

Analisis Sentimen Terhadap Pengalaman Pengguna Di Tempat Wisata Melalui Google Bisnis Review: Studi Kasus (Telaga Sarangan)

1st Muhammad Arjun Kurniawan
Sistem Informasi Universitas Telkom
Purwokerto
Universitas Telkom Purwokerto
Purwokerto, Indonesia
arjunkurniawan@student.telkomuniversity.ac.id

2nd Khairun Nisa Meiah Ngafidin,
S.Pd., M.kom
Sistem Informasi Universitas Telkom
Purwokerto
Universitas Telkom Purwokerto
Purwokerto, Indonesia
nisameiah@telkomuniversity.ac.id

3rd Mahazam Afrad, S.Kom., M.Kom
Sistem Informasi Universitas Telkom
Purwokerto
Universitas Telkom Purwokerto
Purwokerto, Indonesia
mahazama@telkomuniversity.ac.id

Abstrak — Meningkatnya penggunaan platform digital seperti Google Maps membuat ulasan pengguna terhadap destinasi wisata menjadi sumber informasi penting. Namun, banyaknya ulasan yang tersebar secara acak dan tidak terstruktur menyebabkan informasi berharga tentang kepuasan pengunjung belum dimanfaatkan secara optimal oleh pengelola wisata. Telaga Sarangan, sebagai salah satu destinasi unggulan di Kabupaten Magetan, menerima banyak ulasan dari pengunjung yang berpotensi untuk dianalisis lebih lanjut. Penelitian ini bertujuan untuk mengklasifikasikan sentimen pengguna terhadap Telaga Sarangan melalui ulasan Google Maps menggunakan metode Support Vector Machine (SVM). Proses analisis dilakukan melalui tahapan text preprocessing, pelabelan sentimen (positif, negatif, netral), ekstraksi fitur menggunakan TF-IDF, penyeimbangan data menggunakan SMOTE, serta evaluasi dengan confusion matrix. Dataset yang dianalisis terdiri dari 919 ulasan. Hasil penelitian menunjukkan bahwa metode SVM memiliki performa yang baik dalam klasifikasi sentimen, dan berhasil mengidentifikasi faktor-faktor yang memengaruhi kepuasan serta ketidakpuasan pengunjung. Temuan ini diharapkan dapat menjadi bahan evaluasi strategis bagi pengelola wisata dalam meningkatkan kualitas layanan secara berkelanjutan.

Kata kunci— Analisis Sentimen, Telaga Sarangan, Google Maps, Support Vector Machine, TF-IDF, SMOTE

I. PENDAHULUAN

Berkembangnya teknologi informasi sekarang ini memudahkan seseorang dalam mengakses berbagai hal, di zaman sekarang ini ketersediaan data dan informasi sangatlah berlimpah. Data dapat diperoleh dari berbagai *platform* seperti *social media*, *search engine* maupun televisi. *Search engine* atau lebih dikenal sebagai mesin pencari, yang sering di manfaatkan dalam kegiatan sehari – hari yaitu Google terdapat berbagai layanan seperti Google Search, Google Foto, Google Maps dan masih banyak layanan lain yang terdapat di dalamnya. Pemanfaatan Google Maps dalam kegiatan sehari – hari berupa pencarian suatu tempat atau lokasi seperti restoran, tempat tinggal, kantor, pusat perbelanjaan dan masih banyak tempat – tempat yang lainnya, selain di gunakan untuk mencari lokasi Google Maps juga

dapat digunakan untuk memberi ulasan dari seseorang yang telah mengunjungi lokasi tersebut [1]. Berkembangnya teknologi informasi sekarang ini memudahkan seseorang dalam mengakses berbagai hal, di zaman sekarang ini ketersediaan data dan informasi sangatlah berlimpah. Data dapat diperoleh dari berbagai *platform* seperti *social media*, *search engine* maupun televisi. *Search engine* atau lebih dikenal sebagai mesin pencari, yang sering di manfaatkan dalam kegiatan sehari – hari yaitu Google terdapat berbagai layanan seperti Google Search, Google Foto, Google Maps dan masih banyak layanan lain yang terdapat di dalamnya. Pemanfaatan Google Maps dalam kegiatan sehari – hari berupa pencarian suatu tempat atau lokasi seperti restoran, tempat tinggal, kantor, pusat perbelanjaan dan masih banyak tempat – tempat yang lainnya, selain di gunakan untuk mencari lokasi Google Maps juga dapat digunakan untuk memberi ulasan dari seseorang yang telah mengunjungi lokasi tersebut [1]. Google Maps merupakan aplikasi panduan berbasis internet yang tersedia gratis milik Google. Google Maps dapat diakses melalui browser dari personal computer maupun ponsel. Google Maps diakses untuk mencari suatu alamat yang akan dituju atau hanya sekedar melihat tempat serta ulasan yang dimiliki [2]. Hal ini dapat membantu seseorang ketika akan mendatangi lokasi tersebut, khususnya pada lokasi – lokasi seperti tempat wisata, tak jarang seseorang perlu mempertimbangkan berbagai kebutuhan sebelum berkunjung ke tempat wisata seperti layanan, kebersihan, fasilitas dan lain sebagainya. Dengan adanya fitur ini membantu dalam mempertimbangkan kebutuhan di atas. Pada penelitian ini akan memanfaatkan ulasan pada Google Maps “Objek Wisata Telaga Sarangan” menggunakan analisis sentimen. Ulasan yang digunakan merupakan sebuah fitur dari Google yang bernama Google Business Review fitur ini berisikan ulasan yang diberikan oleh pengunjung ketika usai mengunjungi tempat tersebut terkadang setelah mengunjungi suatu tempat Google memberikan notifikasi untuk memberikan ulasan serta penilaian dari tempat yang baru saja atau telah dikunjungi [3].

Telaga Sarangan, juga dikenal sebagai Telaga Pasir adalah telaga alami yang berada di ketinggian 1.200 meter di atas permukaan laut dan terletak di lereng Gunung Lawu,

Kecamatan Plaosan, Magetan. Memiliki luas sekitar 30 hektare dan ke dalam 28 meter. Pada Telaga Sarangan terdapat rumah makan, tempat bermain, pasar wisata, taman dan berbagai wahana lainnya [4]. Berdasarkan wawancara Bersama Kepala Dinas Pariwisata Kabupaten Magetan yang diperoleh pada laman website [5] pengunjung Telaga Sarangan mencapai 23.900 pada h+3 lebaran 2024 jumlah ini merupakan jumlah terbanyak selama ini. Banyaknya wisatawan yang datang tentu menimbulkan beragam pengalaman dan persepsi, baik dalam hal layanan, sarana prasarana, maupun interaksi dengan petugas dan pedagang di sekitar lokasi. Namun demikian, hingga saat ini belum banyak dilakukan kajian sistematis untuk menggali dan mengelompokkan opini publik yang tersebar melalui platform digital seperti Google Maps. Berbagai ulasan dari pengunjung masih tersebar secara acak dan belum dianalisis secara terstruktur untuk memahami sejauh mana kepuasan atau ketidakpuasan mereka. Hal ini menunjukkan bahwa terdapat kesenjangan antara besarnya data opini pengguna yang tersedia secara daring dengan upaya pemanfaatannya untuk evaluasi dan perbaikan layanan wisata. Oleh karena itu, diperlukan pendekatan berbasis analisis sentimen guna mengklasifikasikan ulasan-ulasan tersebut ke dalam kategori positif maupun negatif secara otomatis dan objektif. Berdasarkan data pariwisata yang diperoleh dari website Badan Pusat Statistik Kabupaten Magetan tahun 2023, didapatkan jumlah pengunjung wisata di wilayah tersebut. Dari diagram tersebut dapat disimpulkan bahwa Telaga Sarangan merupakan objek wisata dengan jumlah pengunjung tertinggi, yakni lebih dari 200 ribu wisatawan sepanjang tahun 2023. Hal ini menunjukkan bahwa Telaga Sarangan memiliki daya tarik yang kuat dibandingkan destinasi wisata lainnya di Kabupaten Magetan. Penelitian ini menjadi penting untuk dilakukan mengingat Telaga Sarangan merupakan objek wisata dengan daya tarik paling tinggi di Kabupaten Magetan, sehingga pengelola perlu memahami secara mendalam persepsi dan kepuasan pengunjung sebagai dasar pengambilan keputusan. Di era digital saat ini, ulasan pengguna di platform seperti Google Maps dapat menjadi sumber informasi yang sangat berharga, namun masih belum dimanfaatkan secara maksimal. Melalui penerapan analisis sentimen berbasis metode machine learning seperti Support Vector Machine (SVM), penelitian ini berupaya mengelola dan menginterpretasikan opini pengunjung dalam skala besar secara efisien. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi praktis bagi pengelola wisata dalam meningkatkan kualitas layanan, serta menjadi referensi akademik untuk pengembangan kajian serupa pada destinasi wisata lainnya. Selain itu, penelitian ini juga mendorong pemanfaatan teknologi analisis teks dalam sektor pariwisata daerah, yang masih tergolong minim dan memiliki potensi pengembangan yang sangat besar. Oleh karena itu dengan banyaknya jumlah wisatawan tersebut pastinya menimbulkan beberapa pendapat yang timbul dari pengunjung.

[7]. Dalam penerapan analisis sentimen terdapat beberapa metode yang sering digunakan seperti K-NN, *Naïve Bayes*, *Random Forest*, dan *Support Vector Machine (SVM)* (Arsi & Waluyo, 2021). Dalam konteks pariwisata, analisis sentimen dapat mengidentifikasi faktor-faktor yang mempengaruhi kepuasan pengunjung ketika mendatangi destinasi wisata tersebut (Singgalen, 2022). Pemilihan metode dalam analisis sentimen merupakan hal yang penting

karena metode dalam analisis sentimen akan menentukan seberapa akurat hasil akan diperoleh.

Metode SVM dinilai memiliki tingkat akurasi lebih tinggi dibanding dengan metode yang populer lainnya (Ananda & Suryono, 2024). Dari penelitian (Giovani et al., 2020), algoritma SVM memiliki tingkat akurasi tertinggi dibandingkan dengan penggunaan algoritma lain seperti *Naive Bayes (NB)*, *K-Nearest Neighbour (K-NN)*, dan *feature selection*. Oleh karena itu, pemanfaatan metode SVM dinilai cocok untuk mengklasifikasikan kepuasan dan ketidakpuasan pengunjung wisata Telaga Sarangan.

Dengan beragamnya ulasan yang terdapat dalam Google Maps Telaga Sarangan, selain menentukan sentimen positif dan negatif, penelitian ini digunakan untuk mengklasifikasikan faktor-faktor kepuasan pengunjung objek wisata Telaga Sarangan. Penelitian ini memanfaatkan pengambilan data dari ulasan Google Maps dengan teknik *crawling data*. *Crawling data* adalah proses pengambilan data dari sebuah website dengan memanfaatkan API dari website tersebut (Duei Putri et al., 2022). Data yang diambil berkisar antara bulan Januari 2024 sampai dengan Juni 2024, hal ini dimaksudkan untuk mengetahui opini publik sebelum dan sesudah terjadinya lonjakan wisatawan pada lebaran 2024 yang mencapai 23.900 wisatawan. Dengan memanfaatkan analisis sentimen pada objek wisata Telaga Sarangan, ini berguna untuk menentukan seberapa besar ulasan positif dan negatif objek wisata Telaga Sarangan.

Lebih lanjut, dampak dari adanya ulasan positif dan negatif ini tidak hanya berpengaruh pada kemajuan destinasi wisata itu sendiri, tetapi juga dapat menjadi acuan bagi pemerintah daerah dan pengelola wisata dalam merumuskan kebijakan pengembangan pariwisata yang lebih tepat sasaran. Bagi masyarakat sekitar, meningkatnya kunjungan wisatawan akibat perbaikan layanan juga akan berdampak pada peningkatan pendapatan dan peluang usaha, sehingga secara keseluruhan dapat mendorong pertumbuhan ekonomi lokal.

Pada penelitian [8] dilakukan penerapan algoritma *Support Vector Machine (SVM)* pada dua transportasi online yaitu Gojek dan Grab untuk proses klasifikasi supaya mengetahui tingkat akurasi sentimen terhadap data yang digunakan. Hasil pengujian pada terhadap ke dua transportasi tersebut antara lain, Gojek memperoleh tingkat akurasi sebesar 87%, presisi 93% dan re-call sebesar 84%. Lalu untuk aplikasi Grab memperoleh nilai akurasi sebesar 87%, presisi 93% dan re-call sebesar 83%, pada penelitian ini didapatkan bahwa SVM merupakan algoritma yang paling baik dalam mengklasifikasikan teks.

II. KAJIAN TEORI

A. Analisis Sentimen

Analisis sentimen, yang juga dikenal sebagai opinion mining, adalah salah satu bidang penelitian dalam text mining. Proses ini melibatkan pengolahan data secara otomatis untuk memperoleh informasi. Analisis sentimen digunakan untuk mengidentifikasi pendapat terhadap suatu subjek atau objek, seperti individu, organisasi, produk, atau tempat, dalam kumpulan data tertentu. Jika ada sekumpulan dokumen yang memuat opini terkait suatu objek, tujuan dari opinion mining adalah untuk

mengidentifikasi atribut serta elemen dari objek yang diberikan komentar pada masing-masing dokumen, sekaligus menentukan apakah komentar tersebut mengandung sentimen positif, negatif, atau netral. [10]

B. Tempat Wisata

Tempat wisata merupakan suatu sumber daya alam yang bisa dimanfaatkan oleh suatu organisasi atau pemerintahan bahkan perseorangan (individu) guna menarik wisatawan baik lokal maupun mancanegara untuk datang atau berkunjung ke tempat tersebut sehingga memperoleh suatu keuntungan bagi pengelola tempat tersebut. Tempat wisata yang dikelola memiliki tarif sesuai dengan kebijakan dari organisasi atau individu. Suatu tempat wisata menyediakan berbagai macam layanan serta hiburan untuk mendapatkan kepuasan dari pengunjung tempat tersebut [11].

C. Telaga Sarangan

Telaga Sarangan merupakan salah satu destinasi wisata terpopuler milik pemerintah Kabupaten Magetan. Telaga Sarangan terletak di kaki Gunung Lawu, Kecamatan Plaosan, Kabupaten Magetan. Telaga Sarangan memiliki danau dan sebuah pulau kecil didalamnya, Telaga Sarangan disebut juga dengan telaga pasir. Telaga Sarangan memiliki luas sekitar 30 hektar dan kedalaman telaga sedalam 28 meter [13].

D. Text Preprocessing

Proses pengolahan teks dari dokumen atau kumpulan data mentah dikenal sebagai text preprocessing. Tahap ini bertujuan mengubah data teks yang awalnya tidak terorganisir atau acak menjadi bentuk data yang terstruktur dan rapi [14]. Pada penelitian ini *preprocessing* memiliki beberapa tahapan. Tahapan dari *preprocessing* diuraikan seperti berikut :

1. Cleaning adalah sebuah proses menghapus elemen yang tidak dibutuhkan seperti (.) (,) emoji dan simbol lain yang tidak berhubungan dengan kalimat
2. Case folding merupakan proses menyamakan jenis huruf dalam sebuah dokumen. Tujuan dari proses ini adalah untuk memudahkan pencarian. Karena tidak semua dokumen teks menggunakan huruf kapital secara konsisten, maka *case folding* diperlukan untuk mengubah seluruh teks dalam dokumen menjadi format standar, yaitu huruf kecil (*lowercase*).
4. Tokenizing adalah tahap memecah dokumen menjadi segmen-segmen kecil yang disebut token. Proses ini memecah suatu kalimat yang dipisahkan oleh spasi sehingga kalimat yang sudah terpecah akan berdiri sendiri sebagai suatu kata.
5. Stopword Removal merupakan proses menghapus kata-kata yang tidak memberikan kontribusi berarti terhadap makna utama dokumen. Kata-kata tersebut dihilangkan karena dapat mengganggu atau tidak memberikan manfaat optimal dalam proses *text mining*. Kata-kata yang termasuk ke dalam stopwords dihilangkan karena memberikan pengaruh yang tidak

baik. Contoh kata yang termasuk *stopword* antara lain “dan”, “saya”, “kamu”, “dengan”, “dia”, dan sebagainya.

6. Stemming merupakan tahapan yang dilakukan dengan cara menghapus afiks atau imbuhan dari setiap kata sehingga diperoleh bentuk kata dasar. Proses ini bertujuan untuk menstandarkan penulisan kata agar konsisten, termasuk melakukan pembersihan terhadap kata yang mengalami kesalahan ejaan pada tahap ini, sehingga hasil teks menjadi lebih siap untuk dianalisis pada proses text mining selanjutnya.

E. Weighting

Weighting pembobotan kata merupakan tahap pemberian bobot menggunakan metode *Term Frequency – Inverse Document Frequency* (TF-IDF), di mana setiap kata akan memperoleh nilai antara 0 hingga 1. Pada tahap ini, kata-kata yang telah berbentuk token dan melewati proses preprocessing akan diubah menjadi vektor yang merepresentasikan kata tersebut dalam bentuk numerik. *Term Frequency* (TF) menunjukkan seberapa sering suatu kata muncul dalam sebuah dokumen, sedangkan *Inverse Document Frequency* (IDF) digunakan untuk mengukur tingkat kepentingan kata tersebut dalam keseluruhan kumpulan dokumen. [16].

F. Support Vector Machine

SVM adalah algoritma *machine learning* berjenis *supervised learning* yang berbasis pada konsep vektor. Algoritma ini bekerja dengan mengklasifikasikan data ke dalam dua kelas menggunakan sebuah *hyperplane*. *Hyperplane* tersebut terletak di antara dua kelas dan memiliki jarak tertentu terhadap titik terdekat dari masing-masing kelas. Jarak ini dikenal dengan istilah *margin*, sedangkan titik-titik yang berada pada batas *margin* disebut *support vector*. Tujuan utama dari SVM adalah menemukan *hyperplane* yang optimal sehingga jaraknya terhadap titik terdekat dari kedua kelas tersebut menjadi maksimal. [1]. Ciri dari SVM adalah *low generation error*, mudah untuk diinterpretasikan, namun sensitif pada *tuning parameter* dan pilihan kernel.

G. Confusion Matrix

Confusion Matrix merupakan salah satu metode dalam evaluasi performa suatu model dalam bentuk tabel yang menampilkan hasil prediksi terhadap dua kelas. Adapun fungsi dari Confusion Matrix ini adalah mengukur tingkat akurasi dari suatu model dengan persamaan (Ipmawati et al., 2024).

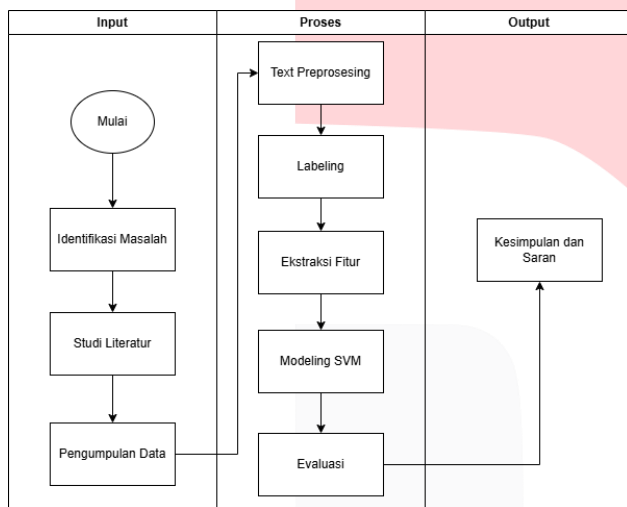
TABEL 1
CONFUSION MATRIX

		Kelas Asli	
		Kelas – 1	Kelas – 2
Kelas Prediksi	Kelas - 1	True Positive (TP)	False Negative (FN)
	Kelas - 2	False Positive (FP)	True Negatif (TN)

III. METODE

A. Sistematika Penyelesaian Masalah

Penyelesaian masalah digambarkan dalam diagram alir penelitian berisi langkah-langkah dari penyelesaian penelitian “ANALISIS SENTIMEN TERHADAP PENGALAMAN PENGGUNA DI TEMPAT WISATA MELALUI GOOGLE BISNIS REVIEW : STUDI KASUS (TELAGA SARANGAN)” mulai dari identifikasi masalah, melakukan studi literatur, proses pengumpulan data yang dilanjutkan dengan melakukan Handling Duplication, Labeling, Text Prerosecing, Weighting, kemudian permodelan menggunakan algoritma *Support Vector Machine* serta membuat kesimpulan dan saran. Penjelasan lebih lanjut alur proses penelitian yang akan dilakukan, dapat diamati secara lebih rinci melalui Gambar 1



GAMBAR 1
DIAGRAM ALIR PENELITIAN

a. Identifikasi Masalah

Proses ini dilakukan pada para pengunjung objek wisata Telaga Sarangan. Hal ini dilakukan untuk mengetahui faktor dari kepuasan pengunjung objek wisata Telaga Sarangan yang di tinjau dari ulasan Google Maps. Sering kali pada suatu objek wisata ditemukan berbagai permasalahan mengenai kepuasan pengunjung baik dari sarana dan prasarana, layanan yang diberikan bahkan perilaku pedagang pada objek wisata tersebut apakah berbuat curang dengan menaikkan harga kepada pengunjung. Oleh karena itu, analisis sentimen penting untuk mengetahui sebaik mana pengalaman pengunjung pada objek wisata Telaga Sarangan. Pada tahap ini, juga dimulai penelusuran referensi literatur penelitian sebelumnya yang relevan dengan penelitian yang akan dilakukan

b. Studi Literatur

Langkah ini dilakukan proses pengumpulan referensi literatur yang serupa dan relevan dengan penelitian yang akan

dilakukan yaitu literatur yang berkaitan dengan topik analisis sentimen. Literatur tersebut antara lain seperti mengenai machine learning, text mining, Telaga Sarangan, scraping data, Support Vector Machine, serta pengolahan data analisis sentimen.

c. Pengumpulan Data

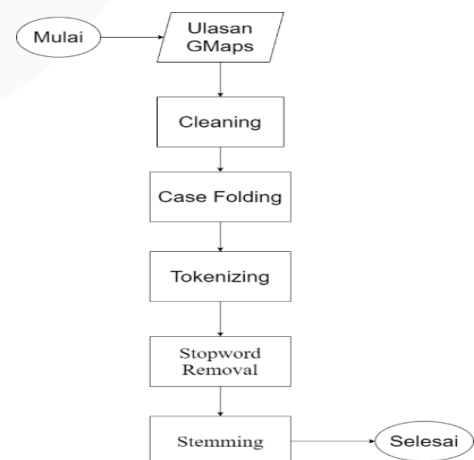
Pada proses ini pengumpulan data dibantu dengan Google Colaboratory dengan mengambil ulasan pada Google Maps Telaga Sarangan. Data didapatkan berupa hasil ulasan dari pengunjung objek wisata Telaga Sarangan dari berbagai Rating. Berikut hasil dari crawling data menggunakan phyton dari ulasan Google Maps Telaga Sarangan :

TABEL 2
HASIL CRAWLING

Nama	Rating	Review
Suachiyat. Sulaiman	3	Pemandangan indah, naik boat n kuda kesukaan anak2
AlFarid SFMyusic	3	Tempat penghilang penat paling rekomendasi 🙌

d. Text Preprocecing

Pada langkah *text preprocessing*, dilakukan proses pembersihan terhadap kalimat-kalimat dalam data teks dengan tujuan untuk menghilangkan elemen-elemen yang tidak relevan, menyederhanakan struktur kalimat, dan menyiapkan data agar lebih mudah dianalisis oleh algoritma pemrosesan bahasa alami. Proses ini penting dilakukan guna memastikan bahwa data yang digunakan dalam tahap analisis memiliki kualitas yang baik, terstruktur, dan dapat merepresentasikan makna yang sebenarnya dari opini atau ulasan yang diberikan oleh pengguna. Berikut merupakan langkah-langkah dari *text preprocessing*, antara lain:



GAMBAR 1
DIAGRAM ALIR PREPROCECING

- a. Cleaning digunakan untuk menghapus simbol, tagar, nama pengguna, emoji dan semua unsur yang tidak berkaitan. Berikut ini contoh dari hasil cleaning data

TABEL 1
TABEL HASIL CLEANING

Data Awal	Data sesudah cleaning
Tempat nya nyaman Sejuk. Tapi ke toilet aj suruh bayar 8 k □	Tempat nya nyaman Sejuk Tapi ke toilet aj suruh bayar 8 k

- b. Case folding digunakan untuk mengubah setiap kata menjadi huruf kecil semua. Berikut ini contoh dari hasil case folding

TABEL 2
TABEL HASIL CASEFOLDING

Data Awal	Data sesudah case folding
Tempat nya nyaman Sejuk Tapi ke toilet aj suruh bayar 8 k	tempat nya nyaman sejuk tapi ke toilet aj suruh bayar 8 k

- c. Tokenizing digunakan untuk memisahkan kalimat menjadi per kata. Berikut hasil dari Tokenizing

TABEL 3
TABEL HASIL TOKENIZING

Data Awal	Data sesudah Tokenizing
tempat nya nyaman sejuk tapi ke toilet aj suruh bayar 8 k	["tempat" "nya" "nyaman" "sejuk" "tapi" "ke" "toilet" "aj" "suruh" "bayar" "8" "k"]

- d. Stopworld Removal berperan dalam mengeliminasi kata-kata yang tidak memiliki makna penting. Berikut hasil dari Stopworld Removal

TABEL 4
TABEL HASIL STOPWORD

Data Awal	Data sesudah Stopworld Removal
["tempat" "nya" "nyaman" "sejuk" "tapi" "ke" "toilet" "aj" "suruh" "bayar" "8" "k"]	["tempat" "nyaman" "sejuk" "tapi" "toilet" "suruh" "bayar"]

- e. Stemming untuk memperbaiki kata dengan mengembalikannya ke bentuk dasarnya, dengan menghapus imbuhan kata dibagian depan dan belakang setiap kata. Berikut hasil dari Stemming

TABEL 5
TABEL HASIL STEMMING

Data Awal	Data sesudah Stemming
["tempat" "nyaman" "sejuk" "tapi" "toilet" "suruh" "bayar"]	["tempat" "nyaman" "sejuk" "tapi" "toilet" "suruh" "bayar"]

- e. Labeling

Pelabelan dikelompokkan yakni pada kelas positif, negatif, dan netral. Untuk menyederhanakan pengklasifikasian, kelas positif diberi label 2 dan negatif 0 sedangkan label netral adalah 1. Berikut hasil pelabelan dari data ulasan Telaga Sarangan.

TABEL 6
TABEL HASIL LABELING

No	Ulasan	Label
1	Tempatnya dingin sejuk dan Bersih .Suasannya tenang. Disediakan banyak pilihan tempat penginapan pula. Bagi anda yang bermalam , saya sarankan membawa jaket .	Positif
2	Tempat nya nyaman Sejuk. Tapi ke toilet aj suruh bayar 8 k □	Negatif
3	Kesini pertama kali, tempatnya cocok untuk wisata keluarga harga tiket masuknya per orang 20 rb wisata domestik belum termasuk parkir, suasanaanya adem ada wahana naik boat per satu putaran 80rb 2 putaran 150 rb, 3 putaran 200rb, harga tersebut dihitung per boat bisa disi s.d 4 orang, selain itu ada wahana dayung bebek dan naik kuda harganya saya kurang tahu sekitar tempat wisata juga banyak kuliner khususnya sate kelinci harganya bervariasi 15 - 20 rb, buah buahan, sayur sayuran segar, handycraf dll dihari libur atau weekend pengunjungnya selalu rame, memang cocok untuk wisata keluarga	Netral

- f. Ekstraksi Fitur

Ekstraksi dilakukan dengan pembobotan TF-IDF Tahap pembobotan merupakan tahap dimana pemberian bobot pada setiap kata dengan menggunakan perhitungan Term Frequency-Invert Document Frequency (TF-IDF) [16]. Pemboboftan TF-IDF ini digunakan untuk melihat suatu kata sering muncul dalam

sebuah dokumen yang sedang dianalisis. Dari data yang telah melalui preprosesing setelah itu data akan dihitung menggunakan perhitungan TF-IDF

g. Modeling Support Vector Machine

Support Vector Machine (SVM) adalah salah satu algoritma machine learning yang kuat dan banyak digunakan dalam klasifikasi dan regresi. Pada penelitian ini, SVM digunakan untuk model analisis sentimen terhadap tempat wisata. Klasifikasi data komentar dengan menggunakan metode Support Vector Machine, baik yang berbentuk positif dan negatif. Setelah data selesai melewati tahapan preprosesing, labeling dan ekstraksi fitur maka implementasi algoritma SVM merupakan langkah terakhir dari analisis sentimen terhadap objek wisata Telaga Sarangan

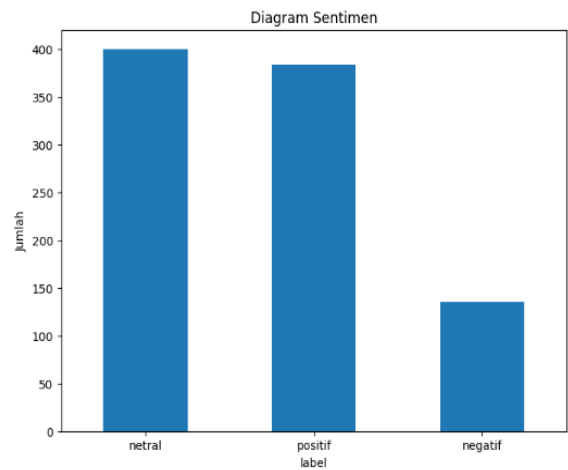
h. Evaluasi

Evaluasi dilakukan dengan menggunakan Confusion Matrix pada machine learning berfungsi untuk memudahkan penelitian dalam menentukan apakah terdapat kesalahan dalam klasifikasi atau tidak. Cara kerja metode ini yaitu matriks dari prediksi akan dibandingkan dengan kelas asli yang berisi informasi nyata dan prediksi nilai klasifikasi Kondisi untuk melakukan pengujian klasifikasi menggunakan metode Confusion Matrix.

IV. HASIL DAN PEMBAHASAN

A. Hasil Labeling

Data yang terkumpul kemudian diberi label sentiment berupa 'Positif', 'Negatif' dan 'Netral'. Proses pelabelan menggunakan cara consensus dengan bantuan 3 orang. Jika suatu data berupa kalimat mengkritik atau keluhan kalimat tersebut akan di klasifikasikan ke dalam label 'Negatif', sebaliknya apabila sebuah data berupa kalimat memuji dan mendukung data tersebut telah diklasifikasikan ke dalam label 'Positif' lalu sebuah data diklasifikasikan menjadi 'Netral' apabila kalimat tidak mengandung opini emosional atau mengandung pujian dan keluhan secara seimbang. Lalu hasil dari consensus dihitung berdasarkan suara terbanyak. Berikut ini visualisasi dari hasil labeling

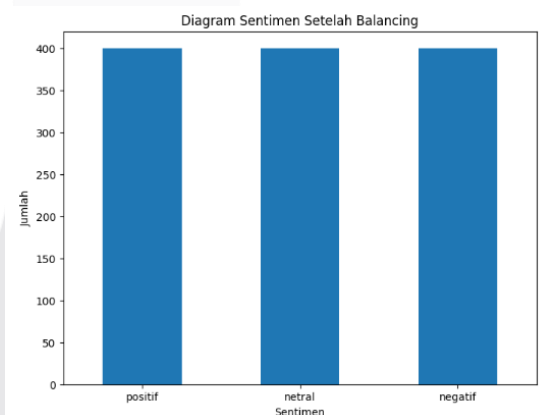


GAMBAR 3
DIAGRAM HASIL LABELING

Terlihat pada diagram di gambar 4.1 Bahwa jumlah label positif dan negatif memiliki selisih, dengan label positif sejumlah 383 sedangkan label negatif berjumlah 135 lalu pada label netral sebanyak 400.

B. Balancing SMOTE

Pada proses ini SMOTE dilakukan untuk menyeimbangkan data. Karena label yang diperoleh memiliki selisih yang besar maka SMOTE digunakan untuk menyeimbangkan label sentimen negatif dan positif. Supaya hasil evaluasinya memiliki nilai akurasi yang optimal.



GAMBAR 4
DIAGRAM HASIL BALANCING

Pada gambar 4 merupakan hasil dari proses balancing menggunakan metode SMOTE. Terlihat bahwa label pada sentimen positif, netral dan negatif telah memiliki nilai yang sama yaitu label positif sejumlah 400 dan Negatif berjumlah 400 dan netral berjumlah 400.

C. TF-IDF

Setelah pelabelan dan preprosesing telah dilakukan dilanjutkan dengan tahapan TF-IDF. Tahapan ini digunakan untuk memberikan bobot pada suatu kata dalam sebuah dokumen yang nantinya digunakan untuk mengetahui seberapa sering sebuah kata tersebut muncul dalam sebuah kalimat. Sehingga setiap kata akan memunculkan suatu nilai hasil perhitungan TF-IDF dapat dilihat pada gambar 4. 3 setelah dilakukan perhitungan

dengan menggunakan bantuan sklearn berupa TfidfVectorizer pada library yang disediakan oleh Bahasa pemrograman Python.

	term	idf	tf0	tf1	tf2	tf3	tf4	tf5	tf6	tf7	...	tf900	tf909	tf910	tf911	tf912	tf913	tf914	tf915	tf916	tf917
0	aa	6.129050	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	aaaaa	6.129050	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	abadi	6.129050	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	abis	5.435903	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	absss	6.129050	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3340	yogyakarta	6.129050	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3341	you	3.989884	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3342	your	5.212759	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3343	ysk	6.129050	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3344	zaman	6.129050	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

GAMBAR 5
HASIL TF-IDF

D. Splitting Data

Setelah data melalui proses labeling, preprocessing, balancing dan TF-IDF, langkah selanjutnya ialah membagi data menjadi data *training* dan data *testing*. Splitting data ini dilakukan guna memisahkan kedua data menjadi data *trainig* dan data *testing*. Data *training* dilakukan untuk melatih algoritma untuk mengenali sebuah label pada suatu kalimat. Sedangkan data *testing* berguna untuk menguji seberapa akurat model yang digunakan dalam mengklasifikasikan sebuah kalimat. Semakin besar data *training* yang digunakan maka semakin optimal sebuah model dalam mengkalsifikasikan sebuah label. Pada tahap ini menggunakan 80% data *training* dan 20% data *testing*. Hasil dari splitting data dapat dilihat pada gambar 4.4

```
Jumlah data setelah SMOTE dan Split:
Jumlah data latih : 840
Jumlah data pengujian: 360
```

```
Distribusi label pada data latih setelah SMOTE:
Counter({0: 280, 1: 280, 2: 280})
```

```
Distribusi label pada data pengujian setelah SMOTE:
Counter({1: 120, 0: 120, 2: 120})
```

GAMBAR 6
HASIL SPLIT DATA

E. Klasifikasi Support Vector Machine

Pada tahapan ini memanfaatkan library sklearn dengan kelas LinierSVC . Pada tahap ini algoritma SVM digunakan untuk menguji model. LinierSVC adalah salah satu implementasi dari Support Vector Machine (SVM) untuk klasifikasi yang menggunakan garis lurus (linear) sebagai pemisah antar kelas. LinierSVC merupakan bagian dari library scikit-learn, dan cocok digunakan untuk klasifikasi dalam bentuk teks. LinierSVC bekerja dengan mencari garis atau hyperplane terbaik yang memisahkan data antar kelas dengan jarak terjauh ke titik terdekat dari masing-masing kelas. LinierSVC memiliki kecepatan dalam mengolah data dengan skala besar, cocok dalam pengklasifikasian dalam bentuk teks dan mudah diatur menggunakan parameter C.

```
from sklearn.SVM import LinearSVC
```

```
SVM_model = LinearSVC()
SVM_model.fit(X_train, y_train)
y_pred = SVM_model.predict(X_test_resampled)
```

Berdasarkan baris kode di atas. Pertama memasukan kelas LinearSVC dari library Scikit-learn, kemudian menginisiasikan model dari linearSVC dengan nama SVM_model. Setelah model terinisiasi Langkah selanjutnya melatih model menggunakan data training berupa x_train_resampled sebagai fitur dan y_train_resampled sebagai label, dengan kedua data tersebut merupakan hasil dari resampling menggunakan SMOTE Balancing. Kedua data tersebut dilatih menggunakan fungsi 'fit()' supaya dapat mengenali pola prediksi kedepannya. Kemudian setelah model dilatih dilakukan tahap uji dengan data x_test_resampled yang hasil pengujiannya akan disimpan dalam variabel y_pred pengtesan ini menggunakan fungsi 'predict'.

F. Evaluasi Model

Setelah memodelkan dengan algoritma *Support Vector Machine* maka perlu dilakukan evaluasi terhadap model berupa pengukuran akurasi, precision, recall, f1-score kemudian dilanjut dengan visualisai menggunakan *Confusion Matrix*. berikut ini merupakan hasil dari evaluasi model yang dapat dilihat pada gambar 4.

```
--- Evaluasi SVM ---
Akurasi: 0.7971014492753623
```

Classification Report:				
	precision	recall	f1-score	support
negatif	1.00	0.68	0.81	41
netral	0.77	0.84	0.80	120
positif	0.78	0.79	0.79	115
accuracy			0.80	276
macro avg	0.85	0.77	0.80	276
weighted avg	0.81	0.80	0.80	276

GAMBAR 7
HASIL EVALUASI SVM

Dari gambar di atas diperoleh dari perhitungan algoritma *Support Vector Machine* sebagai berikut :

a. Akurasi (Accuracy)

$$Accuracy = \frac{28 + 101 + 91}{28 + 7 + 6 + 0 + 101 + 19 + 0 + 24 + 91} = \frac{220}{276} \approx 79.71\%$$

b. Presisi (Precision)

$$Precision = \frac{TP}{TP + FP} = \frac{91}{91 + 25} = \frac{91}{116} \approx 78.45\%$$

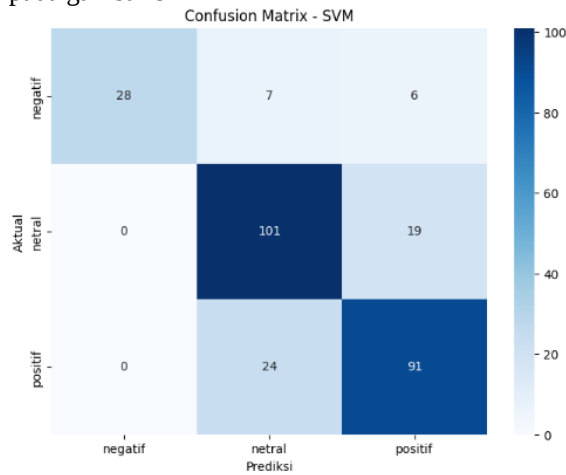
c. Recall

$$Recall = \frac{TP}{TP + FN} = \frac{91}{91 + 24} = \frac{91}{115} = 79.13\%$$

d. F1-Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.7651 \times 0.8417}{0.7651 + 0.8417} \approx 80.14\%$$

kemudian untuk hasil dari *Confusion Matrix* dapat dilihat pada gambar 8



GAMBAR 8
CONFUSION MATRIX SVM

Dari gambar *Confusion Matrix* dijelaskan bahwa model algoritma *Support Vector Machine* dapat mengklasifikasikan sebuah sentimen negatif sebanyak 28 lalu netral sebanyak 101 dan positif sebanyak 91 lalu prediksi yang tidak sesuai dengan aktualnya bernilai 13 pada sentimen negatif 19 netral dan 24 positif

G. Wordcloud Visualisasi

Pada tahap ini menampilkan kata apa saja yang sering muncul dalam label sentimen positif dan juga negatif. Memanfaatkan library dari Python berupa wordcloud untuk menampilkan kata apa yang sering muncul pada sebuah label untuk sentiment positif dapat dilihat pada gambar 9



GAMBAR 9
WORDCLOUD SENTIMEN POSITIF

Sedangkan untuk label negatif dapat dilihat pada gambar 10.



GAMBAR 10
WORDCLOUD SENTIMEN NEGATIF

Lalu berikut ini merupakan wordcloud untuk sentimen netral



GAMBAR 11
WORDCLOUD SENTIMEN NETRAL

H. Perbandingan dengan Algoritma lain

Pada tahap ini membanding algoritma SVM dengan multinomialnaivebaye dan *logisticregresion* guna mengetahui algoritma mana yang lebih akurat dalam pengklasifikasian dalam bentuk teks dengan menggunakan sklearn *MultinomialNB* dan sklearn *logisticregression*. Hasil dari algoritma *Logisticregression* dapat dilihat pada gambar 4.10

```

=== Logistic Regression ===
Akurasi: 0.7282608695652174

```

	precision	recall	f1-score	support
negatif	1.00	0.32	0.48	41
netral	0.68	0.88	0.77	120
positif	0.76	0.71	0.74	115
accuracy			0.73	276
macro avg	0.81	0.64	0.66	276
weighted avg	0.76	0.73	0.71	276

GAMBAR 12
HASIL AKURASI ALGRITMA LOGISTIC REGRESI

Kemudian untuk hasil dari algoritma Naïve Bayes dapat dilihat pada gambar 4.11 berikut ini.

```

=== Naive Bayes ===
Akurasi: 0.6557971014492754

```

	precision	recall	f1-score	support
negatif	1.00	0.02	0.05	41
netral	0.74	0.65	0.69	120
positif	0.60	0.89	0.72	115
accuracy			0.66	276
macro avg	0.78	0.52	0.49	276
weighted avg	0.72	0.66	0.61	276

GAMBAR 13
HASIL AKURASI ALGORITMA NAIVE BAYES

V. KESIMPULAN

Berdasarkan hasil penelitian mengenai “Analisis Sentimen terhadap Pengalaman Pengguna di Tempat Wisata melalui Review Google Maps: Studi Kasus Telaga Sarangan”, diperoleh bahwa analisis sentimen berhasil mengklasifikasikan 919 ulasan Google Maps ke dalam tiga kategori, yaitu positif, negatif, dan netral. Setelah melalui tahapan preprocessing dan pelabelan, mayoritas ulasan diketahui bernada positif. Sentimen positif umumnya berkaitan dengan keindahan alam, udara yang sejuk, serta pengalaman menyenangkan, sedangkan sentimen negatif banyak mencerminkan keluhan terhadap fasilitas umum, pengelolaan parkir, dan kenyamanan lingkungan.

Faktor-faktor kepuasan pengunjung meliputi panorama alam yang indah, udara yang sejuk, ketersediaan kuliner lokal seperti sate kelinci, serta keberagaman aktivitas wisata seperti speed boat dan berkuda. Sementara itu, faktor ketidakpuasan mencakup fasilitas umum yang kurang memadai seperti toilet berbayar dan tempat duduk yang rusak, penataan parkir kendaraan yang kurang baik, serta pengelolaan kawasan wisata yang dinilai belum profesional.

Metode Support Vector Machine (SVM) berhasil diimplementasikan untuk melakukan klasifikasi sentimen secara otomatis. Dengan dukungan ekstraksi fitur Term Frequency-Inverse Document Frequency (TF-IDF) dan penyeimbangan data menggunakan Synthetic Minority Oversampling Technique (SMOTE), model SVM menunjukkan performa yang baik dengan akurasi sekitar 80%. Uji pembandingan menggunakan algoritma Logistic Regression dan Naïve Bayes menunjukkan bahwa SVM memiliki performa yang lebih stabil dan akurat, baik dari segi akurasi, precision, maupun recall.

Dengan tercapainya seluruh tujuan penelitian, dapat disimpulkan bahwa analisis sentimen berbasis ulasan Google Maps merupakan pendekatan yang efektif untuk mengevaluasi kualitas pelayanan dan pengalaman wisatawan pada destinasi wisata seperti Telaga Sarangan.

REFERENSI

- [1] J. Ipmawati, S. Saifulloh, and K. Kusnawi, “Analisis Sentimen Tempat Wisata Berdasarkan Ulasan pada Google Maps Menggunakan Algoritma Support Vector Machine,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 1, pp. 247–256, 2024, doi: 10.57152/malcom.v4i1.1066.
- [2] A. E. Dian Siti Utami, “Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 6, no. 1, pp. 418–427, 2022.
- [3] F. U. Haq, “Penggunaan Google Review Sebagai Penilaian Kepuasan Pengunjung Dalam Pariwisata,” *Tornare*, vol. 2, no. 1, p. 10, 2020, doi: 10.24198/tornare.v2i1.25826.
- [4] Wikipedia contributors, “Telaga Sarangan,” Wikipedia. Accessed: Jun. 15, 2024. [Online]. Available: https://id.wikipedia.org/wiki/Telaga_Sarangan#cite_note-e-Sarang-1
- [5] Kompas.com, “Tembus 24.000 Pengunjung Sehari, Destinasi Wisata Magetan Sumbang PAD Rp 2 M,” Kompas. Accessed: Jun. 15, 2024. [Online]. Available: <https://regional.kompas.com/read/2024/04/15/090616578/tembus-24000-pengunjung-sehari-destinasi-wisata-magetan-sumbang-pad-rp-2-m>
- [6] C. F. Hasri and D. Alita, “Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter,” *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 3, no. 2, pp. 145–160, 2022, doi: 10.33365/jatika.v3i2.2026.
- [7] A. Rifa’i, H. Sujaini, and D. Prawira, “Sentiment Analysis Objek Wisata Kalimantan Barat Pada Google Maps Menggunakan Metode Naive Bayes,” *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 3, p. 400, 2021, doi: 10.26418/jp.v7i3.48132.
- [8] E. Suryati, Styawati, and A. A. Aldino, “Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM),” *J. Teknol. Dan Sist. Inf.*, vol. 4, no. 1, pp. 96–106, 2023, [Online]. Available: <https://doi.org/10.33365/jtsi.v4i1.2445>
- [9] D. A. Putri, D. A. Kristiyanti, E. Indrayuni, A. Nurhadi, and D. R. Hadinata, “Comparison of Naive Bayes Algorithm and Support Vector Machine using PSO Feature Selection for Sentiment Analysis on E-Wallet Review,” *J. Phys. Conf. Ser.*, vol. 1641, no. 1, 2020, doi: 10.1088/1742-6596/1641/1/012085.
- [10] S. Styawati, N. Hendrastuty, and A. R. Isnain, “Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine,” *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp. 150–155, 2021, doi: 10.30591/jpit.v6i3.2870.
- [11] D. Ananda and R. R. Suryono, “Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia dengan Metode Support Vector Machine dan Naïve Bayes,” vol. 8, no. April, pp. 748–757, 2024, doi: 10.30865/mib.v8i2.7517.
- [12] P. Arsi and R. Waluyo, “Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM),” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 147, 2021, doi: 10.25126/jtiik.0813944.
- [13] J. Adiputra and D. Mahdiana, “Analisis Sentimen Dengan Algoritma Support Vector Machine Terhadap Penyakit Hepatitis Akut Misterius,” *IDEALIS Indones. J. Inf. Syst.*, vol. 6, no. 1, pp. 1–8, 2023, doi: 10.36080/idealism.v6i1.2985.
- [14] A. Karami, M. Lundy, F. Webb, and Y. K. Dwivedi, “Twitter and Research: A Systematic Literature Review through Text Mining,” *IEEE Access*, vol. 8, pp. 67698–67717, 2020, doi: 10.1109/ACCESS.2020.2983656.
- [15] O. Alsemaree, A. S. Alam, S. S. Gill, and S. Uhlig, “Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions,” *Heliyon*, vol. 10, no. 9, p. e27863, 2024, doi: 10.1016/j.heliyon.2024.e27863.
- [16] D. Alita and A. R. Isnain, “Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier,” *J. Komputasi*, vol. 8, no. 2, pp. 50–58, 2020, doi: 10.23960/komputasi.v8i2.2615.
- [17] G. K. Pati and E. Umar, “Analisis Sentimen Komentar Pengunjung Terhadap Tempat Wisata Danau Weekuri Menggunakan Metode Naive Bayes Classifier Dan K-Nearest Neighbor,” *J. Media Inform. Budidarma*, vol. 6, no. 4, p. 2309, 2022, doi: 10.30865/mib.v6i4.4635.
- [18] Ardiansyah, Adika Sri Widagdo, Krisna Nuresa Qodri, F. E. N. Saputro, and Nisrina Akbar Rizky P, “Analisis sentimen terhadap pelayanan Kesehatan berdasarkan ulasan Google Maps menggunakan BERT,” *J. Fasilkom*, vol. 13, no. 02, pp. 326–333, 2023, doi: 10.37859/jf.v13i02.5170.

- [19]B. Suyono, ““Analisis Kepuasan Wisatawan Telaga Sarangan Magetan Berdasarkan Tourism Experience,”” vol. 9, no. November 2018, pp. 39–43, 2020.
- [20]A. Erfina and N. Resti Wardani, ““Analisis Sentimen Perguruan Tinggi Termewah Di Indonesia Menurut Ulasan Google Maps Menggunakan Support Vector Machine (Svm),”” J. Manaj. Inform. Sist. Inf., vol. 5, pp. 77–85, 2022.
- [21]Suyanto, “Machine Learning: Tingkat Dasar dan Lanjut,” Edisi 1. Bandung: Informatika, 2018.

