

Klasifikasi Kepuasan Pelanggan Aplikasi Tvri Klik Berdasarkan Analisis Sentiment Ulasan Menggunakan Algoritma SVM dan Naive Bayes

Tabri Eko Miyardi^{1*}, Yohanes Setiawan²

^{1,2}Teknik Informatika, ISB Atma Luhur Pangkalpinang, Kepulauan Bangka Belitung, Indonesia
2311520013@mahasiswa.atmaluhur.ac.id, ysetiawanj@atmaluhur.ac.id

(* : corresponding author)

Abstrak

Ulasan pengguna pada Google Play Store merupakan data tidak terstruktur yang bernilai untuk mengevaluasi kualitas layanan aplikasi streaming penyiaran publik. Penelitian ini membandingkan kinerja Support Vector Machine (SVM) dan Multinomial Naive Bayes dalam mengklasifikasikan sentimen dari 2.006 ulasan pengguna aplikasi TVRI Klik yang dikumpulkan melalui web scraping. Label sentimen ditentukan berdasarkan rating bintang, yaitu empat dan lima sebagai positif, tiga sebagai netral, serta satu dan dua sebagai negatif. Pelabelan menghasilkan distribusi kelas yang tidak seimbang dengan kelas negatif mendominasi data. Pra-pemrosesan meliputi case folding, tokenisasi, penghapusan stopword, dan stemming menggunakan Sastrawi, dilanjutkan dengan ekstraksi fitur TF-IDF. SMOTE diterapkan pada data latih untuk menyeimbangkan ketiga kelas sentimen. Dengan skema pembagian data 80:20, SVM memperoleh akurasi data uji tertinggi sebesar 95,13%, sedangkan Naive Bayes mencapai 86,59%. Wordcloud menunjukkan ulasan negatif didominasi keluhan teknis, sedangkan ulasan positif menonjolkan kemudahan akses siaran nasional. Temuan menunjukkan bahwa SVM dengan SMOTE lebih stabil dan akurat untuk memantau sentimen publik terhadap aplikasi streaming lembaga penyiaran pemerintah.

Kata kunci: Analisis Sentimen, Naive Bayes, Support Vector Machine, SMOTE, TVRI Klik

Abstract

User reviews on the Google Play Store are unstructured data that is valuable for evaluating the quality of public broadcasting streaming application services. This study compares the performance of Support Vector Machine (SVM) and Multinomial Naive Bayes in classifying sentiment from 2,006 user reviews of the TVRI Klik application collected through web scraping. The sentiment labels are determined based on star ratings, namely four and five as positive, three as neutral, and one and two as negative. Labeling results in an unbalanced distribution of classes with negative classes dominating the data. Pre-processing includes case folding, tokenization, stopword removal, and stemming using Sastrawi, followed by TF-IDF feature extraction. SMOTE is applied to the training data to balance all three sentiment classes. With an 80:20 data sharing scheme, SVM obtained the highest test data accuracy of 95.13%, while Naive Bayes reached 86.59%. Wordcloud shows that negative reviews are dominated by technical complaints, while positive reviews highlight the ease of access to national broadcasts. The findings show that SVM with SMOTE is more stable and accurate for monitoring public sentiment towards streaming applications of government broadcasting institutions.

Keywords: Naive Bayes, Sentiment Analysis, Support Vector Machine, SMOTE, TVRI Klik

1. PENDAHULUAN

Layanan penyiaran publik di Indonesia terus bertransformasi mengikuti kebutuhan masyarakat yang semakin mengakses konten televisi melalui perangkat bergerak. TVRI, sebagai lembaga penyiaran publik nasional, meluncurkan aplikasi TVRI Klik agar pengguna dapat menonton siaran nasional maupun siaran olahraga, termasuk pertandingan Piala Dunia 2026, secara gratis melalui smartphone. Peningkatan jumlah pengguna aplikasi streaming semacam ini secara langsung berdampak pada bertambahnya volume ulasan yang ditinggalkan pengguna di Google Play Store, baik berupa apresiasi maupun keluhan atas kendala teknis yang dialami. Ulasan tersebut merupakan data tidak terstruktur yang sesungguhnya sangat bernilai karena mencerminkan persepsi pengguna secara langsung dan real time terhadap kualitas layanan digital yang mereka gunakan sehari-hari, sehingga dibutuhkan pendekatan komputasional yang mampu mengekstraksi pola sentimen secara otomatis dan konsisten dibanding pembacaan manual satu per satu.

Analisis sentimen menjadi salah satu cabang text mining yang bertujuan mengenali polaritas opini pengguna ke dalam kategori positif, netral, atau negatif, dan telah banyak digunakan untuk mengevaluasi kualitas aplikasi berbasis ulasan Google Play Store pada berbagai domain, mulai dari aplikasi pendidikan hingga aplikasi pemerintahan digital [1], layanan kesehatan nasional [2], hingga platform hiburan streaming [3]. Algoritma machine learning klasik seperti Naive Bayes dan Support Vector Machine (SVM) tetap menjadi pilihan yang kompetitif untuk tugas klasifikasi sentimen berskala kecil hingga menengah karena kompleksitas komputasinya relatif rendah dan hasilnya lebih mudah diinterpretasikan dibanding model deep learning yang membutuhkan data latih jauh lebih besar [4], [5]. Keunggulan pendekatan klasifikasi konvensional semacam ini turut ditemukan pada domain non-tekstual, misalnya klasifikasi tingkat kematangan buah berbasis citra digital yang menunjukkan performa kompetitif algoritma KNN dibandingkan SVM pada data berskala kecil [6], sehingga menegaskan relevansi algoritma machine learning klasik pada berbagai jenis data, baik tekstual maupun citra. Studi terhadap aplikasi Google Classroom, misalnya, menunjukkan bahwa SVM mampu memberikan akurasi klasifikasi yang tinggi pada ulasan berbahasa Indonesia setelah dikombinasikan dengan seleksi fitur yang tepat [7], sementara perbandingan Naive Bayes dan SVM pada ulasan aplikasi musik streaming turut menegaskan bahwa kedua algoritma tersebut masih relevan untuk kasus serupa [8]. Perbandingan serupa turut dilaporkan pada klasifikasi ulasan aplikasi Netflix [3] serta analisis sentimen opini publik terhadap kebijakan pemerintah di platform X [9], yang keduanya secara konsisten menunjukkan keunggulan akurasi SVM dibandingkan Naive Bayes pada data ulasan maupun opini publik berbahasa Indonesia.

Salah satu tantangan utama dalam klasifikasi sentimen ulasan aplikasi adalah ketidakseimbangan distribusi kelas, di mana satu kelas sentimen mendominasi data sehingga model cenderung bias memprediksi kelas mayoritas dan mengabaikan kelas minoritas [10]. Kondisi ini semakin nyata apabila label sentimen diturunkan langsung dari rating bintang yang diberikan pengguna, sebab pengguna aplikasi cenderung lebih sering memberikan rating ekstrem berupa bintang satu atau bintang lima dibanding rating tengah [11]. Synthetic Minority Oversampling Technique (SMOTE) banyak diadopsi untuk mengatasi persoalan ini dengan membangkitkan data sintesis pada kelas minoritas melalui interpolasi terhadap tetangga terdekatnya di ruang fitur, alih-alih sekadar menduplikasi data yang sudah ada [10]. Penelitian pada ulasan aplikasi Google Play Store lain melaporkan bahwa penerapan SMOTE mampu meningkatkan akurasi SVM secara signifikan meski tidak selalu memberikan dampak yang sama pada Naive Bayes, sehingga interaksi antara teknik penyeimbangan kelas dan pemilihan algoritma klasifikasi masih perlu diuji secara spesifik pada tiap studi kasus [12]. Penerapan SMOTE pada ulasan pengguna terkait perangkat True Wireless Stereo (TWS) turut menunjukkan peningkatan akurasi SVM yang lebih besar dibandingkan Naive Bayes setelah penyeimbangan kelas dilakukan [13]. Studi lain pada data yang tidak seimbang bahkan menegaskan bahwa penerapan oversampling yang tidak hati-hati, misalnya dilakukan sebelum proses validasi silang, berpotensi menimbulkan bias evaluasi yang menyesatkan sehingga prosedur pengujian perlu dirancang secara cermat [14].

Tantangan lain pada ulasan berbahasa Indonesia adalah gaya bahasa yang informal, penuh singkatan, dan variasi ejaan, yang dapat menurunkan kualitas representasi fitur apabila tahapan pra-pemrosesan tidak dilakukan dengan tepat. Penggunaan pustaka Sastrawi untuk stemming Bahasa Indonesia terbukti efektif memperbaiki hasil pra-pemrosesan pada berbagai tugas klasifikasi dokumen dan telah menjadi praktik umum pada penelitian analisis sentimen berbahasa Indonesia [15], [16], sedangkan pemilihan algoritma stemming yang tepat juga terbukti berpengaruh terhadap kualitas pengolahan teks pada berbagai aplikasi pemrosesan bahasa alami [17]. Selain itu, representasi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) tetap menjadi pendekatan yang banyak dipakai karena mampu menekan bobot kata yang terlalu umum sekaligus memperkuat kontribusi kata yang lebih spesifik dan informatif dalam membedakan sentimen [4].

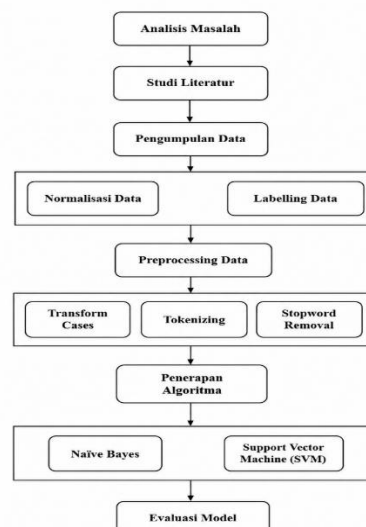
Studi kasus pada aplikasi layanan publik berbasis pemerintah menunjukkan bahwa ulasan pengguna dapat mencerminkan permasalahan nyata yang dihadapi masyarakat, sebagaimana

terlihat pada evaluasi sentimen aplikasi Sirekap milik Komisi Pemilihan Umum yang membandingkan SVM, Naive Bayes, dan Random Forest untuk menangkap opini publik pasca Pemilihan Presiden 2024 [5]. Pola serupa turut ditemukan pada aplikasi streaming komersial, di mana perbandingan SVM dan Random Forest pada ulasan aplikasi Vidio menunjukkan bahwa keluhan teknis terkait buffering dan kegagalan pemutaran menjadi sumber sentimen negatif yang dominan. Namun demikian, penelitian yang secara spesifik mengevaluasi sentimen pengguna terhadap aplikasi streaming milik lembaga penyiaran publik seperti TVRI Klik masih sangat terbatas, padahal aplikasi ini menjadi kanal penting bagi masyarakat untuk mengakses siaran nasional secara gratis, termasuk pada momen strategis seperti penyiaran Piala Dunia 2026. Berdasarkan isi ulasan yang terkumpul pada penelitian ini, pengguna banyak mengeluhkan permasalahan teknis seperti layar kosong, kegagalan memuat siaran, serta tampilan yang tidak berfungsi optimal, sehingga persoalan ini merupakan problem nyata yang perlu dipahami secara sistematis oleh pengembang aplikasi guna meningkatkan kualitas layanan digital yang bersifat publik.

Penelitian ini berfokus pada perbandingan dua algoritma klasifikasi, yaitu SVM dan Multinomial Naive Bayes, untuk menganalisis sentimen ulasan pengguna aplikasi TVRI Klik dengan mempertimbangkan dua aspek yang sering luput pada penelitian sejenis. Pertama, penelitian ini menerapkan SMOTE secara eksplisit untuk menangani ketidakseimbangan kelas yang muncul akibat pelabelan berbasis rating bintang, dan membandingkan distribusi data sebelum maupun sesudah proses penyeimbangan tersebut. Kedua, penelitian ini melakukan pengujian akurasi pada skema pembagian data, 80% dan 20% untuk melihat keunggulan salah satu algoritma. Kontribusi penelitian ini terletak pada penyajian evaluasi menyeluruh terhadap sensitivitas SVM dan Naive Bayes terhadap ukuran data latih pada kasus ulasan aplikasi streaming penyiaran publik berbahasa Indonesia, sekaligus memberikan gambaran praktis mengenai jenis keluhan yang paling banyak diutarakan pengguna melalui visualisasi wordcloud pada tiap kelas sentimen. Hasil penelitian ini diharapkan dapat menjadi rujukan metodologis sekaligus bahan evaluasi teknis bagi pengembang aplikasi TVRI Klik dalam meningkatkan kualitas layanan streaming ke depan.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain eksperimen komputasional berbasis supervised machine learning. Alur penelitian terdiri atas pengumpulan data, pelabelan otomatis berbasis rating bintang, pra-pemrosesan teks, ekstraksi fitur TF-IDF, penyeimbangan kelas menggunakan SMOTE, pelatihan model, serta evaluasi performa pada berbagai skema pembagian data latih dan data uji.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Data penelitian berupa ulasan pengguna aplikasi TVRI Klik yang diperoleh dari Google Play Store menggunakan teknik web scraping dengan pustaka google-play-scraper pada bahasa pemrograman Python, sebagaimana lazim digunakan pada penelitian sejenis untuk mengumpulkan ulasan aplikasi secara otomatis [5]. Atribut yang berhasil dikumpulkan meliputi identitas ulasan, nama pengguna, isi ulasan (content), skor rating (score), jumlah suka (thumbsUpCount), versi aplikasi, serta tanggal ulasan dibuat (at). Dari seluruh atribut tersebut, hanya kolom content, score, dan at yang digunakan pada penelitian ini, sedangkan data dengan nilai kosong dibuang agar tidak mengganggu proses pelabelan dan pelatihan model. Setelah proses seleksi, diperoleh 2.006 ulasan yang siap diproses lebih lanjut, dengan rentang waktu ulasan mencakup periode sebelum hingga sesudah peluncuran fitur siaran Piala Dunia 2026.

2.2. Pelabelan Data Berbasis Rating Bintang

Pelabelan sentimen pada penelitian ini dilakukan secara otomatis berdasarkan nilai rating bintang (score) yang diberikan pengguna, mengikuti pendekatan yang umum digunakan pada penelitian analisis sentimen ulasan aplikasi Google Play Store [18]. Ulasan dengan rating 4 dan 5 diberi label sentimen positif (dikodekan 1), rating 3 diberi label netral (dikodekan 0), sedangkan rating 1 dan 2 diberi label negatif (dikodekan -1). Pendekatan berbasis rating dipilih karena tidak membutuhkan anotasi manual dan mampu diterapkan pada skala data yang besar, meskipun perlu diakui pendekatan ini memiliki keterbatasan karena rating bintang tidak selalu selaras sepenuhnya dengan sentimen yang tersirat pada teks ulasan, sehingga berpotensi menimbulkan noise pada sebagian kecil label dibanding anotasi manual yang diverifikasi langsung [19].

2.3. Pra-pemrosesan Teks

Tahapan pra-pemrosesan dilakukan untuk menormalkan data dan mengurangi noise sebelum ekstraksi fitur. Tahapan ini meliputi konversi huruf menjadi huruf kecil (case folding), penghapusan karakter non-alfabet dan simbol, tokenisasi menggunakan RegexpTokenizer dari pustaka NLTK, penghapusan stopword Bahasa Indonesia, serta stemming menggunakan Sastrawi Stemmer untuk mengembalikan setiap kata ke bentuk dasarnya. Penggunaan Sastrawi pada tahap stemming dipilih karena telah terbukti secara konsisten memperbaiki kualitas pra-pemrosesan dokumen berbahasa Indonesia dibanding tanpa stemming [16], sedangkan pemilihan algoritma stemming yang sesuai karakteristik bahasa turut memengaruhi akurasi klasifikasi teks pada tahap selanjutnya [20]. Setelah proses stemming, token-token pada setiap ulasan digabungkan kembali menjadi satu rangkaian teks bersih (text_string) yang menjadi masukan bagi tahap ekstraksi fitur.

2.4. Ekstraksi Fitur dengan TF-IDF

Representasi numerik teks dibangun menggunakan Term Frequency-Inverse Document Frequency (TF-IDF), yaitu metode pembobotan yang menekan pengaruh kata yang terlalu umum muncul pada banyak dokumen dan memperkuat kontribusi kata yang lebih spesifik dan informatif dalam membedakan sentimen [4]. Pembobotan TF-IDF dipilih karena secara konsisten memberikan performa klasifikasi yang lebih baik dibanding representasi frekuensi kata sederhana seperti Bag-of-Words pada berbagai tugas klasifikasi sentimen teks.

2.5. Penanganan Ketidakseimbangan Kelas dengan SMOTE

Hasil pelabelan berbasis rating bintang menghasilkan distribusi kelas yang sangat tidak seimbang, dengan kelas negatif mendominasi data. Untuk mengatasi hal ini, penelitian ini menerapkan *Synthetic Minority Oversampling Technique* (SMOTE) pada data latih. SMOTE bekerja dengan membangkitkan sampel sintetis pada kelas minoritas melalui interpolasi linear antara sebuah sampel minoritas dengan salah satu tetangga terdekatnya di ruang fitur, sehingga data baru yang dihasilkan tetap berada dalam sebaran yang masuk akal dan tidak sekadar menduplikasi data yang sudah ada. Pendekatan ini banyak digunakan pada penelitian klasifikasi sentimen ulasan aplikasi untuk meningkatkan kemampuan model dalam mengenali kelas minoritas [12], meskipun penerapannya perlu dilakukan hanya pada data latih agar evaluasi pada data uji tetap merepresentasikan distribusi data yang sesungguhnya dan terhindar dari bias optimistis pada hasil pengujian.

2.6. Model Klasifikasi

Penelitian ini membandingkan dua algoritma *Supervised Machine Learning*, yaitu *Support Vector Machine* (SVM) dan Multinomial Naive Bayes. SVM bekerja dengan mencari *hyperplane* pemisah optimal antar kelas pada ruang fitur berdimensi tinggi dan telah terbukti tangguh pada berbagai tugas klasifikasi teks dengan dimensi fitur yang besar. Multinomial Naive Bayes bekerja berdasarkan teorema probabilitas Bayes dengan asumsi independensi antar fitur, sehingga sederhana dan efisien secara komputasi untuk data frekuensi kata berdimensi tinggi [5]. Kedua algoritma ini konsisten menjadi pembanding utama pada berbagai penelitian klasifikasi sentimen ulasan aplikasi Google Play Store [1], sehingga dipilih sebagai model utama pada penelitian ini.

a. Naïve Bayes

Dalam metode *Machine Learning*, salah satu algoritma klasifikasi yang banyak digunakan adalah *Naive Bayes*, khususnya dalam bidang *Natural Language Processing* (NLP). Algoritma ini didasarkan pada Teorema Bayes yang diperkenalkan oleh Thomas Bayes, seorang ilmuwan asal Inggris. Teorema Bayes digunakan untuk memperkirakan probabilitas suatu kejadian berdasarkan informasi atau bukti yang diperoleh dari kejadian sebelumnya. Dalam penerapannya pada pembelajaran mesin, informasi dari data historis dianalogikan sebagai data pelatihan, sedangkan data yang akan diklasifikasikan digunakan sebagai dasar untuk melakukan prediksi atau evaluasi [21]. Algoritma *Naive Bayes* mengasumsikan bahwa setiap atribut atau fitur dalam suatu data bersifat independen satu sama lain. Asumsi tersebut memungkinkan proses klasifikasi dilakukan dengan menggunakan pendekatan probabilistik dan statistik yang relatif sederhana [7]. Berdasarkan prinsip tersebut, keberadaan suatu karakteristik pada suatu kelas dianggap tidak bergantung pada karakteristik lainnya. Salah satu keunggulan *Naive Bayes* adalah struktur algoritmanya yang sederhana, proses komputasinya cepat, serta mampu memberikan tingkat akurasi yang baik. Meskipun memiliki mekanisme yang relatif sederhana, *Naive Bayes* sering menghasilkan kinerja klasifikasi yang baik, terutama pada permasalahan klasifikasi teks. Secara matematis, algoritma *Naive Bayes* dapat dinyatakan melalui rumus berikut:

$$P(H | X) = \frac{P(X | H)P(H)}{P(X)} \quad (1)$$

Keterangan:

X = Data yang memiliki kelas yang tidak diketahui

H = Asumsidata X melambangkan kelas tertentu atau spesifik

P(H|X) = Peluang asumsi H berlandaskan keadaan X (posterior probability)

P(H) = Peluang asumsi H (prior probability)

P(X|H) = Peluang asumsi X berlandaskan keadaan H (likelihood)

P(X) = Peluangdari X(predictor prior probability)

b. SVM

Support Vector Machine (SVM) merupakan salah satu algoritma *machine learning* yang dapat digunakan untuk menyelesaikan permasalahan klasifikasi maupun regresi. Prinsip dasar SVM pada awalnya dikembangkan untuk menangani permasalahan klasifikasi linear, yaitu kondisi ketika data dari beberapa kelas dapat dipisahkan secara linear. Namun, SVM juga mampu menangani permasalahan klasifikasi nonlinier melalui penerapan konsep *kernel* yang memungkinkan data dipetakan ke dalam ruang fitur berdimensi lebih tinggi [9]. Dalam proses klasifikasi, *hyperplane* digunakan sebagai batas pemisah antarkelas. Apabila data dapat dipisahkan secara linear, proses klasifikasi dapat dilakukan menggunakan pendekatan *linear support vector classifier*, sedangkan untuk data yang tidak dapat dipisahkan secara linear, SVM memanfaatkan fungsi *kernel* untuk memperoleh pemisahan yang optimal pada ruang fitur yang lebih tinggi.

Secara geometris, SVM merepresentasikan *hyperplane* sebagai bidang pemisah yang membedakan data dari kelas positif dan kelas negatif. Prinsip utama SVM adalah menentukan *hyperplane* optimal dengan memaksimalkan *margin*, yaitu jarak antara *hyperplane* dan titik data terdekat dari masing-masing kelas. Titik-titik data yang berada paling dekat dengan *hyperplane*

tersebut disebut *support vectors* dan memiliki peran penting dalam menentukan posisi batas keputusan. Dengan memaksimalkan *margin*, SVM berupaya memperoleh batas klasifikasi yang memiliki kemampuan generalisasi yang lebih baik terhadap data baru [7].

Dalam konteks klasifikasi, *Support Vector Classifier* (SVC) merupakan pendekatan yang digunakan untuk menentukan *hyperplane* pemisah dengan mempertimbangkan kesalahan klasifikasi serta batas *margin*. Berbeda dengan pendekatan klasifikasi yang hanya berfokus pada minimisasi kesalahan prediksi, SVM mengoptimalkan kedua aspek tersebut melalui formulasi optimasi matematis. Dengan demikian, model tidak hanya berupaya mengurangi kesalahan klasifikasi, tetapi juga mencari *hyperplane* yang memberikan *margin* optimal antarkelas. Secara matematis, persamaan yang digunakan untuk menentukan *hyperplane* dapat dinyatakan sebagai berikut:

$$w \cdot x + b = 0 \rightarrow w_0 + w_1x_1 + w_2x_2 \quad (2)$$

Keterangan:

$W = \{w_1, w_2, \dots, w_n\} \rightarrow$ bobot

n = Jumlah atribut

b = Skalar yang digunakan sebagai bias

x = Nilai atribut

Dalam penelitian ini, ialah menerapkan model pendekatan kernel linier atau sama saja dengan Support Vector Classifier (SVC). Rumusnya adalah sebagai berikut:

$$k(x, z) = \varphi(x)^T \varphi(z) \quad (3)$$

Keterangan:

k = Fungsi kernel untuk semua vektor input dan a

φ = Fungsi pemetaan dari area input ke area fitur

2.7. Skema Pengujian dan Evaluasi

Evaluasi performa model dilakukan dengan menggunakan proporsi 80% data latih dan 20% data uji. Pada skema ini, proses SMOTE diterapkan hanya pada data latih, sedangkan data uji tetap menggunakan distribusi kelas asli agar hasil evaluasi mencerminkan performa model pada kondisi data riil. Metrik evaluasi utama yang digunakan adalah akurasi (*accuracy*), yaitu proporsi prediksi yang benar dari keseluruhan data uji, sebagaimana lazim dilaporkan pada penelitian perbandingan SVM dan Naive Bayes untuk klasifikasi sentimen [12].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1-score = \frac{2 \times precision \times recall}{precision + recall} \quad (7)$$

3. HASIL DAN PEMBAHASAN

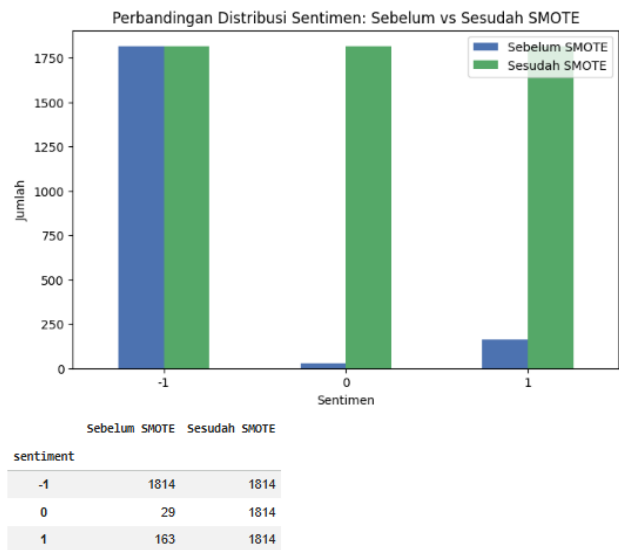
3.1. Distribusi Kelas Sentimen Sebelum dan Sesudah SMOTE

Hasil pelabelan otomatis berbasis rating bintang terhadap 2.006 ulasan menunjukkan distribusi kelas yang sangat tidak seimbang. Kelas negatif mendominasi data dengan 1.814 ulasan, jauh melampaui kelas positif sebanyak 163 ulasan dan kelas netral yang hanya berjumlah 29 ulasan. Ketimpangan ini sejalan dengan temuan pada penelitian ulasan aplikasi lain yang melabeli sentimen berdasarkan rating, di mana pengguna cenderung lebih aktif menuliskan ulasan ketika memberikan rating rendah sebagai bentuk keluhan dibanding ketika memberi rating menengah.

Setelah SMOTE diterapkan pada data latih, jumlah data pada kelas netral dan positif ditingkatkan secara sintetis hingga menyamai jumlah kelas negatif, yaitu masing-masing menjadi 1.814 data, sehingga distribusi ketiga kelas menjadi seimbang sempurna sebagaimana disajikan pada Tabel 1 dan divisualisasikan pada Gambar 1.

Tabel 1. Distribusi Kelas Sentimen Sebelum dan Sesudah SMOTE

Kelas Sentimen	Jumlah Sebelum SMOTE	Jumlah Sesudah SMOTE
Negatif (-1)	1.814	1.814
Netral (0)	29	1.814
Positif (1)	163	1.814
Total	2.006	5.442



Gambar 2. Perbandingan Distribusi Sentimen Sebelum dan Sesudah SMOTE

```
# Perbandingan jumlah data sebelum vs sesudah SMOTE
before_counts = y.value_counts().sort_index()
after_counts = pd.Series(y_sm).value_counts().sort_index()

comparison_df = pd.DataFrame({'Sebelum SMOTE': before_counts, 'Sesudah SMOTE': after_counts})
comparison_df.plot(kind='bar', figsize=(8,5), color=['#4C72B0', '#55A868'])
plt.title('Perbandingan Distribusi Sentimen: Sebelum vs Sesudah SMOTE')
plt.xlabel('Sentimen')
plt.ylabel('Jumlah')
plt.xticks(rotation=0)
plt.legend()
plt.show()

comparison_df
```

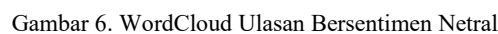
Gambar 3. Pseudocode Perbandingan Distribusi Sentimen Sebelum dan Sesudah SMOTE

Dominasi kelas negatif pada penelitian ini berbeda dengan pola pada beberapa penelitian sejenis yang justru didominasi ulasan positif, dan hal ini konsisten dengan karakteristik ulasan aplikasi TVRI Klik yang banyak diunggah pengguna pada saat mengalami kendala teknis, misalnya ketika hendak menyaksikan siaran langsung pertandingan Piala Dunia 2026. Ketimpangan kelas yang ekstrem seperti ini, apabila tidak ditangani, berpotensi menyebabkan model klasifikasi menjadi bias terhadap kelas mayoritas dan gagal mengenali pola pada kelas minoritas, sehingga penerapan SMOTE pada tahap pelatihan menjadi langkah penting sebelum data digunakan untuk membangun model klasifikasi.

Untuk memperoleh gambaran awal mengenai topik yang paling banyak diutarakan pengguna, kata yang paling sering muncul pada tiap kelas sentimen divisualisasikan menggunakan wordcloud, sebagaimana ditunjukkan pada Gambar 3 hingga Gambar 5. Visualisasi ini membantu memvalidasi kualitas hasil pra-pemrosesan sekaligus memberikan gambaran kualitatif mengenai isu yang paling dominan pada tiap kelas sebelum data digunakan untuk pemodelan.

[illegible]

Sebaliknya, ulasan bersentimen negatif didominasi kata seperti "kosong", "aplikasi", "sampah", "jelek", "buka", "download", "fungsi", dan "muncul", yang menunjukkan bahwa keluhan utama pengguna berpusat pada persoalan teknis, yaitu layar yang tampil kosong, kegagalan membuka atau memuat siaran, serta fitur yang tidak berfungsi sebagaimana mestinya.



Tabri Eko Miyardi | Page 370

mengindikasikan bahwa sebagian ulasan pada kelas ini berisi masukan atau permintaan perbaikan yang disampaikan secara netral tanpa nada pujian maupun kekecewaan yang tegas.

3.3. Perbandingan Akurasi pada Skema Pembagian 80:20

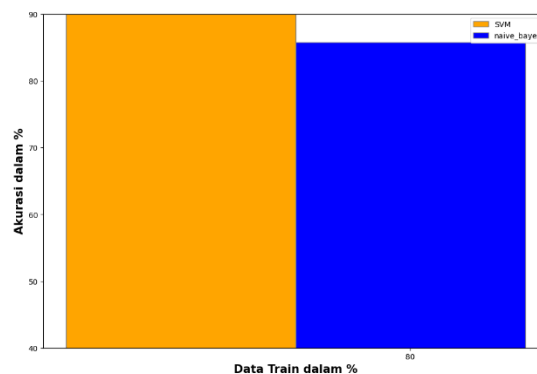
Pada skema pembagian data latih dan data uji sebesar 80:20, SVM memperoleh akurasi data uji sebesar 95,13%, jauh mengungguli Multinomial Naive Bayes yang memperoleh akurasi 86,59%, sebagaimana disajikan pada Tabel 2. Selisih akurasi sebesar 8,54 poin persentase ini menunjukkan bahwa kemampuan SVM dalam menemukan hyperplane pemisah yang optimal pada ruang fitur TF-IDF berdimensi tinggi memberikan keunggulan yang cukup signifikan dibanding asumsi independensi antar fitur yang mendasari Naive Bayes, sejalan dengan temuan pada penelitian perbandingan SVM dan Naive Bayes pada ulasan aplikasi Google Play Store lain yang juga melaporkan keunggulan SVM secara konsisten [2], [20].

Tabel 2. Hasil Akurasi Model pada Skema Pembagian Data 80:20

Data Latih (%)	Data Uji (%)	Akurasi SVM (%)	Akurasi Naive Bayes (%)
80	20	95,13	86,59

3.4. Proporsi Data Latih terhadap Akurasi Model

Untuk menguji konsistensi keunggulan SVM, penelitian ini membandingkan akurasi kedua model pada data latih 80% dan data uji 20%, sebagaimana divisualisasikan pada Gambar 7. SVM menunjukkan performa yang unggul dengan akurasi 95,13% dibandingkan dengan Naive Bayes yang hanya 86,59%. Pola ini mengindikasikan bahwa SVM sudah mampu membangun batas keputusan yang cukup baik meski hanya dilatih dengan sebagian kecil data, dan performanya relatif tidak sensitif terhadap penambahan volume data latih setelah melewati titik tertentu.



Gambar 7. Perbandingan Akurasi SVM dan Naive Bayes pada Berbagai Proporsi Data Latih

```
#sorting berdasarkan data train ascending
df_akurasi = df_akurasi.sort_values(by='data_train', ascending=True)
# set width of bar
barwidth = 0.3
fig = plt.subplots(figsize=(12, 8))

# Set position of bar on X axis
br1 = np.arange(len(df_akurasi['SVM']))
br2 = [x + barwidth for x in br1]

# Make the plot
plt.bar(br1, df_akurasi['SVM'], color='orange', width=barwidth,
        edgecolor='grey', label='SVM')
plt.bar(br2, df_akurasi['naive_bayes'], color='b', width=barwidth,
        edgecolor='grey', label='naive_bayes')

plt.ylim(40, 90)
# Adding Xticks
plt.xlabel('Data Train dalam %', fontweight='bold', fontsize=15)
plt.ylabel('Akurasi dalam %', fontweight='bold', fontsize=15)
plt.xticks([r + barwidth for r in range(len(df_akurasi['SVM']))],
           df_akurasi['data_train'])
```

Gambar 8. Pseudo Code Perbandingan Akurasi SVM dan Naive Bayes

Berbeda dengan SVM, Naive Bayes menunjukkan pola akurasi 86,59%. Hasil ini mengindikasikan bahwa performa Naive Bayes lebih sensitif terhadap komposisi data latih yang terbentuk pada tiap pembagian, kemungkinan besar karena asumsi independensi antar fitur pada Naive Bayes membuat model ini lebih mudah terpengaruh oleh variasi distribusi kata pada subset data latih yang berbeda-beda.

Selain itu, konsistensi performa SVM pada berbagai ukuran data latih memberikan implikasi praktis yang penting, yaitu model SVM yang telah dilatih pada penelitian ini berpotensi tetap andal digunakan meski volume ulasan yang terkumpul relatif terbatas, sebagaimana kondisi umum pada aplikasi yang baru dirilis atau memiliki basis pengguna yang belum terlalu besar. Temuan ini sejalan dengan penelitian lain yang menunjukkan bahwa penerapan teknik penyeimbangan kelas seperti SMOTE dapat meningkatkan kinerja SVM secara lebih signifikan dibanding Naive Bayes pada data ulasan yang tidak seimbang [8], sehingga kombinasi SVM dan SMOTE dapat menjadi pendekatan yang direkomendasikan untuk sistem pemantauan sentimen ulasan aplikasi streaming milik lembaga penyiaran publik ke depan.

4. KESIMPULAN DAN SARAN

Penelitian ini berhasil membandingkan kinerja Support Vector Machine dan Multinomial Naive Bayes untuk klasifikasi sentimen ulasan pengguna aplikasi TVRI Klik berdasarkan 2.006 ulasan yang dikumpulkan dari Google Play Store, dengan label sentimen diturunkan dari rating bintang dan ketidakseimbangan kelas ditangani menggunakan SMOTE pada tahap pelatihan. Hasil pengujian pada skema pembagian data 80:20 menunjukkan SVM mencapai akurasi tertinggi sebesar 95,13%, jauh mengungguli Naive Bayes yang mencapai 86,59%. Visualisasi wordcloud turut mengungkap bahwa keluhan pengguna pada ulasan negatif didominasi persoalan teknis seperti layar kosong dan kegagalan memuat siaran, sementara ulasan positif menonjolkan kemudahan mengakses siaran nasional secara gratis. Dengan demikian, SVM yang dikombinasikan dengan SMOTE dapat direkomendasikan sebagai model utama untuk sistem pemantauan sentimen ulasan aplikasi TVRI Klik ke depan. Penelitian selanjutnya disarankan untuk memperluas periode pengumpulan data, membandingkan hasil pelabelan berbasis rating dengan anotasi manual guna memvalidasi kualitas label, menambahkan algoritma pembanding seperti Random Forest atau XGBoost, serta menjajaki penggunaan model berbasis representasi kontekstual seperti IndoBERT untuk melihat sejauh mana peningkatan performa yang dapat dicapai pada data ulasan aplikasi streaming penyiaran publik berbahasa Indonesia.

DAFTAR PUSTAKA

- [1] E. Eskiyaturrofikoh and R. R. Suryono, "Analisis Sentimen Aplikasi X Pada Google Play Store Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine (SVM)," *JIPi (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 3, pp. 1408–1419, 2024, doi: 10.29100/jipi.v9i3.5392.
- [2] T. Sugihartono and R. R. C. Putra, "Penerapan Metode Support Vector Machine Dalam Classifikasi Ulasan Pengguna Aplikasi Mobile Jkn," *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 7, no. 2, pp. 144–153, 2024, doi: 10.36080/skanika.v7i2.3193.
- [3] N. Khoirunnisaa, K. N. N. Kesuma, S. Setiawan, and A. Y. P. Yusuf, "Klasifikasi Teks Ulasan Aplikasi Netflix Pada Google Play Store Menggunakan Algoritma Naïve Bayes Dan SVM," *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 7, no. 1, pp. 64–73, 2024, doi: 10.36080/skanika.v7i1.3138.
- [4] S. Jatmiko and C. Dometian, "Analisis Sentimen Ulasan Google Play Store: Studi Komparatif Algoritma SVM, Naïve Bayes, dan Logistic Regression," *J. Fasilkom*, vol. 15, no. 3, pp. 610–620, 2026, doi: 10.37859/jf.v15i3.10016.
- [5] D. A. TARIGAN, Z. Situmorang, and R. Rosnelly, "Analisis Sentimen Aplikasi Playstore Sirekap 2024 Pasca Pilpres Dengan Perbandingan Metode Support Vector Machine (SVM), Naïve Bayes Classifier Dan Random Forest.," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 3, pp. 661–670, 2025, doi: 10.25126/jtiik.22025129608.

- [6] F. J. Putri *et al.*, “Perbandingan Kinerja Algoritma Knn dan SVM Dalam Klasifikasi Kematangan Buah Jeruk Medan Berdasarkan Citra Digital,” *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 9, no. 1, pp. 159–170, 2026, doi: 10.36080/skanika.v9i1.3661.
- [7] G. A. Mursianto, D. Widiyanto, and B. T. Wahyono, “Analisis Sentimen Ulasan Pengguna Pada Aplikasi Google Classroom Menggunakan Metode SVM Dan Seleksi Fitur PSO,” *Inform. J. Ilmu Komput.*, vol. 18, no. 3, p. 221, 2022, doi: 10.52958/iftk.v18i3.4685.
- [8] G. Ginabila and A. Fauzi, “Analisis Sentimen Terhadap Pemutar Musik Online Spotify Dengan Algoritma Naive Bayes dan Support Vector Machine,” *J. Ilm. Ilk. - Ilmu Komput. Inform.*, vol. 6, no. 2, pp. 111–122, 2023, doi: 10.47324/ilkominfo.v6i2.180.
- [9] M. F. Kurniawan, and A. H. Azizah, “Analisis Sentimen Kebijakan Pemerintahan Prabowo di Platform X Menggunakan Naïve Bayes dan SVM,” *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 9, no. 2, pp. 248–257, 2026, doi: 10.36080/skanika.v9i2.3726.
- [10] H. A. Raihana, P. H. Gunawan, and N. Aquarini, “Breaking Class Imbalance: Machine Learning Solutions for Stunting Detection,” *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 16, no. 2, p. 03, 2025, doi: 10.24843/LKJITI.2025.v16.i02.p03.
- [11] R. L. Pratiwi, Z. I. Alfianti, A. Fauzi, and G. Ginabila, “Penerapan Algoritma Naive Bayes dan SVM untuk Analisis Sentimen terhadap Penggunaan True Wireless Stereo (TWS),” *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 8, no. 2, pp. 257–268, 2025, doi: 10.36080/skanika.v8i2.3535.
- [12] L. Qadrini, H. Hikmah, and M. Megasari, “Oversampling, Undersampling, Smote SVM dan Random Forest pada Klasifikasi Penerima Bidikmisi Sejava Timur Tahun 2017,” *J. Comput. Syst. Informatics*, vol. 3, no. 4, pp. 386–391, 2022, doi: 10.47065/josyc.v3i4.2154.
- [13] E. Y. Boateng, J. Otoo, and D. A. Abaye, “Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review,” *J. Data Anal. Inf. Process.*, vol. 08, no. 04, pp. 341–357, 2020, doi: 10.4236/jdaip.2020.84020.
- [14] A. Demircioğlu, “Applying oversampling before cross-validation will lead to high bias in radiomics,” *Sci. Rep.*, vol. 14, no. 1, pp. 1–11, 2024, doi: 10.1038/s41598-024-62585-z.
- [15] M. A. Rosid, *et al.*, “Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 874, no. 1, 2020, doi: 10.1088/1757-899X/874/1/012017.
- [16] M. D. Purbolaksono, *et al.*, “Indonesian text classification using back propagation and sastrawi stemming analysis with information gain for selection feature,” *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 10, no. 1, pp. 234–238, 2020, doi: 10.18517/ijaseit.10.1.8858.
- [17] Y. B. Blikon, A. G. Sooi, and P. A. N. Samane, “Perbandinga Kinerja Pengklasifikasi Citra Buah Kakao Sakit Dan Sehat Menggunakan Support Vector Machine (Svm) Dan K-Nearest Neighbors (Knn),” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 1, pp. 1–8, 2023, doi: 10.24176/simet.v14i1.9012.
- [18] S. Rizkiyah *et al.*, “Analisis Sentimen Ulasan Aplikasi Maxim Merchant dengan Support Vector Machine (SVM) dan Random Forest,” *JDMIS J. Data Min. Inf. Syst.*, vol. 4, no. 1, pp. 1–9, 2026, doi: 10.54259/jdmis.v4i1.4765.
- [19] J. A. Munif, A. I. Pradana, and D. Hartanti, “Analisis Diskrepansi Antara Rating Dan Sentimen Teks Pada Ulasan Aplikasi Seabank Menggunakan Model Deep Learning Indobert,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 10, no. 3, pp. 5159–5165, 2026, doi: 10.36040/jati.v10i3.18815.
- [20] E. Hokijuliandy, H. Napitupulu, and Firdaniza, “Application of SVM and Chi-Square Feature Selection for Sentiment Analysis of Indonesia’s National Health Insurance Mobile Application,” *Mathematics*, vol. 11, no. 17, 2023, doi: 10.3390/math11173765.
- [21] M. S. Hadi, J. Akbar, and M. F. Zulkarnain, “Analisis Sentimen Wisata Air Terjun Di Kabupaten Lombok Tengah Menggunakan Metode Support Vector Machine (Svm),” *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 8, no. 2, pp. 318–329, 2025, doi: 10.36080/skanika.v8i2.3578.