

Evaluation Of Naive Bayes And SVM Performance In Predicting The Sentiment Of User Reviews On The Pinterest Application

[Evaluasi Performa Naive Bayes Dan SVM Dalam Memprediksi Sentimen Ulasan Pengguna Aplikasi Pinterest]

Indri Arianti Wibowo¹⁾, Uce Indahyanti^{*2)}

¹⁾Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

²⁾Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

*Email Penulis Korespondensi: uceindahyanti@umsida.ac.id

Abstract. *Advances in digital technology have driven an increase in the use of mobile applications, including Pinterest, which generates many user reviews as a source of satisfaction information. However, unstructured text reviews make manual analysis difficult and have the potential to cause bias due to sentiment data imbalance. This study aims to compare the performance of the Naive Bayes and Support Vector Machine (SVM) algorithms in classifying the sentiment of Pinterest user reviews. The research methods included pre-processing, labelling, feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF), data division at a ratio of 80:20 and 70:30, and handling of unbalanced data using Random Oversampling (ROS). Evaluation was carried out through a confusion matrix with accuracy, precision, recall, and F1-score metrics. The results show that SVM in the 80:20 Imbalanced ratio condition obtained the highest accuracy of 89.01% and an F1-score of 88.71%, while Naive Bayes with a 70:30 ROS ratio produced the highest negative class recall of 87.03%. The study emphasises the importance of evaluating performance per class, particularly negative class recall.*

Keywords - Sentiment Analysis; Pinterest; Naive Bayes; Support Vector Machine; Random Oversampling

Abstrak. *Kemajuan teknologi digital mendorong peningkatan penggunaan aplikasi mobile, termasuk Pinterest yang menghasilkan banyak ulasan pengguna sebagai sumber informasi kepuasan. Namun, ulasan berbentuk teks tidak terstruktur menyulitkan analisis manual dan berpotensi menimbulkan bias akibat ketidakseimbangan data sentimen. Penelitian ini bertujuan membandingkan performa algoritma Naive Bayes dan Support Vector Machine (SVM) dalam klasifikasi sentimen ulasan pengguna Pinterest. Metode penelitian meliputi pre-processing, pelabelan, ekstraksi fitur menggunakan Term Frequency–Inverse Document Frequency (TF-IDF), pembagian data rasio 80:20 dan 70:30, serta penanganan data tidak seimbang menggunakan Random Oversampling (ROS). Evaluasi dilakukan melalui confusion matrix dengan metrik akurasi, precision, recall, dan F1-score. Hasil menunjukkan SVM pada kondisi Imbalanced rasio 80:20 memperoleh akurasi tertinggi 89,01% dan F1-score 88,71%, sedangkan Naive Bayes dengan ROS rasio 70:30 menghasilkan recall kelas negatif tertinggi sebesar 87,03%. Penelitian menekankan pentingnya evaluasi performa per kelas, khususnya recall kelas negatif.*

Kata Kunci - Analisis Sentimen; Pinterest; Naive Bayes; Support Vector Machine; Random Oversampling

I. PENDAHULUAN

Seiring berkembangnya teknologi digital, penggunaan aplikasi mobile seperti Pinterest terus meningkat, dengan lebih dari 522 juta pengguna aktif bulanan di dunia pada tahun 2024 [1]. Pinterest merupakan platform media sosial berbasis visual yang memberi kesempatan pengguna untuk menjelajah, mengoleksi, dan membagikan ide serta inspirasi dalam format gambar atau video [2].

Banyaknya pengguna menghasilkan ulasan beragam yang dapat memengaruhi persepsi dan ekspektasi terhadap kualitas aplikasi [3]. Namun, jumlah ulasan yang sangat besar membuat analisis manual tidak efisien [4]. Sehingga diperlukan analisis sentimen untuk mengetahui tingkat kepuasan pengguna serta mengolah teks tidak terstruktur menjadi informasi yang bermanfaat dalam pengambilan keputusan [5]. Berdasarkan permasalahan tersebut, penelitian ini menganalisis perbandingan performa algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Pinterest, guna mengidentifikasi opini ke dalam sentimen positif dan negatif serta mengevaluasi performa per kelas, khususnya pada sentimen negatif.

Beberapa penelitian telah melakukan perbandingan performa algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) dalam klasifikasi sentimen ulasan pengguna dan menunjukkan hasil yang bervariasi. Pada penelitian platform *e-commerce* Tokopedia dan TikTok Shop, bahwa *Naive Bayes* dilaporkan lebih unggul dengan akurasi 97,50% dan *F1-score* 84,00%, sedangkan *Support Vector Machine* (SVM) memperoleh akurasi 94,90% dan *F1-score*

76,80% [6]. Penelitian lain pada aplikasi Ceria by BRI menghasilkan bahwa *Support Vector Machine* (SVM) memiliki performa akurasi lebih tinggi yaitu 88% daripada *Naive Bayes* yang hanya 84% [7]. Selain itu, penelitian yang menganalisis sentimen publik terhadap kinerja pemerintah dalam penanganan serangan ransomware pada algoritma *Support Vector Machine* (SVM), *Random Forest*, dan *Naive Bayes* pada 6.484 komentar *youtube*. Pengujian dilakukan dengan dan tanpa SMOTE, hasilnya menunjukkan bahwa penerapan dengan SMOTE meningkatkan performa seluruh model, di mana *Support Vector Machine* (SVM) mencapai akurasi tertinggi yaitu 96%, disusul oleh *Random Forest* 95% dan *Naive Bayes* 89% [8]. Penelitian lain juga menunjukkan algoritma *Support Vector Machine* (SVM) mampu menghasilkan performa yang optimal, seperti pada prediksi pemilihan karir bagi alumni UMSIDA dengan akurasi mencapai 97% [9]. Sementara itu, pada analisis aplikasi Weverse, algoritma *Naive Bayes* rasio 80:20 menunjukkan performa optimal dengan akurasi sebesar 86,81% dan *F1-score* sebesar 83,88% [10].

Meskipun berbagai penelitian telah menganalisis sentimen pada aplikasi mobile, belum ditemukan penelitian yang secara khusus mengkaji perbandingan kedua algoritma pada ulasan pengguna aplikasi Pinterest. Dengan demikian, penelitian ini dilakukan guna mengisi celah tersebut.

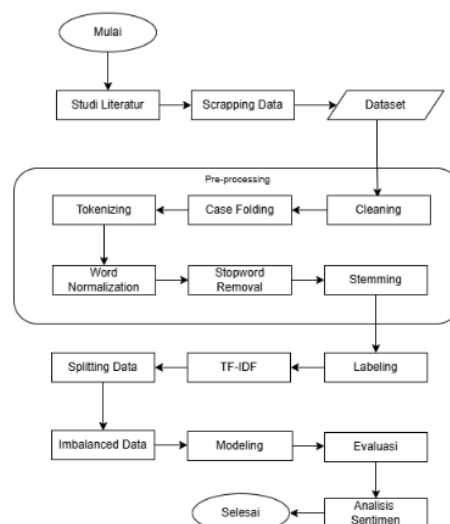
Untuk mengatasi ketidakseimbangan data, penelitian ini diawali dengan *pre-processing* teks dan ekstraksi fitur menggunakan TF-IDF. Data latih diseimbangkan menggunakan *Random Oversampling* (ROS) agar model tidak bias terhadap kelas mayoritas. Klasifikasi dilakukan dengan algoritma *Naive Bayes* dan *Support Vector Machine* (SVM). Performa model dievaluasi menggunakan *confusion matrix* melalui metrik akurasi, *precision*, *recall*, dan *F1-score*. Penelitian ini diharapkan mampu menyajikan gambaran yang lebih mendalam terkait persepsi pengguna serta mendukung peningkatan kualitas aplikasi Pinterest ke depannya.

II. METODE

Penelitian ini membandingkan performa *Naive Bayes* dan *Support Vector Machine* (SVM). *Naive Bayes* merupakan algoritma klasifikasi yang menggunakan pendekatan probabilistik dengan asumsi bahwa setiap fitur bersifat independen, berdasarkan Teorema Bayes [11]. Sedangkan, *Support Vector Machine* (SVM) merupakan algoritma klasifikasi yang menentukan *hyperplane* terbaik sebagai batas pemisah antar kelas dengan memaksimalkan jarak terhadap *support vector* terdekat [12]. Data bersumber dari ulasan pengguna aplikasi Pinterest di *Google Play Store*, yang kemudian dipublikasikan secara terbuka dan dapat diakses melalui platform *Kaggle* pada tautan berikut: <https://www.kaggle.com/datasets/indriarianti/dataset-pinterest>. Selain itu, *source code* yang digunakan dalam proses pengolahan dan analisis data diunggah pada repositori *GitHub* yang tersedia pada: <https://github.com/indriarianti12/SentimenPinterest>.

II.1 Tahap Penelitian

Gambar 1 dijelaskan tentang proses analisis sentimen ini mencakup tahapan seperti scraping data, *pre-processing*, labeling, ekstraksi fitur (TF-IDF), splitting data, *Imbalanced data*, modeling dan evaluasi.



Gambar 1. Tahap Penelitian

II.2 Studi Literatur

Studi literatur dilakukan untuk mengumpulkan berbagai sumber ilmiah yang relevan, seperti buku dan jurnal, guna memperkuat landasan teori serta mendukung analisis terhadap permasalahan penelitian.

II.3 Scraping Data

Data penelitian berasal dari ulasan pengguna aplikasi Pinterest yang diambil dari *Google Play Store*. Hasil pengumpulan data tersebut kemudian diunggah ke *Kaggle* sebagai repositori data publik agar dapat diakses dan digunakan kembali oleh peneliti selanjutnya. Proses tahapan pengolahan data, pemodelan, serta evaluasi dilakukan dengan menggunakan *Google Colab* dengan alur pengujian yang bersifat umum dan dapat diterapkan pada dataset lain dengan karakteristik serupa. Data yang digunakan terdiri dari beberapa atribut, yaitu id, rating, ulasan, dan *date*.

II.4 Pre-processing

Pre-processing adalah tahapan dalam pengolahan data guna mengoptimalkan kualitas data mentah agar analisis yang dilakukan menjadi akurat serta mengurangi potensi permasalahan. Tahapan ini meliputi *cleaning*, *case folding*, *tokenizing*, *word normalization*, *stopword removal* dan *stemming* [13]. Dalam penelitian ini, data yang digunakan difokuskan pada ulasan berbahasa Indonesia.

II.5 Labeling

Pelabelan dilakukan untuk mengategorikan sentimen ulasan pengguna menjadi positif atau negatif. Proses ini memanfaatkan rating pengguna dan metode berbasis leksikon (*lexicon-based*) [14]. Metode ini memberikan label secara otomatis dengan mencocokkan kata dalam ulasan dengan kamus sentimen yang dikelompokkan menjadi positif dan negatif [15]. Pelabelan dilakukan secara semi-otomatis, dengan rating 1-2 sebagai negatif, rating 4-5 sebagai positif, dan rating 3 dikategorikan sebagai ambigu. Oleh karena itu, ulasan dengan rating 3 diberi label menggunakan pendekatan *lexicon-based*, di mana skor sentimen dihitung dari bobot kata, skor > 0 diklasifikasikan sebagai positif dan skor ≤ 0 sebagai negatif. Namun, pendekatan ini memiliki keterbatasan yang dapat menimbulkan bias karena rating tidak selalu selaras dengan isi ulasan, serta metode *lexicon-based* kurang mampu menangkap konteks, ironi, dan makna implisit, sehingga berpotensi menimbulkan bias dalam hasil klasifikasi.

II.6 Ekstraksi Fitur (TF-IDF)

Term Frequency-Inverse Document Frequency (TF-IDF) ialah metode yang mentransformasikan teks dalam format numerik dengan mempertimbangkan frekuensi kemunculan kata dalam suatu dokumen (TF) serta tingkat kelangkaan kemunculan kata pada seluruh kumpulan dokumen (IDF) [16]. Dengan demikian, dapat membantu model dalam mengenali kata-kata yang paling berpengaruh terhadap sentimen setiap ulasan [17]. Berikut perhitungan *Term Frequency* (TF):

$$TF_{t,d} = \frac{\text{Jumlah kemunculan kata } t \text{ dalam dokumen } d}{\text{Jumlah total kata dalam dokumen } d} \quad (1)$$

Rumus *Inverse Document Frequency* (IDF):

$$IDF_t = \log \left(\frac{\text{Jumlah total dokumen } (N)}{\text{Jumlah dokumen yang mengandung kata } t (DF_t)} \right) \quad (2)$$

Hasil dari TF-IDF:

$$TF - IDF_{t,d} = TF_{t,d} \times IDF_t \quad (3)$$

II.7 Splitting Data

Dataset dibagi dengan memisahkan ke dalam dua subset yaitu data latih (*training data*) dan data uji (*testing data*) menggunakan rasio pembagian sebesar 80:20 dan 70:30 [18]. Rasio 80:20 memberikan proporsi lebih besar untuk data latih, sehingga model dapat mempelajari pola dengan baik. Sementara rasio 70:30 menyediakan data uji lebih banyak untuk menguji kemampuan generalisasi model. Kedua rasio digunakan untuk menganalisis perbedaan pembagian data terhadap performa model pada tiap kelas sentimen, terutama dalam mendeteksi kelas negatif sebagai kelas minoritas, serta mengevaluasi dampak *Random Oversampling* (ROS) terhadap *recall* dan *F1-score*.

II.8 Imbalanced Data

Pada banyaknya dataset terutama dalam klasifikasi, sering terjadi *Imbalanced* data, di mana satu kelas memiliki jumlah sampel yang lebih banyak. Untuk mengatasi hal ini, *Random Oversampling* (ROS) digunakan untuk menambah sampel pada kelas minoritas. Dengan demikian, distribusi data pada kelas menjadi lebih merata [19].

II.9 Modeling dan Evaluasi

Tahap pemodelan, *Naive Bayes* dan *Support Vector Machine* (SVM) diterapkan untuk melakukan klasifikasi sentimen pada ulasan pengguna. Dengan hanya dua kategori sentimen yaitu positif dan negatif, penelitian ini

menggunakan pendekatan klasifikasi biner. Serta, model dievaluasi melalui *confusion matrix* guna mengukur performa algoritma menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Akurasi menilai persentase prediksi benar dari total keseluruhan, *precision* merepresentasikan tingkat akurat model terhadap klasifikasi kelas positif, *recall* mengevaluasi performa model menangkap seluruh data pada kelas minoritas, dan *F1-score* metrik yang menyeimbangkan antara *precision* dan *recall* [20].

III. HASIL DAN PEMBAHASAN

III.1 Scraping Data

Sebanyak 10.000 ulasan dikumpulkan dalam rentang waktu Januari 2020 hingga Januari 2025 sebagai data penelitian. Data ini menjadi data mentah yang selanjutnya diproses pada tahap *pre-processing* sebelum dilakukan analisis sentimen.

Tabel 1. Hasil *Scraping Data*

Id	Rating	Ulasan	Date
c6a3f68c-b438-4e51-856a-32284abfa451	1	Ga bisa daftar tolong diperbaiki!!	2025-01-06
d1532942-9f9e-4bdb-aed7-35cdfed7e746	2	Kenapa akun saya di nonaktifkan:((2024-12-15
8688c0c7-67e8-4c62-bac9-88c1bbe429fe	3	Bagus tapi ada kurangnya	2023-12-04
41f736a0-41fa-4576-a3f5-ebaae33ec6d6	4	Ini sangat keren	2022-12-04
342a1020-7d80-43a0-b0b0-f4cbe8c8d981	5	Aplikasi ini bagus pake banget, bisa bikin nyari inspirasi baru...	2021-12-23

III.2 Pre-processing

Tahap ini adalah *pre-processing*. Data yang digunakan hanya pada ulasan berbahasa Indonesia. Proses *scraping* mengumpulkan 10.000 ulasan, kemudian dilakukan penyaringan dengan hanya mengambil ulasan berbahasa Indonesia dan menghapus ulasan dalam bahasa lainnya. Sehingga, diperoleh 5.681 ulasan berbahasa Indonesia yang selanjutnya digunakan dalam tahap analisis. Tahap ini meliputi *cleaning* yaitu membersihkan karakter yang bukan relevan, termasuk tanda baca, angka, emoji dan simbol, *case folding* yaitu menyeragamkan huruf dalam teks menjadi huruf kecil, *tokenizing* yaitu tahap pemrosesan dengan membagi kalimat menjadi unit-unit lebih kecil, *word normalization* yaitu mengubah singkatan, kata tidak baku, typo, dan slang menjadi kata baku, dengan menggunakan *dictionary* berupa dataset kamus slang dari platform *Kaggle* yang tersedia pada tautan: <https://www.kaggle.com/data-sets/fornigulo/kamus-slag> [21]. *stopword removal* yaitu membuang kata yang tidak memiliki informasi penting, *stemming* yaitu mengembalikan kata ke bentuk kata dasarnya, dengan memanfaatkan *library* Sastrawi.

Tabel 2. Hasil *Cleaning*

Sebelum	Sesudah
Aplikasi ini bagus pake banget, bisa bikin nyari inspirasi baru...	Aplikasi ini bagus pake banget bisa bikin nyari inspirasi baru

Tabel 3. Hasil *Case Folding*

Sebelum	Sesudah
Aplikasi ini bagus pake banget bisa bikin nyari inspirasi baru	aplikasi ini bagus pake banget bisa bikin nyari inspirasi baru

Tabel 4. Hasil *Tokenizing*

Sebelum	Sesudah
aplikasi ini bagus pake banget bisa bikin nyari inspirasi baru	['aplikasi', 'ini', 'bagus', 'pake', 'banget', 'bisa', 'bikin', 'nyari', 'inspirasi', 'baru']

Tabel 5. Hasil *Word Normalization*

Sebelum	Sesudah
['aplikasi', 'ini', 'bagus', 'pake', 'banget', 'bisa', 'bikin', 'nyari', 'inspirasi', 'baru']	['aplikasi', 'ini', 'bagus', 'pakai', 'banget', 'bisa', 'bikin', 'mencari', 'inspirasi', 'baru']

Tabel 6. Hasil *Stopword Removal*

Sebelum	Sesudah
['aplikasi', 'ini', 'bagus', 'pakai', 'banget', 'bisa', 'bikin', 'mencari', 'inspirasi', 'baru']	['aplikasi', 'bagus', 'pakai', 'banget', 'bikin', 'mencari', 'inspirasi', 'baru']

Tabel 7. Hasil *Stemming*

Sebelum	Sesudah
['aplikasi', 'bagus', 'pakai', 'banget', 'bikin', 'mencari', 'inspirasi', 'baru']	['aplikasi', 'bagus', 'pakai', 'banget', 'bikin', 'cari', 'inspirasi', 'baru']

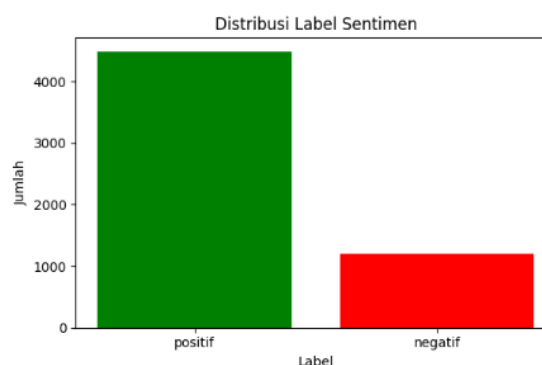
III.3 Labeling

Tahap labeling bertujuan menandai setiap ulasan ke dalam sentimen positif atau negatif dengan memanfaatkan kombinasi rating pengguna.

Tabel 8. Sesudah Pre-Processing dan Labeling

Rating	Ulasan Awal	Ulasan Akhir	Date	Label
1	Ga bisa daftar tolong diperbaiki!!	daftar perbaiki	2025-01-06	Negatif
2	Kenapa akun saya di nonaktifkan:(akun nonaktif	2024-12-15	Negatif
3	Bagus tapi ada kurangnya	bagus kurang	2023-12-04	Positif
4	Ini sangat keren	sangat keren	2022-12-04	Positif
5	Aplikasi ini bagus pake banget, bisa bikin nyari inspirasi baru...	aplikasi bagus pakai banget bikin cari inspirasi baru	2021-12-23	Positif

Distribusi label sentimen dapat dilihat pada Gambar 2, dimana jumlah ulasan positif jauh lebih banyak, yaitu sebanyak 4.482 data dibanding dengan ulasan negatif yang berjumlah 1.199 data. Kondisi ini memperlihatkan distribusi data yang tidak merata, di mana kelas positif menjadi kelas mayoritas dan negatif sebagai kelas minoritas.



Gambar 2. Distribusi Label Sentimen

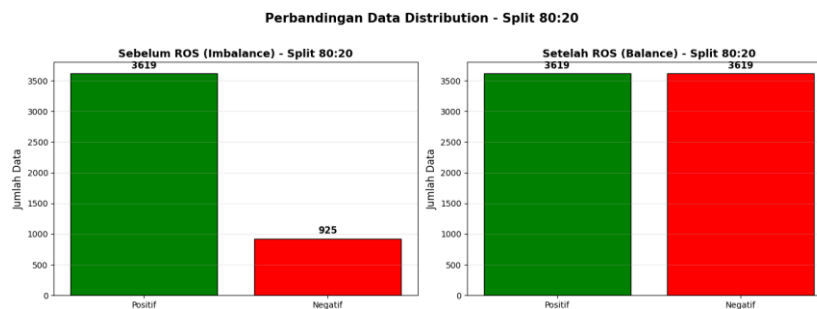
III.4 Penyeimbangan Data

Penyeimbangan data dilakukan dengan metode *Random Oversampling* (ROS) untuk mengatasi ketidakseimbangan antar kelas. Dengan metode ini, sampel kelas minoritas digandakan secara acak hingga sebanding dengan kelas mayoritas, sehingga proporsi data antar kelas menjadi lebih seimbang.

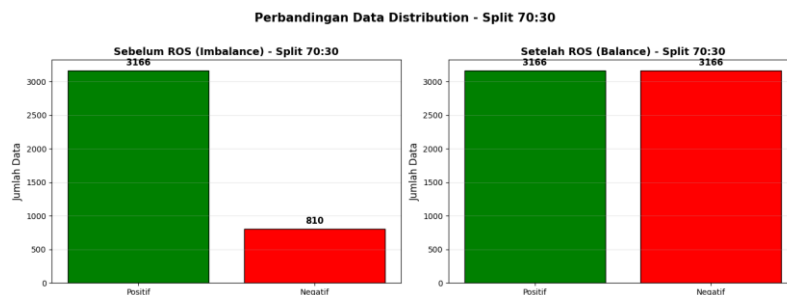
Tabel 9. Distribusi dan Penyeimbangan Data

Rasio Split	Kelas	Data Sebelum ROS	Data Setelah ROS	Data Test (Asli)
80:20	Positif (Mayoritas)	3.619	3.619	905
	Negatif (Minoritas)	925	3.619	232
	Total data train	4.544	7.238	1.137
70:30	Positif (Mayoritas)	3.166	3.166	1.358
	Negatif (Minoritas)	810	3.166	347
	Total data train	3.976	6.332	1705

Berdasarkan Gambar 3 dan Gambar 4, terlihat rasio pembagian data 80:20 dan 70:30 sebelum penerapan ROS menunjukkan kondisi data yang tidak seimbang, data kelas negatif jauh lebih sedikit dibanding kelas positif. Setelah ROS diterapkan, jumlah kelas negatif disetarakan sehingga setara dengan kelas positif. Sehingga, distribusi data menjadi seimbang. Penyeimbangan ini penting agar model dapat dilatih secara seimbang pada kedua kelas dan bias terhadap kelas mayoritas dapat dikurangi.



Gambar 3. Perbandingan Data Rasio 80:20



Gambar 4. Perbandingan Data Rasio 70:30

III.5 Pengujian Algoritma

Pengujian dilakukan dengan membandingkan dua algoritma, yaitu *Naive Bayes* dan *Support Vector Machine* (SVM) menggunakan 2 pembagian data yaitu rasio 80:20 dan 70:30 serta 2 kondisi data latih yaitu *Imbalanced* dan *Balanced*. Tujuan pengujian ini adalah untuk membandingkan efektivitas kedua algoritma serta mengevaluasi dampak penyeimbangan data latih menggunakan metode *Random Oversampling* (ROS) terhadap metrik. Pengujian dengan rasio 80:20 dan 70:30 disajikan pada Tabel 10 dan Tabel 11.

Tabel 10. Hasil Pengujian Rasio 80:20

Algoritma	Kondisi Train	Akurasi	Precision	Recall	F1-score
Naive Bayes	Imbalanced	86,72%	86,07%	86,72%	85,30%
Naive Bayes	Balanced	85,31%	88,44%	85,31%	86,18%
SVM	Imbalanced	89,01%	88,60%	89,01%	88,71%
SVM	Balanced	88,30%	88,89%	88,30%	85,54%

Tabel 11. Hasil Pengujian Rasio 70:30

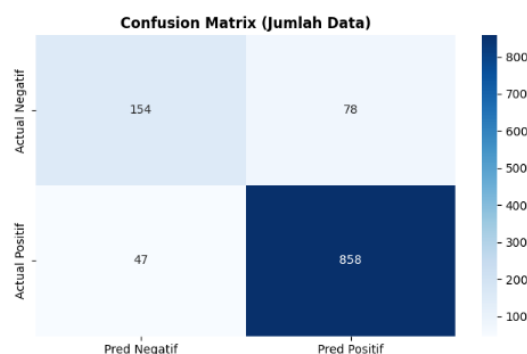
Algoritma	Kondisi Train	Akurasi	Precision	Recall	F1-score
Naive Bayes	Imbalanced	86,22%	85,45%	86,22%	84,66%
Naive Bayes	Balanced	86,69%	89,41%	86,69%	87,43%
SVM	Imbalanced	87,97%	88,59%	88,97%	88,70%
SVM	Balanced	88,21%	88,92%	88,21%	88,48%

III.6 Analisis Perbandingan Performa Model

Analisis perbandingan performa model menunjukkan bahwa *Support Vector Machine* (SVM) rasio pembagian data 80:20 dengan kondisi data *Imbalanced* menjadi model dengan performa global terbaik, ditunjukkan oleh akurasi sebesar 89,01% dan F1-score 88,71%. Penggunaan dua rasio pembagian data, yaitu 80:20 dan 70:30, bertujuan guna menyeimbangkan kebutuhan data latih dalam mempelajari pola sentimen serta data uji dalam mengevaluasi kemampuan generalisasi model. Meskipun *Support Vector Machine* (SVM) unggul pada akurasi global, evaluasi per kelas memperlihatkan bahwa model masih kurang optimal dalam mendeteksi sentimen negatif sebagai kelas minoritas. Oleh karena itu, *Random Oversampling* (ROS) diterapkan untuk mengatasi ketidakseimbangan data, yang terbukti meningkatkan *recall* dan *F1-score* kelas negatif pada kedua algoritma. Hal ini ditunjukkan pada Tabel 12, di mana peningkatan paling signifikan terjadi pada *Naive Bayes* rasio 70:30 dengan *recall* kelas negatif meningkat dari 44,38% menjadi 87,03%, meskipun disertai *trade-off* berupa potensi penurunan akurasi global. Selain itu, analisis *confusion matrix* pada Gambar 5 menunjukkan bahwa tingginya akurasi global dipengaruhi dominasi prediksi kelas positif, sementara kesalahan klasifikasi masih terjadi pada kelas negatif akibat ketidakseimbangan data. Dengan demikian, evaluasi performa model tidak cukup hanya berdasarkan akurasi, tetapi perlu mempertimbangkan metrik per kelas, khususnya kemampuan model dalam mengidentifikasi sentimen negatif sebagai fokus utama pada klasifikasi data tidak seimbang [22].

Tabel 12. Perbandingan Performa Kelas Negatif

Algoritma	Rasio Split	Kondisi Train	Recall	F1-score
Naive Bayes	80:20	Imbalanced	46,55%	58,86%
		Balanced (ROS)	85,34%	70,34%
	70:30	Imbalanced	44,38%	56,72%
		Balanced (ROS)	87,03%	72,68%
SVM	80:20	Imbalanced	66,38%	71,13%
		Balanced (ROS)	77,16%	72,91%
	70:30	Imbalanced	66,86%	77,17%
		Balanced (ROS)	77,81%	72,87%



Gambar 5. Confusion Matrix

IV. SIMPULAN

Penelitian ini membandingkan performa algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) dalam klasifikasi sentimen ulasan pengguna aplikasi Pinterest pada kondisi data tidak seimbang. Hasil penelitian menunjukkan bahwa *Support Vector Machine* (SVM) pada rasio pembagian data 80:20 menghasilkan performa global

terbaik dengan akurasi 89,01% dan F1-score 88,71%. Sementara itu, *Naive Bayes* dengan penerapan ROS pada rasio 70:30 memberikan kemampuan terbaik dalam mendeteksi sentimen negatif dengan *recall* sebesar 87,03%. Penerapan ROS terbukti mampu meningkatkan performa deteksi kelas minoritas meskipun menimbulkan *trade-off* berupa potensi penurunan akurasi global. Evaluasi model pada data tidak seimbang sebaiknya tidak sekedar menilai performa model dari sisi akurasi keseluruhan, melainkan juga mempertimbangkan performa per kelas, khususnya kemampuan model dalam mengidentifikasi sentimen negatif. Hasil analisis juga menunjukkan bahwa ulasan positif didominasi aspek inspirasi dan kegunaan fitur, sedangkan ulasan negatif berkaitan dengan permasalahan teknis aplikasi. Penelitian selanjutnya disarankan untuk menggunakan metode penyeimbangan data lain seperti SMOTE serta menerapkan model *deep learning*, misalnya LSTM, BiLSTM, atau IndoBERT.

UCAPAN TERIMA KASIH

Penulis menyampaikan rasa terima kasih kepada Universitas Muhammadiyah Sidoarjo sebagai institusi yang telah memberikan fasilitas, lingkungan akademik, serta dukungan sepanjang pelaksanaan kegiatan penelitian. Ucapan terima kasih juga penulis sampaikan kepada dosen pembimbing yang telah membimbing serta memberikan berbagai saran dan masukan secara berkelanjutan sehingga proses penelitian dapat berjalan dan diselesaikan dengan lancar. Selain itu, penulis turut mengucapkan terima kasih seluruh pihak yang telah membantu melalui dukungan, saran, maupun motivasi selama proses penelitian dan penyusunan penelitian ini.

REFERENSI

- [1] B. of Apps, "Pinterest Revenue and Usage Statistics," *Business of Apps*. Accessed: Jan. 02, 2026. [Online]. Available: <https://www.businessofapps.com/data/pinterest-statistics/>
- [2] B. Uddin, D. D. F. Basam, and E. M. L. Wijayadi, "Efektivitas Pemanfaatan Pinterest Terhadap Kreativitas Pengguna," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 7, no. 4, pp. 678–684, 2024.
- [3] A. Nurian and B. N. Sari, "Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naive Bayes," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3s1, pp. 829–835, 2023, doi: 10.23960/jitet.v11i3s1.3348.
- [4] Y. T. R. Sholiha, L. A. M. Nabilah, and Imron, "Analisis Sentimen Aplikasi Liputan6.Com pada Ulasan Pengguna di Google Playstore dengan Menggunakan Algoritma Support Vector Machine (Svm) dan Naive Bayes," *Saturnus J. Teknol. dan Sist. Inf.*, vol. 3, no. 3, pp. 30–51, 2025, doi: <https://doi.org/10.61132/saturnus.v3i3.867>.
- [5] P. Anggraini and Winarsih, "KOMPARASI NAIVE BAYES, SUPPORT VECTOR MACHINE, DAN RANDOM FOREST DALAM ANALISIS SENTIMEN APLIKASI SHOPEE DI GOOGLE PLAY STORE," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 3, pp. 4451–4457, 2025.
- [6] A. Nabilla M. P., A. Adrian Y. L., A. Indra J., Y. Arya S., Sumanto, and A. Diah H., "Komparasi Naive Bayes dan SVM untuk Analisis Sentimen Pada E-Commerce Seller Center," *J. Sains Dan Teknol.*, vol. 5, no. 3, pp. 199–208, 2025.
- [7] M. Aji, A. Maldini, and S. Andryana, "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI PERBANKAN," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 3, pp. 4098–4105, 2025.
- [8] M. I. Prayugah, U. Indahyanti, and N. Ariyanti, "ANALISIS SENTIMEN PUBLIK PADA PEMERINTAH DALAM SERANGAN RANSOMWARE DENGAN PENDEKATAN SMOTE," *JOISIE (Journal Inf. Syst. Informatics Eng.)*, vol. 8, no. 2, pp. 333–343, 2024.
- [9] M. D. Qur'ani, H. Setiawan, and I. A. Kautsar, "PENERAPAN METODE SUPPORT VECTOR MACHINE (SVM) UNTUK MEMPREDIKSI PEMILIHAN KARIR BAGI ALUMNI UMSIDA," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 4, pp. 2906–2916, 2025.
- [10] A. R. Annjani, U. Indahyanti, A. W. Azinar, and R. Dijaya, "PENERAPAN ALGORITMA MACHINE LEARNING UNTUK MEMPREDIKSI SENTIMEN KEPUASAN PENGGUNA WEVERSE," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 10, no. 1, pp. 784–790, 2026.
- [11] F. Damayanti, A. Rahim, and N. A. Verdikha, "ANALISIS SENTIMEN ULASAN PENGGUNA TERHADAP APLIKASI K24KLIK," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 2, pp. 3224–3230, 2025.
- [12] J. A. Putra, A. Dharmawan, and J. Gondohanindijo, "SENTIMEN ANALISIS APLIKASI DIGITALENT MOBILE MENGGUNAKAN NAIVE BAYES DAN SVM DENGAN EKSTRAKSI FITUR TF-IDF," *J. Inf. Technol. Comput. Sci.*, vol. 7, 2024.
- [13] R. A. P. Sari, S. Kacung, and B. Santoso, "ANALISIS SENTIMEN LAYANAN KESEHATAN BPJS MENGGUNAKAN METODE SVM," *J. Inform. Teknol. dan Sains*, vol. 7, no. 2, pp. 878–885, 2025.
- [14] S. A. S. Mola, D. L. B. Baun, I. O. Nunes, and M. M. A. R. Sani, "ANALISIS SENTIMEN APLIKASI HALO BCA DI GOOGLE PLAY STORE MENGGUNAKAN METODE NAIVE BAYES , SUPPORT VECTOR

- MACHINE DAN RANDOM FOREST PENDAHULUAN,” *HOAQ J. Teknol. Inf.*, vol. 15, no. 2, pp. 69–79, 2024.
- [15] D. Fitriano, R. Indriati, and A. Ristyawan, “Analisis Sentimen Ulasan Aplikasi Chatgpt Menggunakan Algoritma Support Vector Machine dan Lexicon Based,” *JSITIK*, vol. 3, no. 2, pp. 101–111, 2025.
- [16] D. N. N. Husnina, D. E. Ratnawati, and B. Rahayudi, “Analisis Sentimen Pengguna Aplikasi RedBus berdasarkan Ulasan di Google Play Store menggunakan Metode Naïve Bayes,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 2, 2023.
- [17] D. W. Bhatara and R. R. Suryono, “ANALISIS SENTIMEN APLIKASI BCA MOBILE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 4, pp. 1907–1917, 2024.
- [18] F. A. Soraya and A. D. Indriyanti, “Perbandingan Algoritma Naïve Bayes dan Support Vector Machine dalam Analisis Sentimen Aplikasi Teman Bus,” *JEISBI (Journal Emerg. Inf. Syst. Bus. Intell.*, vol. 05, no. 02, pp. 118–125, 2024.
- [19] N. D. H. Sadikin and S. Susanti, “Analisis Sentimen Publik Terhadap Kampanye Pengurangan Sampah Plastik Menggunakan Algoritma Naïve Bayes,” *J. FASILKOM*, vol. 15, no. 2, pp. 202–212, 2025.
- [20] A. Y. Lubisa and M. Y. H. Setyawan, “Analisis Sentimen Aplikasi Pospay di PlayStore Menggunakan Algoritma Support Vector Machine & Naive Bayes,” *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 6, no. 3, pp. 514–521, 2024.
- [21] Fornieli, “kamus_slang,” Kaggle. Accessed: Feb. 26, 2026. [Online]. Available: <https://www.kaggle.com/datasets/fornigulo/kamus-slag>
- [22] H. P. Jelita, M. I. Saad, and Wahyuni, “Penerapan Algoritma Naïve Bayes Dalam Analisis Sentimen Masyarakat Terhadap STMIK Widya Cipta Dharma,” *Bull. Inf. Technol.*, vol. 6, no. 2, pp. 148–160, 2025, doi: 10.47065/bit.v5i2.2029.

Conflict of Interest Statement:

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.