Analisis Sentimen Ulasan Pengguna Tentang Pantai Widarapayung Pada Google Maps Menggunakan Algoritma Naïve Bayes

1st Muhammad Alfarissi Nurdin
Sistem Informasi, Fakultas Rekayasa Industri
Universitas Telkom Purwokerto
Purwokerto, Indonesia
alfarissin@student.telkomuniversity.ac.id

2nd Sena Wijayanto
Sistem Informasi, Fakultas Rekayasa Industri
Universitas Telkom Purwokerto
Purwokerto, Indonesia
senawijayanto@telkomuniversity.ac.id

Abstrak — Pantai Widarapayung di Kabupaten Cilacap merupakan destinasi wisata unggulan yang mengalami peningkatan jumlah pengunjung setiap tahunnya. Di era digital, ulasan pengguna di platform seperti Google Maps menjadi sumber informasi penting dalam menilai kualitas suatu tempat wisata. Seiring dengan meningkatnya jumlah ulasan membuat analisis manual terhadap ulasan menjadi tidak efektif dan tidak efisien. Oleh karena itu, diperlukan metode otomatis untuk mengolah dan memahami opini publik secara sistematis. Penelitian ini bertujuan untuk mengimplementasikan analisis sentimen terhadap ulasan pengguna tentang Pantai Widarapayung menggunakan algoritma *Multinomial Naïve Bayes* untuk mendukung pengambilan keputusan berbabis data secara efektif dan efisien. Data dikumpulkan melalui web scraping, kemudian diproses melalui tahapan preprocessing, pembobotan TF-IDF, dan penyeimbangan data dengan SMOTE. Hasil klasifikasi menunjukkan akurasi model sebesar 84,22% dengan *F1-score* tertinggi pada sentimen positif (88,86%). Mayoritas ulasan menunjukkan sentimen positif, namun aspek negatif terkait kebersihan dan fasilitas juga teridentifikasi. Temuan ini memberikan wawasan berbasis data yang penting bagi pengelola dalam perbaikan layanan.

***Kata kunci*— analisis sentimen, pantai widarapayung, *naïve bayes*, google maps, TF-IDF**

I. PENDAHULUAN

Pantai Widarapayung di Kabupaten Cilacap merupakan salah satu destinasi wisata bahari yang menawarkan keindahan alam dan kebersihan yang masih terjaga, sehingga menarik minat banyak wisatawan [1]. Peningkatan jumlah pengunjung setiap tahun menjadikan pantai ini sebagai aset pariwisata yang perlu dikelola secara berkelanjutan [2]. Dalam merencanakan kunjungan, wisatawan cenderung mencari referensi melalui ulasan digital seperti yang tersedia di Google Maps, yang menyajikan informasi berupa komentar, foto, dan penilaian pengguna [3], [4]. Namun, banyaknya ulasan yang tersedia membuat analisis secara manual menjadi tidak efisien [5].

Untuk mengatasi permasalahan tersebut, analisis sentimen dapat digunakan sebagai pendekatan otomatis dalam mengelompokkan opini pengguna ke dalam sentimen positif, negatif, atau netral [6]. Salah satu algoritma yang banyak digunakan adalah Naïve Bayes, karena kemampuannya dalam menangani frekuensi kata secara efisien serta kemudahan implementasinya [7], [8]. Beberapa studi sebelumnya juga menunjukkan bahwa algoritma ini mampu memberikan hasil akurasi yang tinggi, baik dalam topik COVID-19 [9] dan juga ulasan wisata pantai [10]. Berdasarkan hal tersebut, penelitian ini dilakukan untuk menganalisis sentimen ulasan pengguna terhadap Pantai Widarapayung melalui Google Maps menggunakan algoritma Naïve Bayes, guna memberikan gambaran persepsi pengunjung dan menjadi dasar pertimbangan peningkatan layanan wisata.

II. KAJIAN TEORI

A. Analisis Sentimen

Analisis sentimen merupakan teknik dalam bidang pemrosesan bahasa alami yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini atau emosi dalam suatu teks menjadi kategori sentimen positif, negatif, atau netral. Pendekatan analisis sentimen dibagi menjadi dua: pendekatan berbasis leksikon dan pendekatan berbasis machine learning. Pendekatan leksikon mengandalkan kamus opini, sedangkan pendekatan machine learning membutuhkan dataset berlabel untuk melatih model klasifikasi [6], [11].

B. Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi yang berdasarkan pada Teorema Bayes dan asumsi independensi antar fitur. Varian *Multinomial Naïve Bayes* sangat umum digunakan untuk klasifikasi teks karena kemampuannya dalam menangani data diskrit seperti frekuensi kata. Algoritma ini menghitung probabilitas suatu teks termasuk ke dalam kategori sentimen tertentu berdasarkan kemunculan

kata-kata yang relevan [12]. Berikut persamaan *Multinomial Naïve Bayes* [13]:

$$P(c|d) \propto P(c) \prod_{i=1}^{n_d} P(w_I|c) \quad (1)$$

Dimana :

$P(c|d)$: Probabilitas posterior kelas c pada dokumen d

$P(c)$: Probabilitas prior kelas c

$P(w_I|c)$: Probabilitas likelihood untuk kata pada kelas c

C. TF-IDF

Term Frequency-Inverse Document Frequency adalah metode pembobotan teks yang mengukur seberapa penting suatu kata dalam sebuah dokumen relatif terhadap kumpulan dokumen lainnya. TF menunjukkan frekuensi kata dalam dokumen, sedangkan IDF mengukur tingkat kekhususan kata di seluruh dokumen. Pembobotan TF-IDF memungkinkan algoritma klasifikasi untuk lebih fokus pada kata-kata yang relevan terhadap konteks [14]. Nilai TF dapat ditentukan dengan formula berikut:

$$\mathrm{TF}(t,d) = \frac{f_{t,d}}{N_d} \quad (2)$$

Dimana:

$f_{t,d}$: jumlah kemunculan kata t dalam dokumen d

N_d : total jumlah kata dalam dokumen d

Sedangkan persamaan TF-IDF sebagai berikut [15]:

$$\mathrm{IDF}(t) = \log \frac{\mathrm{N}}{df_t} \quad (2)$$

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

Dimana:

N : total jumlah dokumen

df_t : jumlah dokumen yang mengandung kata *t*

$\mathrm{TF}(t,d)$: jumlah kemunculan kata *t*

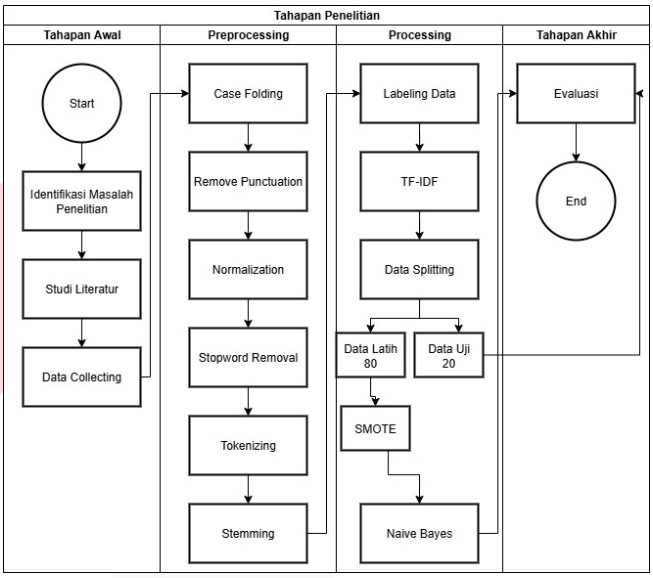
D. Google Maps

Google Maps merupakan layanan pemetaan digital yang tidak hanya menyediakan informasi lokasi dan navigasi, tetapi juga fitur ulasan pengguna berupa teks, foto, dan rating. Fitur ini memungkinkan pengguna membagikan pengalaman mereka terhadap suatu tempat, termasuk destinasi wisata. Dalam konteks penelitian, ulasan dari Google Maps menjadi sumber data yang relevan karena mencerminkan opini publik secara langsung dan dapat diolah menggunakan teknik analisis sentimen untuk mendukung pengambilan keputusan berbasis data [16]–[18].

E. SMOTE

Synthetic Minority Over-sampling Technique digunakan untuk mengatasi permasalahan ketimpangan distribusi data dalam proses klasifikasi. Metode ini bekerja dengan membuat data sintetis pada kelas yang jumlahnya lebih sedikit, sehingga model dapat belajar secara seimbang dan mengurangi kecenderungan bias terhadap kelas mayoritas, serta membantu meningkatkan akurasi model [19], [20].

III. METODE



GAMBAR 1
DIAGRAM ALIR PENELITIAN

Penelitian ini terdiri dari empat tahapan utama, yaitu tahapan awal, preprocessing, pemrosesan, dan tahapan akhir. Tahapan awal dimulai dengan identifikasi masalah penelitian, yakni perlunya analisis otomatis terhadap ulasan pengguna di Google Maps terkait Pantai Widarapayung. Setelah itu, dilakukan studi literatur untuk memperkuat landasan teori, serta pengumpulan data (data collecting) melalui teknik web scraping untuk memperoleh ulasan dalam bentuk teks.

Tahapan selanjutnya adalah preprocessing, yaitu proses pembersihan dan penyiapan data teks agar layak digunakan untuk pelatihan model. Langkah-langkah yang dilakukan meliputi case folding, remove punctuation, normalization, stopword removal, tokenizing, dan stemming.

Setelah preprocessing selesai, data memasuki tahapan processing. Pertama, dilakukan proses labeling data untuk mengelompokkan ulasan ke dalam tiga kategori sentimen: positif, negatif, dan netral. Kemudian, dilakukan pembobotan teks menggunakan metode TF-IDF untuk mengubah teks menjadi representasi numerik yang dapat dibaca oleh algoritma. Data kemudian dibagi menjadi dua bagian menggunakan teknik data splitting dengan rasio 80:20, yaitu 80% sebagai data latih dan 20% sebagai data uji. Untuk mengatasi ketidakseimbangan jumlah data antar kelas sentimen, diterapkan teknik SMOTE pada data latih. Proses klasifikasi dilakukan menggunakan algoritma *Multinomial Naïve Bayes*.

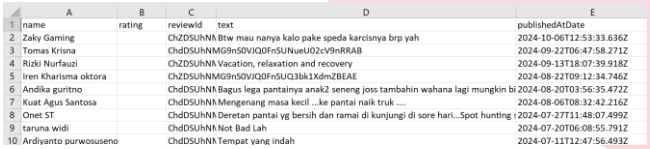
Pada tahapan akhir, model yang telah dibangun dievaluasi menggunakan metrik performa melalui *confusion*

matrix dan teknik cross-validation. Evaluasi ini dilakukan untuk menilai sejauh mana model mampu mengklasifikasikan sentimen dengan tepat, sekaligus memastikan bahwa model yang dikembangkan memiliki tingkat keandalan yang baik dalam menganalisis ulasan pengguna.

IV. HASIL DAN PEMBAHASAN

A. *Data Colecting*

Data dikumpulkan dari platform Google Maps menggunakan teknik *web scraping* pada website *Apify*. Data yang diperoleh berupa ulasan teks dari pengguna yang telah mengunjungi Pantai Widarapayung. Jumlah data yang terkumpul sebanyak 4031 ulasan. Proses ini menghasilkan dataset mentah yang siap diproses lebih lanjut dalam tahapan preprocessing. Hasil *scraping* dapat dilihat pada gambar 2.



	A	B	C	D	E
1	name	rating	reviewId	text	publishedAtDate
2	Zaky Gaming		ChZDSUhNM	Btw mau nanya kalo pake speda karcisnya brp yah	2024-10-06T12:53:33.636Z
3	Tomas Krisna		ChdDSUhNMG	9nS0VJQ0FnSUNueU02cV9nRRAB	2024-09-22T06:47:58.271Z
4	Rizki Nurfauzi		ChZDSUhNM	Vacation, relaxation and recovery	2024-09-13T18:07:39.918Z
5	Iren Kharisma oktora		ChZDSUhNMG	9nS0VJQ0FnSUQ1bk1XdmZBEAE	2024-08-22T09:12:34.746Z
6	Andika guritno		ChdDSUhNM	Bagus lega pantainya anak2 seneng joss tambahin wahana lagi mungkin bi	2024-08-20T03:56:35.472Z
7	Kuat Agus Santosa		ChdDSUhNM	Mengenang masa kecil ...ke pantai naik truk	2024-08-06T08:32:42.216Z
8	Onet ST		ChdDSUhNM	Deretan pantai yg bersih dan ramai di kunjungi di sore hari...Spot hunting :	2024-07-27T11:48:07.499Z
9	taruna widi		ChdDSUhNM	Not Bad Lah	2024-07-20T06:08:55.791Z
10	Ardiyanto purwosuseno		ChdDSUhNM	Tempat yang indah	2024-07-11T12:47:56.493Z

GAMBAR 2
HASIL SCRAPING DALAM BENTUK EXCEL

B. *Preprocessing*

Data yang telah terkumpul kemudian diproses melalui tahapan preprocessing untuk membersihkan dan menyiapkan data teks untuk digunakan dalam proses selanjutnya. Tahap ini memiliki beberapa proses seperti case folding, remove punctuation, normalization, tokenizing, stopword removal, dan stemming. Berikut hasil dari berbagai prosesnya:

1) Case Folding

Proses ini dilakukan untuk menyeragamkan format huruf dengan mengubah seluruh huruf menjadi huruf kecil. Hal ini bertujuan agar sistem tidak membedakan kata yang sama namun ditulis dengan huruf kapital, seperti “Bagus” dan “bagus”

TABEL 1
HASIL CASE FOLDING

Input	Output
Bagus lega pantainya anak2 seneng joss tambahin wahana lagi mungkin biar tambah ramai	bagus lega pantainya anak2 seneng joss tambahin wahana lagi mungkin biar tambah ramai

2) Remove Punctuation

Seluruh tanda baca, simbol, angka, dan karakter khusus seperti tanda seru, tanda tanya, dan angka dihapus karena dianggap tidak memiliki makna penting dalam analisis sentimen.

TABEL 2
HASIL REMOVE PUNCTUATION

Input	Output
bagus lega pantainya anak2 seneng joss tambahin wahana lagi mungkin biar tambah ramai	bagus lega pantainya anak seneng joss tambahin wahana lagi mungkin biar tambah ramai

3) Normalization

Pada tahap ini, kata-kata tidak baku atau alay disesuaikan ke dalam bentuk baku. Misalnya kata “bgt” diubah menjadi “banget”. Proses ini dilakukan dengan mencocokkan kata-kata dengan kamus normalisasi agar model dapat lebih memahami makna kata sebenarnya.

TABEL 3
HASIL NORMALIZATION

Input	Output
bagus lega pantainya anak seneng joss tambahin wahana lagi mungkin biar tambah ramai	bagus lega pantainya anak senang mantap tambahin wahana lagi mungkin biar tambah ramai

4) Tokenizing

Teks yang telah dibersihkan kemudian dipecah menjadi satuan kata atau token. Tahap tokenizing penting untuk membentuk representasi data dalam bentuk yang dapat diolah pada tahap selanjutnya, seperti pembobotan dan klasifikasi.

TABEL 4
HASIL TOKENIZING

Input	Output
bagus lega pantainya anak senang mantap tambahin wahana lagi mungkin biar tambah ramai	[bagus, lega, pantainya, anak, senang, mantap, tambahin, wahana, lagi, mungkin, biar, tambah, ramai,]

5) Stopword Removal

Stopword seperti "yang", "di", "ke", "dan", dan kata umum lainnya dihapus dari data karena tidak memiliki kontribusi signifikan terhadap penentuan sentimen. Ini dilakukan untuk mengurangi dimensi data dan fokus pada kata-kata bermakna.

TABEL 5
HASIL STOPWORD REMOVAL

Input	Output
[bagus, lega, pantainya, anak, senang, mantap, tambahin, wahana, lagi, mungkin, biar, tambah, ramai,]	[bagus, lega, pantainya, anak, senang, mantap, tambahin, wahana, mungkin, biar, tambah, ramai,]

6) Stemming

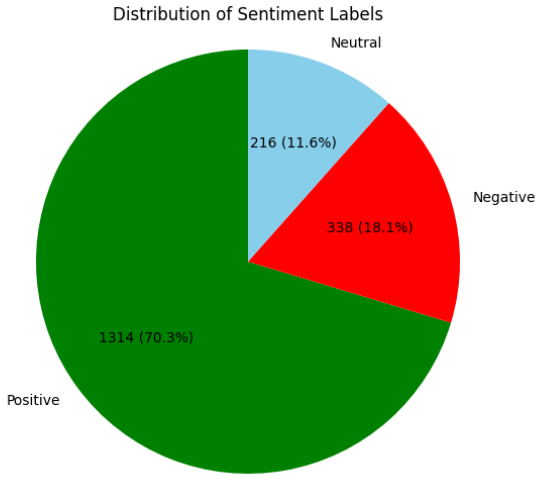
Proses stemming mengubah kata ke bentuk dasarnya dengan menghapus imbuhan, misalnya “bermain”, “bermainan”, menjadi “main”.

TABEL 6
HASIL STEMMING

Input	Output
[bagus, lega, pantainya, anak, senang, mantap, tambahin, wahana, lagi, mungkin, biar, tambah, ramai,]	[bagus, lega, pantai, anak, senang, mantap, tambahin, wahana, lagi, mungkin, biar, tambah, ramai,]

C. *Labeling Data*

Proses labeling dilakukan secara otomatis menggunakan kamus SenticNet, dengan mencocokkan kata-kata dalam ulasan terhadap nilai polaritas yang ada di dalam kamus. Nilai sentimen tiap kata dirata-ratakan untuk menentukan label ulasan secara keseluruhan: positif, negatif, atau netral.



GAMBAR 3
DISTRIBUSI HASIL PELABELAN

GAMBAR 3 menunjukkan distribusi label sentimen dari total 1.868 ulasan pengguna terhadap Pantai Widarapayung. Mayoritas ulasan memiliki sentimen positif sebanyak 1.314 data (70,3%), diikuti oleh sentimen negatif sebanyak 338 data (18,1%), dan sentimen netral sebanyak 216 data (11,6%).

TABEL 7
HASIL LABELING DATA

Teks	Polarity Score	Label Sentimen
lumayan nyaman rekreasi keluarga	2.16	Positif
umum masalah baru duduk ada lima menit tarik uang sewa harga makan warung mahal	-0.422	Negatif
mau tanya pakai sepeda karcis berapa	0	Netral

D. *TF-IDF*

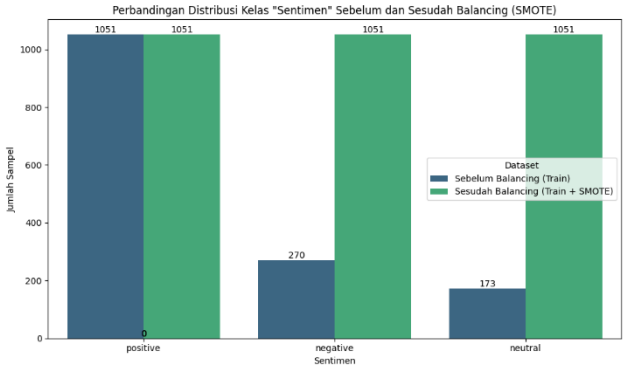
Pembobotan dilakukan menggunakan metode TF-IDF untuk mengubah kumpulan kata menjadi vektor numerik. TF menghitung frekuensi kata dalam dokumen, sedangkan IDF menghitung seberapa penting kata tersebut secara keseluruhan. Dengan metode ini, kata-kata penting memiliki bobot lebih tinggi.

E. *Data Splitting*

Data dibagi menjadi dua bagian, yaitu 80% sebagai data latih dan 20% sebagai data uji. Tujuannya adalah untuk melatih model pada sebagian besar data, dan menguji performa model pada data yang belum pernah dilihat sebelumnya, sehingga hasil evaluasi menjadi lebih objektif.

F. *SMOTE*

Karena distribusi data tidak seimbang, digunakan SMOTE untuk menambahkan data sintetis pada kelas minoritas (negatif dan netral). Ini dilakukan agar model tidak bias terhadap kelas mayoritas (positif), sehingga klasifikasi menjadi lebih adil dan akurat.



GAMBAR 4
PERBANDINGAN DISTRIBUSI SEBELUM DAN SESUDAH SMOTE

G. *Naïve Bayes*

Data latih yang telah seimbang dilatih menggunakan algoritma Multinomial Naïve Bayes. Algoritma ini dipilih karena efisien dan efektif untuk klasifikasi teks. Naïve Bayes menghitung probabilitas tiap kelas berdasarkan frekuensi kata dan menetapkan kategori sentimen untuk masing-masing ulasan.

H. *Evaluasi*

Evaluasi dilakukan untuk mengukur performa model klasifikasi Multinomial Naïve Bayes dalam mengkategorikan ulasan pengguna menjadi sentimen positif, negatif, dan netral. Pengujian dilakukan dalam dua skenario: sebelum dan sesudah penyeimbangan data menggunakan SMOTE. Selain itu, digunakan juga teknik *cross-validation* untuk menilai kestabilan model.

```
Classification Report (Sebelum SMOTE):
              precision    recall  f1-score   support

    negative       0.50      0.03      0.06        68
     neutral       0.00      0.00      0.00        43
    positive       0.71      0.99      0.82       263

    accuracy                           0.70       374
   macro avg       0.40      0.34      0.29       374
weighted avg       0.59      0.70      0.59       374
```

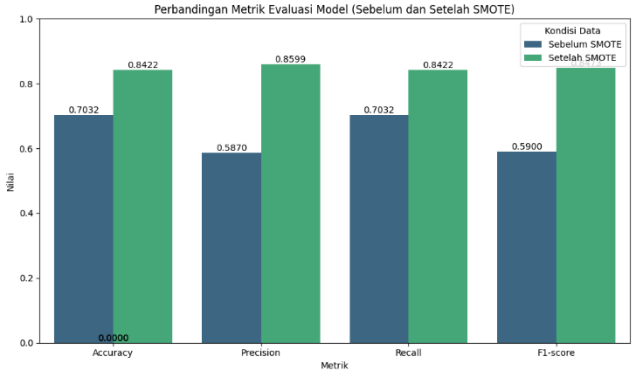
GAMBAR 5
PERFORMA MODEL SEBELUM SMOTE

```
Classification Report (Setelah SMOTE):
              precision    recall  f1-score   support

    negative       0.64      0.82      0.72        68
     neutral       0.77      0.84      0.80        43
    positive       0.93      0.85      0.89       263

    accuracy                           0.84       374
   macro avg       0.78      0.84      0.80       374
weighted avg       0.86      0.84      0.85       374
```

GAMBAR 6
PERFORMA MODEL SETELAH SMOTE

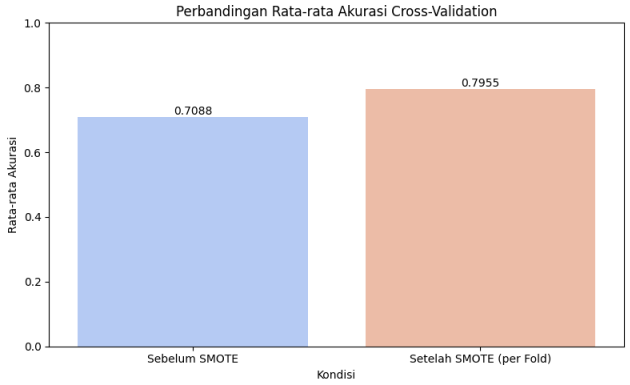
Sebelum penerapan SMOTE, model menghasilkan akurasi sebesar 79,95% dengan F1-score tertinggi pada sentimen positif yang dapat dilihat pada GAMBAR 5. Namun, performa pada sentimen negatif dan netral relatif rendah karena distribusi data yang tidak seimbang. Setelah menerapkan SMOTE pada data latih, model menunjukkan peningkatan performa signifikan, dengan akurasi meningkat menjadi 84,22% dan F1-score sentimen positif sebesar

88,86%, serta sentimen negatif meningkat menjadi 71,67% yang dapat dilihat pada GAMBAR 6.



GAMBAR 7
PERBANDINGAN PERFORAMODEL SEBELUM VS SETELAH SMOTE

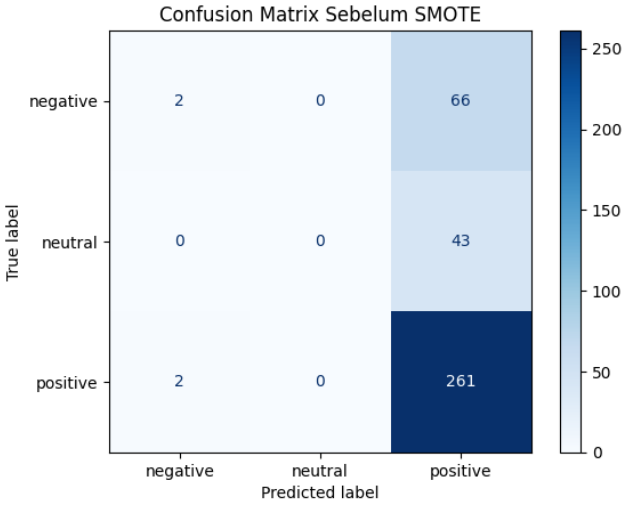
Terlihat pada GAMBAR 7, performa yang dihasilkan setelah penggunaan SMOTE mengalami peningkatan yang signifikan.



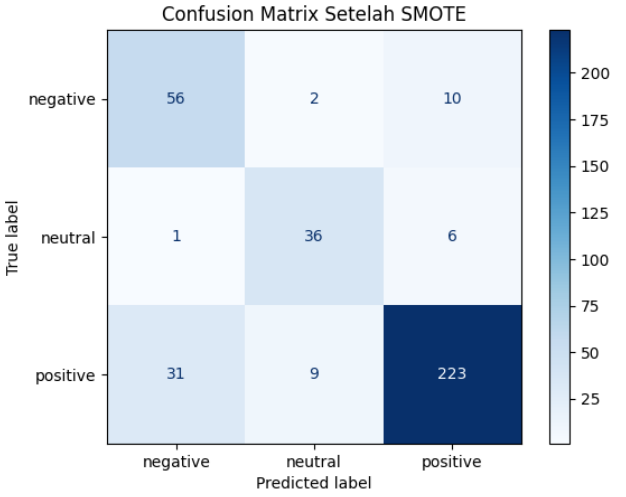
GAMBAR 8
PERBANDINGAN AKURASI CROSS VALIDATION SEBELUM VS SETELAH SMOTE

GAMBAR 8 menunjukkan perbandingan rata-rata akurasi hasil cross-validation sebelum dan sesudah penerapan SMOTE. Sebelum SMOTE, rata-rata akurasi model berada pada angka 0,7088 (70,88%), yang menunjukkan performa klasifikasi yang masih kurang optimal akibat ketidakseimbangan data antar kelas sentimen. Setelah dilakukan oversampling menggunakan SMOTE, akurasi model meningkat secara signifikan menjadi 0,7955 (79,55%). Peningkatan ini menunjukkan bahwa teknik SMOTE efektif dalam memperbaiki distribusi data latih dan membantu model mengenali pola sentimen dari kelas minoritas (negatif dan netral) dengan lebih baik. Hasil ini memperkuat kesimpulan bahwa penyeimbangan data sangat berperan dalam meningkatkan kinerja klasifikasi berbasis teks.

GAMBAR 9 menunjukkan hasil evaluasi model sebelum penerapan SMOTE. Terlihat bahwa model sangat dominan memprediksi semua ulasan sebagai sentimen positif, dengan total 261 prediksi benar untuk kelas positif. Namun, prediksi untuk kelas negatif dan netral sangat buruk, di mana sebagian besar data negatif (66) dan seluruh data netral (43) justru salah diklasifikasikan sebagai positif. Hanya 2 data negatif yang berhasil diklasifikasikan dengan benar. Hal ini menunjukkan bahwa model mengalami overfitting pada kelas mayoritas (positif) karena distribusi data yang tidak seimbang, sehingga gagal mengenali pola pada kelas minoritas.



GAMBAR 9
CONFUSION MATRIX SEBELUM SMOTE



GAMBAR 10
CONFUSION MATRIX SETELAH SMOTE

GAMBAR 10 menunjukkan hasil evaluasi model setelah penerapan SMOTE. Model mengalami peningkatan signifikan dalam mengenali ketiga kelas sentimen. Sentimen negatif berhasil diklasifikasikan dengan benar sebanyak 56 data, netral sebanyak 36 data, dan positif sebanyak 223 data. Jumlah kesalahan klasifikasi juga berkurang dibanding sebelum SMOTE, terutama pada kelas minoritas. Hasil ini menunjukkan bahwa SMOTE berhasil mengurangi ketimpangan data dan meningkatkan kemampuan model dalam mendeteksi sentimen secara lebih seimbang dan akurat.

V. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma Multinomial Naïve Bayes untuk melakukan analisis sentimen terhadap ulasan pengguna tentang Pantai Widarapayung pada Google Maps. Hasil menunjukkan bahwa model memiliki performa terbaik setelah dilakukan penyeimbangan data menggunakan

SMOTE, dengan akurasi mencapai 84,22% dan F1-score tertinggi pada sentimen positif sebesar 88,86%. Penerapan preprocessing dan pembobotan TF-IDF berkontribusi dalam meningkatkan akurasi klasifikasi. Secara keseluruhan, metode ini efektif untuk memetakan opini pengunjung secara otomatis dan dapat dijadikan sebagai referensi untuk pengambilan keputusan berbasis data bagi pihak pengelola.

REFERENSI

[1] L. Syafirullah *et al.*, "Implementasi Multimedia Development Life Cycle Menggunakan Augmented Reality Berbasis Android dalam Pengenalan Landmark Pantai Widarapayung Guna Meningkatkan Daya Dukung Pariwisata Cilacap," *J. Sains dan Manaj.*, vol. 10, no. 2, pp. 1–12, 2022, [Online]. Available: https://ejournal.bsi.ac.id/ejurnal/index.php/evolusi/article/view/13765

[2] M. Yusuf *et al.*, "Penerapan Photovoltaic (PV) Sebagai Energi Listrik Alternatif Pada Rintisan Desa Wisata Widarapayung Wetan Cilacap," *LOGISTA - J. Ilm. Pengabdi. Kpd. Masy.*, vol. 6, no. 1, p. 36, Jun. 2022, doi: 10.25077/logista.6.1.36-44.2022.

[3] F. A. Nasution, Citra Safira, and Yofiendi Indah Indainanto, "Strategi Digital Marketing Dalam Meningkatkan Kunjungan Wisatawan Objek Wisata Tebing Breksi Pasca-Pandemi," *J. Bus. Adm. Entrep. Creat. Ind.*, vol. 2, no. 1, pp. 38–47, 2023, doi: 10.32734/jba.v2i1.10744.

[4] W. Khofifah, D. N. Rahayu, and A. M. Yusuf, "Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps," *J. Interkom J. Publ. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 16, no. 4, pp. 28–38, 2022, doi: 10.35969/interkom.v16i4.192.

[5] I. Afdhal, R. Kurniawan, I. Iskandar, R. Salambue, E. Budianita, and F. Syafria, "Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 1, pp. 122–130, 2022, [Online]. Available: http://ojs.serambimekkah.ac.id/jnkti/article/view/4004/pdf

[6] E. Salim and M. Syafrullah, "Analisis Sentimen Pada Ulasan Pelayanan Suku Dinas Kependudukan Dan Pencatatan Sipil Kota Administrasi Jakarta Barat Menggunakan Algoritme K-Nearest Neighbor," *Bit (Fakultas Teknol. Inf. Univ. Budi Luhur)*, vol. 20, no. 1, p. 58, Apr. 2023, doi: 10.36080/bit.v20i1.2186.

[7] A. Gerliandeva, Y. H. Chrisnanto, and H. Ashaury, "Optimasi Klasifikasi Sentimen pada Komentar Online menggunakan Multinomial Naïve Bayes dan Ekstraksi Fitur TF-IDF serta N-grams Optimization of Sentiment Classification on Online Comments using Multinomial Naïve Bayes and TF-IDF Feature Extraction and N-g," no. X, pp. 259–272, 2024, doi: 10.56873/jpkm.v9i2.5585.

[8] I. F. Rahman, A. N. Hasanah, and N. Heryana, "Analisis Sentimen Ulasan Pengguna Aplikasi Samsat Digiital Nasional (Signal) Dengan Menggunakan Metode Naïve Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, pp. 963–969, 2024, doi: 10.23960/jitet.v12i2.4073.

[9] D. Era, S. Andryana, and A. Rubhasy, "Perbandingan Algoritma Naïve Bayes Dan K-Nearest Neighbor pada Analisis Sentimen Pembukaan Pariwisata Di Masa Pandemi Covid 19," *J. Sains Komput. Inform. (J-SAKTI*, vol. 7, no. 1, pp. 263–272, 2023.

[10] J. C. Aponno, "Penerapan Algoritma Sentimen Analysis dan Naïve Bayes terhadap opini pengunjung di tempat wisata pantai Pintu Kota, Kota Ambon," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 4, pp. 3180–3188, 2022, doi: 10.35957/jatisi.v9i4.2697.

[11] S. Ulya, A. Ridwan, W. Cholid Wahyudin, and F. M. Hana, "Text Mining Sentimen Analisis Pengguna Aplikasi Marketplace Tokopedia Berdasar Rating dan Komentar Pada Google Play Store," *J. Bisnis Digit. dan Sist. Inf.*, vol. 3, no. 2, pp. 33–40, 2022, [Online]. Available: https://ejr.umku.ac.id/index.php/BIDISFO/article/view/1799

[12] R. Wati, S. Ernawati, and H. Rachmi, "Pembobotan TF-IDF Menggunakan Naïve Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH," *J. Manaj. Inform.*, vol. 13, no. 1, pp. 84–93, 2023, doi: 10.34010/jamika.v13i1.9424.

[13] N. Agustina, D. H. Citra, W. Purnama, C. Nisa, and A. R. Kurnia, "Implementasi Algoritma Naive Bayes untuk Analisis Sentimen Ulasan Shopee pada Google Play Store," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 2, no. 1, pp. 47–54, 2022, doi: 10.57152/malcom.v2i1.195.

[14] A. R. Hidayati, A. S. Fitrani, M. A. Rosid, F. Sains, and D. Teknologi, "Analisa Sentimen Pemilu 2019 Pada Judul Berita Online Menggunakan Metode Logistic Regression," *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 4, no. 2, pp. 298–305, 2023, [Online]. Available: http://www.pkm.tunasbangsa.ac.id/index.php/kesatria/article/view/164

[15] P. R. Alvita Wagiswari D, I. Susilawati, and A. Witanti, "Analisis Sentimen pada Komentar Aplikasi MyPertamina dengan Metode Multinomial Naive Bayes," *ForAI J. Informatics Artif. Intell. J.*, vol. 1, no. 1, pp. 10–19, 2023, [Online]. Available: https://jurnal.forai.or.id/index.php/forai/article/view/4

[16] A. M. Khair, "Rancang Bangun Aplikasi Pencarian Bengkel Aktif Berbasis Google Maps Api di Wilayah Kota Makassar," *J. Manaj. Inform. Sist. Inf. dan Teknol. Komput.*, vol. 1, no. 1, pp. 32–42, 2022, doi: 10.70247/jumistik.v1i1.7.

[17] C. Prakoso and A. Hermawan, "Perbandingan Model Machine Learning dalam Analisis Sentimen Ulasan Pengunjung Keraton Yogyakarta pada Google Maps," *Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 3, pp. 1292–1302, 2023, doi: 10.30865/klik.v4i3.1419.

[18] H. V. Sahetapy *et al.*, "Pelatihan dan Pendampingan Pembuatan Tag Lokasi Usaha Berbasis Google Maps sebagai Upaya Promosi UMKM di Kota Ambon," *J. Pengabdi. Pada Masy.*, vol. 9, no. 2, pp. 534–541,

2024, doi: 10.30653/jppm.v9i2.853.
[19] A. Arafa, N. El-Fishawy, M. Badawy, and M. Radad, “RN-SMOTE: Reduced Noise SMOTE based on DBSCAN for enhancing imbalanced data classification,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 8, pp. 5059–5074, 2022, doi: 10.1016/j.jksuci.2022.06.005.
[20] L. Qadrini, H. Hikmah, and M. Megasari, “Oversampling, Undersampling, Smote SVM dan Random Forest pada Klasifikasi Penerima Bidikmisi Sejawa Timur Tahun 2017,” *J. Comput. Syst. Informatics*, vol. 3, no. 4, pp. 386–391, 2022, doi: 10.47065/josyc.v3i4.2154.