameter C (*regularization*) diatur sebesar 1.0 yang merupakan default, dengan kernel linear. Sedangkan untuk model K-NN, nilai K (jumlah tetangga) dipilih 7 berdasarkan eksperimen awal mengacu pada kisaran optimal [13]. Metrik jarak Euclidean digunakan dalam menghitung kedekatan antar ulasan pada model K-NN, dengan data fitur yang dinormalisasi.

A. *Naïve Bayes Evaluation Metrics*

Pada bagian ini juga, dilakukan evaluasi kinerja algoritma Naive Bayes berdasarkan metrik *accuracy*, *precision*, *recall*, dan *f1-score*.

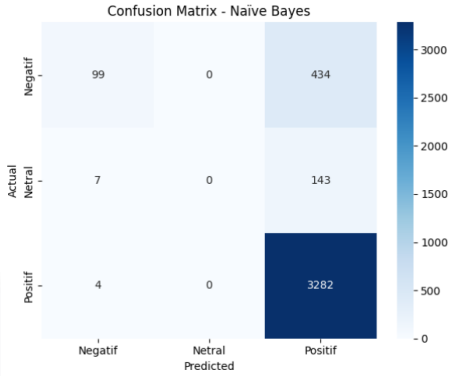
```
Akurasi Model Naïve Bayes: 0.85
Model Precision: 0.82
Model Recall   : 0.85
Model F1-Score : 0.80
              precision    recall  f1-score   support

     Negatif       0.90      0.19      0.31       533
      Netral       0.00      0.00      0.00       150
     Positif       0.85      1.00      0.92      3286

    accuracy                           0.85      3969
   macro avg       0.58      0.39      0.41      3969
weighted avg       0.82      0.85      0.80      3969
```

Gambar 2. Evaluasi Model Naive Bayes

Visualisasi hasil klasifikasi ditampilkan dalam bentuk *confusion matrix* pada Gambar 3, yang menggambarkan distribusi prediksi model terhadap label sebenarnya.



Gambar 3. Confusion Matrix Naive Bayes

Pada model *Naive Bayes*, meskipun akurasinya cukup tinggi (85%), analisis lebih lanjut mengungkap adanya ketidakseimbangan performa antar kelas sentimen. NB sangat baik dalam mengenali ulasan positif, terbukti dengan *precision* dan *recall* kelas positif mencapai 0.85 dan 1.00 (nyaris semua ulasan positif terdeteksi benar). Namun, model ini kesulitan membedakan ulasan negatif dan netral. Dari *confusion matrix*, terpantau bahwa dari 533 ulasan berlabel negatif, hanya 99 ulasan (18.6%) yang berhasil dikenali sebagai negatif, sedangkan mayoritas (434 ulasan) justru salah diklasifikasikan sebagai positif.

Lebih parah lagi, tidak satu pun ulasan netral berhasil diprediksi benar oleh NB. Seluruh ulasan netral (150 data) diklasifikasikan sebagai positif. Akibat bias kuat ke kelas positif ini, *precision* kelas negatif hanya sekitar 0.19 dan netral 0.00, dengan *recall* negatif 0.19 dan netral 0.00. Hal ini menjelaskan mengapa *macro-average* F1 NB hanya 0.41 meskipun *weighted* F1 cukup tinggi (0.80), model sangat mengutamakan akurasi pada kelas dominan (positif) sehingga kinerja pada kelas minoritas (negatif, netral)

terabaikan. Bias seperti ini umum terjadi ketika distribusi kelas tidak seimbang dan model terlalu "percaya diri" pada pola mayoritas. Pada skenario nyata, kelemahan NB ini berarti model cenderung gagal mendeteksi adanya keluhan atau sentimen negatif jika proporsi ulasan positif jauh lebih banyak.

B. SVM *Evaluation Metrics*

Pada bagian ini, dilakukan evaluasi kinerja algoritma SVM berdasarkan metrik *accuracy*, *precision*, *recall*, dan *f1-score*.

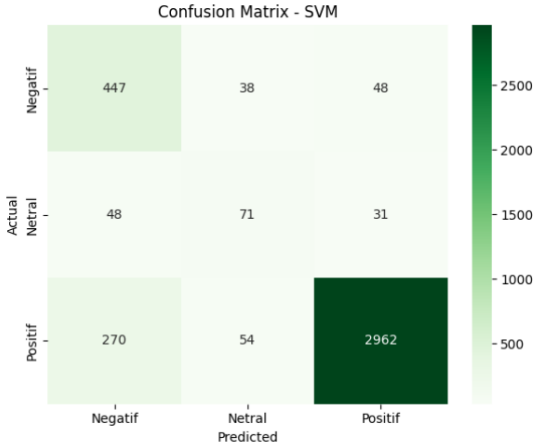
```
Evaluasi Model SVM:
Akurasi Model SVM: 0.88
Model Precision: 0.90
Model Recall   : 0.88
Model F1-Score : 0.88
              precision    recall  f1-score   support

     Negatif       0.58      0.84      0.69       533
      Netral       0.44      0.47      0.45       150
     Positif       0.97      0.90      0.94      3286

    accuracy                           0.88      3969
   macro avg       0.66      0.74      0.69      3969
weighted avg       0.90      0.88      0.88      3969
```

Gambar 4. Evaluasi Model SVM

Visualisasi hasil klasifikasi ditampilkan dalam bentuk *confusion matrix* pada Gambar 5, yang menggambarkan distribusi prediksi model terhadap label sebenarnya.



Gambar 5. *Confusion Matrix* SVM

SVM tampil sebagai model terbaik dengan kinerja paling seimbang. Dengan akurasi 88%, SVM tidak hanya unggul dalam nilai agregat, tetapi juga mampu mengenali ketiga kelas sentimen dengan lebih baik. Precision *weighted-average* SVM mencapai 0.90 dan recall 0.88, menandakan proporsi prediksi benar yang tinggi secara keseluruhan. Dilihat per kelas, SVM memiliki *precision* 0,97 dan *recall* 0,90 untuk kelas positif, sedikit lebih tinggi dari NB dalam mengenali ulasan positif, namun yang paling menonjol adalah perbaikan signifikan pada kelas negatif dan netral. *Precision* kelas negatif SVM sekitar 0,58 dengan *recall* 0,84, jauh lebih baik dibanding NB maupun K-NN, artinya model SVM berhasil menangkap sebagian besar ulasan negatif meskipun masih terdapat tidak sesuai (precision 58% menunjukkan ada prediksi negatif yang seharusnya bukan).

Untuk kelas netral, SVM juga menunjukkan kemampuan yang lebih baik daripada model lain denga *precision* 0,44 dan *recall* 0,47. Meskipun angka ini masih relatif rendah, setidaknya model SVM mampu mengenali hampir separuh ulasan netral dengan benar, berbanding nyaris 0% pada NB dan K-NN. *Macro F1-score* SVM tercatat 0.69, tertinggi di antara ketiga model, menandakan kinerja antar kelas yang lebih merata. Hasil ini menjadikan SVM sebagai model paling andal dalam tugas klasifikasi sentimen ulasan wisata pada penelitian ini.

C. K-NN Evaluation Metrics

Evaluasi kinerja algoritma K-NN berdasarkan metrik *accuracy*, *precision*, *recall*, dan *f1-score* sebagai berikut.

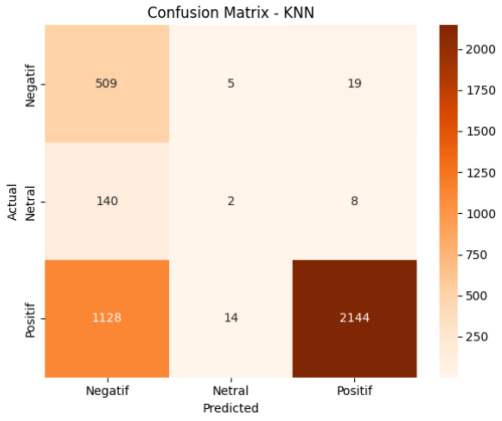
```
Evaluasi Model KNN:
Model Precision: 0.86
Model Recall   : 0.67
Model F1-Score : 0.71
Akurasi: 0.67
              precision    recall  f1-score   support

     Negatif       0.29      0.95      0.44       533
      Netral       0.10      0.01      0.02       150
     Positif       0.99      0.65      0.79      3286

    accuracy                           0.67      3969
   macro avg       0.46      0.54      0.42      3969
weighted avg       0.86      0.67      0.71      3969
```

Gambar 2. Evaluasi Model K-NN

Visualisasi hasil klasifikasi ditampilkan dalam bentuk *confusion matrix* pada Gambar 7, yang menggambarkan distribusi prediksi model terhadap label sebenarnya.



Gambar 3. Confusion Matrix K-NN

Model *K-Nearest Neighbors* juga menunjukkan pola bias serupa terhadap kelas mayoritas, meskipun dengan karakteristik yang sedikit berbeda. K-NN menghasilkan *precision* rata-rata yang cukup tinggi (0.86) karena kinerjanya pada kelas positif lumayan baik (*precision* 0.87, *recall* 0.95). Artinya, sebagian besar prediksi positif oleh K-NN memang benar-benar positif. Namun, *recall* kelas negatif dan netral tergolong rendah (sekitar 0.66 untuk negatif, dan mendekati 0 untuk netral), menunjukkan banyak ulasan negatif/netral gagal terdeteksi. *Confusion matrix* K-NN mengonfirmasi hal ini dari total 533 ulasan negatif, K-NN berhasil mengklasifikasikan benar 509 di antaranya (*recall* negatif 0.95, lebih baik dari NB), namun untuk ulasan netral, model hanya mengenali 2 dari 150 (nyaris semua netral dikira sebagai kelas lain).

Kesulitan K-NN dalam membedakan nuansa netral vs negatif disebabkan karena algoritma ini sangat bergantung pada kedekatan di ruang fitur. Jika ulasan netral memiliki kemiripan kata yang lebih dekat ke klaster ulasan positif, model akan cenderung memberi label positif. Hal ini terlihat dari cukup banyaknya *false positive* dimana ulasan seharusnya netral/negatif dikira positif (tercatat 1.128 ulasan yang sebenarnya negatif/netral salah diklasifikasi sebagai positif). *F1-score* K-NN secara keseluruhan (weighted) sekitar 0.71, lebih rendah dari NB, yang menandakan penurunan performa global. Meski demikian, *macro-average*

F1 K-NN (0.75) sedikit lebih tinggi daripada NB karena K-NN mampu mengenali lebih banyak ulasan negatif (*precision* pada negatif 0.87) dibanding NB.

D. Pembahasan

Hasil pengujian menunjukkan bahwa di antara ketiga algoritma yang dibandingkan, SVM memberikan performa paling unggul dalam mengklasifikasikan sentimen ulasan wisata. Model SVM mencapai nilai akurasi sebesar 88% pada data uji, sedikit lebih tinggi dibandingkan Naive Bayes yang meraih akurasi sekitar 85%, sedangkan K-NN tertinggal dengan akurasi di kisaran 67%. SVM unggul karena dapat memisahkan kelas dengan lebih jelas di ruang fitur TF-IDF berdimensi tinggi. Naive Bayes mendekati SVM pada akurasi namun gagal mengenali sentimen minoritas (netral/negatif) akibat asumsi independence yang sederhana sehingga pola minor kurang terdeteksi. K-NN tampil paling lemah karena rentan bias ke kelas dominan dan kinerjanya sangat dipengaruhi sebaran data (dalam kasus ini tidak seimbang).

V. KESIMPULAN

Penelitian ini telah menghasilkan sebuah sistem analisis sentimen berbasis machine learning yang mampu mengolah ulasan wisatawan dari Google Maps secara otomatis dan menyajikan hasilnya dalam bentuk informasi yang berguna bagi pengelola pariwisata. Dengan menerapkan algoritma SVM, *Naive Bayes*, dan K-NN serta teknik TF-IDF, sistem dapat mengklasifikasikan sentimen ulasan ke dalam kategori positif, netral, dan negatif dengan akurasi tertinggi dicapai oleh model SVM (88%), diikuti oleh *Naive Bayes* (85%) dan K-NN (67%). Hasil analisis pada studi kasus destinasi wisata di Kabupaten Bandung menunjukkan bahwa mayoritas wisatawan memiliki sentimen positif, namun terdapat pula masukan negatif yang konstruktif terkait aspek fasilitas dan layanan. Model SVM terbukti paling andal dan seimbang dalam mengenali ulasan positif, netral, maupun negatif, sementara K-NN cenderung bias ke sentimen positif dan kurang akurat membedakan ulasan negatif/netral.

Kontribusi utama dari penelitian ini adalah membuktikan bahwa teknik analisis sentimen dapat memberikan wawasan objektif dan berbasis data mengenai persepsi wisatawan. Informasi tersebut berperan penting sebagai landasan bagi pemerintah daerah dan pengelola destinasi dalam merumuskan kebijakan dan strategi peningkatan kualitas pariwisata yang lebih tepat sasaran. Untuk penelitian selanjutnya, sistem ini dapat dikembangkan dengan memasukkan analisis aspek (*aspect-based sentiment analysis*) untuk menggali opini wisatawan pada elemen spesifik destinasi, atau mengintegrasikan model bahasa terbaru (seperti BERT) untuk meningkatkan akurasi pemahaman konteks ulasan. Secara keseluruhan, penelitian ini berkontribusi pada pemanfaatan teknologi *machine learning* dan *data analytics* dalam mendukung pengambilan keputusan di sektor pariwisata, menuju pengelolaan destinasi wisata yang lebih cerdas dan berkelanjutan.

REFERENSI

[1] F. C. Manosso and T. C. D. Ruiz, "Using sentiment analysis in tourism research: A systematic, bibliometric, and integrative review," *Journal of Tourism, Heritage and Services Marketing*, vol. 7, no. 2, pp. 16–27, 2021, doi: 10.5281/zenodo.5548426.

[2] J. Ipmawati, S. Saifulloh, and K. Kusnawi, "Analisis Sentimen Tempat Wisata Berdasarkan Ulasan pada Google Maps Menggunakan Algoritma Support Vector Machine," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 247–256, 2024, doi: 10.57152/malcom.v4i1.1066.

[3] S. Jiang *et al.*, "Explainable Text Classification via Attentive and Targeted Mixing Data Augmentation," *in Proc. Int. Joint Conf. Artificial Intelligence* (IJCAI), 2023, pp. 5085–5094, doi: 10.24963/ijcai.2023/565.

[4] M. C. Pramuji, R. Purnamasari, Y. Eliskar, A. R. Thaha, "Analisis Sentimen Menggunakan Algoritma Naive Bayes Classifier Pada Ulasan Aplikasi PLN Mobile di Google Play Store," *e-Proceeding of Engineering*, vol. 11, no. 6, pp. 5700–5706, 2024.

[5] I. W. B. Suryawan, N. W. Utami, and K. Q. Fredlina, "Analisis Sentimen Review Wisatawan pada Objek Wisata Ubud Menggunakan Algoritma Support Vector Machine," *Jurnal Informatika Teknologi dan Sains*, vol. 5, no. 1, pp. 133–140, 2023.

[6] W. Rafdinal, "Is smart tourism technology important in predicting visiting tourism destination? Lessons from West Java, Indonesia," *Journal of Tourism Sustainability*, vol. 1, no. 2, pp. 102–115, 2021, doi: 10.35313/jtos.v1i2.20.

[7] F. Akbar, H. Hadiyanto, and C. E. Widodo, "Sentiment Analysis of Data on Google Maps Reviews Regarding Tourism on Keraton Kasepuhan Cirebon Using the Lexicon Based Method," in *Proc. Int. Conf. on Applied Information System and Data Analytics (ICAISD)*, 2024, pp. 19–24, doi: 10.5220/0012440100003848.

[8] J. P. Mellinas and M. Sicilia, "Comparing Google reviews and TripAdvisor to help researchers select the more appropriate information source," *Consumer Behavior in Tourism and Hospitality*, vol. 19, no. 4, pp. 646–655, 2024, doi: 10.1108/CBTH-01-2024-0039.

[9] N. A. Semary, W. Ahmed, K. Amin, P. Pławiak, and M. Hammad, "Enhancing machine learning-based sentiment analysis through feature extraction techniques," *PLOS ONE*, vol. 19, no. 2, e0294968, 2024, doi: 10.1371/journal.pone.0294968.

[10] A. A. Mudding, "Mengungkap Opini Publik: Pendekatan BERT-based-caused untuk Analisis Sentimen pada Komentar Film," *Journal of System and Computer Engineering (JSCE)*, vol. 5, no. 1, pp. 36–43, 2024, doi: 10.61628/jsce.v5i1.1060.

[11] N. Wulandari, Y. Cahyana, Rahmat, and H. H. Handayani, "Sentiment Analysis on the Relocation of the National Capital (IKN) on Social Media X Using Naive Bayes and K-Nearest Neighbor (KNN) Methods," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 3, pp. 724–731, Jun. 2025.

[12] F. Sulianta, Basic Data Mining From A to Z. Bandung, Indonesia. Jul. 2023, Accessed: Jun. 19, 2025. [Online]. Available:

https://www.researchgate.net/publication/382274667
[13] R. Kurniawan, H. O. L. Wijaya, and R. P. Aprisusanti, “Sentiment Analysis of Google Play Store User Reviews on Digital Population Identity App Using K-Nearest Neighbors,” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 13, no. 2, pp. 170–178, 2024, doi: 10.32736/sisfokom.v13i2.2071.
[14] M. Gunawan, M. Zarlis, and R. Roslina, “Analisis Komparasi Algoritma Naïve Bayes dan K-Nearest Neighbor Untuk Memprediksi Kelulusan Mahasiswa Tepat Waktu,” *Jurnal Media Informatika Budidarma*, vol. 5, no. 2, pp. 513–520, Apr. 2021, doi: 10.30865/mib.v5i2.2925.
[15] D. Chi, T. Huang, Z. Jia, and S. Zhang, “Research on sentiment analysis of hotel review text based on BERT-TCN-BiLSTM-attention model,” *Array*, vol. 25, no. February, p. 100378, 2025, doi: 10.1016/j.array.2025.100378.
[16] S. A. Rutba and S. Pramana, “Aspect-based Sentiment Analysis and Topic Modelling of International Media on Indonesia Tourism Sector Recovery,” *Indonesian Journal of Tourism and Leisure*, vol. 6, no. 1, 2025, doi: 10.36256/ijtl.v6i1.502.
[17] M. S. Syahlan *et al.*, “Analisis sentimen terhadap tempat wisata dari komentar pengunjung menggunakan metode Support Vector Machine (SVM) (studi kasus: Taman Air Mancur Sri Baduga Purwakarta),” *Jurnal Sistem Informasi dan Teknik Komputer*, vol. 8, No. 2, no. 2, pp. 315–319, Oct. 2023, Accessed: Jun. 19, 2025. [Online]. Available: https://ejournal.catursakti.ac.id/index.php/simtek/article/view/281/215