

Aspect-Based Sentiment Analysis of Reviews for Pandawa Beach Using Naive Bayes and SVM Methods

Made Ayu Asri Oktarini Putri ^{1*}, I Wayan Sumarjaya ^{2*}, I Gusti Ngurah Lanang Wijayakusuma ^{3*}

* Matematika, Universitas Udayana

ayuoktarini29@gmail.com ¹, sumarjaya@unud.ac.id ², lanang_wijaya@unud.ac.id ³

Article Info

Article history:

Received 2025-01-18

Revised 2025-01-29

Accepted 2025-01-30

Keyword:

Sentiment analysis,
Google Maps,
SVM,
Naive Bayes,
Pandawa Beach.

ABSTRACT

The presence of digital technology, especially online platforms such as Google Maps, has changed the way people search for information about tourist destinations, including reviews and ratings from previous visitors. Aspect-based sentiment analysis becomes a very useful tool to understand people's views and feelings towards a place or product based on the reviews given and identify aspects of interest to tourists visiting Pandawa Beach, by utilizing Naive Bayes and Support Vector Machine (SVM) methods. The main objective of this research is to identify sentiment patterns based on aspects such as attraction, accessibility, amenities, and ancillary. Data was collected and labeled according to sentiment and aspects, then processed using preprocessing techniques, extracted by bag-of-words method, and chi-square feature selection. The model evaluation results showed that SVM produced the highest F1-Score value of 79,625%, while the Naive Bayes method reached 73.29%.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Di Indonesia, pariwisata menjadi salah satu sumber pendapatan yang cukup penting, terutama di Bali, salah satu tempat wisata terkenal di dunia [1]. Pada tahun 2023, Bandara Ngurah Rai mencatat 5,25 juta kunjungan wisatawan mancanegara, menunjukkan daya tarik yang terus meningkat [2]. Pantai merupakan salah satu daya tarik utama yang berhasil diidentifikasi sebagai faktor penting bagi wisatawan mancanegara untuk berkunjung ke Bali, di samping delapan elemen menarik lainnya yang telah diteliti [3]. Di antara berbagai destinasi di Bali, Pantai Pandawa menonjol dengan keindahan alamnya yang memukau, tebing kapur yang menjulang, serta patung-patung besar dari Pandawa Lima yang menjadi daya tarik utama. Kehadiran teknologi digital, khususnya *platform online* seperti Google Maps, telah mengubah cara orang dalam mencari informasi tentang destinasi wisata, termasuk ulasan dan penilaian dari pengunjung sebelumnya. Pengguna dapat meninggalkan ulasan dan penilaian tentang pengalaman pengunjung di suatu tempat, memberikan wawasan yang berharga bagi calon pengunjung lainnya.

Namun, meskipun popularitasnya, pemahaman tentang pengalaman dan sentimen pengunjung terhadap Pantai

Pandawa masih terbatas. Memiliki nilai sentimen positif, negatif atau netral adalah cara analisis sentimen menangani opini polaritas [4].

Berdasarkan penelitian yang dilakukan oleh Yutika et al. pada tahun 2021 mengenai analisis sentimen berbasis aspek pada *review female daily* menggunakan TF-IDF dan *naive Bayes*, diperoleh nilai *F1-Score* sebesar 62,81% untuk data yang diterjemahkan ke dalam bahasa Inggris kemudian ke dalam bahasa Indonesia dan tidak menggunakan *stopword removal* [5]. Selain itu, penelitian oleh Mustakim dan Priyanta pada tahun 2022 juga melakukan analisis sentimen berbasis aspek pada ulasan pengguna KAI *Access* dengan menggunakan metode *naive Bayes classifier* (NBC) dan *support vector machine* (SVM). Hasil pengujian menunjukkan model terbaik dari metode SVM dengan rata-rata skor akurasi 91,63%, *F1-Score* 75,55%, presisi 77,60%, dan *recall* 74,47% [6].

Dari penelitian sebelumnya, peneliti menggunakan metode *naive bayes* dan SVM dalam analisis sentimen berbasis aspek. Oleh karena itu, penggunaan analisis sentimen berbasis aspek dengan kata lain *aspect-based sentiment analysis* (ABSA) digunakan untuk mengetahui aspek yang menjadi daya tarik bagi wisatawan berkunjung ke Pantai Pandawa dengan pendekatan *naive Bayes* dan SVM. *Naive Bayes* adalah

algoritma probabilitas yang efisien dalam pengklasifikasian teks, sementara SVM adalah metode yang mencari *hyperplane* maksimum untuk membuat pemisahan maksimum antar kelas [7]. Kelebihan *naive* Bayes adalah dapat mencapai nilai akurasi yang tinggi dengan jumlah data latihan yang relatif kecil [8]. Kelemahan *naive* Bayes yaitu ketiadaan hubungan antar fitur atau independen, meskipun keterkaitan ini harus ada dan tidak dapat dimodelkan oleh *naive* Bayesian Classifier [9]. SVM bekerja dengan mencari fungsi pemisah yang optimal untuk membedakan antara kelas-kelas [10]. Kelebihan metode SVM adalah kemampuan untuk menemukan *hyperplane* yang berbeda, yang memungkinkan untuk mengoptimalkan margin dari berbagai kelas [11]. Kekurangan metode SVM termasuk masalah dengan data dengan properti yang sama, yang dapat memengaruhi nilai akurasi [12].

Penelitian ini berfokus pada ABSA menggunakan metode *naive* Bayes dan SVM (studi kasus: *review* Pantai Pandawa) dengan batasan masalah menggunakan dua kelas sentimen positif dan negatif dan empat kelas dalam klasifikasi aspek dengan menerapkan seleksi fitur *chi-square*. Seleksi fitur ini diharapkan dapat mengurangi fitur yang kurang penting dan meningkatkan performa model algoritma dalam memprediksi kelas sentimen dan aspek. Metode *naive* Bayes dan SVM akan dibandingkan untuk mengetahui algoritma mana yang lebih baik dalam klasifikasi sentimen serta mengetahui aspek berdasarkan sentimen.

Tujuan dari analisis sentimen berbasis aspek adalah untuk mengidentifikasi sentimen yang terkait dengan entitas atau aspeknya. Penelitian ini menjadi relevan mengingat dampaknya yang besar bagi industri pariwisata di pulau Bali. Informasi mengenai sentimen pengunjung dapat memberikan wawasan berharga bagi pemerintah daerah, pemangku kepentingan pariwisata, dan pengelola destinasi wisata untuk meningkatkan kualitas layanan, mengatasi masalah yang mungkin timbul, dan mengoptimalkan pemasaran destinasi wisata.

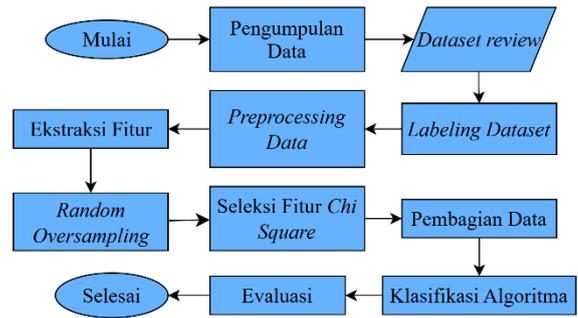
II. METODE

A. Tahapan Penelitian

Penelitian ini akan menggunakan metode pendekatan kualitatif deskriptif dengan mengumpulkan data dari ulasan pada aplikasi Google Maps. Bahasa pemrograman yang digunakan yaitu bahasa pemrograman Python 3.10.12 pada Google Colab. Dengan melakukan ABSA maka metode *naive* Bayes dan SVM mengklasifikasikan sentimen dan aspek dilakukan secara otomatis. Rancangan penelitian tersedia dalam Gambar 1.

B. Pengumpulan Data

Penelitian ini mengumpulkan data dengan *scrapping* menggunakan bahasa pemrograman Python dari ulasan Google Maps di Pantai Pandawa. Pengumpulan data menggunakan Selenium, BeautifulSoup, WebDriver Manager dan Longdetect.



Gambar 1. Diagram Alur Penelitian

Adapun populasi pada penelitian ini yaitu ABSA Pantai Pandawa Bali melibatkan semua wisatawan yang pernah mengunjungi Pantai Pandawa Bali dan memberikan ulasan tentang pengalaman pengunjung pada aplikasi Google Maps. Sedangkan sampel pada penelitian ini yaitu sebagian dari populasi yang dipilih dari ABSA Pantai Pandawa Bali adalah wisatawan yang pernah mengunjungi dan memberikan ulasan tentang pengalaman pengunjung Pantai Pandawa dalam bahasa Indonesia.

C. Pelabelan Data

Setelah data dikumpulkan, label dapat diberikan pada setiap ulasan. Pada penelitian ini hanya menggunakan dua kelas sentimen yaitu positif dan negatif. Pelabelan data berdasarkan sentimen dengan jumlah bintang. Bintang satu, dua dan tiga menunjukkan sentimen negatif, sementara bintang empat dan lima menunjukkan sentimen positif. Selanjutnya akan dilakukan pemberian label berdasarkan aspek secara manual berdasarkan aspek menurut *attraction* (daya tarik), *accessability* (aksesibilitas), *amenities* (amenitas), dan *ancillary* (pelayanan tambahan) [13]. Karena pelabelan aspek dilakukan secara manual, maka terlebih dahulu harus mengidentifikasi kata kunci untuk setiap aspek dengan indikator pada variabel aspek mengenai faktor yang dapat memengaruhi perkembangan pariwisata pantai [14]. Berikut daftar kata kunci dalam identifikasi aspek pada tabel 1.

TABEL I
IDENTIFIKASI ASPEK KATEGORI

Aspek Kategori	Kata
<i>attraction</i>	Pantai, gelombang, laut, pasir, air, pemandangan, tebing, ombak, panorama
<i>accessability</i>	jalan
<i>amenities</i>	Fasilitas, musala, restoran, kafe, warung, toilet, hotel, parkir, penginapan, kano, selancar, menyelam
<i>ancillary</i>	Sekuriti, penjaga, pedagang, penjual

D. Preprocessing Data

Preprocessing adalah proses menyediakan data mentah untuk analisis atau penggunaan model pembelajaran mesin [15]. Tujuan dari tahap persiapan *preprocessing* data adalah untuk mendapatkan data yang bersih dan siap digunakan untuk penelitian [16]. Mengurangi jumlah teks yang dihapus memungkinkan proses selanjutnya dalam menambang

informasi dari dokumen atau teks. Tahapan ini mencakup beberapa langkah:

1) *Data cleansing*

Proses ini melakukan pembersihan data dari berbagai *noise* seperti data kosong, mengecek duplikat, melakukan penghapusan tanda baca, *emoticon*, angka, *hashtag*, *mention*, dan karakter,

2) *Case folding*

Proses menggabungkan semua teks dalam kumpulan data menjadi huruf kecil. Tujuannya untuk mengurangi variasi dalam data.

3) *Normalisasi*

Proses ini dengan mengubah kata-kata pada ulasan Google Maps yang tidak baku atau kata-kata cakapan (slang) menjadi kata-kata yang baku sesuai dengan Kamus Besar Bahasa Indonesia. Proses normalisasi menggunakan daftar bahasa alay dan perbaikan kata dalam formal kata yang dikembangkan oleh Salsabila et al [17].

4) *Stopwords removing*

Proses ini untuk mengeliminasi kata-kata yang dianggap tidak memberikan kontribusi signifikan terhadap makna kalimat seperti kata penghubung [18]. Proses ini berdasarkan daftar *stopword* yang dikembangkan oleh Tala [16].

5) *Stemming*

Proses ini mengubah kata-kata yang ada dalam teks menjadi bentuk dasar atau akarnya. Tujuan dari *stemming* adalah menghilangkan imbuhan dan berpengaruh terhadap pengurangan jumlah fitur.

E. *Ekstraksi Fitur*

Metode ekstraksi fitur yang digunakan adalah *bag-of-words* (BoW). Model BoW mengumpulkan kosakata dari semua dokumen dan kemudian menghitung berapa kali setiap kata muncul untuk memodelkan setiap dokumen [19]. Proses ekstraksi fitur ini akan memastikan bahwa data yang digunakan dalam analisis sentimen disajikan dengan cara yang jelas dan memberikan informasi yang bermanfaat [20].

F. *Random Oversampling*

Random oversampling (ROS) adalah teknik yang melibatkan penambahan data dari kelas minoritas ke dalam data training secara acak [21]. Proses *oversampling* diperlukan ketika dalam pengolahan data terdeteksi adanya ketidakseimbangan yang signifikan antara kelas-kelas yang akan digunakan sebagai data latih dalam penelitian ini [22].

G. *Chi-Square*

Seleksi fitur *chi-square* memanfaatkan prinsip-prinsip statistika untuk menguji apakah suatu *term* bersifat independen terhadap kategorinya. Seleksi fitur yang didasarkan pada teori statistika melibatkan dua peristiwa, yaitu munculnya fitur dan munculnya kategori [23]. Dalam proses ini, peneliti memutuskan untuk mempertahankan 80% dari jumlah fitur asli berdasarkan penelitian Pramesti et al [24].

H. *Naïve Bayes*

Metode klasifikasi *naïve Bayes* adalah pendekatan yang didasarkan pada teorema Bayes [25]. Metode klasifikasi *naïve Bayes* melalui dua langkah dalam proses klasifikasi teks, yaitu tahap pelatihan dan tahap klasifikasi. Pada tahap pelatihan, dilakukan pengolahan terhadap sampel data yang diupayakan untuk menjadi representasi yang baik dari keseluruhan data. Selanjutnya, probabilitas prior untuk setiap kategori ditentukan berdasarkan sampel tersebut. Pada tahap klasifikasi, nilai kategori dari suatu data ditentukan berdasarkan *term* yang muncul dalam data yang sedang diklasifikasikan [26].

I. *Support Vector Machine*

Pada penelitian ini, analisis berdasarkan sentimen menggunakan klasifikasi biner dalam metode SVM dengan mengamsusikan dapat dipisahkan secara linear. SVM adalah sistem pembelajaran yang memanfaatkan ruang hipotesis yang berbentuk fungsi linear dalam ruang fitur berdimensi tinggi [27]. SVM merupakan metode klasifikasi yang berfokus pada pencarian *hyperplane* optimal. Tahapan awal dalam proses SVM adalah mengubah setiap ulasan menjadi vektor fitur yang mewakili frekuensi kemunculan kata-kata dalam ulasan yang berkaitan dengan sentimen tersebut. Kemudian melatih model dengan *hyperplane* optimal yang memisahkan kedua kelas ditentukan dengan cara mengukur margin *hyperplane* dan mencari nilai maksimum dari margin tersebut [28], dan selanjutnya melakukan pengujian dengan menggunakan fungsi keputusan dengan *hyperplane* yang telah didapatkan sebelumnya dan terakhir melakukan pengambilan keputusan untuk klasifikasi dari setiap kelasnya. Untuk analisis berdasarkan aspek menggunakan *multiclass classification (One-vs-Rest)* dengan tahapan yang sama dengan kalsifikasi biner pada metode SVM.

J. *Evaluasi Model*

Penilaian kinerja model klasifikasi penting untuk menilai seberapa efektif model tersebut. Evaluasi kinerja model dapat dilakukan dengan menggunakan *confusion matrix* dan nilai AUC.

Confusion matrix yang juga disebut sebagai *error matrix*, menyajikan informasi yang membandingkan hasil klasifikasi dari sistem (model) dengan klasifikasi yang sebenarnya [29]. Kolom dalam *confusion matrix* menggambarkan hasil kelas yang diprediksi, sementara baris menunjukkan hasil kelas yang sebenarnya, sehingga dapat menghitung semua kemungkinan kasus dalam masalah klasifikasi [30]. Nilai keberhasilan klasifikasi yang diperoleh dapat dilakukan pengujian pada nilai yang diperoleh melalui penggunaan metode ini dengan empat tipe dikombinasikan secara berbeda-beda dari nilai hasil algoritma prediksi dan nilai yang didapat secara aktual [31].

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Sensitivity (recall) = \frac{TP}{TP+FN} \tag{3}$$

$$F1-Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

Selain *confusion matrix*, diperlukan metrik yang dapat mempertimbangkan keempat nilai yang terdapat dalam *confusion matrix* untuk menilai kinerja model secara keseluruhan dan proporsional pada kedua kelas, baik positif maupun negatif, yaitu *matthews correlation coefficient* (MCC) [32], yang didefinisikan sebagai berikut,

$$\frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP) \cdot (TP+FN) \cdot (TN+FP) \cdot (TN+FN)}} \tag{5}$$

Metrik ini tidak dipengaruhi oleh masalah ketidakseimbangan data, dengan nilai yang berkisar antara -1 hingga +1. Nilai -1 menunjukkan adanya kesalahan klasifikasi yang sempurna, sementara +1 menunjukkan klasifikasi yang sempurna [32].

Penentuan kinerja klasifikasi model terbaik juga dapat ditentukan dari nilai *area under the curve* (AUC) yang merupakan luas di bawah kurva *receiver operator characteristic* (ROC). ROC adalah plot kurva antara nilai *true positive rate* (TPR) atau *sensitivity* dengan nilai *false positive rate* (FPR). Nilai AUC berada dalam rentang 0 hingga 1, sehingga semakin mendekati 1, semakin baik kinerja model klasifikasi tersebut [33].

III. HASIL DAN PEMBAHASAN

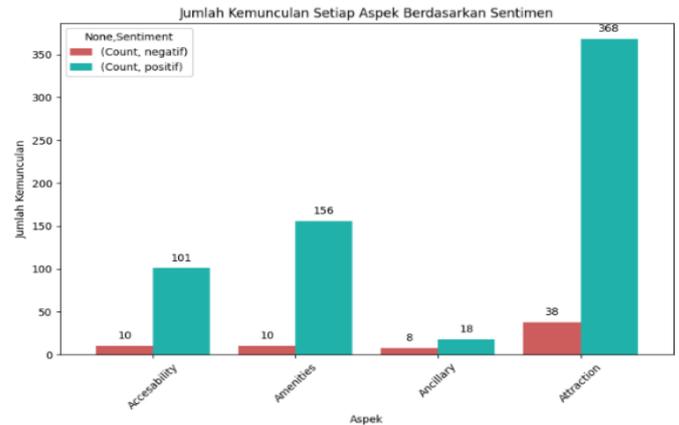
Hasil pengumpulan data sebanyak 795 ulasan dalam bahasa Indonesia yang terdiri dari 796 baris dan 4 kolom yaitu *username*, *content*, *bintang*, dan *date*. Data yang digunakan adalah kolom *content* yang berisi teks ulasan dan bintang untuk menentukan label sentimen.

TABEL II
CONTOH HASIL PENGAMBILAN DATA

Username	Content	Bintang	Date
bonardo marpaung	Sejauh mata memandang, yang ada hanya laut dan laut. Pemandangan yang luar biasa. Mengunjunginya kembali setelah 11 tahun yang lalu, fasilitas jauh lebih bagus. Dan orang sekitar masih menyambut dengan senyum dan sapaan yang ramah, khas Bali.	5 bintang	Seminggu lalu

Penelitian ini menggunakan pendekatan *supervised learning*. Pendekatan *supervised learning* adalah pendekatan yang bergantung pada data latih yang diberi label atau kelas [34]. Salah satu algoritma yang digunakan dalam analisis sentimen pada penelitian ini adalah *naïve Bayes* dan SVM, sehingga data yang dikumpulkan kemudian diberi label

berdasarkan aspek dan sentimennya. Gambaran distribusi pemberian label dapat dilihat pada Gambar 2.



Gambar 2. Jumlah Kemunculan Setiap Aspek Berdasarkan Sentimen

Berdasarkan analisis terhadap empat aspek penting dalam pengembangan Pantai Pandawa, dapat disimpulkan bahwa sektor pariwisata di lokasi ini memiliki potensi yang besar untuk berkembang. Aksesibilitas yang baik, dengan 101 respon positif, menunjukkan bahwa pengunjung merasa mudah untuk mencapai pantai, yang merupakan faktor penting dalam menarik lebih banyak wisatawan. Selain itu, fasilitas yang memadai, yang mendapatkan 156 respon positif, juga berkontribusi pada kenyamanan pengunjung, meskipun masih terdapat peluang untuk meningkatkan layanan tambahan, yang hanya memperoleh 18 respon positif. Daya tarik pantai yang sangat tinggi, dengan 368 respon positif, menegaskan bahwa Pantai Pandawa menawarkan banyak hal menarik. Oleh karena itu, pengembangan lebih lanjut pada aspek layanan tambahan dan peningkatan fasilitas akan sangat penting untuk meningkatkan pengalaman pengunjung, memperkuat daya tarik, dan mendukung pertumbuhan sektor pariwisata yang berkelanjutan di Pantai Pandawa.

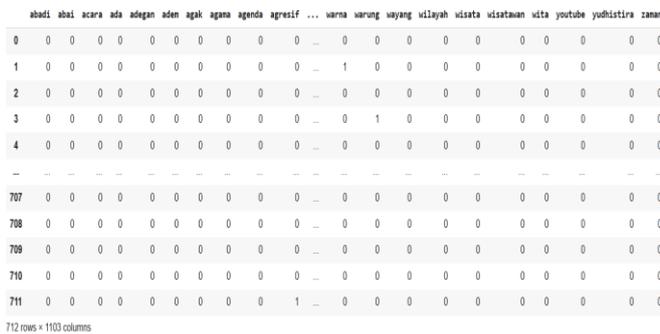
Data yang telah dilabeli berdasarkan aspek dan sentimen akan melalui tahap *preprocessing* untuk mempermudah proses analisis yang dapat dilihat pada Tabel 3.

TABEL III
HASIL PREPROCESSING DATA

Proses	Hasil Proses
Data	Pantai yg cantik sebaiknya dtg kesini bs lebih pagi karna panas sekali, tiket masuk 8k parkir 2k, makanan di warung pinggir pantai jg harga msh wajar okelah pokoknya 🍷 ...
Data cleansing	Pantai yg cantik sebaiknya dtg kesini bs lebih pagi karna panas sekali tiket masuk parkir makanan di warung pinggir pantai jg harga msh wajar okelah pokoknya
Case folding	pantai yg cantik sebaiknya dtg kesini bs lebih pagi karna panas sekali tiket masuk parkir makanan di warung pinggir pantai jg harga msh wajar okelah pokoknya
Normalisasi	pantai yang cantik sebaiknya datang kesini bisa lebih pagi karena panas sekali tiket masuk parkir

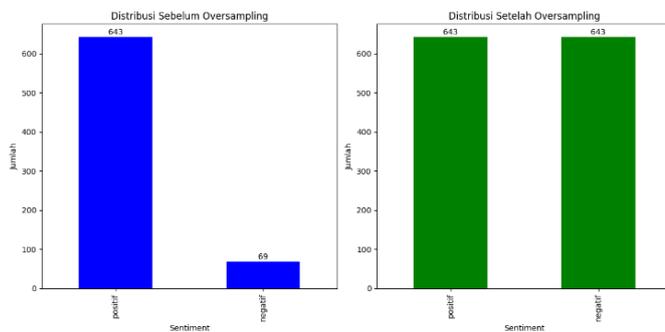
	makanan di warung pinggir pantai juga harga masih wajar ok pokoknya
<i>Stpowords removing</i>	Pantai cantik datang pagi panas tiket masuk parkir makanan warung pinggir pantai harga wajar pokoknya
<i>Stemming</i>	Pantai cantik datang pagi panas tiket masuk parkir makan warung pinggir pantai harga wajar pokok

Data hasil *preprocessing* akan dilakukan tahap ekstraksi fitur dengan metode BoW yaitu menghitung frekuensi kemunculan kata dalam setiap dokumen. Setelah menerapkan metode BoW, jumlah fitur yang diperoleh adalah 1.103 fitur.

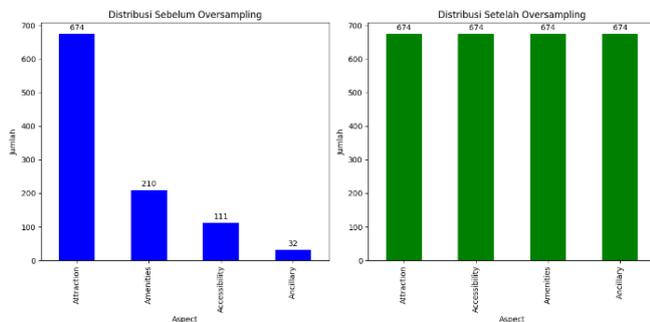


Gambar 3. Hasil Ekstraksi Fitur

Dalam penelitian ini, terdapat masalah distribusi data yang tidak seimbang, sehingga diperlukan langkah *random oversampling* untuk menduplikasi data minoritas hingga mencapai keseimbangan.



Gambar 4. Distribusi *Oversampling* Berdasarkan Sentimen



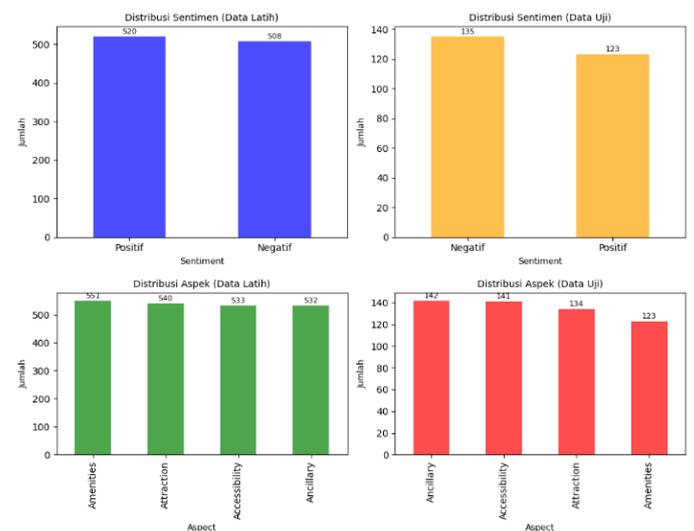
Gambar 5. Distribusi *Oversampling* Berdasarkan Aspek

Selanjutnya, dilakukan seleksi fitur menggunakan metode *chi-square*, yang mempertahankan 80% dari jumlah fitur asli, sehingga menghasilkan total 882 fitur yang tersisa untuk analisis lebih lanjut.

TABEL IV
HASIL CHI-SQUARE

<i>Feature</i>	<i>Chi</i> ²	<i>p-value</i>
indah	69,8505	$6,40 \times 10^{-17}$
putih	58,67442	$1,86 \times 10^{-14}$
luas	56	$7,25 \times 10^{-14}$
ombak	53,76744	$2,26 \times 10^{-13}$
pandawa	50,9887	$9,29 \times 10^{-13}$
biaya	48,48993	$3,32 \times 10^{-12}$
parkir	47,62791	$5,15 \times 10^{-12}$
jernih	47	$7,10 \times 10^{-12}$
tenang	44	$3,28 \times 10^{-11}$
buruk	41,08889	$1,45 \times 10^{-10}$

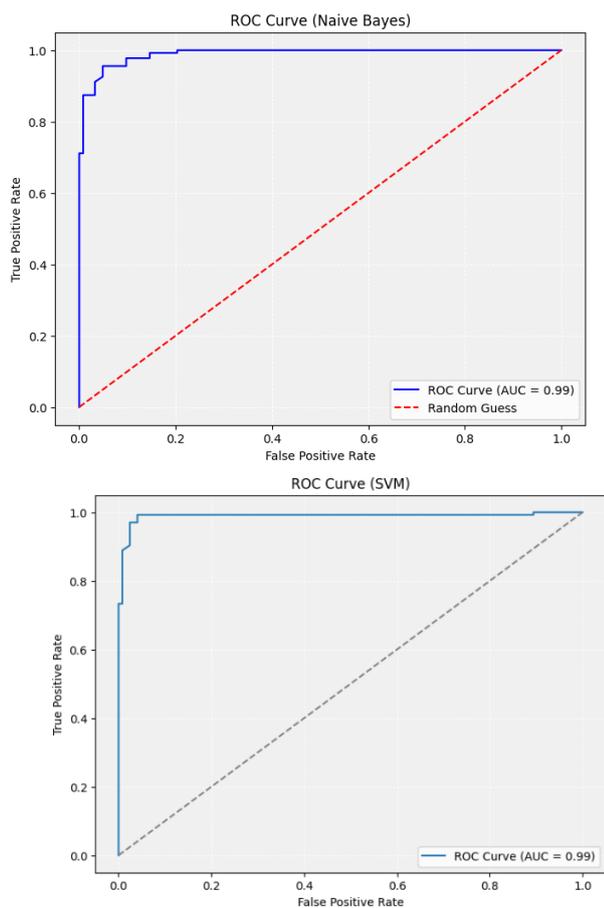
Setelah melalui tahapan seleksi fitur, *dataset* yang telah diproses kemudian dibagi menjadi dua bagian dari total data untuk keperluan pelatihan model klasifikasi. Pembagian data ini dilakukan dengan memisahkan 80% sebagai data latih, dan 20% sebagai data uji.



Gambar 6. Pembagian Data

A. Hasil Berdasarkan Sentimen

Analisis berdasarkan sentimen ini terdiri dari dua kelas, yaitu positif dan negatif. Berikut kurva ROC dari kedua metode dapat dilihat pada Gambar 7.



Gambar 7. Kurva ROC Model *naive* Bayes dan SVM Berdasarkan Sentimen

Hasil perhitungan evaluasi model dapat dilihat pada Tabel 5, yang memberikan gambaran tentang kinerja model dalam mengklasifikasikan berdasarkan aspek.

TABEL V
HASIL ANALISIS SENTIMEN

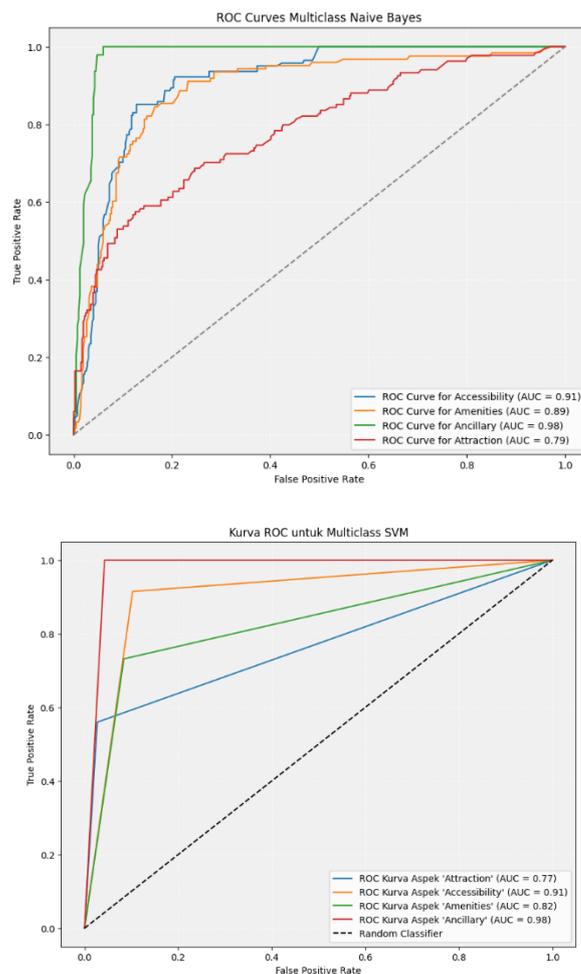
Model	Accuracy	Precision	Recall	F1-Score	MCC	AUC
Naive Bayes	93,02%	96,80%	89,63%	93,08%	86%	99%
SVM	93,37%	99,17%	88,15%	93,36%	87%	99%

Hasil evaluasi kinerja model klasifikasi menunjukkan bahwa kedua algoritma, *naive* Bayes dan SVM, memiliki performa yang sangat baik dalam hal akurasi dan metrik lainnya. *naive* Bayes mencatat akurasi sebesar 93,02%, dengan *precision* yang tinggi mencapai 96,80% dan *recall* sebesar 89,63%, menghasilkan *F1-Score* sebesar 93,08%. Sementara itu, SVM sedikit lebih unggul dengan akurasi 93,37%, *precision* yang sangat tinggi sebesar 99,17%, dan *recall* 88,15%, yang menghasilkan *F1-Score* 93,36%. Meskipun SVM memiliki *precision* yang lebih baik, *naive* Bayes menunjukkan keseimbangan yang baik antara *precision* dan *recall*. Kedua model juga menunjukkan nilai

matthews correlation coefficient (MCC) yang cukup tinggi, yaitu 86% untuk *naive* Bayes dan 87% untuk SVM, serta nilai AUC yang mencapai 99% untuk keduanya, menandakan bahwa kedua model memiliki kemampuan yang sangat baik dalam membedakan antara kelas positif dan negatif. Secara keseluruhan, hasil ini menunjukkan bahwa kedua algoritma dapat diandalkan untuk aplikasi klasifikasi berdasarkan sentimen.

B. Hasil Berdasarkan Aspek

Analisis berdasarkan aspek ini terdiri dari empat kelas, yaitu *accessibility*, *amenities*, *ancillary*, dan *attraction*. Berikut kurva ROC dari kedua metode terdapat pada Gambar 8.



Gambar 8. Kurva ROC Model *naive* Bayes dan SVM Berdasarkan Aspek

Hasil perhitungan evaluasi model dapat dilihat pada Tabel 6, yang memberikan gambaran tentang kinerja model dalam mengklasifikasikan berdasarkan aspek, dengan keterangan *Accessibility* (A1), *Amenities* (A2), *Ancillary* (A3), *Attraction* (A4).

TABEL VI
HASIL PERHITUNGAN UNTUK MULTI-LABEL METODE NAÏVE BAYES

Aspect	Accuracy	Precision	Recall	F1-Score	MCC	AUC
A1	86,3%	70,06%	82,98%	75,8%	67%	91%
A2	85,56%	68,59%	67,57%	67,99%	59%	89%
A3	91,67%	86,87%	97,89%	92,88%	89%	98%
A4	81,85%	69,57%	47,76%	56,5%	47%	79%
Average	86,345%	73,77%	74,05%	73,29%		

Hasil evaluasi kinerja model klasifikasi berdasarkan empat aspek menunjukkan perbedaan yang mencolok dalam performa masing-masing. Aspek *ancillary* mencatat hasil terbaik dengan akurasi 91,67%, *precision* 86,87%, dan *recall* yang sangat tinggi mencapai 97,89%, menghasilkan F1-Score 92,88% dan MCC 89%, yang menunjukkan efektivitas model dalam mengidentifikasi kelas positif. Sementara itu, aspek *accessibility* memiliki akurasi 86,3% dan F1-Score 75,8%, tetapi dengan *precision* yang lebih rendah di 70,06%, menunjukkan adanya ruang untuk perbaikan. Aspek *amenities* menunjukkan kinerja yang lebih rendah dengan akurasi 85,56% dan F1-Score 67,99%, serta *precision* dan *recall* yang tidak terlalu tinggi, mencerminkan tantangan dalam mengidentifikasi kelas positif. Aspek *attraction* memiliki performa terendah dengan akurasi 81,85% dan *recall* yang sangat rendah di 47,76%, menunjukkan kesulitan dalam mendeteksi kelas positif. Rata-rata keseluruhan untuk semua aspek menunjukkan akurasi 86,345% dan F1-Score 73,29%, yang menunjukkan bahwa meskipun ada variasi dalam kinerja, model secara keseluruhan masih menunjukkan hasil yang cukup baik. Namun, perhatian perlu diberikan pada aspek-aspek dengan kinerja lebih rendah, terutama *amenities* dan *attraction*, agar model dapat ditingkatkan efektivitasnya di masa depan.

TABEL VII
HASIL PERHITUNGAN UNTUK MULTI-LABEL METODE SVM

Aspect	Accuracy	Precision	Recall	F1-Score	MCC	AUC
A1	90,2%	75,9%	91,5%	83,2%	90%	91%
A2	87,4%	72%	73,2%	72,6%	82%	82%
A3	96,9%	89,2%	100%	94,4%	97%	98%
A4	87%	87,2%	55,9%	68,3%	76%	77%
Average	90,375%	81,075%	80,15%	79,625%		

Hasil analisis pada tabel 7, evaluasi kinerja model klasifikasi berdasarkan empat aspek menunjukkan variasi yang signifikan dalam performa masing-masing. Aspek *ancillary* mencatat hasil terbaik dengan akurasi 96,9%, *precision* 89,2%, dan *recall* yang sempurna mencapai 100%, menghasilkan F1-Score 94,4% dan MCC 97%. Ini menunjukkan bahwa model pada aspek *ancillary* sangat efektif dalam mengidentifikasi kelas positif. Aspek *accessibility* juga menunjukkan kinerja yang baik dengan akurasi 90,2% dan F1-Score 83,2%, meskipun *precision*-nya lebih rendah di 75,9%, yang menunjukkan adanya ruang untuk perbaikan dalam mengurangi kesalahan klasifikasi positif. Di sisi lain, aspek *amenities* memiliki akurasi 87,4%

dan F1-Score 72,6%, dengan *precision* dan *recall* yang cukup rendah, mencerminkan tantangan dalam mengidentifikasi kelas positif secara efektif. Aspek *attraction* menunjukkan performa yang lebih rendah dengan akurasi 87% dan *recall* hanya 55,9%, yang menunjukkan kesulitan dalam mendeteksi kelas positif. Rata-rata keseluruhan untuk semua aspek menunjukkan akurasi 90,375% dan F1-Score 79,625%, yang menunjukkan bahwa meskipun ada variasi dalam kinerja, model secara keseluruhan masih menunjukkan hasil yang cukup baik.

C. Pembahasan

Metode evaluasi yang akan diterapkan mencakup *accuracy*, *precision*, *recall*, dan F1-Score. Namun, nilai evaluasi yang akan dibandingkan hanya F1-Score, karena metrik ini lebih efektif untuk menangani data yang tidak seimbang dibandingkan dengan *accuracy* [5].

Berdasarkan nilai F1-Score, MCC, AUC dari kedua model, dapat disimpulkan bahwa keduanya menunjukkan kinerja yang sangat baik dalam klasifikasi. Model SVM memiliki F1-Score yang sedikit lebih tinggi (93,36%) dibandingkan *naïve* Bayes (93,08%), yang menunjukkan bahwa SVM lebih baik dalam mencapai keseimbangan antara *precision* dan *recall*. Selain itu, nilai MCC untuk SVM (87%) juga sedikit lebih tinggi dibandingkan *naïve* Bayes (86%), menandakan bahwa SVM memiliki kemampuan yang lebih baik dalam mengklasifikasikan data dengan benar, terutama dalam konteks data yang tidak seimbang. Kedua model memiliki nilai AUC yang sangat tinggi (99%), yang menunjukkan bahwa keduanya sangat efektif dalam membedakan antara kelas positif dan negatif. Secara keseluruhan, meskipun kedua model menunjukkan performa yang sangat baik, SVM sedikit lebih unggul dalam hal F1-Score dan MCC, menjadikannya pilihan yang lebih baik untuk aplikasi klasifikasi berdasarkan sentimen.

Berdasarkan analisis terhadap kedua metode berdasarkan aspek, yaitu *naïve* Bayes dan SVM, dapat disimpulkan bahwa SVM menunjukkan performa yang lebih baik dan konsisten secara keseluruhan. Meskipun *naïve* Bayes mencatat hasil terbaik pada aspek *ancillary*, dengan F1-Score sebesar 92,88%, MCC 89%, dan AUC 98%, kinerjanya mengalami penurunan yang signifikan pada aspek *attraction*, di mana F1-Score hanya mencapai 56,5%, MCC 47%, dan AUC 79%. Sebaliknya, SVM menunjukkan kinerja yang lebih stabil di semua aspek, dengan F1-Score tertinggi pada *ancillary* sebesar 94,4%, MCC 97%, dan AUC 98%, meskipun hasilnya sedikit lebih rendah pada *attraction*. Rata-rata performa *naïve* Bayes adalah 73,29%, sedangkan SVM mencapai 79,625%. Dengan demikian, SVM dapat dianggap sebagai model yang lebih unggul dan lebih dapat diandalkan dibandingkan *naïve* Bayes.

Aspek *ancillary* menunjukkan performa tertinggi dibandingkan dengan aspek lainnya. Hal ini disebabkan oleh distribusi data pada aspek *ancillary* yang lebih sedikit, sehingga menciptakan dataset yang lebih sederhana untuk dipelajari oleh model, yang dapat menyebabkan model terlalu

- [15] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *KomtekInfo*, vol. 10, no. 1, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [16] G. Radiena and A. Nugroho, "Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Kai Access Menggunakan Metode Support Vector Machine," *J. Pendidik. Teknol. Inf.*, vol. 6, no. 1, pp. 1–10, 2023, doi: 10.37792/jukanti.v6i1.836.
- [17] N. A. Salsabila, Y. A. Winatmoko, A. A. Septiandri, and A. Jamal, "Colloquial Indonesian Lexicon," *Inst. Electr. Electron. Eng.*, 2018, doi: 10.1109/IALP.2018.8629151.
- [18] M. D. Hendriyanto, A. A. Ridha, and U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine," *J. Inf. Technol. Comput. Sci.*, vol. 5, no. 1, pp. 1–7, 2022, doi: https://doi.org/10.31539/intecom.v5i1.3708.
- [19] Deepu, P. Raj, and Rajaraajeswari, "A Framework for Text Analytics using the Bag of Words (BoW) Model for Prediction," *Int. J. Adv. Netw. Appl.*, pp. 320–321, 2016.
- [20] D. Atmajaya, A. Febrianti, and H. Darwis, "Metode SVM dan Naive Bayes untuk Analisis Sentimen ChatGPT di Twitter," *Indones. J. Comput. Sci.*, vol. 12, no. 4, pp. 2173–2181, 2023, doi: 10.33022/ijcs.v12i4.3341.
- [21] S. Diantika, "Penerapan Teknik Random Oversampling Untuk Mengatasi Imbalance Class Dalam Klasifikasi Website Phishing Menggunakan Algoritma LightGBM," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 19–25, 2023, doi: https://doi.org/10.36040/jati.v7i1.6006.
- [22] A. Ahmad, W. Gata, and S. Panggabean, "Sentimen Analisis dengan Long Short-Term Memory dan Synthetic Minority Over Sampling Technic Pada Aplikasi Digital Perbankan," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 8, no. 4, pp. 973–984, 2024, doi: https://doi.org/10.35870/jti.k.v8i4.2320.
- [23] A. Z. Amrullah, A. Sofyan Anas, and M. A. J. Hidayat, "Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square," *J. BITE J. Bumigora Inf. Technol.*, vol. 2, no. 1, pp. 40–44, 2020, doi: 10.30812/bite.v2i1.804.
- [24] N. L. K. I. Pramesti, M. A. Raharja, N. A. S. ER, and I. G. A. Wibawa, "Analisis Sentimen Ulasan Aplikasi Solusi Kota Cerdas Menggunakan Algoritma Naïve Bayes dan Support Vector Machine (SVM) dengan Seleksi Fitur Chi-Square," *J. Elektron. Ilmu Komput. Udayana*, vol. 13, no. 1, pp. 1045–1058, 2024.
- [25] D. Aryanti, "Analisis Sentimen Ibukota Negara Baru Menggunakan Metode Naïve Bayes Classifier," *J. Inf. Syst. Res.*, vol. 3, no. 4, pp. 524–531, 2022, doi: 10.47065/josh.v3i4.1944.
- [26] D. Rusdaman and D. Rosiyadi, "Analisa Sentimen Terhadap Tokoh Publik Menggunakan Metode Naive Bayes Classifier dan Support Vector Machine," *J. Comput. Eng. Syst. Sci.*, vol. 4, no. 2, pp. 230–235, 2019, doi: https://doi.org/10.24114/cess.v4i2.13796.
- [27] R. Noviana and I. Rasal, "Penerapan Algoritma Naive Bayes dan SVM untuk Analisis Sentimen Boy Band BTS Pada Media Sosial Twitter," *J. Tek. dan Sci. dan Sci.*, vol. 2, no. 2, pp. 51–60, 2023, doi: https://doi.org/10.56127/jts.v2i2.791.
- [28] A. Faadhilah and H. Nugroho, "Pemetaan Daerah Rawan Longsor di Kabupaten Bandung Barat menggunakan Metode Machine Learning dengan Teknik SVM," *J. Teknol. Ramah Lingkungan*, vol. 8, no. 2, pp. 185–199, 2024, doi: https://doi.org/10.26760/jrh.v8i2.185-199.
- [29] M. A. Ramadhan and R. Andarsyah, *Klasifikasi Text Spam Menggunakan Metode Support Vector Machine Dan Naïve Bayes*. Bandung: Penerbit Buku Pedia, 2022.
- [30] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," *Inf. Sci. (Ny)*, vol. 507, pp. 772–794, 2020, doi: 10.1016/j.ins.2019.06.064.
- [31] A. S. Sinaga, S. Ramen, and S. Mulyani, "Teknik Ensemble Dalam Machine Learning Untuk Menentukan Tingkat Akurasi Perkembangan Motorik Bayi 0-12 Bulan," *SAINTIKOM*, vol. 23, no. 2, pp. 418–425, 2024, doi: https://doi.org/10.53513/jis.v23i2.10059.
- [32] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 1–13, 2020, doi: 10.1186/s12864-019-6413-7.
- [33] M. Bekkar, H. K. Djemaa, and T. A. Alitouche, "Evaluation Measures for Models Assessment over Imbalanced Data Sets," *J. Inf. Eng. Appl.*, vol. 3, no. 10, pp. 27–38, 2013.
- [34] M. Cindo, D. P. Rini, and Ermatita, "Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) Literatur Review: Metode Klasifikasi Pada Sentimen Analisis," *SAINTEKS 2019*, pp. 66–70, 2019.