JITET (Jurnal Informatika dan Teknik Elektro Terapan)
Vol. 13 No. 1, pISSN: 2303-0577 eISSN: 2830-7062 http://dx.doi.org/10.23960/jitet.v13i1.5748

IMPLEMENTASI AKURASI MODEL NAIVE BAYES MENGGUNAKAN SMOTE DALAM ANALISIS SENTIMEN PENGGUNA APLIKASI BRIMO

Muhammad Andi Hermawan[1*], Ahmad Faqih[2], Gifthera Dwilestari[3]

[1,2] Teknik Informatika, STMIK IKMI Cirebon, Indonesia
[3] Sistem Informasi, STMIK IKMI Cirebon, Indonesia

Received: 18 Desember 2024
Accepted: 14 Januari 2025
Published: 20 Januari 2025

Keywords:
Ketidakseimbangan kelas, Model Smote, Naïve bayes, Ulasan pengguna, Akurasi Model, Recal, F1-Score, Confusion Matrix.

Corespondent Email:
andibronze297@gmail.com

Abstrak. Aplikasi Brimo dari Bank Rakyat Indonesia (BRI) menjadi salah satu platform utama interaksi nasabah dengan layanan perbankan. Analisis sentimen ulasan pengguna aplikasi ini pentring untuk memahami pendapat dan menaikkan kualitas pelayanan. Penelitian ini menggunakan algoritma Naïve bayes dengan menerapkan model Smote (Synthetic Minority Over-sampling Technique) untuk menangani ketidakseimbangan kelas antara kelas positif dan negatif dalam data ulasan pengguna. Dataset yang di peroleh mencapai 1.000 ulasan Play store yang di proses melalui tahap pengumpulan, pra-pemrosesan teks, dan evaluasi menggunakan Confusion Matrix. penelitian menunjukkan bahwa metode SMOTE secara signifikan meningkatkan kinerja model, dengan recall untuk sentimen negatif meningkat dari 0,55 menjadi 0,87 dan F1-score dari 0,71 menjadi 0,84. Akurasi model juga naik dari 93% menjadi 95%, dengan pengurangan False Negatives. Temuan ini membuktikan efektivitas SMOTE dalam meningkatkan akurasi dan representasi model untuk memahami opini pengguna BRImo secara lebih baik.

Abstract. *The Brimo application from Bank Rakyat Indonesia (BRI) is one of the main platforms for customer interaction with banking services. Sentiment analysis of user reviews of this application is important to understand opinions and improve service quality. This study uses the Naïve Bayes algorithm by implementing the Smote (Synthetic Minority Over-sampling Technique) model to handle class imbalances between positive and negative classes in user review data. The dataset obtained reached 1,000 Play Store reviews which were processed through the collection stage, text pre-processing, and evaluation using the Confusion Matrix. The study showed that the SMOTE method significantly improved model performance, with recall for negative sentiment increasing from 0.55 to 0.87 and F1-score from 0.71 to 0.84. Model accuracy also increased from 93% to 95%, with a reduction in False Negatives. These findings prove the effectiveness of SMOTE in improving model accuracy and representation to better understand BRImo user opinions.*

1. PENDAHULUAN

Di era digital yang terus berkembang, aplikasi mobile menjadi platform utama untuk

interaksi antara pengguna dan layanan perbankan. BRImo, aplikasi milik Bank Rakyat Indonesia (BRI), menawarkan berbagai fitur perbankan. Analisis sentimen pengguna terhadap aplikasi ini penting untuk memahami persepsi dan kebutuhan mereka dalam menggunakan BRImo. Teknik SMOTE digunakan untuk mengatasi ketidakseimbangan dalam data kelas minorita. Model Naive Bayes dilatih dan dievaluasi sebelum dan sesudah penerapan SMOTE. Hasil evaluasi menunjukkan bahwa akurasi model meningkat dari 76,5% menjadi 78,0% setelah penggunaan SMOTE, meskipun terjadi penurunan pada recall kelas positif. Integrasi SMOTE secara efektif meningkatkan performa model dalam mengidentifikasi kelas negatif [1]. Analisis sentimen membantu mengidentifikasi masalah yang dihadapi pengguna, seperti bug atau fitur yang tidak berfungsi, ulasan pengguna memberi Bank BRI wawasan tentang kebutuhan dan harapan mereka dari aplikasi BRImo. Informasi ini penting untuk mengembangkan aplikasi yang lebih baik. Dengan menangani masalah yang ditemukan, Bank BRI bisa membuat aplikasi lebih handal dan mudah digunakan, yang meningkatkan pengalaman pengguna. Peningkatan terus-menerus berdasarkan umpan balik pengguna membantu BRImo tetap kompetitif, menarik pengguna baru, dan meningkatkan loyalitas nasabah. Merespons umpan balik juga dapat meningkatkan rating aplikasi di platform unduhan. Analisis sentimen membantu Bank BRI menemukan tren dan inovasi untuk menjaga BRImo tetap relevan. Naïve Bayes adalah algoritma yang sering digunakan untuk analisis sentimen karena sifatnya yang sederhana dan berbasis probabilistik. Algoritma ini bekerja dengan menghitung probabilitas berdasarkan pola data yang ada. Dalam analisis sentimen pengguna aplikasi Grab, metode ini berhasil mencapai tingkat akurasi 87%, presisi 86%, dan recall 97%. Sebelum proses klasifikasi, dilakukan tahap preprocessing data, seperti case folding, tokenisasi, penghapusan kata-kata yang tidak penting (stopwords), dan stemming. Langkah-langkah ini bertujuan untuk meningkatkan kualitas data sehingga hasil analisis lebih akurat [2]. Evaluasi kinerja model dilakukan menggunakan Confusion Matrix di Google Colaboratory, dan hasilnya menunjukkan bahwa model mampu mengidentifikasi sentimen dengan sangat baik. Variasi nilai akurasi, presisi, dan recall disebabkan oleh performa model yang kurang seimbang dalam mengklasifikasikan data. Hal ini terlihat dari hasil implementasi model pada 200 data uji yang digunakan[3]. Penelitian ini memberikan gambaran mendalam tentang persepsi pengguna terhadap aplikasi BRImo, termasuk kekuatan, kelemahan, serta kebutuhan dan harapan yang belum terpenuhi. Wawasan ini dapat membantu Bank BRI dalam menyesuaikan fitur aplikasi, meningkatkan efisiensi, dan menciptakan pengalaman yang lebih ramah pengguna. Dengan memahami kebutuhan pengguna, Bank BRI dapat memperbaiki layanan, mengurangi keluhan, serta memperkuat posisi BRImo di pasar mobile banking [4]. Data ulasan pengguna dikumpulkan dari platform seperti Google Play Store dan Apple App Store. Tahap awal melibatkan prapemrosesan data, seperti pembersihan teks, penghapusan elemen tidak relevan (misalnya tanda baca dan angka), serta stemming. Algoritma Naïve Bayes digunakan untuk menganalisis sentimen, mengelompokkan ulasan ke dalam kategori positif, negatif, atau netral [5]. Melalui pendekatan ini, penelitian tidak hanya membantu Bank BRI dalam meningkatkan kualitas aplikasi tetapi juga memperkuat penerapan teknologi modern seperti Google Colaboratory dalam analisis data [6]. Hasilnya dapat menjadi panduan untuk pengembangan aplikasi yang lebih baik di masa depan. Dengan memahami sentimen pengguna, Bank BRI dapat mengembangkan aplikasi BRImo agar lebih responsif dan sesuai dengan kebutuhan pengguna. Perbaikan teknis dan penambahan fitur berdasarkan hasil penelitian akan meningkatkan kualitas layanan, membuat aplikasi lebih andal, fungsional, dan mudah digunakan. Penelitian ini juga memberikan wawasan untuk mengoptimalkan antarmuka pengguna agar lebih intuitif, sehingga dapat meningkatkan kepuasan dan kenyamanan pengguna [7]. Hasil analisis dapat dimanfaatkan sebagai dasar strategi pemasaran yang lebih efektif, dengan menonjolkan fitur-fitur yang paling dihargai oleh pengguna. Selain itu, wawasan dari analisis sentimen membantu Bank BRI membuat keputusan bisnis yang lebih terarah dalam menghadapi persaingan di pasar mobile banking. Langkah-langkah perbaikan berdasarkan temuan penelitian akan

memperkuat loyalitas pengguna dan menarik lebih banyak pelanggan baru. Dengan pendekatan ini, aplikasi BRImo dapat terus dikembangkan agar tetap relevan dan memenuhi ekspektasi pengguna yang terus berkembang, sehingga memperkuat posisinya di pasar mobile banking yang kompetitif.

2. TINJAUAN PUSTAKA

2.1. Sentimen Analisis

Analisis sentimen merupakan metode yang digunakan untuk memahami pendapat masyarakat mengenai berbagai topik, seperti layanan publik, isu sosial, kinerja pemerintah, atau hal lainnya. Metode ini berfungsi sebagai alat evaluasi untuk menilai kualitas layanan yang telah diberikan. Salah satu cara yang umum dilakukan dalam analisis sentimen adalah dengan mengumpulkan opini dari platform media sosial [3].

2.2. Aplikasi BRI Mobile

Aplikasi mobile kini menjadi platform utama bagi pengguna untuk berinteraksi dengan layanan perbankan. BRImo, sebagai salah satu aplikasi yang ditawarkan oleh Bank BRI, menyediakan beragam fitur perbankan untuk memenuhi kebutuhan penggunanya. Analisis sentimen pengguna terhadap aplikasi ini menjadi sangat penting bagi Bank BRI untuk memahami persepsi serta kebutuhan mereka, sehingga dapat meningkatkan kualitas layanan dan pengalaman pengguna [8].

2.3. Data Mining

Data mining adalah proses menganalisis dan menggali pola, informasi, atau pengetahuan berharga dari kumpulan data yang besar. Proses ini menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk menemukan hubungan yang tersembunyi dalam data dan membuat prediksi atau keputusan yang lebih baik [9].

2.4. Naïve bayes

Naive Bayes adalah salah satu algoritma machine learning yang digunakan untuk menyelesaikan masalah klasifikasi. Algoritma ini didasarkan pada teorema Bayes dengan asumsi bahwa semua fitur dalam dataset bersifat independen satu sama lain (inilah mengapa disebut "naive" atau sederhana). Meskipun asumsi independensi jarang terpenuhi secara sempurna di dunia nyata, algoritma ini tetap bekerja dengan sangat baik dalam banyak kasus, terutama untuk dataset yang besar [6].

2.5. Google coleb

Google Colab adalah layanan online yang memungkinkan kita menulis dan menjalankan kode Python langsung dari browser, tanpa perlu menginstal software tambahan. Layanan ini dirancang khusus untuk mempermudah pekerjaan seperti analisis data, machine learning, deep learning, dan membuat visualisasi data. Dengan Google Colab, kita juga bisa menggunakan perangkat keras canggih seperti GPU dan TPU secara gratis. Karena mudah digunakan, bisa diakses dari mana saja, dan mendukung kolaborasi seperti Google Docs, platform ini sangat populer di kalangan mahasiswa, peneliti, dan praktisi data science [2].

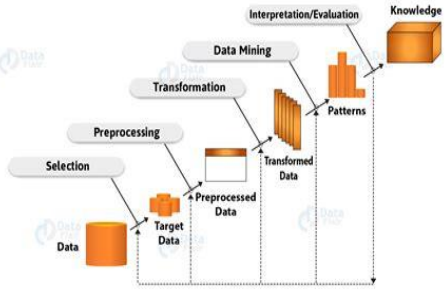
2.6. Model Smote

SMOTE (Synthetic Minority Oversampling Technique) adalah teknik yang digunakan untuk mengatasi masalah ketidakseimbangan kelas dalam pembelajaran mesin. Ketidakseimbangan kelas terjadi ketika salah satu kelas memiliki jumlah data yang jauh lebih sedikit dibandingkan kelas lainnya, yang dapat menyebabkan model bias terhadap kelas yang dominan. SMOTE bekerja dengan menciptakan data sintetik untuk kelas minoritas, yang tidak hanya menggandakan data yang ada, tetapi juga menghasilkan data baru dengan cara interpolasi antar data asli. Dengan menambahkan data baru yang beragam, SMOTE membantu menciptakan dataset yang lebih seimbang, sehingga model dapat lebih baik mengenali pola dari semua kelas [10].

3. METODE PENELITIAN

Analisis sentimen dilakukan menggunakan algoritma Multinomial Naïve Bayes untuk mengelompokkan ulasan menjadi sentimen positif, negatif, atau netral. Untuk mengatasi ketidakseimbangan kelas dalam data, metode SMOTE (Synthetic Minority Over-sampling Technique) diterapkan, yang membantu meningkatkan akurasi model. Proses pelabelan data dipercepat dengan menggunakan skor ulasan otomatis dari Google Play Store [11]. Data dikumpulkan melalui web scraping menggunakan Google Play Scraper, lalu diolah menggunakan metode Text Mining. Bobot kata dihitung dengan Term Frequency - Inverse

Document Frequency (TF-IDF). Knowledge Discovery in Databases (KDD) adalah proses sistematis untuk menganalisis data dan mengekstraksi pengetahuan yang berharga. Proses dimulai dengan memahami tujuan penelitian dan mempersiapkan data, termasuk menangani nilai yang hilang atau data yang tidak konsisten [12]. Analisis Deskriptif dan visualisasi dilakukan dalam tahap Exploratory Data Analysis (EDA) untuk memahami pola dan karakteristik data. Berikut adalah gambar 4.1 KDD.



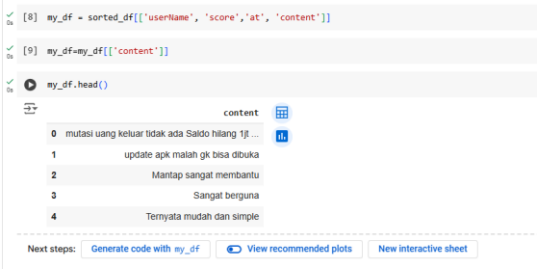
Gambar 3.1 KDD

4. HASIL DAN PEMBAHASAN

SMOTE diterapkan untuk mengatasi ketidakseimbangan kelas dalam data, di mana jumlah data sentimen negatif lebih sedikit dibandingkan dengan sentimen positif. Penelitian ini mengikuti tahapan Knowledge Discovery in Databases (KDD), yang merupakan proses eksplorasi data bertahap untuk mendapatkan wawasan berharga dari dataset yang besar dan kompleks. Proses pengolahan data dalam penelitian ini menggunakan algoritma Naive Bayes yang dikombinasikan dengan SMOTE (Synthetic Minority Over-sampling Technique) untuk menyeimbangkan jumlah data sentimen positif dan negatif. Tujuannya adalah meningkatkan akurasi model klasifikasi [10]. Naive Bayes, algoritma berbasis probabilitas dan Teorema Bayes, sering digunakan dalam analisis sentimen karena kemampuannya menangani klasifikasi berbasis teks.

4.1. Selection Data

Pada penelitian ini, bertujuan untuk memastikan hanya data yang relevan dan berkualitas digunakan dalam analisis sentimen aplikasi BRImo. Proses ini meliputi pemilihan variabel utama, yaitu kolom content yang memuat ulasan pengguna dan kolom label yang menunjukkan sentimen (positif atau negatif), serta penghapusan kolom yang tidak relevan. Selain itu, data yang tidak valid, seperti baris kosong, ulasan tanpa opini jelas, atau ulasan yang tidak terkait langsung dengan aplikasi BRImo, disaring untuk mengurangi noise [13]. Dengan data yang bersih dan terfokus, penelitian dapat melanjutkan ke tahap-tahap berikutnya, seperti prapemrosesan teks, ekstraksi fitur, dan penerapan SMOTE, guna meningkatkan akurasi model Naive Bayes dalam analisis sentimen. Berikut adalah gambar 2 data selection



```
[8] my_df = sorted_df[['userName', 'score','at', 'content']]
[9] my_df=my_df[['content']]
    my_df.head()
```

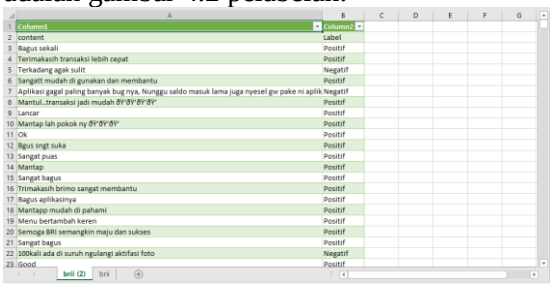
	content
0	mutasi uang keluar tidak ada Saldo hilang 1jt ...
1	update apk malah gk bisa dibuka
2	Mantap sangat membantu
3	Sangat berguna
4	Ternyata mudah dan simple

Next steps: Generate code with my_df | View recommended plots | New interactive sheet

Gambar 4.1 Data Selection

4.2. Labeling

Setelah data latih dibersihkan melalui proses text processing, langkah berikutnya adalah pelabelan data secara manual. Proses ini dilakukan dengan membaca setiap ulasan satu per satu, lalu memberi label sentimen negatif jika terdapat kata-kata bernada emosi negatif, dan sentimen positif jika berisi kata-kata pujian [8]. Berikut adalah hasil data yang telah dibersihkan dan dilabeli secara manual. Berikut adalah gambar 4.2 pelabelan.



	A	B
1	Column1	Column2
2	content	Label
3	Bagus sekali	Positif
4	Terimakasih transaksi lebih cepat	Positif
5	Terkadang agak sulit	Negatif
6	Sangatt mudah di gunakan dan membantu	Positif
7	Aplikasi gagal paling banyak bug nya, Nunggu saldo masuk lama juga nyesel gw pake ni aplik	Negatif
8	Mantul...transaksi jadi mudah 😍😍😍😍	Positif
9	Lancar	Positif
10	Mantap lah pokok ny 😍😍😍	Positif
11	Ok	Positif
12	Bgus sngt suka	Positif
13	Sangat puas	Positif
14	Mantap	Positif
15	Sangat bagus	Positif
16	Trimakasih brimo sangat membantu	Positif
17	Bagus aplikasinya	Positif
18	Mantapp mudah di pahami	Positif
19	Menu bertambah keren	Positif
20	Semoga BRI semangkin maju dan sukses	Positif
21	Sangat bagus	Positif
22	100kali ada di suruh ngulangi aktifasi foto	Negatif
23	Good	Positif

Gambar 4.2 Labeling

4.3. Case folding

Case folding adalah langkah penting dalam analisis sentimen karena membantu menyederhanakan teks dengan mengurangi variasi yang tidak diperlukan. Misalnya, kata "Aplikasi" dan "aplikasi" akan dianggap sama setelah diubah menjadi huruf kecil. Proses ini juga membantu mengurangi dimensi data, sehingga model hanya perlu menangani representasi kata yang relevan [14]. Hasilnya, akurasi klasifikasi dapat meningkat karena

variasi yang tidak signifikan telah dihilangkan. Secara keseluruhan, case folding mendukung pembersihan data yang lebih efisien dan memberikan teks yang lebih konsisten untuk tahap analisis berikutnya. Berikut adalah gambar 4.3 dan Tabel 4.1 case folding.

Gambar 4.3 Case folding

Tabel 4.1 Case folding

Text	Case folding
Terkadang agak sulit	terkadang agak sulit
Sangat mudah di gunakan dan membantu	sangat mudah di gunakan dan membantu

4.4. Stop words

Penghapusan stop words bertujuan untuk menyederhanakan teks dengan menghilangkan kata-kata yang tidak memiliki pengaruh signifikan terhadap analisis makna atau sentimen ulasan, seperti "yang," "di," atau "dan." Proses ini memungkinkan analisis lebih fokus pada kata-kata yang benar-benar relevan dan memengaruhi interpretasi sentimen [15]. Selain meningkatkan relevansi data, penghapusan stop words juga membantu mengurangi dimensi dataset, sehingga mempercepat proses pelatihan model dan meningkatkan akurasi klasifikasi. Berikut adalah gambar 4.4 dan tabel 4.2 stop words.

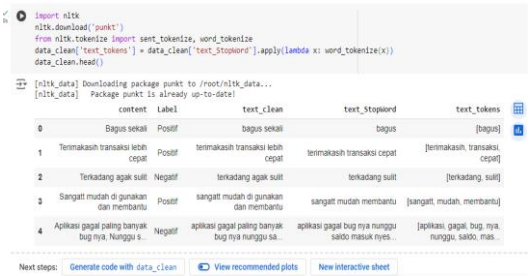
Gambar 4.4 Stop words

Tabel 4.2 Stop words

Text	Stop words
Mantapp mudah di pahami	mantapp mudah pahami
Menu bertambah keren	menu bertambah keren

4.5. Tokenize

Langkah penting berikutnya dalam analisis adalah tokenization. Proses ini memecah teks ulasan menjadi bagian-bagian kecil yang disebut tokens, seperti kata atau frasa. Tokenization memungkinkan model untuk menganalisis setiap kata secara terpisah, sehingga dapat mengidentifikasi pola atau hubungan antara kata-kata yang memengaruhi sentimen positif atau negatif [16]. Berikut adalah contoh gambar 4.5 dan tabel 4.3 Tokenize.



Gambar 4.5 Tokenize

Tabel 4.3 Tokenize

Text	Tokenize
Mantap brimo bagus 👏 👏 🙏	['mantap', 'brimo', 'bagus']
Gak bisa click lanjutkan	['gak', 'click', 'lanjutkan']

4.6. Stemming

Langkah berikutnya dalam analisis adalah stemming. Stemming adalah proses mengubah kata-kata menjadi bentuk dasarnya atau akar kata, seperti "membeli," "pembelian," dan "membelikan" yang direduksi menjadi "beli." Stemming sangat penting karena memungkinkan model untuk mengenali makna yang sama dari berbagai bentuk kata, sehingga meningkatkan efisiensi pengolahan data dan mengurangi kompleksitas yang harus dipelajari oleh model [17]. Jika model harus mempertimbangkan semua variasi kata, hal ini dapat menciptakan kebingungan dan menurunkan akurasi klasifikasi. Dengan menyamakan bentuk kata, stemming membantu menghasilkan representasi teks yang lebih konsisten dan fokus. Berikut adalah gambar 7 dan tabel 4 stemming.

Gambar 4.6 Stemming

Tabel 4.4 Stemming

Text	Stemming
cara ganti email di brimo gimana ya,email saya udah di hack orang lain jadi kalo mau masuk gak bisa	ganti email brimo gimana yaemail udah hack orang kalo masuk gak
Beli pulsa ga perlu kekonter,bayar listrik ga perlu ke loket pembayaran listrik.cukup dr rumah.	beli pulsa ga kekonterbayar listrik ga loket bayar listrikcukup dr rumah

4.7. Transfomation

Tahap data transformation bertujuan untuk mengubah ulasan teks menjadi format numerik agar dapat diproses oleh model. Salah satu metode yang sering digunakan adalah TF-IDF (Term Frequency-Inverse Document Frequency), yang mengonversi teks menjadi vektor berbobot. Bobot ini dihitung berdasarkan frekuensi kata dalam dokumen tertentu dan seberapa unik kata tersebut di seluruh dataset, sehingga kata-kata yang lebih relevan dalam menentukan sentimen akan mendapatkan bobot yang lebih tinggi [2]. Untuk meningkatkan efisiensi dan mencegah overfitting, jumlah fitur yang digunakan dapat dibatasi dengan parameter max_features pada TF-IDF.

Gambar 4.7 Transformation

Gambar 8 diatas memastikan bahwa hanya kata-kata yang paling penting dipertimbangkan, sehingga model dapat lebih fokus pada informasi utama yang memengaruhi sentimen. Pendekatan ini mendukung analisis yang lebih terarah dan akurat.

4.8. Data Mining

Tahap data mining bertujuan untuk mengungkap informasi berharga dari dataset ulasan pengguna aplikasi BRImo, yang mencakup sentimen positif dan negatif. Proses ini dimulai dengan penerapan algoritma Naive Bayes, metode klasifikasi probabilitas yang efektif untuk analisis teks. Algoritma ini menggunakan statistik untuk menghitung kemungkinan sebuah ulasan termasuk dalam kategori tertentu berdasarkan fitur yang dihasilkan dari text vectorization, seperti TF-IDF. Untuk mengatasi masalah ketidakseimbangan kelas, di mana ulasan positif mungkin lebih dominan dibandingkan ulasan negatif, teknik SMOTE (Synthetic Minority Over-sampling Technique) digunakan. SMOTE menciptakan contoh sintetis dari kelas minoritas, sehingga distribusi data menjadi lebih seimbang dan model dapat mempelajari kedua kelas dengan lebih baik [9]. Proses ini tidak hanya bertujuan untuk menghasilkan model klasifikasi yang akurat, tetapi juga untuk mendapatkan wawasan mendalam tentang pengalaman pengguna dengan aplikasi BRImo. Penelitian ini diharapkan dapat memberikan rekomendasi untuk pengembangan aplikasi dan peningkatan kualitas layanan di masa depan.

4.9. Evaluation

Algoritma SMOTE berperan penting dalam membantu model Naïve Bayes mengenali sentimen negatif (kelas minoritas) dengan lebih baik, terlihat dari peningkatan recall pada kelas tersebut. Hal ini penting agar model tidak mengabaikan sentimen negatif secara berlebihan. Meski akurasi yang diperoleh dari penerapan model dengan SMOTE mencapai 95%, penting untuk tidak hanya bergantung pada metrik ini, terutama dalam situasi data yang tidak seimbang. Evaluasi model harus mempertimbangkan precision, recall, dan f1-score untuk memberikan gambaran yang lebih komprehensif mengenai performa model pada setiap kelas. Berikut ini adalah gambar hasil penerapan model Naïve Bayes yang dikombinasikan dengan SMOTE. Berikut

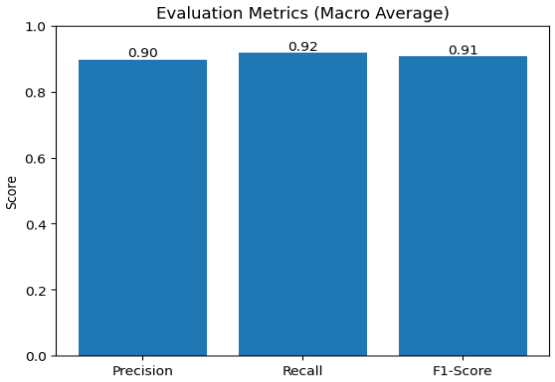
adalah gambar 4.8 hasil penerapan model naïve bayes dan Smote.

```
Accuracy Score: 0.95

Classification Report:
              precision    recall  f1-score   support

     negatif       0.82      0.87      0.84        31
     positif       0.98      0.96      0.97       169

    accuracy                           0.95       200
   macro avg       0.90      0.92      0.91       200
weighted avg       0.95      0.95      0.95       200


Confusion Matrix:
 [[ 27   4]
 [  6 163]]
```

Gambar 4.8 Penerapan Model Naïve bayes dan Smote

4.10. Pembahasan

Hasil evaluasi model Naïve Bayes menunjukkan akurasi keseluruhan 93%, menunjukkan performa yang cukup baik dalam klasifikasi sentimen. Namun, model ini tidak seimbang antara kelas positif dan negatif, dengan precision untuk kelas negatif mencapai 1.00 (semua ulasan negatif yang diprediksi benar), tetapi recall hanya 0.55 (hanya 55% ulasan negatif yang dikenali), sehingga F1-score untuk kelas negatif hanya 0.71, lebih rendah dari kelas positif. Untuk kelas positif, model memiliki recall 1.00 dan F1-score 0.96, menunjukkan performa yang sangat baik pada kelas mayoritas. Penerapan SMOTE pada data pelatihan berhasil meningkatkan performa model, menghasilkan akurasi 95%. Dengan SMOTE, precision untuk kelas negatif menjadi 0.82, recall meningkat menjadi 0.87, dan F1-score mencapai 0.84, menunjukkan peningkatan yang signifikan dalam mengenali ulasan negatif. Untuk kelas positif, precision tetap tinggi di 0.98 dan recall 0.96. Penerapan SMOTE membantu model mengatasi ketidakseimbangan kelas, sehingga dapat mengenali ulasan negatif dengan lebih baik tanpa mengorbankan performa pada ulasan positif. Hasil ini penting untuk analisis sentimen pada aplikasi BRI mobile, memberikan pemahaman lebih dalam tentang keluhan pengguna, dan sejalan dengan penelitian sebelumnya yang menunjukkan manfaat SMOTE dalam menyeimbangkan data dan meningkatkan performa model. Berikut adalah gambar 4.9 diagram batang hasil dari penerapan model smote.



Gambar 4.9 Diagram Model Smote

5. KESIMPULAN

Penerapan SMOTE pada model Naive Bayes memberikan peningkatan signifikan pada performa kelas negatif, terutama pada metrik recall dan F1-score. Recall untuk kelas negatif meningkat dari 0.55 menjadi 0.87, menunjukkan bahwa model lebih baik dalam mengenali ulasan negatif. Peningkatan F1-score pada kelas negatif juga menunjukkan keseimbangan yang lebih baik antara precision dan recall, membuat model lebih andal dalam mengidentifikasi kedua sentimen secara proporsional. Meskipun akurasi keseluruhan model tidak berubah secara drastis, penerapan SMOTE memperbaiki kemampuan model dalam mengidentifikasi kedua kelas, baik minoritas (negatif) maupun mayoritas (positif). Ini menunjukkan bahwa SMOTE tidak hanya mengatasi ketidakseimbangan data, tetapi juga meningkatkan keandalan model dalam mengklasifikasikan sentimen yang sulit dikenali, seperti ulasan negatif.

UCAPAN TERIMA KASIH

Terima kasih atas kesempatan untuk menyusun penelitian ini. Saya menghargai kontribusi dan saran yang telah diberikan selama proses penulisan jurnal ini. Saya juga ingin mengucapkan terima kasih kepada teman-teman saya yang telah memberikan dukungan sepanjang penelitian ini, yaitu Enricco, Aldi, dan Fikri. Selain itu, saya juga berterima kasih kepada Bapak Ahmad Faqih dan Ibu Gifthera Dwilestari selaku dosen pembimbing yang telah memberikan bimbingan dan arahan selama penelitian ini. Semoga penelitian ini dapat memberikan manfaat yang bermanfaat bagi pengembangan ilmu pengetahuan.

DAFTAR PUSTAKA

[1] R. MD, R. D. Restiyan, dan H. Irsyad, "Analisis Sentimen Masyarakat terhadap Perilaku Lawan Arah yang Diunggah pada Media Sosial Youtube Menggunakan Naïve Bayes," *BANDWIDTH: Journal of Informatics and Computer Engineering*, vol. 2, no. 2, hlm. 75–83, Jul 2024, doi: 10.53769/bandwidth.v2i2.706.

[2] A. Suryana, A. I. Purnamasari, dan I. Ali, "Mengoptimalkan Kepuasan Pengguna: Analisis Sentimen Review Aplikasi Grab Di Indonesia," Jl.Perjuangan No.10 B Majasem, Kec.Kesambi, Kota Cirebon Jawa Barat 45135, Indonesia, Jun 2024. Diakses: 29 Mei 2024. [Daring]. Tersedia pada: https://ejournal.itn.ac.id/index.php/jati/article/view/9688/5524

[3] A. Saputra dan F. Noor Hasan, "Analisis Sentimen Terhadap Aplikasi Coffee Meets Bagel Dengan Algoritma Naïve Bayes Classifier," *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, vol. 2, no. 2, hlm. 465–474, Jan 2023, doi: 10.54443/sibatik.v2i2.579.

[4] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, dan W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," *Jurnal Teknoinfo*, vol. 14, no. 2, hlm. 115, Jul 2020, doi: 10.33365/jti.v14i2.679.

[5] B. Z. Ramadhan, R. I. Adam, dan I. Maulana, "Analisis Sentimen Ulasan pada Aplikasi E-Commerce dengan Menggunakan Algoritma Naïve Bayes," *Journal of Applied Informatics and Computing*, vol. 6, no. 2, hlm. 220–225, Des 2022, doi: 10.30871/jaic.v6i2.4725.

[6] A. Nurian, "Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 3s1, Sep 2023, doi: 10.23960/jitet.v11i3s1.3348.

[7] N. Cahyono dan Anggista Oktavia Praneswara, "Analisis Sentimen Ulasan Aplikasi TikTok Shop Seller Center di Google Playstore Menggunakan Algoritma Naive Bayes," *Indonesian Journal of Computer Science*, vol. 12, no. 6, Des 2023, doi: 10.33022/ijcs.v12i6.3473.

[8] M. K. Khoirul Insan, U. Hayati, dan O. Nurdiawan, "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, hlm. 478–483, Mar 2023, doi: 10.36040/jati.v7i1.6373.

[9] A. Safira dan F. N. Hasan, "Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier," *ZONAsi: Jurnal Sistem Informasi*, vol. 5, no. 1, hlm. 59–70, Jan 2023, doi: 10.31849/zn.v5i1.12856.

[10] A. Andreyestha dan Q. N. Azizah, "Analisa Sentimen Kicauan Twitter Tokopedia Dengan Optimalisasi Data Tidak Seimbang Menggunakan Algoritma SMOTE," *Infotek : Jurnal Informatika dan Teknologi*, vol. 5, no. 1, hlm. 108–116, Jan 2022, doi: 10.29408/jit.v5i1.4581.

[11] H. Hidayatullah, P. Purwantoro, dan Y. Umaidah, "Penerapan Naïve Bayes Dengan Optimasi Information Gain Dan Smote Untuk Analisis Sentimen Pengguna Aplikasi Chatgpt," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 3, hlm. 1546–1553, Okt 2023, doi: 10.36040/jati.v7i3.6887.

[12] A. Septiani, A. Voutama, S. Siska, A. Andri Hendriadi, dan N. Heryana, "Analisis Sentimen Menggunakan Algoritma Naïve Bayes Terhadap Regulasi Tiktok Shop Pada Media Sosial X (Twitter)," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 4, hlm. 5729–5735, Jun 2024, doi: 10.36040/jati.v8i4.10040.

[13] T. A. Dewi dan E. Mailoa, "Perbandingan Implementasi Metode Smote Pada Algoritma Support Vector Machine (Svm) Dalam Analisis Sentimen Opini Masyarakat Tentang Mixue," *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*, vol. 4, no. 3, hlm. 849–855, Sep 2023, doi: 10.35870/jimik.v4i3.289.

[14] R. Aryanti, T. Misriati, dan A. Sagiyanto, "Analisis Sentimen Aplikasi Primaku Menggunakan Algoritma Random Forest dan SMOTE untuk Mengatasi Ketidakseimbangan Data," *Journal of Computer System and Informatics (JoSYC)*, vol. 5, no. 1, hlm. 218–227, Nov 2023, doi: 10.47065/josyc.v5i1.4562.

[15] D. Darwis, N. Siskawati, dan Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *Jurnal Tekno Kompak*, vol. 15, no. 1, hlm. 131, Feb 2021, doi: 10.33365/jtk.v15i1.744.

[16] R. Sari dan R. Y. Hayuningtyas, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Pada Wisata TMII Berbasis Website," *Indonesian Journal on Software Engineering (IJSE)*, vol. 5, no. 2, hlm. 51–60, Des 2019, doi: 10.31294/ijse.v5i2.6957.

[17] W. Wahyudi, R. Kurniawan, dan A. Y. Wijaya, "Analisis Sentimen Pengguna Terhadap Aplikasi Blu Bca Di Playstore Mengunakan Algoritma Naïve Bayes," vol. 8, no. 3, hlm. 2511–2517, Mei 2024.