Seminar Nasional Teknologi dan Inovasi (SITASI)
Vol 1, Nomor 1, Juli 2025, hal. 310-324
P-ISSN: XXXX-XXXX | E-ISSN: XXXX- XXXX
Penerbit Institut Bisnis dan Teknologi Indonesia

Evaluasi Sentimen Wisatawan terhadap Pura Tanah Lot dengan Metode *Support Vector Machine* (SVM)

Gusti Ayu Shinta Dwi Astari[1] , I Kadek Dwi Gandika Supartha[2], Putu Satria Udyana[3] Putra, I Kadek Windu Ardiana[4], Ni Made Ari Adnyani[5]

[1*,2,3, 4,5] *Institut Bisnis dan Teknologi Indonesia, Denpasar, Bali*
[1] *shinta.astari@instiki.ac.id,* [2]*gandika.supartha@instiki.ac.id,* [3]*satria@instiki.ac.i,* [4]*winduardiana@gmail.com,* [5]*ariadnyani2101@gmail.com*

INFO ARTIKEL

Article history:
Received Juni 2025
Accepted Juli 2025
Published Juli 2025

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen terhadap destinasi Pura Tanah Lot berdasarkan ulasan wisatawan yang tersedia di Google Review. Metode yang digunakan untuk menganalisis sentimen adalah *Support Vector Machine* (SVM). Klasifikasi yang dilakukan pada penelitian ini dilakukan dengan membagi 3 kelas yakni positif, negatif dan netral. Data yang digunakan dalam penelitian ini berjumlah 4906 *review* yang sudah dilakukan preprocessing. kemudian, fitur yang relevan diekstraksi dan diproses menggunakan teknik TF-IDF untuk membentuk vektor fitur yang dapat digunakan dalam model SVM. Hasil penelitian menunjukkan bahwa metode SVM mampu mengklasifikasikan sentimen pengunjung dengan accuracy 90,94%, precision di 85% lalu recall di 75,49%, dan terakhir F1-Score di 78,67%. Angka tersebut merupakan hasil terbaik yang dapat dihasilkan SVM dengan menggunakan kernel linier dengan pembagian data di 80:20. Sentimen yang dominan terhadap Pura Tanah Lot cenderung positif, dengan pengunjung memberikan apresiasi tinggi terhadap keindahan alam yang ada.

Kata Kunci: Pura Tanah Lot, Review, Sentimen Analisis. TF-IDF, *Support Vector Machine* (SVM)

ABSTRAK

This research aims to conduct a sentiment analysis of Tanah Lot Temple as a tourist destination, based on reviews available on Google Reviews. The sentiment analysis was performed using the Support Vector Machine (SVM) method. Sentiments in this study were classified into three categories: positive, negative, and neutral. A total of 4,906 reviews, which had undergone preprocessing, were utilized as the dataset. Relevant features were extracted and processed using the TF-IDF technique to construct feature vectors for the SVM model. The results demonstrate that the SVM method achieved an accuracy of 90.94%, a precision of 85%, a recall of 75.49%, and an F1-Score of 78.67%. These results represent the optimal performance obtained using an SVM with a linear kernel and an 80:20 data split. The analysis reveals that sentiments toward Tanah Lot Temple are predominantly positive, with visitors highly appreciating its natural beauty.

Keywords: Tanah Lot Temple, Reviews, Sentiment Analysis, TF-IDF, Support Vector Machine (SVM)

1. Pendahuluan

Perkembangan teknologi saat ini telah diikuti oleh revolusi industri 4.0, yang merupakan dasar dari teknologi digital yang canggih (Indra et al, 2023). Perkembangan ini tentu memberikan dampak luas terhadap berbagai sektor termasuk sektor pariwisata (Djamil et al, 2023). Dalam pengembangan sektor pariwisata teknologi menjadi sarana yang sangat memberikan keuntungan di bidang informasi. Salah satu contoh nyata teknologi membantu wisatawan mencari segala informasi terkait destinasi yang akan di kunjungi yang dalam hal ini tentu membantu industri pariwisata semakin berkembang (Nuryananda et al, 2023). Pariwisata adalah salah satu sektor andalan dan juga menjadi sektor yang menjadi prioritas untuk dikembangkan di Indonesia (Winoto et al, 2024). Salah satu daerah yang paling terkenal akan pariwisatanya adalah Pulau Bali. Bali adalah salah satu provinsi yang menjadi ikon pariwisata di Indonesia (Kameda et al, 2023). Bali memiliki potensi wisata alam yang sangat indah serta seni budaya yang khas. Karena keindahan alam dan keunikan seni budayanya, Bali mampu menarik wisatawan baik dari dalam negeri maupun mancanegara (Suryawan et al, 2023). Berdasarkan catatan (Badan Pusat Statistik Provinsi Bali, 2023), menerangkan bahwa pada periode Januari–Februari 2024, jumlah wisatawan mancanegara (wisman) yang datang ke Bali mencapai 874.838 kunjungan, naik 33,50% dibandingkan periode yang sama pada 2023. Pada tahun 2023, total kunjungan wisman ke Bali mencapai 5.273.258, dengan rata-rata 439.438 pengunjung per bulan. Bulan Juli 2023 mencatat jumlah pengunjung tertinggi, dengan 541.353 kunjungan. Hal ini tentu menjadi tolak ukur bahwa memang destinasi wisata dibali menjadi salah satu daya tarik wisatawan untuk datang ke Bali.

Ada beragam destinasi wisata yang dapat dinikmati di bali salah satunya adalah wisata budaya. Bali memiliki budaya yang kaya dan unik. Salah satu daya tarik utama Bali adalah pura-puranya yang megah dan sakral (Liyushiana, 2024). Pura ini tidak hanya berfungsi sebagai tempat ibadah bagi umat Hindu, tetapi juga merupakan destinasi wisata yang populer bagi wisatawan lokal dan mancanegara (Viona, 2023). Ada cukup banyak Pura dengan destinasi wisata seperti Pura Uluwatu, Pura Besakih, Pura Rambut Siwi, Pura Lempuyang Luhur, Pura Ulun Danu Batur, Pura Ulun Danu Beratan dan Pura Tanah Lot. Berdasarkan *review* google maps bulan juli 2024 Pura dengan destinasi wisata yang memiliki *review* terbanyak dibali adalah Pura Tanah Lot dengan 93,833 *review* dibanding dengan pura Besakih 15,189 review, Pura uluwatu 44,665 review, Pura Rambut Siwi 700 review, Pura Lempuyang Luhur 799 review, dan Pura Ulun Danu Batur 3250 review, dan Pura Ulun Danu Beratan 42,391 review. dengan demikian dapat dikatakan bahwa pura tanah lot merupakan pura yang paling populer sebagai destinasi wisata dibali. Pura Tanah Lot dan rangkaian pura di sekitarnya menjadi salah satu daya tarik wisata utama di Bali Selatan. Selain sebagai tempat pemujaan, areal sekitar pura juga dimanfaatkan sebagai tempat wisata, wahana rekreasi, maupun tempat untuk menampilkan seni pertunjukan pariwisata Bali (Untara et al, 2020). Hal ini tentu mampu menarik wisatawan untuk datang ke Pura Tanah Lot. Berdasarkan catatan Dinas Pariwisata Provinsi Bali Pada tahun 2023, 2,02 juta orang mengunjungi Tanah Lot, yang mana Ini merupakan peningkatan 50,99% dari tahun 2022, ketika 1,3 juta orang mengunjungi destinasi wisata tersebut. Dengan semakin banyaknya jumlah pengunjung setiap tahun, ulasan atau *review* tentang destinasi wisata menjadi sangat penting. Ulasan ini membantu wisatawan membentuk ekspektasi yang lebih realistis tentang pengalaman yang akan mereka dapatkan, berdasarkan pengalaman pengunjung sebelumnya (Mahardika, 2020). Selain itu, ulasan yang jujur dan konstruktif memberikan umpan balik berharga bagi pengelola destinasi wisata untuk meningkatkan layanan dan fasilitas (Husain et al, 2024). Oleh karena itu, analisis sentimen menjadi hal yang penting. Dalam konteks ini, teknologi memainkan peran penting dalam melaksanakan analisis sentimen. Dengan analisis sentimen berbasis teknologi, para pemangku kepentingan di industri pariwisata dapat memahami persepsi dan perasaan wisatawan terhadap destinasi, layanan, dan pengalaman yang ditawarkan, termasuk di destinasi wisata Pura Tanah Lot. Ini memungkinkan pengelola untuk mengetahui sentimen wisatawan yang telah berkunjung ke Pura Tanah Lot, yang berguna baik bagi pengambilan keputusan wisatawan untuk berkunjung, maupun sebagai tinjauan bagi pengelola untuk memajukan destinasi wisata tersebut.

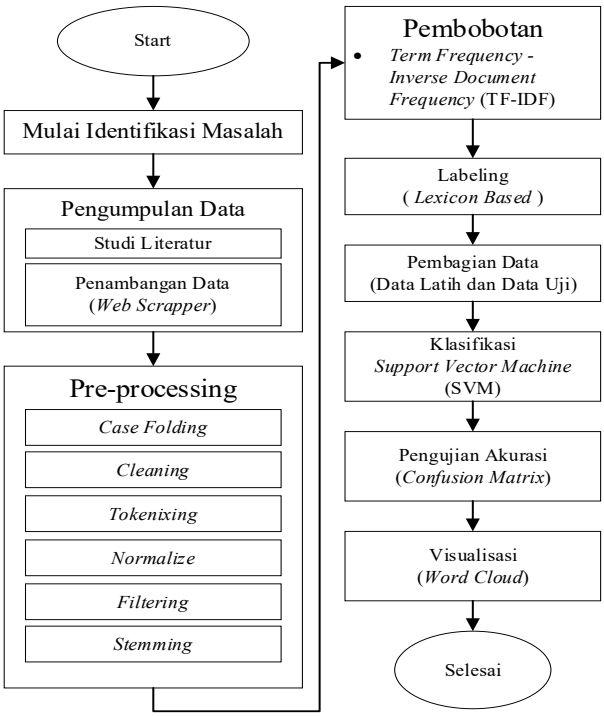
Terdapat beberapa penelitian sebelumnya yang membahas tentang sentimen analisis seperti Penelitian dari (Herlawati et al, 2021). Penelitian ini menggunakan 2 metode yaitu Naïve Bayes dan Support Vector Machine. Dalam penelitiannya akurasi dalam melakukan analisis sentimen pada ulasan Summarecon Mal Bekasi dengan data sebanyak 2.143 komentar dengan akurasi untuk Naïve Bayes dan *Support Vector Machine* berturut-turut memperoleh angka di 80,95% dan 100%. Penelitian lainnya yang membahas tentang sentimen analisis adalah penelitian dari (Suryawan et al, 2023). Pada penelitian ini, analisis sentimen terhadap destinasi wisata di Ubud menghasilkan 551 data dengan sentimen positif dan 118 data dengan

sentimen negatif dari total 669 data uji. Hasil ini menunjukkan nilai positif yang signifikan pada destinasi wisata di Ubud. Selanjutnya, pengujian algoritma *Support Vector Machine* (SVM) menggunakan *Confusion Matrix* menunjukkan hasil yang baik dalam analisis sentimen, dengan akurasi sebesar 84,01%, recall sebesar 89,83%, precision sebesar 90,40%, dan F1-Score sebesar 90,11%.

Berdasarkan dari uraian diatas, sentimen analisis diperlukan untuk menganalisis sentimen wisatawan apakah positif, negatif, atau netral terhadap destinasi wisata Pura Tanah Lot, yang nantinya hasil penelitian ini dapat berguna sebagai tinjauan kepada pengelola untuk memajukan destinasi Wisata Pura tanah Lot. Oleh karenanya penelitian ini akan meneliti sentimen terhadap destinasi wisata Pura Tanah Lot dengan Judul "Analisis Sentimen Terhadap Destinasi Wisata Pura Tanah Lot dengan Metode *Support Vector Machine* (SVM)".

2. Metode Penelitian

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan wisata Pura Tanah Lot yang diperoleh dari Google *Review* menggunakan metode *Support Vector Machine* (SVM). Data dikumpulkan melalui web scraping dan diproses melalui beberapa tahap pre-processing, termasuk case folding, cleaning, tokenizing, normalisasi, filtering, dan stemming untuk meningkatkan kualitas data. Selanjutnya, fitur teks diberi bobot menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) sebelum dilakukan labeling sentimen secara semi-otomatis dengan pendekatan lexicon-based ke dalam tiga kelas: positif, netral, dan negatif. Dataset kemudian dibagi menjadi data latih dan data uji sebelum diklasifikasikan menggunakan SVM. Kinerja model dievaluasi menggunakan *Confusion Matrix* dengan metrik akurasi, presisi, recall, dan F1-score. Hasil analisis divisualisasikan dalam bentuk Word Cloud untuk memahami pola kata yang sering muncul dalam ulasan wisatawan, sehingga dapat memberikan wawasan bagi pengelola destinasi wisata dalam meningkatkan kualitas layanan dan pengalaman wisatawan secara lebih jelas, gambar 1 berikut ini merupakan metodologi yang diusulkan dalam penelitian ini.



Gambar 1. Metode Penelitian

2.1 Identifikasi Masalah

Pada tahap ini, ditentukan permasalahan yang akan dianalisis, yaitu bagaimana persepsi wisatawan terhadap Pura Tanah Lot berdasarkan ulasan yang tersedia di Google Review.

2.2 Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan dengan 2 tahapan, yang pertama adalah studi literatur yang dilakukan dengan mengumpukan data – data dan informasi melalui Jurnal – jurnal, *e-*

book, media *online* dan referensi lain yang berhubungan dengan penelitian ini. Lalu yang kedua adalah pengumpulan data yang akan digunakan untuk analisis sentimen. Data yang dikumpulkan berasal dari ulasan pada *Google Review* dengan kata Kunci Tanah Lot dengan total *review* 92.562. Pengumpulan data ulasan pada penelitian ini hanya sampai di tanggal 20 Juni 2024 disimpan ke dalam file excel dengan format "nama file.xlsx."

Tabel 1. Hasil Web Scrapper

No	Waktu	Komentar
1	10 bulan lalu	Salah satu lokasi wisata di Bali yang harus kamu kunjungi jika baru pertama kali ke Bali.Bila beruntung bisa menyeberang dan naik ke pelataran Pura nya (karena hanya …
2	10 bulan yang lalu	Cantiiikkk banget t4 iniWajib, kudu, mesti kalau ke Bali ke dtw 1 ini 😍 …
3	11 bulan yang lalu	Bagus banget dan bersih tdk bnyk ssmpah..bagus buat yg suka selfi …

2.3 Preprocessing

Preprocessing, dilakukan dengan tujuan memastikan bahwa data yang digunakan untuk analisis adalah data yang representatif dan berkualitas tinggi. Jika data yang digunakan dalam proses penemuan pengetahuan berkualitas rendah, maka hasil pengetahuan yang diperoleh juga akan rendah. (Fadhilah et al, 2023) Tahapan preprocessing yang dilakukan dalam penelitian ini adalah sebagai berikut:

a. *Case folding*, dipergunakan untuk mengubah teks secara keseluruhan dalam dokumen menjadi bentuk standar, (dalam hal ini huruf kecil atau *lowercase*) (Alita et al, 2020).
b. *Cleansing*, atau pembersihan data, adalah proses menghapus karakter atau simbol yang tidak memiliki arti dalam sebuah kalimat (Syah et al, 2022).
c. *Tokenizing*, proses membagi teks menjadi kata-kata lebih kecil yang memiliki arti atau "token" (Wulandari et al, 2024). Secara sederhana proses ini Memecah teks menjadi kata-kata atau frasa-frasa
d. *Normalize*, adalah tahap memperbaiki dan mengubah singkatan menjadi kata yang bermakna sama. (Fadhilah et al, 2023) Sebagai contoh merubah kata "tdk" menjadi "tidak" dan lain sebagainya
e. *Stopword Removal*, adalah bagian dari tahapan filtering. Tahap filtering biasanya menggunakan algoritma *stopword removal*, juga dikenal sebagai algoritma *remove stopword*. (Syah et al, 2022) stopword removal digunakan untuk menghilangkan kata – kata yang tidak berpengaruh signifikan terhadap proses klasifikasi seperti "dan", "atau", "di", "untuk", dan sebagainya.
f. *Stemming* merupakan cara mengembalikan kata ke bentuk semula bentuk dasar (Sinaga et al, 2023). Sebagai contoh mengurangi kata-kata ke bentuk dasar atau akar (misalnya, "berlari" menjadi "lari").

2.4 Pembobotan Kata

Tujuan pembobotan kata merupakan mengubah data yang belum terstruktur menjadi lebih terstruktur, setelah itu data yang telah terstruktur dikategorikan menggunakan metode *classifier* (Ramadhan et al, 2021). Dalam penelitian ini, dua teknik pembobotan kata digunakan untuk menentukan mana yang memberikan nilai frekuensi kata yang optimal. Metode *term frequency – inverse document frequency* (TF-IDF) digunakan untuk memembobot kata (Rabbani et al, 2023).

Metode TF-IDF mengukur tingkat kepentingan kata terhadap dokumen dengan memberikan bobot terhadap kata. Nilai bobot dihitung dengan menghitung berapa kali kata muncul dalam dokumen (Hakim et al, 2024).

Dalam dokumen yang relevan, istilah frekuensi *Term Frequency* (TF) digunakan untuk menghitung frekuensi kemunculannya. Semakin banyak kata yang sering muncul (term) dalam dokumen, semakin besar bobot kata. Menurut metode ini, kata-kata yang memiliki nilai TF tinggi memiliki makna yang

signifikan dalam sebuah dokumen, sedangkan DF menunjukkan berapa kali kata tersebut muncul dalam kumpulan dokumen (Maulana et al, 2023). Berikut Persamaannya (Septiani et al, 2022):

Perhitungan TF (*Term Frequency*)

TF adalah frekuensi kemunculan suatu kata dalam sebuah dokumen. Rumus TF untuk *term t* dalam dokumen *d* adalah:

$$TF\ (t,d) = \frac{\text{frekuensi } t \text{ dalam } d}{\text{Jumlah total kata dalam } d} \quad (1)$$

Selanjutnya, dihitung Inverse Document Frequency (IDF) untuk mengukur seberapa unik kata tersebut dalam koleksi dokumen.

$$IDF(t,D) = \log\left(\frac{N}{df\ (t)}\right) \quad (2)$$

Di mana:

- N adalah jumlah total dokumen dalam corpus.
- df(t) adalah jumlah dokumen yang mengandung term t

Perhitungan TF-IDF

Bobot TF-IDF untuk term t dalam dokumen d adalah hasil kali antara TF dan IDF:

$$\text{TF} - \text{IDF}(t,d,D) = \text{TF}(\text{t},\text{d}) * IDF(t,d) \quad (3)$$

Kata-kata yang sering muncul dalam satu dokumen tetapi jarang muncul dalam dokumen lain akan diberikan bobot lebih besar. Hasil dari perhitungan TF dan IDF ini dikombinasikan untuk menghasilkan skor TF-IDF untuk setiap kata, yang selanjutnya digunakan dalam model analisis sentimen untuk klasifikasi teks.

2.5 Labeling

Labaling adalah proses dimana Data yang telah didapat harus didefinisikan terlebih dahulu sebagai kalimat yang mengandung nilai positif atau nilai negatif. Masing-masing data *review* harus diberi label sebagai *review* positif atau *review* negatif (Syah et al, 2022).

Pada penelitian ini Teknik pelabelan yang digunkan adalah *Lexicon Base*. *Lexicon Based* adalah proses pemilihan kata-kata penting dalam suatu dokumen berdasarkan kamus/kosa kata yang ada. Dalam penerapannya, dua kamus digunakan untuk membuat daftar kata. Satu kamus berisi kumpulan kata-kata dengan emosi positif dan satu kamus berisi kumpulan kata-kata dengan emosi negatif. Daftar kata digunakan dalam proses penyaringan untuk mengambil kata-kata yang memiliki sentimen (Fathullah et al, 2020) (Roiqoh et al, 2023).

2.6 Pembagian Data

Pembagian data adalah proses membagi dataset menjadi dua bagian: satu bagian untuk melatih model (*train data*) dan satu bagian untuk menguji kinerja model (*test data*). Hal ini penting dalam pembangunan model *machine learning* untuk mengukur seberapa baik model dapat menggeneralisasi pola dari data yang belum pernah dilihat sebelumnya. Pada penelitian ini, untuk mencari akurasi terbaik maka pembagian data latih dan data uji dibagi menjadi 3 skenario dengan perbandingan 60:40, 70:30 dan 80:20.

2.7 Klasifikasi

Klasifikasi digunakan dalam data mining untuk mengelompokkan data ke dalam kelas tertentu sesuai dengan kriteria yang telah ditentukan. Proses pengelompokan atau klasifikasi data sangat dipengaruhi oleh kriteria ini, dan dari kriteria ini akan terbentuk hubungan pola yang dapat digunakan untuk menyelesaikan masalah (Wijaya et al, 2023). Pada penelitian ini algoritma yang digunakan untuk klasifikasi adalah *Support Vector Machine* (SVM), sebuah algoritma pembelajaran terawasi yang bekerja

dengan mencari *hyperplane* optimal untuk memisahkan kelas-kelas dalam data. Dengan pendekatan ini, SVM mampu menangani data dengan dimensi tinggi dan menghasilkan klasifikasi yang akurat, terutama ketika dipadukan dengan teknik ekstraksi fitur seperti TF-IDF. Berikut persamaan untuk mencari *Hyperplane* (Muhathir et al, 2021):

$$w.\Phi\,(x) + b = 0 \quad (4)$$

Di mana w adalah normal untuk *hyperplane*, Φ (x) adalah pemetaan fungsi yang digunakan untuk memetakan setiap vektor masukan ke ruang fitur dan b adalah biasnya. Masalah optimasi menemukan *hyperplane* dengan margin terbesar dirumuskan sebagai berikut (Muhathir et al, 2021):

$$\min_{\alpha}(\alpha) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \quad (5)$$
$$K(x_i, x_j)\,\alpha_i\alpha_j - \sum_{j=1}^{N}\alpha_i$$

Persoalan menjadi:

$$\sum_{i=1}^{N} y_i\alpha_j = 0, 0 \le \alpha i \le C, dan\ i = 1, \dots, n \quad (6)$$

Di mana x_(i,) adalah vektor masukan ke-i, y_(i,)adalah korespondensinya kelas, α adalah pengali Lagrange. Fungsi kernel K mungkin linier, *polinomial, eksponensial*, atau jenis lainnya. Hukuman parameter C harus dipilih dengan hati-hati untuk setiap kumpulan data.

2.8 Pengujian Akurasi

Evaluasi kinerja model *Support Vector Machine* (SVM) dalam analisis sentimen dilakukan menggunakan Confusion Matrix, yang mengukur akurasi berdasarkan perbandingan antara prediksi dan label sebenarnya. *Confusion Matrix* terdiri dari empat komponen utama: True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN), yang digunakan untuk menghitung metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Dengan menggunakan metrik ini, dapat dianalisis sejauh mana model mampu mengklasifikasikan sentimen positif, netral, dan negatif dengan baik. Tabel *Confusion Matrix* ditunjukan pada tabel 2 (Awalullaili et al, 2023):

Tabel 2. *Confusion Matrix*

Kondisi sebenarnya	Hasil Prediksi		
	Positif	Negatif	Total Baris
Positif	TP	FN	
Negatif	FP	TN	
Total Kolom	TP+FP	FN+TN	N = TP+TN+FP+FN

Akurasi dapat dihitung dengan rumus *Accuracy*:

$$= \frac{TN+TP}{TN+TP+FN+FP} \quad (7)$$

Sensitivitas dapat dihitung dengan rumus *Recall*:

$$= \frac{TP}{TP+FN} \quad (8)$$

Presisi dihitung dengan rumus *Precision*:

$$= \frac{TN}{TN+FP} \quad (9)$$

Nilai F1-Score dihitung dengan *rumus F1-Score*:

$$= 2 * \frac{Precisio*Recall}{precision+Recall} \quad (10)$$

2.9 Visualisasi

Visualisasi data merupakan langkah penting dalam analisis sentimen untuk memahami pola dan distribusi sentimen dalam ulasan secara intuitif. Dengan visualisasi, informasi yang kompleks dapat disajikan dalam bentuk yang lebih mudah dipahami, sehingga membantu dalam interpretasi hasil klasifikasi. Salah satu teknik yang umum digunakan adalah *word cloud*. *Word Cloud* merupakan representasi visual berdasarkan frekuensi kemunculan kata dalam sekumpulan teks, dimana besar kecilnya huruf menentukan frekuensi kemunculan suatu kata, yaitu semakin besar ukuran huruf maka semakin besar pula frekuensi kemunculannya. kata tersebut dan sebaliknya, semakin kecil hurufnya, semakin rendah frekuensi kemunculan kata tersebut (Julianto, 2022) (Zai et al, 2024). Selain itu, grafik sentimen, seperti diagram batang digunakan untuk menunjukkan proporsi ulasan positif, negatif, dan netral, sehingga memberikan gambaran yang lebih jelas mengenai persepsi wisatawan terhadap destinasi tertentu.

3. Hasil dan Pembahasan

3.1 Dataset

Proses pengumpulan data dilakukan dengan menggunakan layanan APIFY. APIFY merupakan platform otomatisasi berbasis web yang memungkinkan pengambilan data dari berbagai situs web secara efisien. Dalam penelitian ini data maksimal yang dapat diambil oleh Apify untuk *review* destinasi wisata Pura Tanah lot dengan total 7997 *review* yang terbagi menjadi 2 yaitu 5007 data dengan komentar dan 2990 data tanpa komentar. Sehingga dalam melakukan proses sentimen analisis data *review* yang diperlukan adalah *review* yang berupa komentar yaitu sejumlah 5007 komentar.

3.2 Preprocessing

Preprocessing merupakan tahap penting dalam analisis sentimen yang bertujuan untuk membersihkan dan menyiapkan data teks sebelum dilakukan klasifikasi. Dalam penelitian ini, preprocessing dilakukan untuk meningkatkan kualitas data ulasan destinasi wisata Pura Tanah Lot di Google *Review* agar lebih relevan dalam analisis sentimen.

1) Case Folding
Case folding dalam penelitian ini digunakan untuk melakukan konversi seluruh teks ke dalam format huruf kecil (*lowercase*) atau huruf besar (*uppercase*). Berikut hasil dari case *folding*.

Tabel 3. *Case Folding*

Sebelum *Case Folding*	Sesudah *Case Folding*
Jika berlibur ke pulau **B**ali jangan lupa berkunjung ke **T**anah **L**ot.	jika berlibur ke pulau bali jangan lupa berkunjung ke tanah lot.

2) Cleansing
Tahap ini akan menghilangkan tanda baca, menghapus angka, karakter khusus atau simbol, dan menghapus spasi berlebih. Berikut hasil dari *Cleansing*.

Tabel 4. *Cleansing*

Sebelum *Cleansing*	Setelah *Cleansing*
terkagum-kagum. tempat wisata yang indah. terima kasih tanah lot ❤	terkagumkagum tempat wisata yang indah terima kasih tanah lot
luar biasa ✌ ✌	luar biasa
bali....semua pantainya luar biasa indah ❤ ❤ ❤	balisemua pantainya luar biasa indah

3) Tokenizing
Pada penelitian ini akan dilakukan proses pemisahan kata per kata yang mana ini sangat penting karena memungkinkan sistem untuk memahami dan memproses teks dengan cara yang lebih terstruktur. Berikut adalah hasil dari *Tokenizing*.

Tabel 5. *Tokenizing*

Sebelum *Tokenizing*	Setelah *Tokenizing*
acara tour bareng bersama teman puskesmas pelayanannya baik dan tempat wisatanya juga bagus	'acara', 'tour', 'bareng', 'bersama', 'teman', 'puskesmas', 'pelayanannya', 'baik', 'dan', 'tempat', 'wisatanya', 'juga', 'bagus'

4) Normalisation
Pada tahap ini akan dilakukan Normalisasi yang bertujuan untuk mengubah berbagai bentuk teks menjadi bentuk yang konsisten dan standar. Berikut adalah hasil dari *Normalisation*.

Tabel 6. *Normalisation*

Sebelum *Nomalisation*	Sesudah *Nomalisation*
tempat wisata terbaik **d** pulau bali berupa pura **yg** sangat **d** sucikan umat hindu **d** bali	“tempat” “wisata” “terbaik” “**di**” “pulau” “bali” “berupa” “pura” “**yang**” “sangat” “**di**” “sucikan” “umat” “hindu” “**di**” “bali”

5) Stopword Removal
Pada penelitian ini *Stopword Removal* digunakan untuk menghilangkan kata sambung dan kata kata yang tidak memiliki makna signifikan terhadap sentimen. Berikut salah satu hasil *Stopword Removal*.

Tabel 7. *Stopword Removal*

Sebelum *Stopword Removal*	Setelah *Stopword Removal*
tempat yang tidak pernah bosan **untuk** dikunjungi indah **dan** sejuk	bosan, dikunjungi, indah, sejuk

6) Stemming
Stemming dilakukan untuk mengubah kata kedalam bentuk dasarnya. Dengan demikian, proses *stemming* bisa mengubah kata yang lebih panjang menjadi bentuk yang tidak selalu benar, tetapi tetap bisa merepresentasikan makna dasar dari kata tersebut Berikut salah satu hasil *stemming*.

Tabel 8. *Stemming*

Sebelum Proses *Stemming*	Setelah Proses *Stemming*
petugas, pintu, masuk, **menagih**, turis, rupiah, lokal, **mengenakan**, tarif, rupiah, **menerimanya**, turis, **membawa**, uang	**tugas** pintu masuk **tagih** turis rupiah lokal **kena** tarif rupiah **terima** turis **bawa** uang

Setelah dilakukannya proses *stemming* berarti preprocessing sudah selesai dilakukan dengan hasil data bersih yang didapatkan sejumlah 4906 komentar yang nantinya data ini akan dilakukan pembobotan dengan menggunakan TF-IDF.

3.3 Pembobotan Kata

Pembobotan kata dilakukan dengan menggunakan TF-IDF. Proses yang dilakukan pada TF-IDF ini adalah melakukan pembobotan kata. Untuk melakukan pembobotan kata tentu perlu untuk memilih dan memilah kata kata unik pada semua data *review* yang telah dilakukan *Pre-Processing* sebelumnya. yang mana dari semua *review* yang ada didapatkan 3450 kata unik yang akan dilakukan pembobotan pada masing masing *review*. Berikut Contoh beberapa hasil dari proses TF-IDF yang dapat dilihat pada Tabel 9:

Tabel 9. Contoh hasil TF-IDF

	Teks (12)	Teks (13)	Teks (19)	Teks (26)
letak	0	0	0,051757	0
lewat	0,14576	0	0	0
libur	0	0	0	0,654724
licin	0	0,253469	0	0

3.4 Pelabelan Lexicon Base

Dalam penelitian ini pelabelan dilakukan dengan menggunakan kamus kata (*Lexicon*) yang dibuat manual dan dikelompokkan berdasarkan sentimen kata itu sendiri yaitu Negatif dan Positif. Dalam kamus tersebut terdapat 406 kata yang mengandung sentimen negatif dan 420 kata yang mengandung sentiment positif. Teks yang digunakan dalam *labeling* adalah teks yang sudah dilakukan *preprocesing* sehingga kata demi kata terlihat tidak berhubungan. Tabel 10 menunjukkan bagaimana hasil dari proses Lexicon Base.

Tabel 10. Hasil *Lexicon Base*

	Teks	P	N	Sentimen	Kata (+)	Kata (-)
1	**bagus** potret awas batu injak kaki **licin** berhatihatilah **kunjung** jam sore **panas** manis	2	2	netral	bagus, kunjung	licin, panas
2	**bagus luas** parkir **luas bersih** main **layak kunjung** izin wisatawan **kunjung** kuil jelajah **nikmat** pandang **baik** lokasi laut pergi cuaca **bagus suka** alam	11	0	positif	bagus, luas, luas, bersih, layak, kunjung, kunjung, nikmat, baik, bagus, suka	
3	**tipu bayar** masuk situs tisu toilet **sulit** temu sampah	0	3	negatif		tipu, bayar, sulit

Dari hasil diatas dapat dilihat bahwa setiap teks memiliki kata yang mengandung sentimen positif atau negatif. Untuk menentukan komentar itu memiliki sentimen, maka diperlukan kamus yang memuat beragam kata positif dan negatif yang akan di komparasi dengan data yang akan diproses. Lalu pada proses ini kata yang bermakna positif akan diberi nilai 1 dan kata negatif juga akan di beri nilai 1. Nilai ini akan di jumlahkan sesuai sentimen jadi misalkan dalam 1 dokumen ada 5 kata positif maka nilainya 5 dan begitupun untuk kata negatif. Jadi untuk menentukan label berdasarkan skor maka dilakukan dengan menggunakan operator ">" (lebih dari) yang dalam kasus ini jika nilai positif

lebih besar dari nilai negatif maka menghasilkan sentimen positif, lalu jika nilai negatif lebih besar dari nilai positif maka sentimen yang dihasilkan negatif dan jika kedua kondisi tersebut tidak terpenuhi (bernilai sama) maka sentimen yang dihasilkan adalah netral.

3.5 Pembagian Data

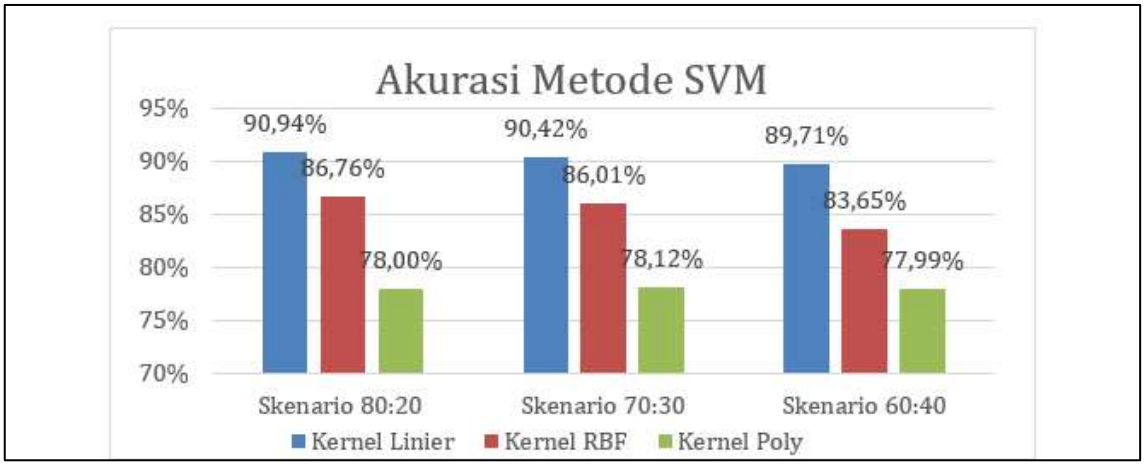
Pembagian data dalam hal ini adalah membagi data yang digunakan untuk melatih model dan data yang dibagi untuk menguji model. Sekenario pembagian data dapat dilihat pada Tabel 11 dibawah ini:

Tabel 11. Pembagian Data

Skenario	Data Latih	Data Uji
60 : 40	2943	1963
70 : 30	3434	1472
80 : 20	3924	982

3.6 Klasifikasi Support Vector Machine

Pengujian ini dilakukan dengan menggunakan 3 Kernel yaitu kernel Linear, kernel RBF (*Radial Basic Function*) dan kernel *polynomial*. Penggunaan masing - masing kernel ini di uji dengan skenario 80:20, 70:30 dan 60:40. Untuk hasil akurasi yang dihasilkan dari masing masing kernel dapat dilihat pada gambar 2.



Gambar 2. Hasil SVM pada 3 Kernel

Pada gambar diatas dapat dilihat bahwa grafik batang yang berwarna biru merupakan kernel linier yang di uji dengan 3 skenario yang menghasilkan akurasi 90,94% di skenario 80:20, lalu pada skenario 70:30 menghasilkan akurasi di 90,42% dan terakhir pada skenario 60:40 menghasilkan nilai akurasi sebesar 89,71%.

Grafik batang yang berwarna merah merupakan kernel RBF (*Radial Basic Function*). Kernel ini menghasilkan nilai akurasi 86,76% pada skenario 80:20, lalu pada skenario 70:30 menghasilkan nilai akurasi di 86,01% dan pada skenario 60:40 menghasilkan nilai akurasi di 83,65%.

Lalu untuk grafik yang berwarna hijau merupakan kernel polynomial yang dilakukan pengujian akurasi yang dimana pada skenario 80:20 kernel ini menghasilkan nilai 78% lalu pada skenario 70:30 kernel ini menghasilkan akurasi 78,12%. Dan terakhir pada skenario 60:40 akurasi yang dihasilkan hanya 77,99%.

Dapat disimpulkan pada proses ini bahwa kernel dan pembagian data uji sangat mempengaruhi akurasi dari proses klasivikasi SVM dan pada kasus ini kernel linear menghasilkan nilai terbaik dibanding dengan kernel lainnya pada 3 skenario yang uji.

3.7 Evaluasi

Pada tahap klasivikasi SVM ditunjukan bahwa adanya pengujian akurasi data. akurasi ini memberikan gambaran umum kinerja model, hasil akurasi SVM tidak cukup untuk menilai kinerja model secara menyeluruh, terutama jika terdapat pendistribusian data kelas yang tidak seimbang (misalnya, lebih banyak data di kelas positif dari pada kelas negatif). Oleh karenanya perlu evaluasi prediksi model dengan menggunakan *Confusion Matrix*.

Data yang akan digunakan pada *Confusion Matrix* adalah data uji yang berasal dari pembagian data yang dilakukan sebelumnya yaitu di 20%, 30% dan 40% dari total data yang masing masing 982, 1472, dan 1963 data. Hasil *Confusion Matrix* akan ditampilkan pada masung masing kernel dengan 3 skenario. Hasil pengujian terhadap masing masing kernel dan scenario terlihat seperti pada tabel dibawah ini.

Tabel 12. Rekap Hasil *Accuracy, Precision, Recall,* dan *F1-Score*

Pengujian *Confusion matrix*		80:20			70:30			60:40		
		Linier	RBF	Poly	Linier	RBF	Poly	Linier	RBF	Poly
Accuracy		**90,94%**	**86,76%**	**78,00%**	**90,42%**	**86,01%**	**78,12%**	**89,71%**	**83,65%**	**77,99%**
Precision	negatif	83,33%	95,40%	100,00%	84,74%	97,05%	100,00%	88,00%	95,40%	100,00%
	netral	75,73%	75,16%	69,23%	74,82%	74,49%	74,60%	74,00%	71,16%	74,35%
	positif	95,94%	88,72%	78,26%	94,61%	87,52%	78,04%	93,61%	84,91%	77,87%
Recall	negatif	48,39%	33,87%	16,12%	51,02%	33,67%	15,30%	50,39%	32,06%	17,55%
	netral	80,83%	59,58%	18,65%	76,45%	52,90%	17,02%	73,59%	41,23%	16,00%
	positif	97,25%	98,48%	99,03%	97,45%	99,00%	99,09%	97,33%	99,04%	99,24%
F1-Score	negatif	61,22%	50,00%	27,76%	63,70%	50,00%	26,53%	64,08%	48,00%	29,86%
	netral	78,20%	66,47%	29,38%	75,63%	61,86%	27,71%	73,80%	52,20%	26,33%
	positif	96,60%	93,34%	87,42%	96,00%	92,90%	87,31%	95,43%	91,43%	87,65%

Tabel di atas menunjukkan hasil Accuracy, Precision, Recall, dan F1-Score berdasarkan tiga skenario yang direncanakan. Pada skenario 80% data latih dan 20% data uji, pengujian dilakukan menggunakan tiga jenis kernel, yaitu Linier, RBF, dan Polynomial. Hasil menunjukkan bahwa kernel linier memiliki nilai accuracy tertinggi, yaitu 90,94%, dan unggul di hampir semua metrik evaluasi, kecuali pada precision kelas negatif, di mana kernel polynomial mencapai nilai 100%. Pada skenario 70% data latih dan 30% data uji, performa model sedikit menurun dibandingkan skenario 80:20, dengan accuracy tertinggi sebesar 90,42%, yang juga dicapai oleh kernel linier. Sementara itu, pada skenario 60% data latih dan 40% data uji, model menghasilkan accuracy terendah dibandingkan skenario lainnya, dengan nilai tertinggi sebesar 89,71%, yang kembali diperoleh menggunakan kernel linier.

Penggunaan kernel dalam proses klasifikasi sangat mempengaruhi performa klasifikasi yang tentu berdampak pada hasil klasifikasi. Juga pembagian data traning dan data uji berperan penting dalam klasifikasi yang dalam prosesnya semakin banyak data yang di training akan menghasilkan performa klasifikasi yang tinggi.

Berdasarkan hasil *Confusion Matrix*, dapat disimpulkan bahwa dari tiga kernel yang digunakan dalam klasifikasi, kernel linier menghasilkan nilai Accuracy, Precision, Recall, dan F1-Score rata-rata tertinggi dibandingkan kernel lainnya. Selain itu, dari tiga skenario yang diuji, skenario 80:20

menunjukkan skor rata-rata tertinggi. Dengan demikian, kombinasi kernel linier dan skenario 80:20 terbukti paling efektif dalam mengklasifikasikan sentimen ke dalam kelas positif, negatif, dan netral.

3.8 Visualisasi

1. WordCloud

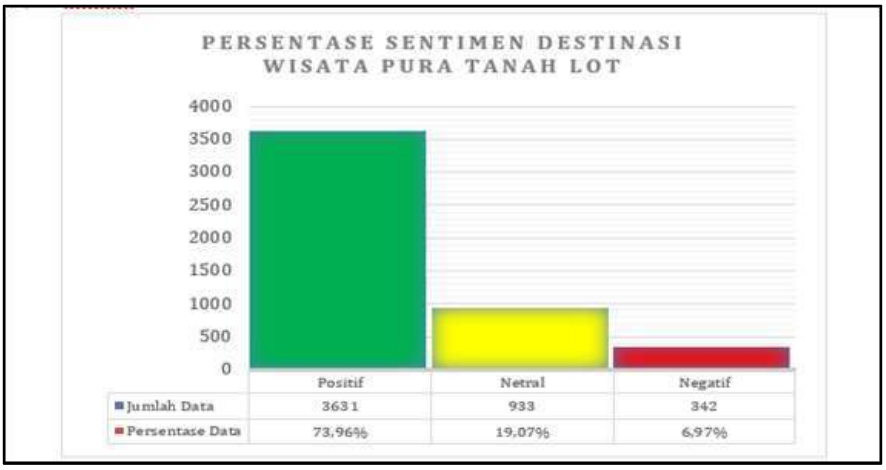
Gambar 3 WordCloud sentimen positif

Pada gambar diatas ditunjukkan bahwa terlihat kata yang sedikit lebih besar dari kata lainnya adalah "matahari", "benam " dan "indah" yang dalam konteks komentarnya bahwa banyak turis yang datang ke tanah lot memuji indahnya pemandangan ketika matahari terbenam. Selain itu ada beberapa kata yang memiliki sentiment positif namun terlihat kecil dibanding yang lainnya seperti "unik", "baik", layak kunjung", "suka" dan lain lain. Ini dikarenakan kata tersebut tidak cukup sering muncul di banding dengan kata lainnya yang terlihat lebih besar dalam visualisasinya.

Gambar 4. WordCloud sentimen negatif

Pada gambar diatas terlihat dengan jelas kata yang paling besar adalah "bayar" yang menunjukan bahwa keluhan dari para turis adalah terkait dengan kata "bayar" baik itu terkait tiket masuk, toilet, pelayanan dan lain lain. Selain itu terlihat ada beberapa kata yang muncul cukup besar yang tidak mengandung sentimen seperti "bali" "pandang", "kuil", "kunjung", "tanah lot" dan lain lain. Ini dikarenakan pada prosesnya data yang di tampilkan di wordcloud itu diambil dari seluruh teks yang memiliki label sentimen yang yang ingin ditampilkan, contoh jika ingin menampilkan WordCloud dengan sentimen positif maka yang diambil adalah seluruh data yang mengandung sentimen positif. Jika ada 100 data memiliki label positif maka seluruh kata yang ada pada 100 data tersebut akan ditampilkan di WordCloud yang sehingga ini menyebabkan beberapa kata yang tidak mengandung sentimen muncul di WordCloud.

2. Grafik



Gambar 5. Persentase Sentimen Destinasi Wisata Pura Tanah Lot

Total data yang diperoleh dalam penelitian ini adalah 4906 data. Pada proses labeling 3631 data kenali sebagai sentimen positif, lalu 342 data dikenali sebagai sentimen negatif dan untuk 933 data komentar ini di kenali sebagai sentimen netral. Dari diagram diatas dapat disimpulkan bahwa 73.96% dari total komentar memiliki sentimen positif. Ini menunjukkan bahwa mayoritas orang memberikan tanggapan yang baik atau mendukung terhadap topik yang dibahas yaitu Pura Tanah Lot. 6.97% dari total komentar memiliki sentimen negatif. Artinya, terdapat sebagian kecil orang yang memberikan tanggapan negatif atau tidak setuju terhadap topik yang dibahas. Dan yang terakhir persentase 19.07% dari sentimen netral ini menunjukkan bahwa ada sebagian orang yang memberikan tanggapan yang tidak terlalu positif maupun negatif, atau mungkin tidak memiliki pendapat yang kuat mengenai topik tersebut.

4. Kesimpulan

Penelitian ini membuktikan bahwa metode *Support Vector Machine* (SVM) efektif dalam melakukan klasifikasi sentimen ulasan pengunjung Pura Tanah Lot. Penggunaan kernel yang berbeda, yaitu linear, RBF, dan polynomial, memberikan hasil yang bervariasi. Kernel linear menunjukkan performa terbaik pada skenario pembagian data latih dan uji 80:20, dengan akurasi sebesar 90,94%. Pemilihan kernel dan skenario pembagian data berpengaruh signifikan terhadap hasil klasifikasi. Pada kasus ini kernel dan scenario terbaik didapatkan dengan menggunakan kernel linier dengan skenerio data latih di 80% dan data uji di 20%.

Hasil analisis sentimen menunjukkan bahwa mayoritas ulasan bersentimen positif (73,96%), mencerminkan pengalaman pengunjung yang cenderung baik. Sentimen netral mencapai 19,07%, menunjukkan adanya ulasan deskriptif tanpa kecenderungan emosional tertentu, sedangkan sentimen negatif hanya 6,97%, yang mengindikasikan pengalaman kurang menyenangkan relatif jarang terjadi. Temuan ini menegaskan bahwa metode SVM dapat menjadi alat yang efektif dalam analisis sentimen di sektor pariwisata. Pemahaman terhadap sentimen pengunjung dapat membantu pengelola Pura Tanah Lot dalam meningkatkan kualitas layanan, memahami kebutuhan wisatawan, serta merumuskan strategi pengembangan destinasi wisata secara lebih tepat.

Daftar Pustaka

A. N. P. Viona Rosalia Berti. (2023). Pura Agung Besakih Temple As A Historical Tourist Attraction In Karangasem Regency, Rendang District. *J. Nirwasita, vol. 4, no. 1, p.* 89, doi: 10.5281/zenodo.7781542.

A. Sinaga and S. P. Nainggolan. (2023). Analisis Perbandingan Akurasi Dan Waktu Proses Algoritma Stemming Arifin-Setiono Dan Nazief-Adriani Pada Dokumen Teks Bahasa Indonesia. *Sebatik, vol. 27, no. 1, pp.* 63–69, doi: 10.46984/sebatik.v27i1.2072.

Badan Pusat Statistik Provinsi Bali. (2023). Statistik Wisatawan Mancanegara ke Provinsi Bali 2022.

C. Zai and A. R. Isnain. (2024). Komparasi Algoritma NaÃ¯ve Bayes dan Support Vector Machine (SVM) pada Analisis Sentimen Capcut. *INOVTEK Polbeng - Seri Inform., vol. 9, no. 1, pp.* 272–285, doi: 10.35314/isi.v9i1.4054.

D. Alita and A. R. Isnain. (2020). Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier. *J. Komputasi, vol. 8, no. 2, pp.* 50–58, doi: 10.23960/komputasi.v8i2.2615.

D. Septiani and I. Isabela. (2022). Analisis Term Frequency Inverse Document Frequency (Tf-Idf) Dalam Temu Kembali Informasi Pada Dokumen Teks. *SINTESIA J. Sist. dan Teknol. Inf. Indones., vol. 1, no. 2, pp.* 81–88.

F. D. Djamil and A. Sulistyo. (2023). Implementasi Virtual Reality Dalam Pemasaran Pariwisata. *J. Inf. Syst. Manag., vol. 5, no. 1, pp.* 33–39, doi: 10.24076/joism.2023v5i1.1084.

F. Indra, E. Suryadi, G. Blessing Gosal Manopo, M. Lee, N. Penulis Korespondensi, and F. Indra. (2023). Etika Profesi Pariwisata Yang Perlu Dimiliki Oleh Sdm Pada Masa Revolusi 4.0: Teknologi Yang Mereduksi Manusia Sebagai Tenaga Kerja. *J. Bangun Manaj., vol. 2, no. 1, pp.* 2830–1862, doi: 10.56854/jbm.v2i1.203.

F. L. Hakim and K. E. Dewi. (2024). KOMPUTA : Jurnal Ilmiah Komputer dan Informatika PRODUK KECANTIKAN MENGGUNAKAN NEIGHBOR WEIGHTED K-NEAREST NEIGHBOR KOMPUTA : *Jurnal Ilmiah Komputer dan Informatika. vol. 13, no. 1, pp.* 1–10.

F. O. Awalullaili, D. Ispriyanti, and T. Widiharih. (2023). Klasifikasi Penyakit Hipertensi Menggunakan Metode Svm Grid Search Dan Svm Genetic Algorithm (Ga). *J. Gaussian, vol. 11, no. 4, pp.* 488–498, doi: 10.14710/j.gauss.11.4.488-498.

H. Herlawati, R. T. Handayanto, P. D. Atika, F. N. Khasanah, A. Y. P. Yusuf, and D. Y. Septia. (2021). Analisis Sentimen Pada Situs Google Review dengan Naïve Bayes dan Support Vector Machine. *J. Komtika (Komputasi dan Inform., vol. 5, no. 2, pp.* 153–163, doi: 10.31603/komtika.v5i2.6280.

H. Syah and A. Witanti. (2022). Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm). *J. Sist. Inf. dan Inform.,* vol. 5, no. 1, pp. 59–67, doi: 10.47080/simika.v5i1.1411.

I. M. N. O. Mahardika. (2020). Literatur Review : Destination Image and Revisit Intention. pp. 54–67.

I. T. Julianto. (2022). Analisis Sentimen Terhadap Sistem Informasi Akademik Institut Teknologi Garut. *J. Algoritm., vol. 19, no. 1, pp.* 449–456, doi: 10.33364/algoritma/v.19-1.1112.

I. W. B. Suryawan, N. W. Utami, and K. Q. Fredlina. (2023). Analisis Sentimen Review Wisatawan pada Objek Wisata Ubud Menggunakan Algoritma Support Vector Machine. *J. Inform. Teknol. dan Sains, vol. 5, no. 1, pp.* 133–140.

Kameda Delen. (2023). Potensi Mice Sebagai Tulang Punggung Pariwisata Bali (Potential As The Backbone Of Bali Tourism). *Gemawisata J. Ilm. Pariwisata, vol. 19, no. 1, pp. 49–54*, doi: 10.56910/gemawisata.v19i1.271.

L. Liyushiana. (2024). Preferensi Wisatawan terhadap Atribut Wisata Budaya Bali. *Home Journal., vol. 6, no. 1, pp.* 1–19, doi: 10.61141/home.v6i1.411.

M. G. S. Untara and W. Supada. (2020). Eksistensi Pura Tanah Lot Dalam Perkembangan Pariwisata Budaya Di Kabupaten Tabanan. *Cultoure, vol. 1, no. 2, pp.* 186–197.

M. Muhathir, M. H. Santoso, and D. A. Larasati. (2021). Wayang Image Classification Using SVM Method and GLCM Feature Extraction. *J. Informatics Telecommun. Eng., vol. 4, no. 2, pp.* 373–382, doi: 10.31289/jite.v4i2.4524.

N. P. Husain, S. Sukirman, and S. SAJIAH. (2024). Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine. *J. Syst. Comput. Eng., vol. 5, no. 1, pp.* 91–102, doi: 10.61628/jsce.v5i1.1105.

N. S. Fathullah, Y. A. Sari, and P. P. Adikara. (2020). Analisis Sentimen Terhadap Rating dan Ulasan Film dengan menggunakan Metode Klasifikasi Naïve Bayes dengan Fitur Lexicon-Based. *J. Pengemb. Teknol. Inf. dan Ilmu Komput., vol. 4, no. 2, pp.* 590–593.

P. F. Nuryananda and A. Q. Al Fitriani. (2023). Permasalahan Kultural dan Pentingnya Kontekstualisasi dalam Penerapan Teknologi dalam Pengembangan Pariwisata Kampung Adat Segunung. *Khasanah Ilmu - J. Pariwisata Dan Budaya, vol. 14, no. 2, pp.* 104–114, doi: 10.31294/khi.v14i2.15931.

P. N. Fadhilah and A. D. Indriyanti. (2023). Analisis Sentimen terhadap Opini Publik Mengenai Childfree dalam Pernikahan pada Twitter Menggunakan K-Nearest Neighbor (K-NN). *J. Informatics Comput. Sci., vol. 05, pp.* 58–62.

R. Maulana, A. Voutama, and T. Ridwan. (2023). Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC. *J. Teknol. Terpadu, vol. 9, no. 1, pp*. 42–48, doi: 10.54914/jtt.v9i1.609.

R. Ramadhan, Y. A. Sari, and P. P. Adikara. (2021). Perbandingan Pembobotan Term Frequency-Inverse Document Frequency dan Term Frequency-Relevance Frequency terhadap Fitur N-Gram pada Analisis Sentimen. *J. Pengemb. Teknol. Inf. dan Ilmu Komput., vol. 5, no. 11, pp*. 5075–5079, [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/10173

S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam. (2023*)*. Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM. *MALCOM Indones. J. Mach. Learn. Comput. Sci., vol. 3, no. 2, pp*. 153–160, doi: 10.57152/malcom.v3i2.897.

S. Roiqoh, B. Zaman, and K. Kartono. (2023). Analisis Sentimen Berbasis Aspek Ulasan Aplikasi Mobile JKN dengan Lexicon Based dan Naïve Bayes. *J. Media Inform. Budidarma, vol. 7, no. 3, pp*. 1582–1592, doi: 10.30865/mib.v7i3.6194.

S. Wulandari and F. N. Hasan. (2024). Analisis Sentimen Masyarakat Indonesia Terhadap Pengalaman Belanja Thrifting Pada Media Sosial Twitter Menggunakan Algoritma Naïve Bayes. *J. Media Inform. Budidarma, vol. 8, no. 2, pp*. 768–776, doi: 10.30865/mib.v8i2.7520.

Y. F. Wijaya and A. Triayudi. (2023). Perbandingan Algoritma Klasifikasi Data Mining Pada Prediksi Penyakit Diabetes. *J. Comput. Syst. Informatics, vol. 5, no. 1, pp*. 165–174, doi: 10.47065/josyc.v5i1.4614.

Y. Winoto, R. K. Anwar, and F. I. Septian. (2024). Pariwisata Keagamaan di Indonesia: Tinjauan Literatur Sistematis dan Analisis Bibliometrik. *J. Ilm. Pariwisata, vol. 29, no. 1, p*. 48, doi: 10.30647/jip.v29i1.1745.