Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi
Jl. Ahmad Yani, K.M. 33,5 - Kampus STMIK Banjarbaru
Loktabat – Banjarbaru (Tlp. 0511 4782881), e-mail: puslit.stmikbjb@gmail.com
e-ISSN: 2685-0893
p-ISSN: 2089-3787

Analisis Sentimen Pengguna Tinder dengan Metode *Synthetic Minority Oversampling Technique*

Irene Tangke Lamba[1*], Sunneng Sandino Berutu[2], Jatmika[3]
Informatika, Universitas Kristen Immanuel, Yogyakarta, Indonesia
*e-mail *Corresponding Author.*irentangkelamba@gmail.com

Abstract

Imbalanced data tends to cause models like Naive Bayes to be biased toward the majority class, which negatively impacts model accuracy. Even if the model achieves high accuracy, it often makes incorrect predictions. The Naive Bayes model calculates class probabilities based on word frequency and the previous probability of classes. When the majority class dominates, the minority class is often overlooked. The SMOTE (Synthetic Minority Over-sampling Technique) method is applied in this study to address the imbalance in the Tinder app review data. SMOTE allows the model to be trained on a more representative distribution by balancing the quantity of samples for each class through the synthesis of novel data in the minority class. Based to the results of the evaluation, the overall model accuracy rose to 71%, while metrics for the majority class, including precision, recall, and f1-score, either stayed the same or even significantly improved. However, despite an improvement in recall for the neutral class (increased to 0.19), the performance remains low (f1-score of 0.15), indicating that the semantic complexity of neutral reviews is not sufficiently captured through synthetic data augmentation alone.

Keywords: *Naive Bayes; Synthetic Minority Over-sampling Technique; Tinder; Precision & recall; f1-score.*

Abstrak

Data yang tidak seimbang cenderung membuat model seperti naive bayes bias terhadap data mayoritas dan memberi dampak pada akurasi model. Meskipun model memiliki akurasi yang tinggi tapi sering keliru. Model *Naive Bayes* menghitung probabilitas kelas berdasarkan frekuensi kata dan prior probabilitas kelas. Jika data mayoritas lebih banyak maka data minoritas akan diabaikan. Penelitian ini bertujuan untuk menerapkan metode SMOTE dalam menangani data ulasan aplikasi Tinder yang tidak seimbang. Dengan menerapkan SMOTE, jumlah data untuk masing-masing kelas diseimbangkan melalui sintesis data baru pada kelas minoritas, sehingga model dapat dilatih pada distribusi yang lebih representatif Berdasarkan pengujian model, diperoleh hasil bahwa akurasi keseluruhan model meningkat menjadi 71%. Sementara itu parameter *precision, f1-score* dan *recall* pada kelas mayoritas tetap stabil bahkan sedikit meningkat. Namun, meskipun performa prediksi terhadap kelas netral membaik dari sisi recall (menjadi 0.19), hasilnya masih rendah (*f1-score* 0.15), yang menandakan bahwa kompleksitas semantik dari ulasan netral tidak cukup ditangkap hanya dengan penambahan data buatan.

Kata Kunci: *Naive Bayes; Synthetic Minority Over-sampling Technique; Tinder; Precision & recall; f1-score.*

1. Pendahuluan

Perkembangan teknologi digital telah mengubah cara manusia berinteraksi dan menjalin hubungan sosial. Perubahan sosial tersebut salah satunya ditunjukkan melalui kepopuleran aplikasi Tinder. Media sosial, secara umum, berfungsi sebagai arena untuk mengekspresikan diri, menyediakan kesenangan sosial, serta memicu rasa ingin tahu dan pengalaman emosional serta hedonis [1].

Saat ini, Tinder menjadi salah satu aplikasi terpopuler di Google Play Store dalam kategori kencan, dengan jutaan unduhan dan ulasan. Namun, tingginya angka penggunaan tersebut tidak selalu mencerminkan tingkat kepuasan pengguna. Banyak ulasan mencerminkan pandangan yang beragam positif, netral, maupun negatif terhadap fitur, kenyamanan, dan efektivitas aplikasi dalam menjembatani relasi sosial. Oleh karena itu, penting untuk mengukur

persepsi pengguna secara lebih objektif melalui analisis sentimen terhadap ulasan pengguna. Permasalahan ini bersifat terukur karena berbasis data ulasan aktual yang dapat diolah dan dianalisis secara kuantitatif. Selain itu dataset yang dikumpulkan biasanya sangat condong di mana beberapa kelas secara signifikan melebihi jumlah kelas lainnya [2].

Untuk menjawab permasalahan tersebut, penelitian ini mengusulkan penerapan analisis sentimen menggunakan algoritma NB yang dikombinasikan dengan metode SMOTE. Klasifikasi NB merupakan algoritma klasifikasi yang menerapkan teorema *Bayes* dan asumsi bahwa semua prediktor saling independen satu sama lain [3]. Algoritma ini bekerja dengan mengasumsikan independensi antar fitur untuk memprediksi kategori data, seperti sentimen positif, netral, atau negatif. Namun, distribusi data sentimen sering kali tidak seimbang, sehingga diperlukan teknik penyeimbang seperti SMOTE.

Penelitian ini bertujuan untuk mengetahui sentimen pengguna terhadap aplikasi Tinder berdasarkan ulasan pengguna Google Play Store, serta mengungkapkan persepsi dominan baik positif, netral, maupun negatif khususnya dari kalangan Generasi Z. Hasil penelitian ini dapat menjadi salah satu masukan bagi pengelola aplikasi untuk memahami opini pengguna dalam meningkatkan fitur layanan yang lebih relevan, sekaligus memberikan kontribusi akademik dalam pengembangan metode analisis sentimen berbasis machine learning.

2. Tinjauan Pustaka

Penelitian sebelumnya telah menerapkan metode NB untuk mengukur opini masyarakat terhadap kinerja presiden Jokowi dalam pelaksanaan pemilihan presiden pada tahun 2024. Berdasarkan rentang waktu yang dianalisis, terjadi peningkatan pada sentimen netral. Sementara itu, penurunan sentimen terjadi pada kategori positif dan negatif [4].

Salah satu studi sebelumnya menerapkan metode *Naive Bayes* untuk menganalisis opini masyarakat terhadap berbagai jenis makanan tradisional. Dari hasil analisis, gudeg memperoleh persentase sentimen positif tertinggi sebesar 57,9%, sedangkan rendang mencatat sentimen negatif tertinggi sebesar 21,9%. Adapun sentimen netral paling dominan ditemukan pada makanan hamburger. Selain itu, evaluasi performa model dengan menggunakan dataset hamburger menunjukkan hasil terbaik, dengan akurasi mencapai 0,72, *precision* sebesar 0,72, dan *recall* sebesar 0,68 dalam memprediksi sentimen [5].

Penelitian lain juga telah melakukan analisis sentimen terhadap persepsi masyarakat terkait penggunaan aplikasi MyPertamina melalui *platform* media sosial Twitter, dengan menerapkan algoritma *Naive Bayes Classifier*. Hasil evaluasi model menunjukkan performa yang cukup tinggi, dengan akurasi sebesar 82,96%, *precision* sebesar 81,17%, *recall* mencapai 86,07%, serta nilai AUC sebesar 0,906 [6].

Hingga saat ini, belum banyak penelitian yang secara khusus mengatasi isu ketidakseimbangan data dalam analisis sentimen, maupun yang menggabungkan metode *Naive Bayes* dengan teknik oversampling seperti SMOTE. Selain itu, objek kajian dalam penelitian-penelitian terdahulu belum mencakup aplikasi kencan daring seperti *Tinder*, yang merefleksikan fenomena interaksi sosial digital di era modern. Oleh karena itu, penelitian ini menghadirkan kebaruan dengan mengintegrasikan metode *Naive Bayes* dan SMOTE untuk meningkatkan kualitas prediksi sentimen pada data yang tidak seimbang, sekaligus menyoroti sentimen pengguna terhadap aplikasi Tinder sebagai studi kasus yang relatif baru dalam literatur analisis sentimen.

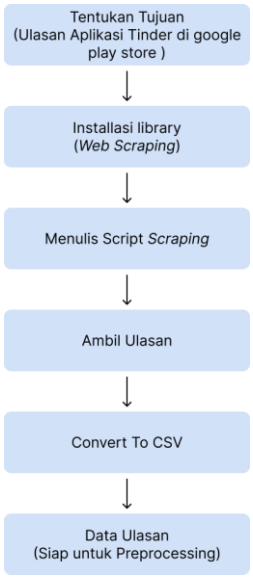
3. Metodologi

Penelitian ini menerapkan pendekatan kuantitatif dengan fokus pada proses pengumpulan serta analisis data untuk menguji hipotesis yang telah ditetapkan. Pendekatan ini dilakukan secara sistematis melalui langkah-langkah mengidentifikasi variabel, mengumpulkan data, memproses data, dan melakukan analisis statistik. metode ini menelaah, menafsirkan, dan menarik kesimpulan dari data numerik.

3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini diambil dari ulasan pengguna aplikasi kencan online Tinder di google palystore. Ulasan-ulasan tersebut mengandung berbagai opini pengguna (positif, negatif, dan netral).Data ulasan dibatasi pada rentang ulasan yang diposting dalam satu tahun terakhir.Pengambilan data ulasan dari Google Play Store dilakukan melalui teknik web

scraping. Metode ini bisa mengambil data secara praktis dan waktu yang singkat.Tahapan pengumpulan data dilakukan seperti pada diagram berikut.



Gambar 1. Diagram proses web scraping dengan python

3.2 Pra-pemrosesan Teks

Tahap Pra-pemroresan (*pre-processing*) data dilakukan dengan data cleaning (menghapus *user*, menghapus *emoticon*, dan menghapus *stop words*), *case folding*, dan *stemming* teks [7]. Tahapan *preprocessing* memiliki peran krusial dalam analisis sentimen, karena bertujuan untuk menyusun dan membersihkan teks dengan cara mengevaluasi pola serta struktur yang terdapat dalam data, baik yang bersifat semi-terstruktur maupun tidak terstruktur [8]. *Preprocessing* merupakan tahap awal yang dilakukan untuk menyiapkan data mentah sebelum dianalisis atau digunakan dalam model pembelajaran mesin. Tujuan dari proses ini adalah mengubah data yang belum terstruktur atau acak menjadi bentuk yang lebih terorganisir, konsisten, dan relevan, sehingga dapat diolah secara efektif dalam proses analisis maupun pemodelan [9]. Beberapa tahapan *Preprocessing data* di data ulasan yang telah diambil sebagai berikut :

Tokenisasi	:	Teks ulasan dipecah menjadi unit-unit kata (token) agar lebih mudah diproses dalam model analisis sentimen.
Stopwords Removal	:	Proses ini bertujuan untuk menghapus stopwords, yaitu kata-kata yang tidak memiliki kontribusi penting dalam analisis sentimen.
Stemming atau *Lemmatization*	:	Mengonversi kata-kata ke bentuk dasarnya guna menjaga konsistensi dalam proses analisis.

3.3 Labeling Sentimen

Labeling Sentimen (Simplifikasi) adalah tahapan penting dalam *preprocessing data* teks yang bertujuan mengubah berbagai bentuk label sentimen yang kompleks menjadi format yang lebih sederhana dan seragam untuk keperluan analisis. Dalam konteks ini, proses dilakukan dengan mengelompokkan berbagai ekspresi sentimen menjadi tiga kategori utama: positif, netral, dan negatif.

3.4 Implementasi SMOTE

Metode *Synthetic Minority Oversampling Technique* (SMOTE) merupakan metode dasar (*baseline*) untuk menyelesaikan masalah data yang tidak seimbang. Konsep kerja dari metode SMOTE adalah menghasilkan pola data sintetis baru dengan melakukan interpolasi linier antar

sampel dari kelas minoritas berdasarkan algoritma k-nearest neighbors. Namun, metode SMOTE memiliki kelemahan, yaitu masalah overgeneralisasi akibat o*versampling* pada data noise dan meningkatnya tumpang tindih (*overlapping*) antar kelas pada area batas keputusan (*decision boundary*), yang berpotensi menghasilkan data noise [10].Penerapan SMOTE berkontribusi pada peningkatan stabilitas dan kinerja model, yang ditunjukkan melalui kenaikan signifikan pada metrik akurasi, *presisi, recall*, dan F1-*score* [11]. SMOTE menawarkan beberapa keunggulan khas ketika diterapkan pada deteksi intrusi. SMOTE memiliki karakteristik yang sederhana dan telah terbukti efektif di berbagai bidang, termasuk dalam deteksi intrusi [12]. Berikut adalah formula SMOTE yang diterapkan pada penelitian ini :

$$x_{new} = x_i + \delta \cdot (x\ z_i - x_i) \quad \ldots\ldots\ldots\ldots (1)$$

- x_i*: contoh data minoritas.*
- x_i*: salah satu dari kkk tetangga terdekat dari* x_i
- $\delta \sim U(0,1)$ *: bilangan acak dari distribusi uniform antara 0 dan 1.*

3.5 Analisis Sentimen

Analisis sentimen merupakan proses pengolahan data berbentuk teks untuk mengidentifikasi opini yang bersifat positif, netral, atau negatif [13]. Penanganan distribusi kelas yang tidak seimbang dalam tahap pra-pemrosesan data dapat memengaruhi performa algoritma *Naïve Bayes*, terutama terhadap nilai akurasi dan *geometric mean* (G-Mean) [14]. Metode *Naïve Bayes Classifier* dinilai efektif dalam pengklasifikasian teks karena mampu menghasilkan tingkat akurasi yang cukup baik [6]. Algoritma ini bekerja berdasarkan prinsip *Teorema Bayes*, yang menggunakan pendekatan probabilitas bersyarat dalam menentukan kelas suatu data. Berikut adalah rumus probalitas bersyarat:

$$\mathrm{P(C|X)}\ \frac{(P(X|C)*P(C)}{P(X)} \quad \ldots\ldots\ldots\ldots\ldots (2)$$

P(C|X) : Probabilitas kelas C (misalnya positif/negatif) diberikan fitur X.
Ini adalah tujuan akhir dari prediksi.
P(X|C) : Probabilitas fitur X muncul jika diketahui kelas C.
Ini disebut *likelihood*, dan dihitung dari data pelatihan.
P(C) : Probabilitas terjadinya kelas C secara umum (tanpa melihat fitur).
Ini disebut *prior probability*.
P(X) : Probabilitas terjadinya fitur X secara umum.
Karena sama untuk semua kelas saat klasifikasi, ini sering diabaikan dalam perbandingan (bisa dihilangkan saat pembobotan relatif antar kelas).

3.6 Interpretasi Hasil

Pengukuran kinerja model klasifikasi di interpretasikan dengant *matrix confusion* yang parameter presisi akurasi, *recall, f1-score*, dan *confusion matrix. Confusion matrix* menghasilkan prediksi model terhadap data uji yang di gunakan, hasil prediksi di presentasikan kedalam 4 kelompok yaitu: *true positive, truenegative, false positive, false negative*. Representasi tersebut digunakan untuk mengidentifikasi kemampuan model dalam menentukan *false positive* dan negatif [15]. *Confusion matrix*, di sisi lain, menyajikan gambaran lebih rinci tentang kinerja model dengan membagi hasil prediksi ke dalam empat kategori: *true positive (TP), truenegative (TN), false positive (FP),* dan *false negative* (FN). Ini membantu mengidentifikasi seberapa baik model dapat membedakan antara kelas positif dan negatif. *Recall*, juga dikenal sebagai sensitivitas atau true *positive rate*, adalah metrik yang mengukur kemampuan model untuk menemukan semua *instance* positif dalam dataset [16]. Berikut adalah rumus masing-masing metrik evaluasi.

$$\mathrm{Accuracy}\ \frac{TP/TN}{TP+TN+FP+FN} \quad \ldots\ldots\ldots\ldots (3)$$

$$\mathrm{Precision}\ \frac{TP}{TP+FP} \quad \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

$$\text{Recall}\frac{TP}{TP+FN} \text{ …….. (5)}$$

$$\text{F1} - \text{Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{Precision+Recall} \text{ ……. (6)}$$

Table 1. Komponen metrik evaluasi

	Prediksi Positif	**Prediksi Negatif**
Aktual Positif	TP (True Positive)	FN (False Negative)
Aktual Negatif	FP (False Positive)	TN (True Negative)

4. Hasil dan Pembahasan

4.1. Hasil

4.1.1 Pengumpulan Data

Pengumpulan data dilakukan menggunakan teknik web scraping terhadap ulasan pengguna aplikasi Tinder di Google Play Store. Sebanyak 1.000 data berhasil diperoleh dengan memanfaatkan pustaka Python bernama google-play-scraper, di mana proses pengambilan data berlangsung dalam waktu kurang dari satu menit. Data ulasan diambil dari dua target yang berbeda yaitu "*us*" untuk ulasan dari *United States* dan "*id*" untuk ulasan dari Indonesia untuk memastikan keberagaman perspektif dan representasi sentimen dari dua kelompok pengguna yang berbeda. Hasil disimpan secara sistematis dalam format excel (.xlsx) agar dapat digunakan untuk tahapan selanjutnya yaitu tahap *preprocessing* dan analisis lebih lanjut.

```
Scraping US...
Scraping ID...
```

	review	rating
0	Although i pay for gold.. Matching is difficul...	2
1	I kind of hate it. What's point of Top Picks i...	2
2	Well, I would give it 0 star. Right after sign...	1
3	it's working for me	1
4	nice	5

Gambar 2. Data yang berhasil dikumpulkan dengan teknik *web scraping*

4.1.2 Pra-Pemrosesan

Dalam implementasi ini, proses *preprocessing* diawali dengan konversi seluruh huruf dalam ulasan menjadi huruf kecil (*lowercasing*). Ini penting untuk menyamakan bentuk kata misalnya, kata "*Love*" dan "*love*" seharusnya dianggap sama oleh model, dan dengan melakukan konversi ini, duplikasi bentuk kata dapat dihindari, sehingga membantu dalam proses tokenisasi dan perhitungan frekuensi kata secara lebih akurat.

Langkah selanjutnya adalah penghapusan angka dan tanda baca menggunakan ekspresi reguler (*regex*). Angka tidak relevan dalam konteks analisis sentiment pada penelitian ini, sehingga dihapus menggunakan *re.sub(r'\d+', '', text)*. Selanjutnya, tanda baca seperti titik, koma, tanda seru, tanda tanya, dan simbol non-alfanumerik lainnya dihilangkan menggunakan fungsi re.sub(r'[^\w\s]', '', text) dari pustaka regular *expression (regex).* Hal ini dilakukan untuk menyederhanakan struktur kalimat menjadi hanya kumpulan kata, yang kemudian memudahkan dalam pemrosesan lanjutan seperti *stemming* atau analisis frekuensi kata. Penghapusan tanda baca juga mengurangi *noise* pada data, sehingga meningkatkan kualitas fitur yang akan diekstraksi dari data teks.

Setelah proses pembersihan tersebut, dilakukan *tokenisasi* menggunakan fungsi *word_tokenize(text)*. Teks yang telah dibersihkan kemudian dipecah menjadi daftar kata-kata individu melalui proses tokenisasi. Langkah ini menjadi dasar dalam berbagai teknik pemrosesan bahasa alami (NLP), karena memungkinkan analisis dilakukan pada tingkat kata, bukan keseluruhan kalimat. Selanjutnya, dilakukan penyaringan terhadap *stopwords,* yaitu kata-kata umum seperti "the", "is", dan "on" yang sering muncul namun tidak memiliki kontribusi berarti

dalam penentuan sentimen. Penghapusan stopwords bertujuan untuk mengurangi beban fitur yang tidak relevan, sehingga model dapat lebih fokus pada kata-kata yang mengandung makna emosional atau opini yang kuat.

Tahapan selanjutnya adalah *stemming*, yaitu proses mengembalikan setiap kata ke bentuk dasarnya menggunakan algoritma *stemmer* yaitu *PorterStemmer*. Contohnya, kata "*running*", "*runs*", dan "*ran*" semuanya akan distemming menjadi "*run*". Hal ini penting agar semua variasi kata kerja atau bentuk jamak/singular dari kata yang sama tidak dianggap sebagai entitas yang berbeda oleh model. Setelah stemming, token-token hasil olahan tersebut digabungkan kembali menjadi sebuah string yang bersih dan sederhana dengan menggunakan *' '.join(stemmed)*. Kalimat hasil olahan ini kemudian disimpan dalam kolom baru bernama *cleaned_review*.Terakhir, data yang telah diproses kemudian diekspor ke dalam file *Excel* (*tinder_reviews_cleaned.xlsx*) menggunakan *df.to_excel(...)* agar bisa digunakan dalam tahapan analisis selanjutnya, seperti labeling sentimen, pelatihan model, dan evaluasi.

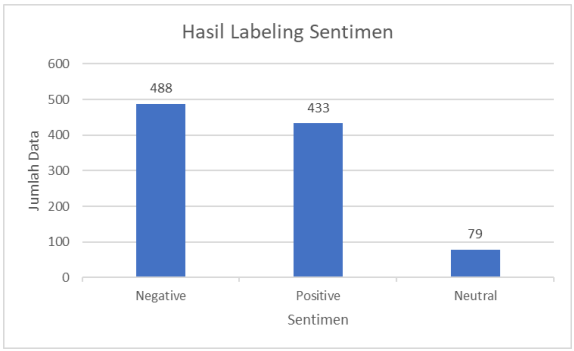
	review	rating	cleaned_review
0	Although i pay for gold.. Matching is difficul...	2	although i pay for gold match is difficult as ...
1	I kind of hate it. What's point of Top Picks i...	2	i kind of hate it what point of top pick if th...
2	Well, I would give it 0 star. Right after sign...	1	well i would give it star right after sign in ...
3	it's working for me	1	it work for me
4	nice	5	nice

Gambar 3. *Output* proses pengumpulan data di *jupyter notebook*.

Hasil akhir dari proses ini seperti yang terlihat di gambar 3, yang memperlihatkan teks ulasan yang telah disederhanakan secara struktur dan makna. Misalnya, kalimat "*Although i pay for gold.. Matching is difficult…*" berubah menjadi "*although i pay for gold match is difficult as*", yang menunjukkan bahwa tanda baca telah dihilangkan, kata kerja disederhanakan, dan kalimat menjadi lebih ringkas serta informatif. Proses ini menghasilkan data teks yang lebih bersih, relevan, dan representatif untuk dilanjutkan ke tahap labeling sentimen maupun pelatihan model analisis sentimen.

4.1.3 Labeling Sentimen

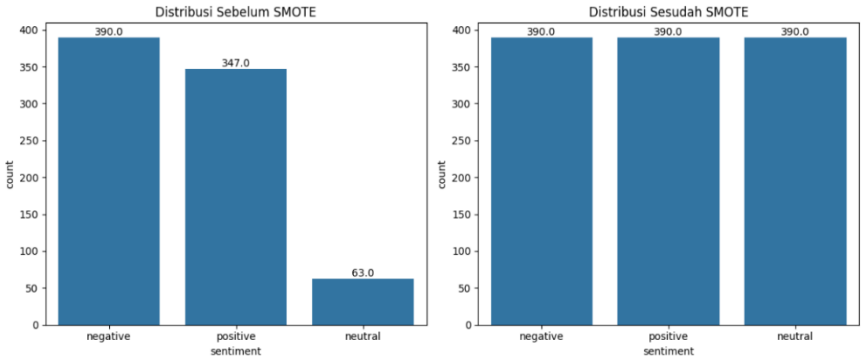
Tahapan berikutnya adalah proses labeling sentimen yang dilakukan dengan menyederhanakan skor rating pengguna menjadi tiga kategori utama: *positive*, *neutral*, dan *negative*. Fungsi label_sentiment digunakan untuk mengklasifikasikan ulasan berdasarkan nilai rating, di mana rating 4 dan 5 dianggap sebagai *positive*, rating 3 sebagai *neutral*, dan rating di bawah 3 (yaitu 1 dan 2) sebagai *negative*. Label ini kemudian ditambahkan ke kolom baru bernama sentiment dalam dataset. Hasil yang terlihat pada gambar 4, ditemukan bahwa sebagian besar ulasan termasuk ke dalam kategori *negative* sebanyak 488 data, diikuti oleh *positive* sebanyak 433, dan *neutral* sebanyak 79. Hal ini menunjukkan bahwa persepsi pengguna terhadap aplikasi Tinder dalam dataset ini cenderung negatif, yang menjadi temuan awal penting dalam analisis sentiment.



Gambar 4. Grafik hasil labeling sentiment.

4.1.4 Impementasi SMOTE

Implementasi modul SMOTE dari *imblearn. over_sampling* digunakan untuk mensintesis data baru pada kelas minoritas, yaitu *neutral*, agar distribusi data menjadi seimbang. SMOTE bekerja dengan membuat contoh baru secara sintetis berdasarkan kedekatan fitur dari sampel yang ada dalam ruang vektor. Proses ini dimulai dengan memisahkan fitur dan label, kemudian data pelatihan diseimbangkan menggunakan *fit_resample()* dari SMOTE, dan hasil distribusinya divisualisasikan melalui diagram batang seperti pada gambar 4.

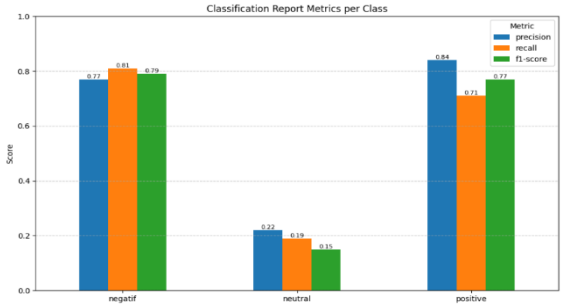


Gambar 5. Grafik hasil implementasi SMOTE.

Hasil visualisasi menunjukkan perbedaan signifikan antara distribusi sebelum dan sesudah penerapan SMOTE. Sebelum SMOTE, kelas *neutral* memiliki jumlah data yang jauh lebih sedikit dibanding *negative* dan *positive*, yang dapat menyebabkan bias model saat melakukan pelatihan. Setelah SMOTE diterapkan, jumlah data pada ketiga kelas menjadi seimbang yaitu *negative*, *positive*, dan *neutral* masing-masing memiliki jumlah data yang setara. Ini penting karena pelatihan model pada data yang seimbang memungkinkan algoritma pembelajaran mesin *Naive Bayes* untuk membentuk representasi yang adil terhadap semua kelas, sehingga meningkatkan kemampuan generalisasi model dan performa prediksi secara keseluruhan, terutama dalam mendeteksi sentimen yang sebelumnya jarang muncul.

4.1.5 Analisis Sentimen

Dalam implementasinya, *Naive Bayes* menghitung frekuensi kata dalam setiap kelas dan menggunakan pendekatan *bag-of-words* atau *TF-IDF* untuk membentuk vektor fitur. Hasil pelatihan model menunjukkan bahwa sebelum diterapkan teknik SMOTE, model lebih condong memprediksi pada kelas mayoritas, yaitu *negative* dan *positive*, sementara kelas *neutral* sering diabaikan karena kurangnya data. Namun, setelah dilakukan penyamaan distribusi data dengan SMOTE, performa model meningkat secara signifikan.Evaluasi performa model melalui metrik seperti akurasi, presisi, *recall*, dan *f1-score* menunjukkan adanya peningkatan, khususnya dalam efektivitas model dalam mengidentifikasi kelas minoritas. Akurasi keseluruhan naik menjadi 71%, yang mengindikasikan bahwa data yang telah diseimbangkan membantu model menghasilkan prediksi yang lebih representatif terhadap ketiga kelas. Selain itu, skor *precision* dan *recall* pada kelas *neutral* meningkat, memperkuat argumen bahwa keseimbangan data berperan penting dalam meningkatkan keadilan dan efektivitas klasifikasi.

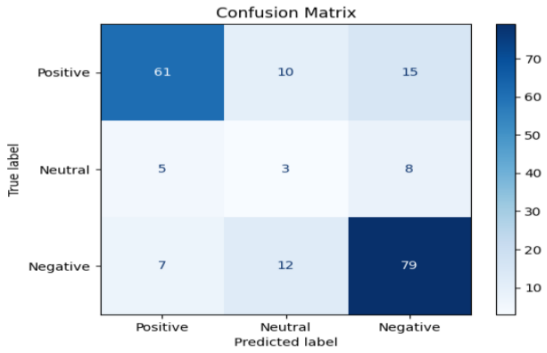


Gambar 6. Hasil analisis sentiment dengan *naïve bayes.*

Setelah penerapan SMOTE, evaluasi terhadap model *Naive Bayes* menunjukkan performa yang cukup baik secara keseluruhan, dengan akurasi mencapai 71%. Berdasarkan analisis metrik per kelas, kinerja tertinggi diperoleh pada kelas positif dengan nilai *precision* sebesar 0,84 dan f1-score sebesar 0,77. Sementara itu, kelas negatif mencatat precision sebesar 0,77 dan f1-score sebesar 0,79, yang juga menunjukkan hasil yang solid. Namun, performa model masih lemah pada kelas *neutral*, yang hanya memperoleh *precision* 0.12 dan *f1-score* 0.15, menandakan bahwa meskipun SMOTE telah diterapkan, model masih kesulitan dalam membedakan sentimen netral dari polaritas lainnya, kemungkinan karena ambiguitas dalam ekspresi netral yang lebih kompleks secara linguistik. Rata-rata *macro f1-score* hanya mencapai 0.57, yang mencerminkan ketidakseimbangan performa antar kelas. Sebaliknya, nilai weighted average *f1-score* sebesar 0,73 mengindikasikan bahwa model memiliki kinerja yang cukup baik secara menyeluruh, dengan mempertimbangkan proporsi jumlah data pada setiap kelas. Kesimpulan yang didapat adalah penerapan SMOTE membantu menyeimbangkan data dan meningkatkan akurasi total, tetapi masih terdapat ruang untuk perbaikan terutama dalam mendeteksi sentimen netral agar model dapat lebih adil dan andal dalam mengklasifikasikan seluruh jenis sentimen.

4.1.6 Interpretasi Hasil

Pada gambar 7, *confusion matrix menyajikan representasi terperinci mengenai performa klasifikasi model, dengan membandingkan antara label sebenarnya dan hasil prediksi yang dihasilkan*. Dari hasil *confusion matrix*, terlihat bahwa model memiliki performa yang cukup baik dalam mendeteksi sentimen *negative* dan *positive*, namun sangat lemah dalam mengenali sentimen *neutral*. Sebagai contoh, dari total 98 data berlabel *negative*, sebanyak 79 di antaranya berhasil diklasifikasikan dengan benar, sedangkan sisanya salah diklasifikasikan sebagai *neutral* dan *positive*. Hal yang serupa terjadi pada kelas *positive*, di mana dari 86 data, 61 di antaranya berhasil dikenali dengan benar. Hal ini menunjukkan bahwa model cukup kuat dalam mengenali emosi yang jelas (baik positif maupun negatif), yang umumnya memiliki kata-kata dengan ekspresi yang lebih eksplisit.Sebaliknya, untuk sentimen *neutral*, performa model sangat rendah. Hanya 3 dari 16 data berlabel netral yang berhasil diklasifikasikan dengan benar, sementara sisanya tersebar ke kelas *negative* dan *positive*. Hal ini mengindikasikan bahwa model kesulitan dalam menangkap makna dari ulasan yang netral atau tidak memiliki kata-kata yang jelas menggambarkan emosi. Dalam konteks algoritma *Naive Bayes*, ini bisa disebabkan oleh probabilitas kata-kata yang netral (misalnya "oke", "biasa", atau "lumayan") terlalu rendah dibandingkan kata-kata kuat seperti "*hate*", "*love*", "*terrible*", atau "*excellent*" yang sering muncul di ulasan negatif dan positif.



Gambar 7. *Confusion Matrix.*

Secara umum, confusion matrix menunjukkan bahwa meskipun model berhasil mencapai akurasi total sebesar 71%, masih terdapat ketimpangan dalam kinerja klasifikasi antar kelas. Hal ini juga didukung oleh nilai *macro average f1-score* yang rendah (0.57), mengindikasikan bahwa model belum optimal dalam memperlakukan semua kelas secara merata. Untuk meningkatkan performa di masa depan, terutama pada kelas neutral, dapat dipertimbangkan pendekatan seperti peningkatan fitur representasi (contohnya menggunakan TF-IDF atau word embedding), penggunaan model yang lebih kompleks seperti LSTM atau BERT, atau bahkan modifikasi dataset untuk menyertakan lebih banyak variasi ekspresi netral. Dengan demikian, model tidak

hanya andal dalam mendeteksi polaritas emosi ekstrem, tetapi juga mampu menangkap nuansa yang lebih halus dalam bahasa alami.

4.2. Pembahasan

Hasil pengujian menunjukkan bahwa penerapan SMOTE pada model *Naive Bayes* menghasilkan performa yang cukup memadai dalam mengklasifikasikan sentimen pengguna, meskipun model tersebut masih memiliki potensi untuk ditingkatkan lebih lanjut. Pada penerapan naive bayes yang dilakukan oleh peneliti terdahulu, naive bayes selalu memberikan hasil yang baik [4][5][6]. Dengan adanya Penerapan SMOTE membedakan hasil dari penelitian ini yang meningkatkan distribusi data ulasan menjadi lebih seimbang, yang secara langsung berdampak pada kemampuan model dalam mengenali berbagai jenis sentimen, termasuk *neutral* yang sebelumnya sulit terdeteksi. Dengan akurasi keseluruhan yang dicapai sebesar 71%, serta peningkatan pada metrik *f1-score* dan *recall*, model ini menunjukkan potensi dalam menyelesaikan tantangan analisis sentimen pada data yang tidak seimbang, yang sebelumnya sering menyebabkan bias klasifikasi dan dominasi oleh kelas mayoritasHasil pengujian dalam penelitian ini menunjukkan bahwa model *Naive Bayes* yang dibangun setelah penerapan SMOTE mampu memberikan kinerja yang cukup baik dalam mengklasifikasikan sentimen pengguna, meskipun masih terdapat ruang untuk perbaikan. Penerapan SMOTE meningkatkan distribusi data ulasan menjadi lebih seimbang, yang secara langsung berdampak pada kemampuan model dalam mengenali berbagai jenis sentimen, termasuk *neutral* yang sebelumnya sulit terdeteksi. Dengan akurasi keseluruhan yang dicapai sebesar 71%, serta peningkatan pada metrik *f1-score* dan *recall*, model ini menunjukkan potensi dalam menyelesaikan tantangan analisis sentimen pada data yang tidak seimbang, yang sebelumnya sering menyebabkan bias klasifikasi dan dominasi oleh kelas mayoritas.

5. Simpulan

Penelitian ini berhasil menunjukkan bahwa penerapan metode *Synthetic Minority Oversampling Technique* (SMOTE) efektif dalam meningkatkan kinerja analisis sentimen terhadap data ulasan pengguna aplikasi *Tinder* yang memiliki distribusi kelas tidak seimbang. Melalui tahapan yang sistematis, mulai dari pengumpulan data ulasan secara *real-time* (500 dari wilayah Indonesia dan 500 dari Amerika Serikat), pra-pemrosesan teks (*tokenisasi*, *stopword removal*, dan *stemming*), pelabelan sentimen berbasis skor rating, hingga implementasi algoritma klasifikasi *Naive Bayes*, penelitian ini memberikan gambaran menyeluruh tentang bagaimana ketidakseimbangan data dapat memengaruhi akurasi dan generalisasi model.

Hasil awal menunjukkan dominasi kelas negatif dan positif terhadap kelas netral, yang berdampak pada rendahnya akurasi dan kesulitan model dalam mengklasifikasikan sentimen minoritas. Dengan penerapan SMOTE, distribusi data berhasil diseimbangkan sehingga model dapat belajar dari semua kelas secara lebih proporsional. Evaluasi model pasca penerapan SMOTE menunjukkan perbaikan pada sejumlah metrik utama seperti akurasi (71%), *precision, recall,* dan *f1-score*. Namun demikian, model masih menghadapi kendala dalam mengidentifikasi sentimen netral dengan tingkat akurasi yang memadai. Visualisasi distribusi sebelum dan sesudah SMOTE serta hasil confusion matrix memperkuat temuan bahwa SMOTE berperan signifikan dalam meningkatkan performa klasifikasi terhadap kelas yang sebelumnya terpinggirkan.

Secara keseluruhan, penelitian ini membuktikan bahwa penerapan SMOTE dalam analisis sentimen merupakan pendekatan yang efektif untuk mengatasi ketidakseimbangan kelas, meningkatkan kinerja model, dan memperluas kemampuan sistem dalam memahami berbagai ekspresi sentimen pengguna. Temuan ini tidak hanya relevan untuk pengembangan model analisis sentimen berbasis teks, tetapi juga memberikan kontribusi pada strategi pengambilan keputusan berbasis data dalam konteks aplikasi berbasis ulasan pengguna, terutama dalam domain yang sensitif terhadap persepsi publik seperti aplikasi kencan online.

Daftar Referensi

[1] H. L. A. Hart, *Punishment and Responsibility: Essays in the Philosophy of Law*. Oxford: OUP Oxford, 2008. [Online]. Available: https://books.google.co.id/books?id=IMTmCwAAQBAJ

[2] M. Mukherjee and M. Khushi, “SMOTE-ENC: A novel SMOTE-based method to generate synthetic data for nominal and continuous features,” *Applied System Innovation*, vol. 4, no. 1, p. 18, 2021.

[3] O. Peretz, M. Koren, and O. Koren, “Naive Bayes classifier—An ensemble procedure for recall and precision enrichment,” *Engineering Applications of Artificial Intelligence*, vol. 136, p. 108972, 2024.

[4] P. H. Nehe, S. S. Berutu, and H. Budiati, “Analisis Sentimen Opini Masyarakat Terhadap Presiden Jokowi Sebelum Dan Sesudah Pilpres 2024 Menggunakan Metode Naive Bayes Classification,” Jutisi J. Ilm. Tek. Inform. dan Sist. Inf., vol. 13, no. 1, p. 451, 2024, doi: 10.35889/jutisi.v13i1.1841.

[5] S. S. Berutu, “Text Mining dan Klasifikasi Sentimen Berbasis Naïve Bayes Pada Opini Masyarakat terhadap Makanan Tradisional,” *J. Sist. Komput. dan Inform.*, vol. 4, no. 2, p. 254, 2022, doi: 10.30865/json.v4i2.5138.

[6] R. Maria, R. U. Umayah, S. Mahardinny, D. Kalana, and D. D. Saputra, “Analisis Sentimen Persepsi Masyarakat Terhadap Penggunaan Aplikasi My Pertamina Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier,” *Jurnal Komputer Antartika*, vol. 1, no. 1, pp. 1–10, 2023.

[7] S. S. Berutu, H. Budiati, J. Jatmika, and F. Gulo, “Data preprocessing approach for machine learning-based sentiment classification,” *Jurnal Infotel*, vol. 15, no. 4, pp. 317–325, 2023, doi: 10.20895/infotel.v15i4.1030.

[8] S. M. Chamzah, M. Lestandy, N. Kasan, A. Nugraha *et al*., “Penerapan Synthetic Minority Oversampling Technique (SMOTE) untuk Imbalance Class pada Data Text Menggunakan KNN,” *Syntax: Jurnal Informatika*, vol. 11, no. 02, pp. 56–67, 2022.

[9] D. Berniawan, A. Amri, and T. Tinaliah, “Implementasi Algoritma Naïve Bayes Untuk Klasifikasi Sentimen Pengguna Twitter Terhadap KEMKOMINFO Di Indonesia,” *MDP Student Conf.*, vol. 2, no. 1, pp. 24–31, 2023, doi: 10.35957/mdp-sc.v2i1.4326.

[10] H. Hairani, T. Widiyaningtyas, and D. Dwi Prasetya, “Addressing Class Imbalance of Health Data: A Systematic Literature Review on Modified Synthetic Minority Oversampling Technique (SMOTE) Strategies,” *JOIV Int. J. Informatics Vis.*, vol. 8, no. 3, pp. 1310–1318, 2024.

[11] T. P. W. Sukma and M. R. Pribadi, “Analisis Sentimen Review Pengguna Viu Pada Play Store Dengan Algoritma Random Forest,” *Jurnal Software Engineering and Computational Intelligence*, vol. 2, no. 01, pp. 9–16, 2024.

[12] H. R. Sayegh, W. Dong, and A. M. Al-madani, “Enhanced Intrusion Detection with LSTM-Based Model, Feature Selection, and SMOTE for Imbalanced Data,” *Appl. Sci.*, vol. 14, no. 2, pp. 1–20, 2024, doi: 10.3390/app14020479.

[13] J. S. Gea and H. Budiati, “Analisis Sentimen Masyarakat Terhadap Direktorat Jenderal Pajak,” *Jurnal Sains Dan Komputer*, vol. 8, no. 01, pp. 30–36, 2024, doi: 10.61179/jurnalinfact.v8i01.466.

[14] M. Sulistiyono, Y. Pristyanto, S. Adi, and G. Gumelar, “Implementasi algoritma synthetic minority over-sampling technique untuk menangani ketidakseimbangan kelas pada dataset klasifikasi,” *Sistemasi: Jurnal Sistem Informasi*, vol. 10, no. 2, pp. 445–459, 2021.

[15] O. Caelen, “A Bayesian interpretation of the confusion matrix,” *Annals of Mathematics and Artificial Intelligence*, vol. 81, no. 3, pp. 429–450, 2017.

[16] W. Wahyudi, R. Kurniawan, and Y. A. Wijaya, “Analisis Sentimen Pengguna Terhadap Aplikasi Blu BCA di Playstore Menggunakan Algoritma Naïve Bayes,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, pp. 2511–2517, 2024.