

# Peningkatan Kinerja Model Naïve Bayes untuk Analisis Sentimen Komentar Terkait “Sound Horeg” Menggunakan SMOTE dan Tuning Parameter

Mustafid Kaisalana, Gustina Alfa Trisnapradika\*

Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia

Email: <sup>1</sup>mustafid.mk@gmail.com, <sup>2,\*</sup>gustina.alfa@dsn.dinus.ac.id

Email Penulis Korespondensi: gustina.alfa@dsn.dinus.ac.id

Submitted: 17/10/2025; Accepted: 06/12/2025; Published: 08/12/2025

**Abstrak**—Fenomena “Sound Horeg” di platform daring memicu beragam sentimen, menjadikan analisis sentimen sebagai alat penting untuk memahami opini publik. Penelitian ini bertujuan untuk mengklasifikasikan sentimen (positif/negatif) pada komentar pengguna terkait “Sound Horeg” menggunakan algoritma Naïve Bayes. Dataset yang digunakan menunjukkan ketidakseimbangan kelas yang signifikan, dengan dominasi sentimen negatif. Metodologi yang diterapkan meliputi serangkaian tahapan preprocessing teks, yaitu case folding, tokenizing, normalization, sentiment labeling berbasis leksikon, stopword removal, stemming, dan penghapusan duplikasi. Proses sentiment labeling dilakukan menggunakan leksikon sentimen bahasa Indonesia yang disusun dari dua sumber yaitu lexicon\_positif.csv dan lexicon\_negatif.csv, yang berisi daftar kata dengan skor sentimen positif dan negatif yang telah ditetapkan berdasarkan kamus opini publik bahasa Indonesia. Selanjutnya, fitur teks diekstraksi menggunakan Term Frequency–Inverse Document Frequency (TF-IDF). Untuk mengatasi ketidakseimbangan data, diterapkan teknik oversampling SMOTE (Synthetic Minority Oversampling Technique) pada data latih guna menyeimbangkan jumlah sampel positif dan negatif. Model Naïve Bayes kemudian dioptimasi menggunakan GridSearchCV untuk menentukan nilai alpha terbaik. Hasil eksperimen menunjukkan bahwa model Naïve Bayes tanpa optimasi memiliki akurasi 73%, namun kesulitan dalam mengklasifikasikan kelas minoritas (sentimen positif) akibat bias data. Setelah penerapan SMOTE dan tuning parameter, kinerja model mengalami peningkatan yang signifikan, membuktikan efektivitas kedua teknik tersebut dalam menghasilkan model yang lebih seimbang dan robust. Penelitian ini menyimpulkan bahwa Naïve Bayes, dengan optimasi berbasis SMOTE dan hyperparameter tuning, efektif untuk analisis sentimen berbahasa Indonesia, terutama pada dataset yang tidak seimbang. Pengembangan di masa depan dapat mencakup eksplorasi algoritma lain dan penggunaan leksikon sentimen yang lebih luas serta fitur linguistik yang lebih kompleks untuk meningkatkan kinerja model.

**Kata Kunci:** Analisis Sentimen; Naïve Bayes; Sound Horeg; Klasifikasi Teks; SMOTE; Hyperparameter Tuning

**Abstract**—The phenomenon of “Sound Horeg” on online platforms has sparked diverse public sentiments, making sentiment analysis an essential tool for understanding public opinion. This study aims to classify user sentiments (positive/negative) related to “Sound Horeg” using the Naïve Bayes algorithm. The dataset used in this research exhibits significant class imbalance, with a predominance of negative sentiments. The methodology involves a series of text preprocessing stages, including case folding, tokenizing, normalization, lexicon-based sentiment labeling, stopword removal, stemming, and duplicate removal. The sentiment labeling process utilizes an Indonesian sentiment lexicon compiled from two sources lexicon\_positif.csv and lexicon\_negatif.csv containing predefined lists of words with positive and negative sentiment scores based on Indonesian public opinion lexicons. Subsequently, text features are extracted using the Term Frequency–Inverse Document Frequency (TF-IDF) method. To address data imbalance, the Synthetic Minority Oversampling Technique (SMOTE) is applied to the training data to balance the number of positive and negative samples. The Naïve Bayes model is then optimized using GridSearchCV to determine the best alpha value. Experimental results show that the unoptimized Naïve Bayes model achieved an accuracy of 73%, but struggled to classify minority classes (positive sentiments) due to data bias. After applying SMOTE and parameter tuning, the model’s performance improved significantly, demonstrating the effectiveness of these techniques in producing a more balanced and robust model. This study concludes that the Naïve Bayes algorithm, when optimized with SMOTE and hyperparameter tuning, is effective for Indonesian-language sentiment analysis, particularly on imbalanced datasets. Future work may include exploring other algorithms and employing broader sentiment lexicons and more complex linguistic features to further enhance model performance.

**Keywords:** Sentiment Analysis; Naïve Bayes; Sound Horeg; Text Classification; SMOTE; Hyperparameter Tuning

## 1. PENDAHULUAN

Fenomena "Sound Horeg", yang merujuk pada penggunaan sistem suara bertenaga tinggi dalam acara publik, telah menjadi subjek diskusi yang intensif di kalangan masyarakat Indonesia, memicu beragam reaksi mulai dari ketidaknyamanan hingga bentuk apresiasi, terutama yang terekam dalam komentar-komentar di platform media sosial [1], [2]. Data opini publik yang bersumber dari media sosial mencerminkan keragaman persepsi dan sikap masyarakat secara beragam terhadap suatu fenomena [3], [4]. Dalam era big data, di mana volume informasi tekstual terus meningkat, analisis sentimen otomatis menjadi instrumen krusial untuk mengidentifikasi, mengklasifikasikan, dan memahami opini, sikap, serta emosi pengguna secara efisien dan sistematis [5], [6]. Oleh karena itu, penelitian ini berupaya menjawab akan pengembangan metode otomatis yang efektif dalam mengklasifikasikan sentimen pada data komentar berbahasa Indonesia yang berkaitan dengan "Sound Horeg" [7].

Penelitian ini secara spesifik diarahkan untuk menganalisis sentimen yang terkandung dalam komentar pengguna mengenai "Sound Horeg" guna memperoleh gambaran kuantitatif mengenai persepsi dan opini masyarakat. Fokus utama penelitian ini adalah mengeksplorasi efektivitas algoritma machine learning, khususnya Naive Bayes, dalam menangani tugas klasifikasi sentimen pada konteks bahasa Indonesia yang memiliki karakteristik linguistik

unik[8]. Penggunaan Naive Bayes sangat relevan dalam konteks ini karena kemampuannya yang efisien dalam melatih model dan melakukan prediksi dengan cepat, menjadikannya pilihan yang akurat untuk klasifikasi teks berskala besar[9], [10]. Algoritma ini juga dikenal unggul dalam menangani data dengan dimensi tinggi, seperti yang umum ditemukan pada data teks yang direpresentasikan sebagai kantung kata (*bag-of-words*)[11], [12]. Meskipun memiliki asumsi independensi fitur yang "naif," Naive Bayes secara empiris tetap menunjukkan performa yang kompetitif bahkan seringkali mengungguli algoritma yang lebih kompleks dalam tugas klasifikasi teks[13]. Lebih lanjut, penelitian ini juga akan menginvestigasi secara mendalam dampak penerapan teknik preprocessing data yang canggih, seperti oversampling menggunakan SMOTE (*Synthetic Minority Oversampling Technique*) untuk mengatasi ketidakseimbangan kelas data, serta tuning parameter model, terhadap peningkatan akurasi dan kinerja keseluruhan model Naive Bayes[14], [15]. Selain itu, Naive Bayes juga relatif tidak sensitif terhadap fitur-fitur yang tidak relevan, yang seringkali ada dalam data tekstual[16]. Ketidakseimbangan kelas merupakan tantangan umum dalam analisis sentimen, terutama ketika data minoritas tidak cukup terwakili, sehingga berdampak pada performa klasifikasi algoritma seperti Naive Bayes[17], [18]. Salah satu pendekatan paling efektif untuk mengatasi masalah ini adalah penggunaan Synthetic Minority Oversampling Technique (SMOTE), yang telah dibuktikan secara teoritis mampu mendekati distribusi data sebenarnya[19]. SMOTE telah terbukti secara konsisten meningkatkan akurasi, presisi, dan F1-score pada klasifikasi emosi dalam data sosial media yang tidak seimbang, melampaui teknik sampling lainnya seperti random under sampling[10].

Sejumlah penelitian sebelumnya telah menunjukkan penggunaan SMOTE untuk mengatasi ketidakseimbangan kelas pada analisis sentimen berbahasa Indonesia[20]. Sementara itu, aspek tuning parameter melalui teknik seperti GridSearch atau meta-heuristik Naive Bayes baru mulai mendapat perhatian[21]. Namun demikian, hingga saat ini belum ditemukan studi yang secara khusus mengombinasikan algoritma Naive Bayes, SMOTE, dan GridSearchCV dalam konteks fenomena sosial budaya seperti "Sound Horeg".

Selain penggunaan teknik penyeimbangan data, optimasi parameter model (*hyperparameter tuning*) juga merupakan aspek penting dalam peningkatan performa algoritma *machine learning*. Proses ini bertujuan untuk menemukan kombinasi nilai parameter yang menghasilkan kinerja terbaik berdasarkan metrik evaluasi tertentu. Dalam konteks penelitian ini, optimasi dilakukan menggunakan metode *GridSearchCV*, yaitu pendekatan sistematis yang mengevaluasi berbagai kombinasi parameter melalui validasi silang (*cross-validation*). Penerapan *GridSearchCV* memungkinkan pemilihan nilai parameter smoothing ( $\alpha$ ) yang paling optimal bagi algoritma *Naive Bayes*, sehingga dapat meminimalkan kesalahan prediksi akibat *overfitting* maupun *underfitting*[22]. Sejumlah penelitian sebelumnya juga menunjukkan bahwa *hyperparameter tuning* dapat memberikan peningkatan signifikan terhadap akurasi dan stabilitas model klasifikasi teks, terutama ketika dikombinasikan dengan teknik penyeimbangan data seperti SMOTE[23]. Berdasarkan latar belakang dan penelitian yang telah diuraikan, dua rumusan masalah utama yang menjadi inti adalah bagaimana performa algoritma Naive Bayes dalam mengklasifikasikan sentimen (positif atau negatif) pada komentar pengguna terkait "Sound Horeg" dalam konteks data yang tidak seimbang? Kedua, sejauh mana implementasi teknik oversampling SMOTE dan tuning parameter model secara sistematis dapat berkontribusi pada peningkatan akurasi klasifikasi sentimen oleh model Naive Bayes? Pertanyaan-pertanyaan ini akan menjadi panduan dalam pelaksanaan eksperimen dan analisis data.

Sejalan dengan perumusan masalah tersebut, penelitian ini memiliki beberapa tujuan yang terukur. Tujuan pertama adalah untuk mengembangkan dan membangun model klasifikasi sentimen yang robust untuk komentar pengguna terkait "Sound Horeg" menggunakan algoritma Naive Bayes. Tujuan kedua adalah untuk melakukan analisis komprehensif dan evaluasi kinerja model Naive Bayes berdasarkan metrik standar seperti akurasi, presisi, recall, dan F1-Score, baik sebelum maupun sesudah optimasi. Terakhir, penelitian ini bertujuan untuk secara empiris mengevaluasi pengaruh signifikan dari penerapan SMOTE dan tuning parameter terhadap peningkatan kinerja klasifikasi sentimen yang dicapai oleh model Naive Bayes, sehingga dapat memberikan kontribusi pada pengembangan metode analisis sentimen yang lebih akurat.

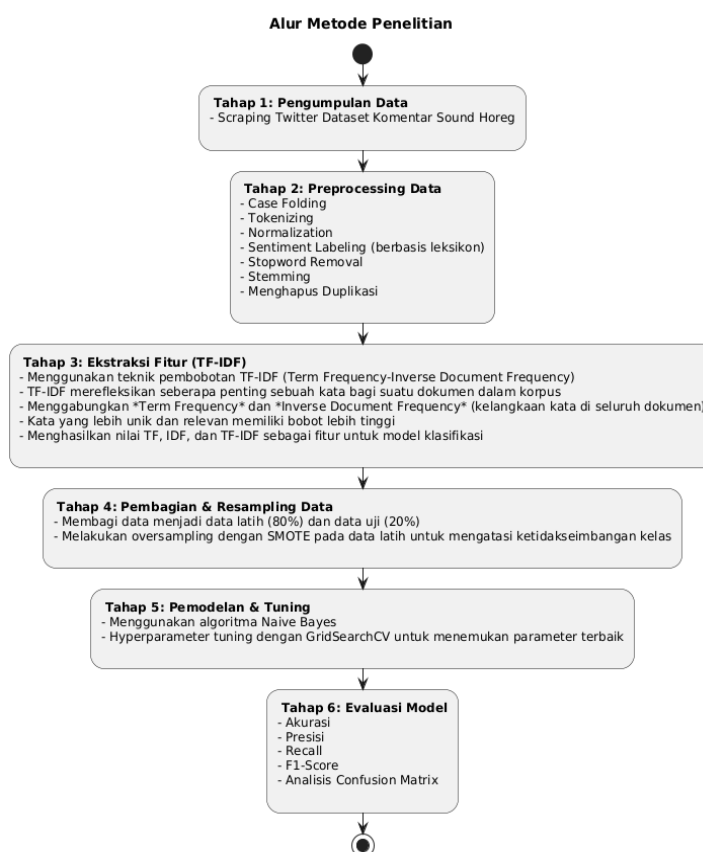
Selain aspek teknis, penelitian ini juga diharapkan memberikan kontribusi praktis terhadap pemahaman fenomena sosial di Indonesia, khususnya mengenai dinamika persepsi publik terhadap "Sound Horeg". Dengan memanfaatkan hasil klasifikasi sentimen, pihak-pihak terkait seperti penyelenggara acara, pemerintah daerah, maupun pembuat kebijakan dapat memperoleh wawasan yang lebih objektif mengenai tingkat penerimaan atau penolakan masyarakat terhadap penggunaan sistem suara bertenaga tinggi[24]. Hasil analisis ini dapat dijadikan dasar dalam merancang kebijakan atau pedoman teknis yang mampu menyeimbangkan kebutuhan hiburan publik dengan kenyamanan lingkungan, sekaligus meminimalkan potensi konflik sosial yang mungkin timbul.

Lebih jauh lagi, penelitian ini diharapkan menjadi kontribusi akademik dalam bidang Natural Language Processing (NLP) berbahasa Indonesia, yang hingga saat ini masih menghadapi tantangan signifikan seperti keterbatasan dataset berkualitas, keragaman dialek, serta kompleksitas morfologi bahasa. Pengembangan model klasifikasi yang dioptimalkan melalui SMOTE dan tuning parameter dapat menjadi referensi bagi peneliti lain yang ingin mengkaji analisis sentimen dalam konteks serupa. Dengan demikian, penelitian ini tidak hanya memberikan solusi terhadap permasalahan yang spesifik pada fenomena "Sound Horeg," tetapi juga mendorong perkembangan metode komputasional yang dapat diaplikasikan pada isu-isu sosial lainnya yang memerlukan analisis opini publik secara sistematis dan berbasis data.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Fenomena “Sound Horeg” yang menjadi objek penelitian ini merupakan praktik penggunaan perangkat pengeras suara secara berlebihan dalam kegiatan hiburan atau perayaan, yang menimbulkan kontroversi dan memicu berbagai reaksi di media sosial. Komentar-komentar pengguna di platform X (Twitter) mencerminkan persepsi publik yang beragam, mulai dari dukungan terhadap kebebasan berekspresi hingga kecaman terhadap gangguan kebisingan yang ditimbulkan. Oleh karena itu, analisis sentimen terhadap komentar-komentar ini menjadi relevan untuk memahami pola opini masyarakat terhadap fenomena tersebut. Dalam konteks ini, Naïve Bayes dipilih karena kemampuannya yang efisien dalam menangani teks berdimensi tinggi serta interpretasinya yang sederhana namun efektif untuk klasifikasi biner seperti positif dan negatif. Penelitian ini mengadopsi pendekatan eksperimental dalam bidang machine learning untuk mengembangkan dan mengevaluasi model klasifikasi sentimen pada komentar pengguna terkait “Sound Horeg” menggunakan algoritma Naive Bayes. Pendekatan ini dipilih karena kemampuannya dalam membangun sistem klasifikasi yang dapat secara otomatis mengidentifikasi polaritas sentimen (positif atau negatif) dari data teks, sekaligus mengatasi tantangan data yang tidak seimbang melalui teknik oversampling. Fokus penelitian ini terbatas pada tahapan preprocessing data, pengembangan model, optimasi parameter, dan evaluasi kinerja model Naive Bayes, tanpa mencakup implementasi sistem secara penuh dalam lingkungan produksi atau analisis linguistik yang lebih mendalam di luar kerangka preprocessing yang telah ditentukan.



**Gambar 1.** Metode Penelitian

Gambar 1 merupakan tahapan utama penelitian analisis sentimen terhadap komentar “Sound Horeg” yang diambil dari platform Twitter. Alur ini menunjukkan proses dari tahap awal pengumpulan data hingga evaluasi model secara sistematis.

### 2.2 Pengumpulan Data

Tahap awal penelitian adalah pengumpulan data dengan teknik scraping dari platform Twitter, khususnya dataset komentar akun Sound Horeg. Data yang dikumpulkan berupa teks komentar pengguna yang mengandung opini atau sentiment berjumlah 5882 komentar. Hasil scraping kemudian disimpan dalam format CSV agar mudah diproses lebih lanjut pada tahap berikutnya.

### 2.3 Preprocessing Data

Tahapan preprocessing dilakukan untuk menyiapkan data komentar yang bersifat tidak terstruktur dan penuh variasi gaya bahasa informal khas media sosial. Mengingat topik “Sound Horeg” sering diungkapkan dengan bahasa

emosional, campuran bahasa daerah, atau singkatan, maka proses normalisasi, stemming, dan pembersihan teks menjadi krusial agar model dapat mengenali pola sentimen secara konsisten.

Data hasil scraping perlu dibersihkan agar siap digunakan dalam analisis. Proses preprocessing yang dilakukan meliputi beberapa langkah:

- a. Case Folding, mengubah seluruh huruf menjadi huruf kecil agar konsisten.
- b. Tokenizing, memecah kalimat menjadi kata-kata.
- c. Normalization, menstandarkan kata agar seragam.
- d. Sentiment Labeling, memberikan label sentimen pada setiap komentar berdasarkan leksikon sentimen.
- e. Stopword Removal, menghapus kata yang tidak memiliki makna signifikan.
- f. Stemming, mengubah kata ke bentuk dasarnya.
- g. Menghapus duplikasi, untuk menghindari bias pada model.

## 2.4 Ekstraksi Fitur TF-IDF

Untuk memberikan pemahaman yang lebih mendalam mengenai proses ekstraksi fitur, bagian ini menyajikan contoh konkret perhitungan TF-IDF (Term Frequency-Inverse Document Frequency) untuk sebuah sampel komentar. TF-IDF adalah teknik pembobotan statistik yang merefleksikan seberapa penting sebuah kata bagi sebuah dokumen dalam suatu korpus. Metode ini tidak hanya mempertimbangkan frekuensi kemunculan kata dalam dokumen tertentu (Term Frequency), tetapi juga seberapa jarang kata tersebut muncul di seluruh koleksi dokumen (Inverse Document Frequency). Dengan demikian, kata-kata yang unik dan relevan pada suatu dokumen akan memiliki bobot yang lebih tinggi, menjadikannya fitur yang lebih informatif untuk model klasifikasi.

## 2.5 Pembagian dan Resampling Data

Data yang telah melalui proses ekstraksi fitur kemudian dibagi menjadi dua bagian, yaitu data latih sebesar 80% dan data uji sebesar 20%. Namun, pada dataset komentar ini, terdapat ketidakseimbangan jumlah antara kelas positif dan negatif.

Untuk mengatasi hal tersebut, digunakan metode SMOTE (Synthetic Minority Over-sampling Technique) yang menghasilkan sampel sintetis pada kelas minoritas. Dengan demikian, model mendapatkan lebih banyak contoh dari kelas positif untuk dipelajari, sehingga mengurangi bias terhadap kelas mayoritas.

## 2.6 Modelling

Pada tahap ini, algoritma Naive Bayes diimplementasikan menggunakan MultinomialNB dari library scikit-learn. Untuk mengoptimalkan kinerja model, dilakukan proses hyperparameter tuning menggunakan GridSearchCV. Proses ini dimulai dengan mendefinisikan model Naive Bayes dan menetapkan rentang nilai hyperparameter alpha yang akan diuji, yaitu [0.001, 0.01, 0.1, 0.5, 1, 2, 5, 10]. GridSearchCV secara sistematis mengeksplorasi setiap kombinasi hyperparameter ini dengan menerapkan 10-fold cross-validation ( $cv=10$ ) pada data latih yang telah diseimbangkan. Penggunaan  $n\_jobs=-1$  memastikan pemanfaatan seluruh core CPU untuk mempercepat proses komputasi.

Sebelum tahap *training* dilakukan, data mengalami proses penyeimbangan kelas (class balancing) menggunakan teknik SMOTE. Masalah ketidakseimbangan kelas merupakan tantangan umum dalam klasifikasi data, di mana jumlah sampel dari satu kelas (kelas mayoritas) jauh lebih banyak dibandingkan kelas lainnya (kelas minoritas). Dalam konteks penelitian ini, data sentimen awal menunjukkan ketidakseimbangan yang signifikan, yaitu terdapat 3.422 komentar berlabel negatif (kelas mayoritas) dan hanya 1.283 komentar berlabel positif (kelas minoritas). Kondisi ini dapat menyebabkan model cenderung bias terhadap kelas mayoritas, sehingga performa klasifikasi pada kelas minoritas menjadi buruk.

Teknik SMOTE bekerja dengan cara menghasilkan sampel sintetis baru dari kelas minoritas, bukan hanya menduplikasi sampel yang sudah ada. Prosesnya melibatkan pemilihan sampel dari kelas minoritas, mengidentifikasi beberapa tetangga terdekatnya berdasarkan fitur, lalu membuat sampel baru di sepanjang garis yang menghubungkan sampel asli dengan tetangga tersebut. Dengan demikian, SMOTE mampu memperluas representasi kelas minoritas secara lebih alami dan realistis.

## 2.7 Evaluasi

Evaluasi model dilakukan untuk menilai kinerja algoritma klasifikasi dalam mengidentifikasi kategori data secara tepat. Proses evaluasi ini menggunakan sejumlah metrik yang umum diterapkan dalam bidang pembelajaran mesin, yaitu precision, recall, f1-score, dan accuracy. Keempat metrik tersebut dipilih karena mampu memberikan gambaran menyeluruh mengenai kemampuan model dalam memprediksi kelas dengan benar, baik dari sisi ketepatan maupun kelengkapan hasil prediksi. Selain itu, pada penelitian ini juga dilakukan penanganan terhadap permasalahan ketidakseimbangan kelas yang terdapat dalam dataset. Ketidakseimbangan ini dapat menyebabkan model cenderung lebih akurat dalam memprediksi kelas mayoritas dan mengabaikan kelas minoritas. Untuk mengatasi permasalahan tersebut, digunakan pendekatan Synthetic Minority Over-sampling Technique (SMOTE) yang berfungsi menghasilkan sampel sintetis pada kelas minoritas, sehingga distribusi data menjadi lebih seimbang dan model dapat belajar secara lebih optimal.

Evaluasi dilakukan melalui dua skenario utama, yaitu tanpa penerapan SMOTE dan dengan penerapan SMOTE disertai proses tuning parameter menggunakan teknik cross-validation. Pendekatan ini bertujuan untuk membandingkan performa model sebelum dan sesudah dilakukan penyeimbangan data.

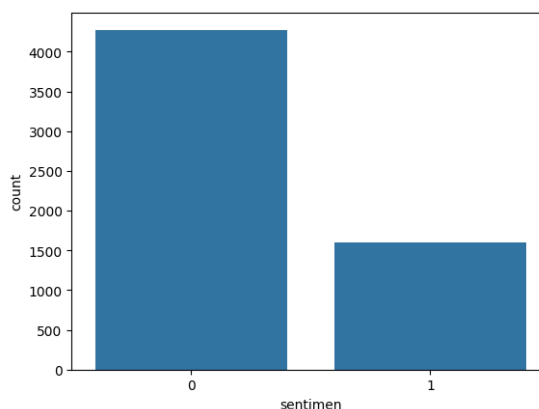
### 3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan temuan-temuan kunci dari eksperimen klasifikasi sentimen komentar pengguna terkait "Sound Horeg" menggunakan algoritma Naive Bayes, serta analisis mendalam terhadap kinerja model sebelum dan sesudah optimasi. Pembahasan akan menguraikan bagaimana setiap tahapan metodologi berkontribusi pada hasil akhir, mengaitkan temuan dengan kerangka teori, dan menyoroti implikasi penelitian ini.

#### 3.1 Pengumpulan data

Data yang digunakan dalam penelitian ini diperoleh melalui proses scraping menggunakan bahasa pemrograman Python dengan bantuan pustaka sncrape untuk mengumpulkan komentar dari media sosial Twitter yang mengandung kata kunci "Sound Horeg". Proses pengumpulan data ini dilakukan dengan tujuan untuk memperoleh data opini publik yang bersifat alami dan terkini mengenai fenomena tersebut.

Setelah dilakukan proses scraping, diperoleh sebanyak 5.882 data komentar, yang terdiri atas 4.276 data dengan sentimen negatif (label 0) dan 1.606 data dengan sentimen positif (label 1). Jumlah data tersebut kemudian digunakan sebagai dasar dalam tahap preprocessing dan analisis selanjutnya.



**Gambar 2 .** Distribusi Data Sentimen

Hasil distribusi data sentimen sebelum dilakukan resampling divisualisasikan pada Gambar 1. Berdasarkan visualisasi tersebut, dapat diketahui bahwa terdapat ketidakseimbangan jumlah data antara kelas sentimen positif dan negatif, di mana kelas negatif jauh lebih dominan.

Ketidakseimbangan data ini dapat berdampak pada performa model klasifikasi, karena algoritma cenderung lebih akurat dalam memprediksi kelas dengan jumlah data yang lebih banyak. Oleh karena itu, pada tahap berikutnya dilakukan proses oversampling menggunakan metode SMOTE (Synthetic Minority Oversampling Technique) untuk menyeimbangkan jumlah data antar kelas.

#### 3.2 Preprocessing Data

- Case Folding:* Tahap ini merupakan langkah awal dalam preprocessing teks yang bertujuan untuk menyeragamkan seluruh teks dengan mengubahnya menjadi huruf kecil (lowercase). Hal ini penting untuk memastikan bahwa sistem tidak memperlakukan kata yang sama sebagai entitas yang berbeda hanya karena perbedaan penggunaan huruf kapital (misalnya, "DPR" dan "dpr" akan dianggap sebagai kata yang sama). Selain itu, pada tahap ini juga dilakukan penghapusan karakter-karakter yang tidak relevan atau tidak informatif untuk analisis sentimen, seperti angka, tanda baca (misalnya, titik, koma, tanda seru), simbol, dan karakter non-alfabet lainnya (misalnya, emoticon atau hashtag yang tidak dianalisis). Proses ini membantu mengurangi kompleksitas data dan fokus pada konten tekstual yang relevan.

**Tabel 1.** Case Folding

Sebelum	Sesudah
@ramaArsyl Testimoni dariku org Jatim yg yg jg jijik sama hal ginian. Anaku pernah disuruh minum obat tidur sama RT biar ga keganggu sama sound horeg gini yg diputer tetangga ibuku. Udh ditegur baik2 toh anaku ga tiap hari dirumah ibuku cuman pas weekend doang. Malah nantangin suruh lapor	testimoni dariku org jatim yg yg jg jijik sama hal ginian anaku pernah disuruh minum obat tidur sama rt biar ga keganggu sama sound horeg gini yg diputer tetangga ibuku udh ditegur baik toh anaku ga tiap hari dirumah ibuku cuman pas weekend doang malah nantangin suruh lapor



- b. **Tokenizing:** Setelah case folding, tahap tokenizing dilakukan untuk memecah atau membagi kalimat menjadi unit-unit kata individual yang lebih kecil, yang dikenal sebagai token. Proses ini mengubah aliran teks kontinu dari format kalimat menjadi daftar atau urutan kata-kata yang terpisah. Setiap kata yang terdeteksi akan menjadi satu token, yang kemudian menjadi dasar untuk analisis lebih lanjut seperti normalisasi, penghapusan stopwords, dan stemming. Tokenisasi yang akurat sangat penting karena kesalahan pada tahap ini dapat memengaruhi kualitas analisis sentimen berikutnya.

**Tabel 2.** Tokenizing

Sebelum	Sesudah
testimoni dariku org jatim yg yg jg jijik sama hal ginian anakku pernah disuruh minum obat tidur sama rt biar ga keganggu sama sound horeg gini yg diputer tetangga ibuku udh ditegur baik toh anaku ga tiap hari dirumah ibuku cuman pas weekend doang malah nantangin suruh lapor	[testimoni, dariku, org, jatim, yg, yg, jg, jijik, sama, hal, ginian, anakku, pernah, disuruh, minum, obat, tidur, sama, rt, biar, ga, keganggu, sama, sound, horeg, gini, yg, diputer, tetangga, ibuku, udh, ditegur, baik, toh, anaku, ga, tiap, hari, dirumah, ibuku, cuman, pas, weekend, doang, malah, nantangin, suruh, lapor]

- c. **Normalisasi Text:** Tahap normalisasi bertujuan untuk mengatasi variasi penulisan kata dalam bahasa informal atau non-standar, seperti singkatan (slang), akronim, atau penulisan yang tidak baku, menjadi bentuk kata yang baku dan standar dalam Bahasa Indonesia. Proses ini sangat krusial untuk memastikan konsistensi ejaan dan makna kata, sehingga model dapat mengenali kata yang sama meskipun ditulis dengan cara yang berbeda. Normalisasi dilakukan dengan merujuk pada kamus normalisasi yang telah dibuat sebelumnya, yang biasanya disimpan dalam format seperti kamus/normalisasi.xlsx, di mana setiap entri berisi pasangan kata tidak baku dan padanannya yang baku.

**Tabel 3.** Normalisasi Text

Sebelum	Sesudah
[testimoni, dariku, org, jatim, yg, yg, jg, jijik, sama, hal, ginian, anakku, pernah, disuruh, minum, obat, tidur, sama, rt, biar, ga, keganggu, sama, sound, horeg, gini, yg, diputer, tetangga, ibuku, udh, ditegur, baik, toh, anaku, ga, tiap, hari, dirumah, ibuku, cuman, pas, weekend, doang, malah, nantangin, suruh, lapor]	[testimoni, dariku, orang, jatim, yang, yang, jg, jijik, sama, hal, ginian, anakku, pernah, disuruh, minum, obat, tidur, sama, rt, biar, tidak, keganggu, sama, sound, horeg, gini, yang, diputer, tetangga, ibuku, sudah, ditegur, baik, toh, anaku, tidak, tiap, hari, dirumah, ibuku, cuman, pas, weekend, doang, malah, nantangin, suruh, lapor]

- d. **Sentiment Labeling (Lexicon-based):** Sebelum memasuki tahap stemming, proses pelabelan sentimen awal dilakukan menggunakan metode lexicon-based. Pada tahap ini, setiap kata dalam komentar yang telah dinormalisasi diberi skor sentimen berdasarkan dua kamus yang telah didefinisikan, yaitu lexicon\_positif.csv dan lexicon\_negatif.csv. Kamus ini berisi daftar kata-kata dengan nilai sentimen yang telah ditentukan (misalnya positif, negatif, atau netral) dan diperoleh dari repositori InSet Lexicon (Indonesia Sentiment Lexicon) yang bersifat open source di platform GitHub[25]. Skor total dari sebuah komentar dihitung dengan menjumlahkan skor sentimen dari setiap kata yang terkandung di dalamnya. Berdasarkan skor total tersebut, sentimen komentar ditentukan sebagai berikut: jika skor total lebih besar atau sama dengan 0 ( $\geq 0$ ), komentar diklasifikasikan sebagai positif, sedangkan jika skor total kurang dari 0 ( $< 0$ ), komentar diklasifikasikan sebagai negatif.

**Tabel 4.** Sentiment Labelling

Text	Sentimen Skor	Sentimen
[testimoni, dariku, orang, jatim, yang, yang, jg, jijik, sama, hal, ginian, anakku, pernah, disuruh, minum, obat, tidur, sama, rt, biar, tidak, keganggu, sama, sound, horeg, gini, yang, diputer, tetangga, ibuku, sudah, ditegur, baik, toh, anaku, tidak, tiap, hari, dirumah, ibuku, cuman, pas, weekend, doang, malah, nantangin, suruh, lapor]	-20	negatif
berikut, list, di, bawah, ini, penyedia, sound, kru, sound, yang, ikut, ngarak, yang, ngerusak, fasum, dan, semua, yang, ikut, serta, dalam, budaya, sound, horeg, goblok, ini, gendang, telinganya, pecah, pendarahan, otak, budeg, mimisan, susah, berak, mampus, lu, pada, bangsaaaat	-39	negatif

- e. **Stopword:** Tahap ini melibatkan penghapusan kata-kata umum yang sering muncul dalam teks namun tidak memberikan kontribusi makna yang signifikan terhadap analisis sentimen. Kata-kata ini, yang dikenal sebagai stopwords, seringkali merupakan kata penghubung, preposisi, atau artikel (misalnya, "dan", "di", "yang", "adalah"). Proses ini menggunakan daftar stopwords standar yang tersedia dari library NLTK untuk Bahasa Indonesia, yang kemudian diperkaya dengan penambahan beberapa kata stopwords manual (ex\_stopword) yang dianggap relevan dan spesifik untuk konteks dataset "Sound Horeg". Penghapusan stopwords membantu mengurangi dimensi data dan meningkatkan efisiensi komputasi tanpa mengorbankan informasi sentimen yang penting.

**Tabel 5.** Stopword

Sebelum	Sesudah
[apa, hubungan, nya, kemajuan, negara, sama, sound, horeg, pekok]	[hubungan, kemajuan, negara, sound, horeg, pekok]
[lebih, rendah, dari, sdm, sound, horeg]	[rendah, sdm, sound, horeg]

- f. **Stemming:** Stemming adalah proses mengubah setiap kata ke bentuk dasarnya (kata dasar) dengan menghilangkan segala bentuk imbuhan (awalan, akhiran, atau sisipan). Tujuan dari stemming adalah untuk mengurangi variasi kata yang memiliki makna dasar yang sama menjadi satu bentuk tunggal, sehingga mengurangi jumlah fitur yang tidak perlu dan meningkatkan konsistensi dalam representasi teks. Proses ini menggunakan library Sastrawi Stemmer yang dirancang khusus untuk Bahasa Indonesia, memastikan bahwa aturan morfologi bahasa Indonesia diterapkan dengan benar.

**Tabel 6.** Stemming

Sebelum	Sesudah
[list, penyedia, sound, kru, sound, ngarak, ngerusak, fasum, budaya, sound, horeg, goblok, gendang, telinganya, pecah, pendarahan, otak, budeg, mimisan, susah, berak, mampus, lu, bangsaaaaat]	list sedia sound kru sound ngarak ngerusak fasum budaya sound horeg goblok gendang telinga pecah darah otak budeg mimisan susah berak mampus lu bangsaaaaat

### 3.3 Ekstraksi Fitur TF-IDF

Berikut adalah representasi nilai TF, IDF, dan TF-IDF untuk sampel komentar: "orang orang kaya rumah sakit jiwa deh jam gin nyalain sound horeg samar suara".

**Tabel 7.** Ekstraksi Fitur

Term	TF	IDF	TF-IDF (TF × IDF)
deh	0.077	4.224	0.325
gin	0.077	5.851	0.450
horeg	0.077	1.011	0.078
jam	0.077	4.724	0.363
jiwa	0.077	7.041	0.542
kaya	0.077	6.153	0.473
nyalain	0.077	6.096	0.469
orang	0.154	3.825	0.588
rumah	0.077	3.742	0.288
sakit	0.077	5.691	0.438
sound	0.077	1.010	0.078
suara	0.077	4.242	0.326

Tabel 7 menunjukkan bagaimana setiap kata dalam komentar tersebut diberi bobot berdasarkan frekuensi kemunculannya dalam dokumen (TF) dan kekhasannya di seluruh koleksi dokumen (IDF), menghasilkan nilai TF-IDF yang merepresentasikan tingkat kepentingannya.

### 3.4 Pembagian Data dan Resampling Data

Pada tahap ini, data hasil ekstraksi fitur dibagi menjadi dua bagian, yaitu data latih sebesar 80% dan data uji sebesar 20%. Pembagian ini bertujuan agar model dapat mempelajari pola dari sebagian besar data yang tersedia melalui data latih, sementara data uji digunakan untuk mengukur kemampuan model dalam melakukan generalisasi terhadap data baru yang belum pernah dilihat sebelumnya. Pembagian dilakukan secara acak untuk memastikan bahwa distribusi data pada masing-masing subset tetap representatif terhadap keseluruhan dataset.

Namun, dari hasil distribusi data diketahui bahwa terdapat ketidakseimbangan antara kelas positif dan negatif, di mana jumlah komentar negatif jauh lebih banyak dibandingkan komentar positif. Kondisi ini dapat menyebabkan model cenderung bias terhadap kelas mayoritas saat proses pelatihan.

Untuk mengatasi hal tersebut, teknik SMOTE (Synthetic Minority Over-sampling Technique) diterapkan pada tahap pemodelan, bukan sebelum pembagian data. Hal ini dilakukan agar proses oversampling hanya mempengaruhi data latih dan tidak menimbulkan kebocoran data (data leakage) pada data uji. Dengan demikian, hasil evaluasi model tetap objektif dan mencerminkan performa sebenarnya terhadap data baru.

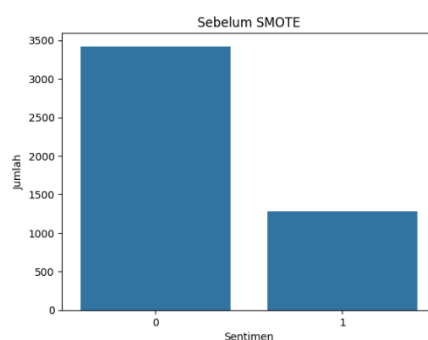
Melalui strategi ini, pembagian data yang proporsional serta penerapan SMOTE di tahap pemodelan diharapkan dapat meningkatkan keadilan model dalam mempelajari pola dari kedua kelas dan menghasilkan performa klasifikasi yang lebih seimbang.

### 3.5 Modelling

Pada kondisi awal, di mana kelas negatif mendominasi dataset, model Naive Bayes cenderung bias dan kesulitan mengidentifikasi sampel dari kelas positif. Hal ini tercermin dari nilai recall yang sangat rendah untuk kelas positif pada model tanpa optimasi. Ketidakmampuan model untuk belajar secara efektif dari kelas minoritas dapat menyebabkan generalisasi yang buruk pada data dunia nyata yang seringkali memiliki distribusi kelas yang tidak merata.

Penerapan SMOTE secara efektif mengatasi masalah ketidakseimbangan data dengan menghasilkan sampel sintetis untuk kelas minoritas. Dengan demikian, model memiliki lebih banyak contoh dari kelas positif untuk dipelajari, mengurangi bias terhadap kelas mayoritas. Hasilnya, kemampuan model untuk mengidentifikasi sentimen positif meningkat drastis, yang terlihat dari peningkatan recall dan F1-Score yang signifikan untuk kelas positif. Selain itu, proses hyperparameter tuning melalui GridSearchCV memungkinkan penemuan nilai alpha optimal untuk MultinomialNB, yang lebih lanjut menyempurnakan kemampuan model dalam membedakan antara kelas-kelas sentimen. Kombinasi SMOTE dan tuning parameter ini memastikan kinerja yang seimbang dan robust di kedua kelas sentimen.

Selanjutnya dilakukan dengan menerapkan SMOTE dan tuning parameter, kinerja model Naive Bayes meningkat secara drastis, mengatasi kelemahan yang ditemukan pada model awal. Proses GridSearchCV menemukan parameter terbaik, yaitu  $\alpha = 0.001$ , yang menghasilkan rata-rata skor cross-validation terbaik sebesar 81.24%

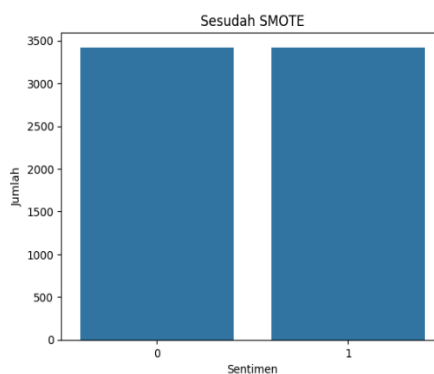


**Gambar 3.** Histogram Sebelum SMOTE

Histogram gambar 2 menggambarkan distribusi sentimen pada dataset sebelum dilakukan proses oversampling menggunakan SMOTE. Dari visualisasi ini, terlihat jelas adanya ketidakseimbangan kelas (class imbalance) yang signifikan antara kedua kategori sentimen. Pada sumbu horizontal (sentimen), nilai '0' merepresentasikan sentimen negatif, sementara nilai '1' merepresentasikan sentimen positif. Sumbu vertikal (Count) menunjukkan jumlah atau frekuensi kemunculan masing-masing sentimen.

Secara spesifik, batang pada nilai '0' (sentimen negatif) memiliki tinggi yang jauh lebih dominan, menunjukkan bahwa jumlah komentar dengan sentimen negatif sangat besar, mencapai lebih dari 3422 sampel. Sebaliknya, batang pada nilai '1' (sentimen positif) jauh lebih rendah, menunjukkan bahwa jumlah komentar dengan sentimen positif berada di bawah 1283 sampel.

Kondisi ketidakseimbangan ini mengindikasikan bahwa kelas sentimen negatif adalah kelas mayoritas, sedangkan kelas sentimen positif adalah kelas minoritas. Apabila model klasifikasi dilatih dengan distribusi data seperti ini, ada risiko tinggi bahwa model akan cenderung bias terhadap kelas mayoritas (negatif) dan kurang mampu mengenali atau mengklasifikasikan sampel dari kelas minoritas (positif) dengan baik. Inilah alasan mengapa teknik seperti SMOTE diperlukan untuk menyeimbangkan distribusi kelas sebelum proses pemodelan.



**Gambar 4.** Histogram Setelah SMOTE

Histogram gambar 3 menunjukkan distribusi sentimen pada dataset setelah dilakukan proses oversampling menggunakan SMOTE (Synthetic Minority Oversampling Technique). Berbeda dengan kondisi sebelumnya yang

menunjukkan ketidakseimbangan kelas yang signifikan, visualisasi ini memperlihatkan bahwa data sentimen kini telah seimbang secara proporsional.

Pada sumbu horizontal (sentimen), nilai '0' tetap merepresentasikan sentimen negatif, dan nilai '1' merepresentasikan sentimen positif. Sementara itu, sumbu vertikal (Count) menunjukkan jumlah atau frekuensi kemunculan masing-masing sentimen. Secara spesifik, kedua batang pada nilai '0' (sentimen negatif) dan '1' (sentimen positif) kini memiliki tinggi yang hampir setara, masing-masing menunjukkan jumlah sampel 3422 untuk positif dan 3422 untuk negatif.

**Tabel 8.** Perbandingan Sebelum dan Setelah SMOTE

	Negatif	Positif
Sebelum SMOTE	3422	1283
Setelah SMOTE	3422	3422

### 3.6 Evaluasi

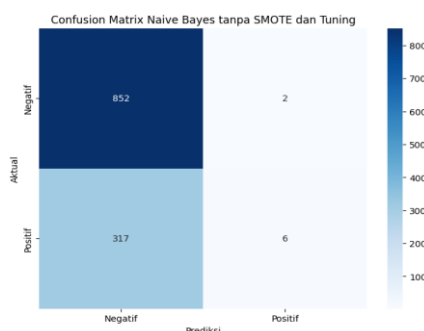
Berdasarkan pemodelan setelah tuning, model mencapai akurasi 71%, yang menunjukkan bahwa model cukup baik dalam memprediksi keseluruhan data. Untuk presisi, model berhasil mencapai 81% untuk kelas negatif (0) tetapi hanya 47% untuk kelas positif (1), yang menunjukkan bahwa model lebih akurat dalam memprediksi kelas mayoritas. Sementara itu, recall mencapai 78% untuk kelas negatif (0) dan 53% untuk kelas positif (1), mengindikasikan bahwa model lebih efektif dalam mengenali kelas mayoritas, meskipun masih kurang dalam mendeteksi kelas minoritas. Di sisi lain, F1-Score model menunjukkan 80% untuk kelas negatif (0) dan 50% untuk kelas positif (1), yang menggambarkan keseimbangan antara presisi dan recall, meskipun performa untuk kelas positif masih perlu ditingkatkan. Secara keseluruhan, macro average untuk presisi, recall, dan F1-Score masing-masing mencapai 64%, sementara weighted average untuk ketiga metrik tersebut adalah 72% (presisi) dan 71% (recall dan F1-Score), mencerminkan kinerja model secara keseluruhan.

**Tabel 9.** Perbandingan Hasil Metrik

	Sebelum SMOTE dan Tuning	Sesudah SMOTE dan Tuning
Metric	Hasil	Hasil
Accuracy	0.73	0.71
Precision	0.74	0.64
Recall	0.51	0.65
F1-Score	0.44	0.65
Support	1177	1177

Model Naive Bayes pada tahap awal (tanpa penerapan SMOTE dan *tuning* parameter) menghasilkan akurasi sebesar 0,73 sebagaimana ditunjukkan pada tabel 9. Namun, model tersebut menunjukkan kelemahan yang cukup signifikan dalam mendeteksi kelas minoritas, dengan nilai *recall* yang sangat rendah. Setelah dilakukan penerapan SMOTE dan *tuning* parameter menggunakan teknik *cross-validation*, performa model terhadap kelas minoritas meningkat secara substansial, yang ditunjukkan dengan kenaikan nilai *recall* dari 0,02 menjadi 0,53. Meskipun demikian, akurasi keseluruhan model mengalami sedikit penurunan menjadi 0,71, atau sekitar 2 persen lebih rendah dibandingkan model awal.

Penurunan akurasi ini tidak dapat diartikan sebagai kemunduran performa model secara keseluruhan. Sebaliknya, hal tersebut merupakan konsekuensi yang wajar dari adanya *trade-off* antara akurasi global dan kemampuan model dalam mengenali kelas minoritas. Peningkatan *recall* pada kelas minoritas menunjukkan bahwa model menjadi lebih sensitif terhadap data dari kategori yang sebelumnya sulit terdeteksi. Dalam konteks dataset yang memiliki ketidakseimbangan kelas, peningkatan sensitivitas terhadap kelas positif jauh lebih penting dibandingkan mempertahankan akurasi tinggi yang didominasi oleh prediksi kelas mayoritas. Oleh karena itu, penurunan akurasi sebesar 2 persen dianggap sebagai biaya yang dapat diterima untuk memperoleh keseimbangan yang lebih baik antara kedua kelas.



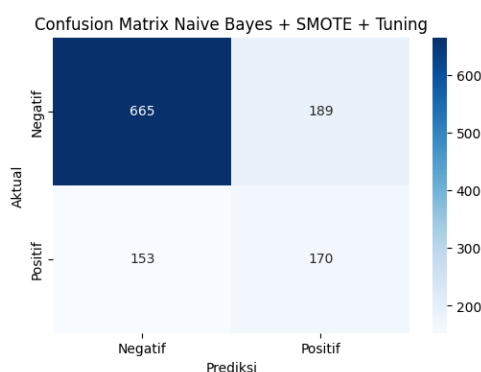
**Gambar 5.** Confusion Matrix Sebelum SMOTE dan Tuning

Confusion Matrix pada Gambar 4 menunjukkan kinerja model klasifikasi Naive Bayes tanpa penerapan SMOTE dan tuning hyperparameter. Dari hasil tersebut, terlihat bahwa model memiliki ketidakseimbangan performa antara kelas negatif dan positif. Model berhasil memprediksi 852 komentar sebagai negatif yang memang benar negatif (True Negative), menandakan kemampuan model yang sangat baik dalam mengenali kelas mayoritas.

Namun, terdapat 317 komentar positif yang diprediksi sebagai negatif (False Negative), menunjukkan bahwa model kesulitan mengenali kelas minoritas. Selain itu, hanya 6 komentar yang benar diprediksi sebagai positif (True Positive), sementara 2 komentar negatif salah diprediksi sebagai positif (False Positive).

Secara keseluruhan, hasil ini memperlihatkan bahwa tanpa teknik oversampling seperti SMOTE dan tanpa tuning parameter, model Naive Bayes cenderung bias terhadap kelas mayoritas (negatif). Meskipun akurasi keseluruhan mungkin terlihat tinggi, hal ini tidak mencerminkan kinerja yang baik dalam mengklasifikasikan kelas minoritas (positif). Oleh karena itu, penerapan SMOTE dan tuning diperlukan untuk meningkatkan kemampuan model dalam mengenali data dengan distribusi yang tidak seimbang.

Confusion Matrix pada Gambar 5 menggambarkan kinerja model klasifikasi Naive Bayes setelah dioptimasi dengan teknik oversampling SMOTE dan hyperparameter tuning, menunjukkan peningkatan performa model yang signifikan dibandingkan kondisi sebelumnya. Model berhasil memprediksi 665 komentar sebagai negatif yang secara aktual memang negatif (True Negatif), menegaskan kekuatan model dalam mengidentifikasi kelas mayoritas. Lebih lanjut, model hanya salah memprediksi 153 komentar sebagai negatif padahal secara aktual positif (False Negatif), serta 189 komentar salah diprediksi sebagai positif padahal negatif (False Positif). Meskipun ada beberapa kesalahan prediksi pada kelas positif, model menunjukkan kemampuan yang cukup baik dalam mengenali sentimen negatif dengan akurat.



**Gambar 6.** Confusion Matrix Sesudah SMOTE dan Tuning

Peningkatan yang paling signifikan terlihat pada kemampuan model dalam mengklasifikasikan kelas minoritas. Berdasarkan matriks, ada sebanyak 153 komentar yang salah diprediksi sebagai negatif padahal secara aktual positif (False Negative = 153). Ini menunjukkan bahwa model belum berhasil mengenali seluruh data positif secara sempurna. Sebaliknya, ada 189 komentar yang salah diprediksi sebagai positif padahal secara aktual negatif (False Positive = 189). Model juga berhasil memprediksi dengan benar 170 komentar positif (True Positive) dan 665 komentar negatif (True Negative). Angka-angka ini menunjukkan kinerja model yang cukup baik, namun dengan beberapa kesalahan klasifikasi.

Secara keseluruhan, Confusion Matrix ini mengonfirmasi bahwa penerapan SMOTE dan tuning parameter telah membantu meningkatkan performa model Naive Bayes dalam mengklasifikasikan kedua kelas sentimen, meskipun masih terdapat ruang untuk perbaikan, terutama dalam mengurangi False Negative.

Performa model sebelum dan sesudah optimasi menunjukkan adanya pergeseran prioritas evaluasi. Meskipun akurasi rata-rata dari *cross-validation* sebelum penerapan SMOTE dan *tuning* mencapai 73.19%, model tersebut cenderung bias terhadap kelas mayoritas, yang kemungkinan besar mengabaikan kelas minoritas. Setelah dilakukan penyeimbangan data menggunakan SMOTE dan optimasi parameter (*tuning*) melalui *GridSearchCV*, performa model secara keseluruhan meningkat, terbukti dari skor *cross-validation* terbaik yang mencapai 81.23%. Peningkatan signifikan ini menunjukkan bahwa model sekarang mampu mengklasifikasikan kedua kelas dengan lebih baik, bukan hanya berfokus pada kelas mayoritas, sehingga menghasilkan model yang lebih efektif.

Secara keseluruhan, hasil ini memberikan wawasan mengenai bagaimana teknik SMOTE dan tuning dapat meningkatkan kinerja model dalam menghadapi masalah ketidakseimbangan data, dengan Naive Bayes memberikan hasil terbaik di antara ketiga model yang diuji.

## 4. KESIMPULAN

Penelitian ini mengeksplorasi dan mengevaluasi model klasifikasi sentimen pada komentar pengguna terkait fenomena “Sound Horeg” menggunakan algoritma Naive Bayes. Pada tahap awal, model yang dilatih dengan dataset tidak seimbang menunjukkan metrik akurasi sebesar 0,73 dan precision sebesar 0,74. Meskipun demikian, dominasi

sentimen negatif sebagai kelas mayoritas menyebabkan model cenderung bias dan kurang mampu mengidentifikasi sentimen positif, yang tercermin dari nilai true positive yang rendah dan false negative yang tinggi pada confusion matrix awal. Setelah penerapan SMOTE dan parameter tuning, akurasi model menurun sedikit menjadi 0,71 dan precision turun menjadi 0,64. Namun, secara substantif, model menjadi lebih seimbang dan adil dalam mendeteksi kedua kelas sentimen. Peningkatan ini terlihat dari kenaikan signifikan pada recall model dari 0,51 menjadi 0,65, serta peningkatan F1-score dari 0,44 menjadi 0,65. Secara rinci, performa kelas negatif (0) mencapai precision 81%, recall 78%, dan F1-score 80%, sementara kelas positif (1) memperoleh precision 47%, recall 53%, dan F1-score 50%. Penurunan akurasi sebesar dua persen perlu dipahami sebagai konsekuensi dari adanya trade-off antara akurasi keseluruhan dan kemampuan model dalam mengenali kelas minoritas. Dalam konteks penelitian ini, recall terhadap sentimen positif memiliki bobot yang lebih penting karena tujuan analisis adalah memahami secara lebih adil persepsi publik yang beragam terhadap fenomena “Sound Horeg.” Dengan demikian, peningkatan sensitivitas model terhadap komentar positif menunjukkan perbaikan yang bermakna secara sosial, bukan hanya secara teknis. Hasil klasifikasi sentimen ini memberikan gambaran kuantitatif mengenai persepsi masyarakat terhadap fenomena “Sound Horeg.” Mayoritas komentar yang terklasifikasi sebagai negatif mengindikasikan adanya kekhawatiran, ketidaknyamanan, atau penolakan terhadap penggunaan sistem suara bertenaga tinggi di ruang publik. Sebaliknya, meskipun proporsinya lebih kecil, komentar positif menunjukkan adanya kelompok masyarakat yang memandang “Sound Horeg” sebagai bentuk ekspresi budaya, hiburan, atau kebanggaan komunitas lokal. Dengan demikian, model yang dioptimalkan melalui SMOTE mampu menampilkan representasi yang lebih seimbang atas persepsi publik, yang sebelumnya terdistorsi oleh ketidakseimbangan data. Secara keseluruhan, penelitian ini tidak hanya berhasil meningkatkan performa model Naive Bayes dalam klasifikasi sentimen berbahasa Indonesia, tetapi juga memberikan kontribusi praktis terhadap pemahaman fenomena sosial di Indonesia, khususnya terkait dinamika penerimaan publik terhadap “Sound Horeg.” Hasil penelitian ini dapat menjadi acuan bagi penyelenggara acara, pemerintah daerah, dan pembuat kebijakan dalam merumuskan pedoman teknis yang menyeimbangkan kebutuhan hiburan publik dengan kenyamanan lingkungan. Untuk penelitian selanjutnya, disarankan eksplorasi algoritma yang lebih kompleks seperti Support Vector Machine (SVM), Random Forest, atau model berbasis Transformer (misalnya BERT) serta pengayaan fitur linguistik dan semantik untuk meningkatkan akurasi klasifikasi sentimen. Dengan demikian, penelitian lanjutan dapat memberikan pemahaman yang lebih mendalam terhadap opini publik berbahasa Indonesia, tidak hanya dalam konteks “Sound Horeg,” tetapi juga terhadap fenomena sosial lainnya yang serupa.

## REFERENCES

- [1] I. H. Saputra, “Analisis Sound Horeg di Jawa Timur: Perspektif Hadis dan Implikasi Medis terhadap Kebisingan dan Etika Sosial,” *J. QURAN HADITH Stud.*, vol. 14, no. 1, Art. no. 1, May 2025, doi: 10.15408/quhas.v14i1.42872.
- [2] F. R. Lail, “Persepsi masyarakat Tentang Fenomena Penggunaan Pengeras Suara dengan Volume Keras Terhadap Kenyamanan Berkehidupan Sosial (Studi Kasus: Desa Mergayu, Kecamatan Bandung, Kabupaten Tulungagung),” *Indones. J. Soc. Stud.*, vol. 6, no. 1, pp. 51–55, Jul. 2023, doi: 10.26740/ijss.v6n1.p51-55.
- [3] M. M. Hossain, M. S. Hossain, M. F. Mridha, M. Safran, and S. Alfahhood, “Multi task opinion enhanced hybrid BERT model for mental health analysis,” *Sci. Rep.*, vol. 15, no. 1, p. 3332, Jan. 2025, doi: 10.1038/s41598-025-86124-6.
- [4] C. Suratnoaji, N. Nurhadi, and I. Arianto, “Measurement of Public Opinion based on Social Media Big Data (Indonesia and Malaysia),” in *Proceedings of the First International Conference on Literature Innovation in Chinese Language, LIONG 2021, 19-20 October 2021, Purwokerto, Indonesia*, Purwokerto, Indonesia: EAI, 2022. doi: 10.4108/eai.19-10-2021.2316598.
- [5] J. P. Venugopal, A. A. V. Subramanian, G. Sundaram, M. Rivera, and P. Wheeler, “A Comprehensive Approach to Bias Mitigation for Sentiment Analysis of Social Media Data,” *Appl. Sci.*, vol. 14, no. 23, Art. no. 23, Jan. 2024, doi: 10.3390/app142311471.
- [6] S. Rahmawati, D. Anggraini, and R. Kurniawan, “Natural Language Processing For Automatic Sentiment Analysis In Social Media Data,” *Int. J. Inf. Eng. Sci.*, vol. 1, no. 1, pp. 16–19, Feb. 2024, doi: 10.62951/ijies.v1i1.54.
- [7] Natasha and R. R. Suryono, “Sentiment Analysis of the Influence of the Korean Wave in Indonesia using the Naive Bayes Method and Support Vector Machine,” *INOVTEK Polbeng - Seri Inform.*, vol. 10, no. 1, Art. no. 1, Mar. 2025, doi: 10.35314/85x4wd90.
- [8] A. Basir, “Analysis of Electronic Wallet User Sentiment on Twitter (x) Social Media Using the Naïve Bayes Classifier Algorithm,” *J. Inform. J. Pengemb. IT*, vol. 10, no. 1, Art. no. 1, Jan. 2025, doi: 10.30591/jpit.v10i1.8180.
- [9] H. Chen, S. Hu, R. Hua, and X. Zhao, “Improved naive Bayes classification algorithm for traffic risk management,” *EURASIP J. Adv. Signal Process.*, vol. 2021, no. 1, p. 30, Jun. 2021, doi: 10.1186/s13634-021-00742-6.
- [10] I. K. Dharmendra, I. M. A. Putra, and Y. P. Atmojo, “Evaluation of the Effectiveness of SMOTE and Random Under Sampling in Emotion Classification of Tweets,” *Inform. Educ. Prof. J. Inform.*, vol. 9, no. 2, Dec. 2024, doi: 10.51211/itbi.v9i2.3183.
- [11] S. Gan, S. Shao, L. Chen, L. Yu, and L. Jiang, “Adapting Hidden Naive Bayes for Text Classification,” *Mathematics*, vol. 9, no. 19, p. 2378, Jan. 2021, doi: 10.3390/math9192378.
- [12] S. Anggina, N. Yudi Setiawan, and F. A. Bachtiar, “Analysis of Customer Reviews Using Multinomial Naïve Bayes Classifier with Lexicon-Based and TF-IDF at Formaggio Coffee and Resto,” *Best Account. Inf. Syst. Inf. Technol. Bus. Enterp.*, vol. 7, no. 1, Sep. 2022, doi: 10.34010/aisthebest.v7i1.7072.
- [13] A. Sauddin, T. A. Nurman, N. Aeni, and S. R. Sudarta, “Klasifikasi Spam Sms Menggunakan Naïve Bayes Classifier Dan K-Nearest Neighbor,” *J. MSA Mat. Dan Stat. Serta Apl.*, vol. 13, no. 1, pp. 101–109, Jun. 2025, doi: 10.24252/msa.v13i1.46192.
- [14] V. Agresia and R. R. Suryono, “Comparison of SVM, Naïve Bayes, and Logistic Regression Algorithms for Sentiment Analysis of Fraud and Bots in Purchasing Concert Ticket,” *INOVTEK Polbeng - Seri Inform.*, vol. 10, no. 2, Art. no. 2, Jul. 2025, doi: 10.35314/npfydh47.



- [15] F. Fahrani and J. Aryanto, "Sentiment Analysis of Public Opinion on the Palestinian-Israeli Conflict using Support Vector Machine and Naïve Bayes Algorithms," *J. Sci. Res. Educ. Technol. JSRET*, vol. 3, no. 4, Art. no. 4, Dec. 2024, doi: 10.58526/jsret.v3i4.606.
- [16] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Syst. Appl.*, vol. 36, no. 3, Part 1, pp. 5432–5435, Apr. 2009, doi: 10.1016/j.eswa.2008.06.054.
- [17] H. Hairani, K. E. Saputro, and S. Fadli, "K-means-SMOTE untuk menangani ketidakseimbangan kelas dalam klasifikasi penyakit diabetes dengan C4.5, SVM, dan naive Bayes," *J. Teknol. Dan Sist. Komput.*, vol. 8, no. 2, pp. 89–93, Apr. 2020, doi: 10.14710/jtsiskom.8.2.2020.89-93.
- [18] H. P. Jelita, M. I. Sa'ad, and Wahyuni, "Penerapan Algoritma Naïve Bayes Dalam Analisis sentiment Masyarakat Terhadap STMIK Widya Cipta Dharma," *Bull. Inf. Technol. BIT*, vol. 6, no. 2, pp. 148–160, Jun. 2025, doi: 10.47065/bit.v6i2.2029.
- [19] F. Kamalov, S. E. Choutri, and A. F. Atiya, "Analytical formulation of synthetic minority oversampling technique (SMOTE) for imbalanced learning," *Gulf J. Math.*, vol. 19, no. 1, Art. no. 1, Jan. 2025, doi: 10.56947/gjom.v19i1.2639.
- [20] J. Saputra, L. Maryani, Rahmadden, D. Wulandari, and W. Eka, "Analisis Performa Naive Bayes Dan Svm Terhadap Sentimen Teks Media Sosial Dengan Word2vec Dan Smote," *J. INSTEK Inform. Sains Dan Teknol.*, vol. 10, no. 1, pp. 143–155, May 2025, doi: 10.24252/instek.v10i1.54889.
- [21] N. Hayatin, G. I. Marthasari, and L. Nuarini, "Optimization of Sentiment Analysis for Indonesian Presidential Election using Naïve Bayes and Particle Swarm Optimization," *J. Online Inform.*, vol. 5, no. 1, pp. 81–88, Jul. 2020, doi: 10.15575/join.v5i1.558.
- [22] Y. I. Kurniawan *et al.*, "Naive Bayes Classifier with SMOTE for Sentiment Analysis of Blibli App Reviews on The Google Play Store," *J. Penelit. Inov.*, vol. 5, no. 3, pp. 2675–2688, Sep. 2025, doi: 10.54082/jupin.1842.
- [23] J. A. Nurcahyo and T. B. Sasongko, "Hyperparameter Tuning Algoritma Supervised Learning untuk Klasifikasi Keluarga Penerima Bantuan Pangan Beras," *Indones. J. Comput. Sci.*, vol. 12, no. 3, 2023, doi: 10.33022/ijcs.v12i3.3254.
- [24] D. A. Novira and H. H. Puspytasari, "Tinjauan Yuridis Pertanggungjawaban Terhadap Ambang Batas Kebisingan Sound Horeg yang Menimbulkan Kerugian," *Indones. J. Contemp. Law*, vol. 1, no. 3, pp. 1–15, Jul. 2025.
- [25] F. Koto and G. Y. Rahmaningtyas, "InSet Lexicon: Indonesia Sentiment Lexicon." 2017. [Online]. Available: <https://github.com/fajri91/InSet>