

# ANALISIS SENTIMEN PRODUK APPLE VISION PRO MENGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE

Meitry Ayu Handini<sup>1)</sup>, Ryan Randy Suryono<sup>\*2)</sup>.

1. Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia.
2. Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia.

## Article Info

**Kata Kunci:** Analisis sentimen, Apple Vision Pro, *Naïve Bayes*, *Support Vector Machine*

**Keywords:** *Apple Vision Pro*, *Naïve Bayes*, *Support Vector Machine*, *Sentiment Analysis*

## Article history:

Received 21 August 2024

Revised 12 September 2024

Accepted 2 October 2024

Available online 1 September 2025

## DOI :

<https://doi.org/10.29100/jipi.v10i3.6495>

\* Corresponding author.

Ryan Randy Suryono

E-mail address:

[ryan@teknokrat.ac.id](mailto:ryan@teknokrat.ac.id)

## ABSTRAK

Dalam era digital, pandangan pengguna kini memengaruhi keputusan konsumen, terutama dalam industri teknologi. Analisis sentimen terhadap produk, seperti yang dilakukan dalam penelitian ini terhadap Apple Vision Pro, menjadi krusial dalam memahami respons pengguna terhadap inovasi teknologi. Sebelum memasuki konteks produk Apple Vision Pro, fenomena ini mencerminkan bagaimana pandangan pengguna secara umum dapat menjadi penentu penting dalam kesuksesan suatu produk teknologi. Penelitian ini menerapkan *Naïve Bayes* dan *Support Vector Machine* (SVM) untuk menganalisis sentimen ulasan produk Apple Vision Pro. Model *Naïve Bayes* menunjukkan konsistensi dengan *precision* 79% (positif) dan 78% (negatif), *recall* 78%, dan *F1-score* 78% untuk kedua kelas. Setelah optimasi dengan SMOTE, model SVM menunjukkan peningkatan dengan *precision* 81% (positif) dan 76% (negatif), *recall* 75% (positif) dan 82% (negatif), serta *F1-score* 78% (positif) dan 79% (negatif). Namun, dengan pertimbangan konsistensi, *Naïve Bayes* lebih diutamakan karena stabilitasnya dalam mengklasifikasikan sentimen. Penelitian ini menyajikan perspektif yang mendalam mengenai penggunaan analisis sentimen untuk memahami tanggapan pengguna terhadap teknologi, terutama dalam konteks produk Apple Vision Pro.

## ABSTRACT

In the digital age, user views now influence consumer decisions, especially in the technology industry. Product sentiment analysis, such as the one conducted in this study on the Apple Vision Pro, is crucial in understanding user responses to technological innovations. Before entering the context of the Apple Vision Pro product, this phenomenon reflects how user views in general can be an important determinant in the success of a technology product. This research applies *Naïve Bayes* and *Support Vector Machine* (SVM) to analyze Apple Vision Pro product review sentiments. The *Naïve Bayes* model showed consistency with 79% (positive) and 78% (negative) precision, 78% re-call, and 78% F1-score for both classes. After optimization with SMOTE, the SVM model showed improvement with 81% (positive) and 76% (negative) precision, 75% (positive) and 82% (negative) recall, and 78% (positive) and 79% (negative) F1-score. However, with consistency considerations, *Naïve Bayes* is preferred due to its stability in classifying sentiments. This research presents an in-depth perspective on the use of sentiment analysis to understand user responses to technology, especially in the context of the Apple Vision Pro product.

## I. PENDAHULUAN

Sering dengan laju evolusi teknologi informasi dan komunikasi, konsumen semakin terlibat dalam proses seleksi dan adaptasi terhadap produk-produk teknologi yang terus berkembang. Tidak hanya sebagai pengguna pasif, namun konsumen saat ini juga menjadi agen yang aktif dalam menggali informasi, membandingkan produk, dan membentuk preferensi berdasarkan kebutuhan dan keinginan mereka. Penelitian ini akan mengarahkan fokusnya pada perusahaan Apple Inc., yang dikenal karena inovasinya yang konsisten dalam industri teknologi dan pengaruhnya yang besar dalam pasar global. Perusahaan Apple Inc, yang berbasis di

Amerika Serikat, dikenal sebagai pemimpin dalam produksi *smartphone* global yang telah memperoleh ketenaran luas. Perusahaan ini selalu berkomitmen untuk melakukan inovasi pada produknya dengan menambahkan fasilitas atau fitur baru yang menarik sesuai dengan kebutuhan konsumen [1].

Apple Vision Pro adalah produk yang sangat dinantikan dalam industri teknologi karena merupakan headset *mixed-reality* yang menjanjikan pengalaman baru dalam interaksi pengguna dengan lingkungan virtual dan realitas augmentasi [2]. Dengan peluncuran produk seperti ini, Apple berusaha untuk memperluas pasar mereka ke domain baru dan menawarkan inovasi yang dapat memengaruhi cara orang berinteraksi dengan teknologi [3]. Analisis sentimen produk adalah pendekatan yang penting dalam strategi pemasaran dan pengembangan produk. Dengan memahami sentimen pelanggan terhadap produk [4], perusahaan dapat mengidentifikasi kekuatan dan kelemahan produk mereka, mengukur kepuasan pelanggan, dan mengidentifikasi peluang untuk peningkatan produk di masa depan [5]. Dalam konteks Apple Vision Pro, analisis sentimen dapat membantu Apple memahami sejauh mana produk ini diterima oleh pasar dan apakah ada aspek tertentu dari produk yang perlu ditingkatkan atau dioptimalkan.

Penelitian terkait metaverse mengumpulkan data dari platform Twitter (X) menggunakan teknik *scraping*, yang difokuskan pada kata kunci atau *keyword metaverse*. Sebanyak 71.233 data tanggapan diperoleh, namun setelah menghapus spam dan iklan, tersisa 62.677 data yang dianalisis lebih lanjut. Pemodelan dilakukan dengan menggunakan algoritma *Support Vector Machine* (SVM) dan *Naive Bayes*. Hasil menunjukkan bahwa SVM mencapai tingkat akurasi tertinggi sebesar 90,32%, sementara *Naive Bayes* memperoleh akurasi tertinggi sebesar 84,23%. Penelitian ini memberikan wawasan mendalam mengenai pandangan masyarakat Indonesia terhadap isu *metaverse* [6]. Studi ini mengkaji opini publik tentang topik yang dibahas di Twitter dalam bahasa Indonesia dan mendukung analisis pasar terkait persepsi publik. Data dikumpulkan melalui proses pelabelan secara manual, *pre-processing*, dan pemodelan menggunakan query "kendaraan listrik" di Kaggle.com. Model klasifikasi dibangun dengan algoritma *Naive Bayes* dan *Support Vector Machine*, dengan akurasi masing-masing mencapai 63,02% dan 70,82%. [7].

Penelitian yang melakukan perbandingan klasifikasi dan akurasi sentimen menggunakan algoritma *Support Vector Machine* (SVM) dan *Naive Bayes* dalam menganalisis pandangan positif dan negatif terhadap informasi dan edukasi aplikasi SATUSEHAT memungkinkan pemahaman yang lebih efisien terhadap perspektif masyarakat. Penelitian menggunakan data dari Twitter dengan kata kunci "SatuSehat", dengan setidaknya 1046 komentar dalam rentang waktu satu tahun pada tahun 2023. Hasil akhir menunjukkan bahwa metode SVM memiliki akurasi 87,95%, dengan 132 data positif dan 947 data negatif dari total 1080 data uji. Metode *Naive Bayes* memiliki akurasi 81,65%. [8]. Penelitian ini bertujuan untuk menerapkan algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) dalam analisis sentimen terhadap ulasan pengguna Apple Vision Pro, dengan fokus pada pemahaman sentimen, identifikasi faktor-faktor yang mempengaruhi sentimen, serta evaluasi keefektifan kedua algoritma dalam menganalisis sentimen terhadap produk teknologi. Studi sebelumnya lebih banyak berfokus pada analisis sentimen terhadap produk teknologi yang sudah mapan atau memiliki penggunaan luas, dan masih minim mengevaluasi persepsi pengguna terhadap produk inovatif seperti Apple Vision Pro.

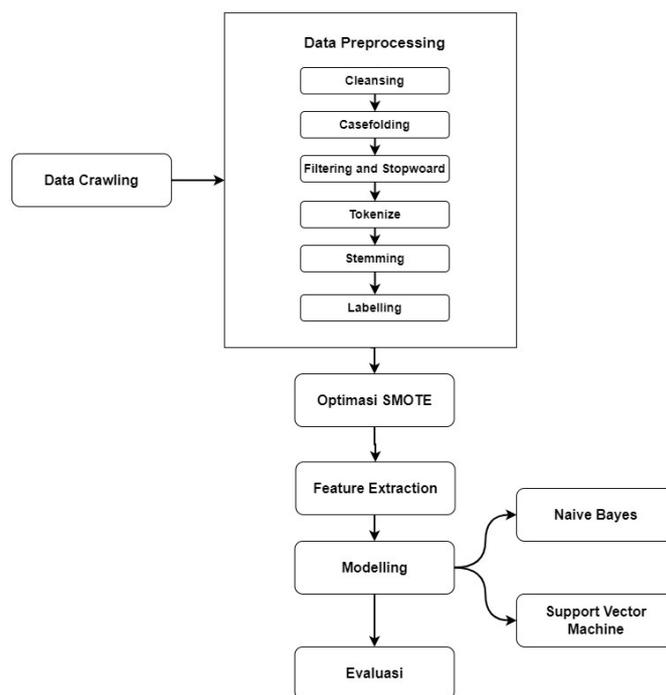
Selain itu, sebagian besar penelitian tidak secara komprehensif membandingkan performa *Naive Bayes* dan SVM dalam konteks ulasan produk teknologi yang kompleks dan baru di pasaran. Gap penelitian ini mencakup kurangnya eksplorasi mendalam mengenai faktor-faktor spesifik yang mempengaruhi sentimen pengguna terhadap produk dengan karakteristik unik seperti Apple Vision Pro, serta belum adanya analisis perbandingan efektivitas kedua algoritma tersebut dalam memahami persepsi dan respons pengguna. Penelitian ini diharapkan dapat mengisi gap tersebut dengan menyediakan wawasan yang lebih mendalam tentang efektivitas pendekatan analisis sentimen dalam konteks produk teknologi inovatif, khususnya Apple Vision Pro.

## II. METODE PENELITIAN

### A. Metode Penelitian

Di dalam metode penelitian ini, diterapkan pendekatan yang komprehensif dalam memproses data, yang dimulai dari tahapan *data crawling* untuk mengumpulkan data mentah. Selanjutnya, dalam proses *data preprocessing*, dilakukan serangkaian langkah penting seperti *cleansing* untuk membersihkan data dari *noise*, *casefolding* untuk standarisasi teks, *filtering and stopword* untuk menghilangkan kata-kata tidak penting, *tokenize* untuk memecah teks menjadi token-token, dan *stemming* untuk mengubah kata-kata menjadi bentuk dasarnya. Setelah itu, tahapan

*labelling* diterapkan untuk memberikan label pada data yang telah diproses. Langkah selanjutnya adalah optimasi menggunakan metode SMOTE untuk menangani ketidakseimbangan kelas, kemudian dilanjutkan dengan *feature extraction* sebagai persiapan data untuk pemodelan menggunakan metode *Naive Bayes* dan *Support Vector Machine*. Pemilihan algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) didasarkan pada keunggulan mereka dalam analisis sentimen. *Naive Bayes* dikenal karena kesederhanaan dan kecepatan komputasinya, ideal untuk pengolahan data teks yang besar namun memiliki keterbatasan pada fitur yang saling bergantung. Sebaliknya, SVM menawarkan kemampuan klasifikasi yang sangat akurat dan efektif pada data non-linear, tetapi memerlukan lebih banyak sumber daya komputasi dan tuning parameter yang cermat. Dengan mengevaluasi dan membandingkan kedua algoritma ini, penelitian ini bertujuan untuk menentukan metode yang paling efektif dalam menganalisis sentimen terhadap ulasan Apple Vision Pro, serta memberikan wawasan yang mendalam mengenai pendekatan terbaik untuk memahami sentimen pada produk teknologi canggih. Terakhir, evaluasi dilakukan untuk mengevaluasi performa model. Tahapan penelitian ini dapat dilihat pada Gambar 1 berikut ini.



Gambar 1. Tahapan penelitian

## B. Data Crawling

Proses *crawling* dalam penelitian ini dilakukan dengan *query* Apple Vision Pro untuk mendapatkan *tweet* dari pengguna Twitter, memanfaatkan library Harvest dari Python serta API yang disediakan oleh Twitter.

## C. Data Preprocessing

Dalam tahapan *text mining*, seleksi dan pembersihan data menjadi fokus utama. Seleksi data melibatkan identifikasi data yang relevan, sedangkan pembersihan data bertujuan menghilangkan elemen yang tidak dibutuhkan. Hal ini penting untuk memastikan analisis yang akurat dan bermakna [9]. Berikut serangkaian langkah yang harus dilakukan dalam tahap *preprocessing* antara lain :

### 1) Cleansing

Proses *cleansing*, atau dikenal juga sebagai pembersihan data, melibatkan analisis mendalam terhadap kualitas data dengan melakukan modifikasi, revisi, atau penghapusan terhadap informasi yang dianggap tidak relevan, tidak lengkap, tidak akurat, atau memiliki format yang tidak sesuai dalam basis data. Tujuannya adalah untuk menghasilkan data yang memiliki kualitas tinggi [10].

### 2) Casefolding

Dalam tahap pengolahan teks, terdapat proses yang disebut *case folding*, yang bertujuan untuk merubah semua huruf dalam suatu kalimat atau kata menjadi huruf kecil. Langkah ini dilakukan untuk menyamakan format teks sehingga mempermudah proses analisis dan pemrosesan lebih lanjut [11].

### 3) *Filtering and Stopword*

Pada tahap *filtering*, dilakukan proses penyaringan terhadap kalimat-kalimat dalam *tweet* dengan tujuan mengurangi kata-kata yang dianggap tidak relevan atau tidak memberikan pengaruh yang signifikan pada kalimat. Dalam proses ini, peneliti menggunakan teknik *filtering* dan penghapusan *stopword*. Penghapusan *stopword* atau biasa dikenal sebagai *stopword removal*, merujuk pada kata-kata umum seperti "ada" atau "ini" yang dihapus karena dinilai tidak memberikan kontribusi yang berarti dalam pemahaman kalimat. Tahap ini sangat penting untuk meningkatkan keakuratan dan kualitas analisis teks yang dilakukan [12].

### 4) *Tokenize*

Tokenisasi merupakan tahap yang bertujuan untuk membagi setiap kata dalam teks menjadi unit yang terpisah. Proses ini dilakukan dengan mengenali karakter spasi atau tanda baca sebagai pemisah antara setiap kata [13].

### 5) *Stemming*

Proses *stemming* digunakan untuk mengonversi kata menjadi bentuk dasar dengan menghapus infleksi dan afiks. Ini mempermudah analisis teks dengan membuat kata-kata yang berakar sama menjadi seragam, sehingga meningkatkan konsistensi dan akurasi [14].

### 6) *Labeling*

Labeling dalam analisis sentimen mengacu pada proses pengklasifikasian teks ke dalam kategori tertentu, seperti "positif" atau "negatif". Proses ini dapat dilakukan secara manual oleh manusia atau otomatis oleh mesin, yang menggunakan teknik pemrosesan bahasa alami untuk mendeteksi sentimen. [15].

## D. *Optimasi SMOTE*

Metode SMOTE (Synthetic Minority Oversampling Technique) merupakan pendekatan yang bertujuan untuk mengatasi ketidakseimbangan kelas dan meningkatkan akurasi algoritma. Fokus penelitian ini adalah menemukan parameter optimal dari Naïve Bayes dan Support Vector Machine, kemudian membandingkan kinerjanya untuk menentukan algoritma yang lebih unggul. [16].

## E. *Feature Extraction*

*Feature extraction* merupakan tahapan krusial dalam mengidentifikasi dan mengambil elemen signifikan dari sebuah dokumen. Salah satu algoritma yang sering diimplementasikan dalam konteks ini adalah *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF adalah metode statistik yang menghitung relevansi sebuah kata dalam sebuah dokumen di antara sekumpulan dokumen dengan menggabungkan dua komponen utama: *Term Frequency* (TF), yang mengukur frekuensi kemunculan kata tersebut dalam sebuah dokumen, dan *Inverse Document Frequency* (IDF), yang menilai seberapa jarang kata tersebut muncul di seluruh dokumen yang ada. Konsep dasar TF-IDF adalah bahwa kata-kata yang sering muncul dalam satu dokumen tertentu namun jarang ditemukan dalam dokumen lainnya memiliki bobot yang lebih besar, karena mereka dapat diidentifikasi sebagai fitur pembeda yang penting dalam proses klasifikasi. Oleh karena itu, TF-IDF memungkinkan untuk mengekstraksi kata-kata kunci yang memiliki peran signifikan dalam merepresentasikan isi dan karakteristik sebuah dokumen. [17].

$$TF(t, d) = \frac{\text{Jumlah kemunculan term } t \text{ dalam dokumen } d}{\text{Total jumlah term dalam dokumen } d} \quad (1)$$

Persamaan 1 ini merujuk pada metode penghitungan *Term Frequency* (TF) dalam analisis teks, yang mengukur seberapa sering sebuah kata muncul dalam sebuah dokumen. Dalam rumus ini,  $TF(t, d)$  adalah frekuensi sebuah kata ( $t$ ) dalam dokumen ( $d$ ), dihitung dengan membagi jumlah kemunculan kata ( $t$ ) dengan total jumlah kata dalam dokumen ( $d$ ). Pendekatan ini memungkinkan penilaian tentang seberapa pentingnya sebuah kata dalam konteks dokumen tersebut, dengan memperhitungkan proporsi kemunculannya dalam seluruh teks. Semakin tinggi nilai  $TF(t, d)$ , semakin signifikan peran kata ( $t$ ) dalam dokumen ( $d$ ).

## F. *Modeling*

Dalam analisis sentimen, tahapan *modeling* merupakan inti dari proses yang membutuhkan pemahaman mendalam terhadap konteks dan kompleksitas teks. Dengan merancang model yang mampu menangkap pola-pola dalam data, kita dapat menciptakan sistem yang lebih mampu dalam menginterpretasikan dan merespons variasi sentimen yang ada dalam teks. Oleh karena itu, fokus pada pengembangan model yang sensitif terhadap konteks dan subtilitas bahasa menjadi kunci untuk menghasilkan hasil analisis sentimen yang lebih akurat dan relevan [18]. Pada penelitian ini model yang digunakan yakni *Naïve Bayes* dan *Support Vector Machine*.

### 1) Naïve Bayes

*Naïve Bayes* merupakan metode klasifikasi yang berlandaskan pada Teorema Bayes yang diperkenalkan oleh Thomas Bayes, menggunakan prinsip probabilitas dan statistik untuk memperkirakan kemungkinan kejadian di masa mendatang dengan mengacu pada data historis. Metode ini dikenal dengan asumsi independensi yang kuat antara kondisi atau variabel yang digunakan dalam proses klasifikasi, yang dapat mempengaruhi performa model. [19]. Persamaan 2 merupakan rumus yang terkandung dalam Teorema Bayes :

$$P(A | B) = (P(B | A) \times P(A)) / (P(B)) \quad (2)$$

Rumus diatas menjelaskan bagaimana probabilitas terjadinya suatu peristiwa A dipertimbangkan setelah peristiwa B terjadi. Penilaian ini didasarkan pada probabilitas terjadinya B jika A sudah terjadi ( $P(B|A)$ ), probabilitas terjadinya A ( $P(A)$ ), dan probabilitas terjadinya B ( $P(B)$ ). Dalam konteks *Naïve Bayes*, rumus ini dimanfaatkan untuk mengukur probabilitas kelas atau label tertentu (A) berdasarkan fitur atau atribut yang diamati (B). Dengan mempergunakan rumus ini, kita dapat mengevaluasi probabilitas suatu peristiwa atau klasifikasi berdasarkan data yang ada. Rumus *Bayes* menjadi pondasi matematis penting bagi algoritma *Naïve Bayes* dan berperan dalam melakukan prediksi atau klasifikasi berdasarkan data yang diamati [20].

### 2) Support Vector Machine

*Support Vector Machine* (SVM) merupakan salah satu algoritma yang sering digunakan dalam klasifikasi dan masuk ke dalam kategori *supervised learning*. Cara kerja SVM ini adalah dengan mencari *hyperplane* atau bidang pemisah yang optimal untuk memisahkan dua kelas. *Hyperplane* ini adalah batas keputusan yang memaksimalkan margin, yaitu jarak antara *hyperplane* dengan titik terdekat dari setiap kelas, sehingga memberikan pemisahan yang paling baik antara dua kelas. SVM mencoba menemukan *hyperplane* yang memiliki margin terbesar, sehingga memungkinkan untuk mengklasifikasikan data baru dengan tepat berdasarkan posisinya terhadap *hyperplane* tersebut [21].

### G. Evaluasi

Performa kedua algoritma dievaluasi dengan mempertimbangkan metrik *accuracy*, *precision*, *recall*, dan *f1-score*. Penilaian klasifikasi bergantung pada pengujian objek yang diklasifikasikan dengan benar dan yang salah. Proses validasi menggunakan *confusion matrix* untuk menentukan model terbaik dengan informasi mengenai hasil klasifikasi aktual yang dapat diprediksi oleh sistem [22]. Pengukuran akurasi model menggunakan *confusion matrix* dapat di lihat pada Tabel 1.

TABEL I.  
RUMUS CONFUSION MATRIX

Aktual	Prediksi		Total
	TRUE	FALSE	
TRUE	TP	FN	P
FALSE	FP	TN	N
Total	P'	N'	P + N

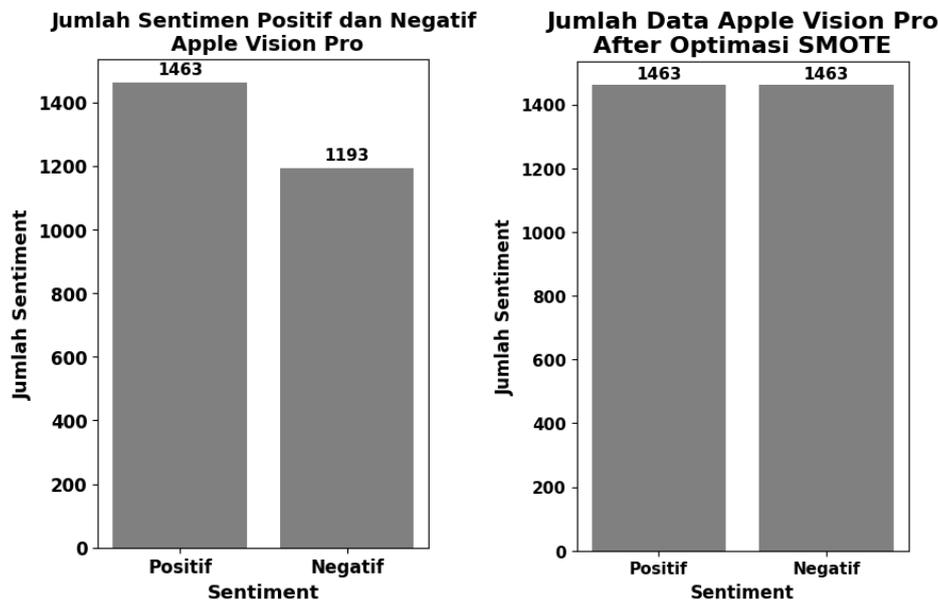
Dalam konteks evaluasi klasifikasi, TP (*True Positive*), TN (*True Negative*), FP (*False Positive*), dan FN (*False Negative*) memiliki arti penting yang mencerminkan keberhasilan dan kegagalan dari suatu algoritma klasifikasi. TP terjadi ketika data yang memang benar-benar positif berhasil diklasifikasikan dengan benar oleh algoritma, sedangkan TN terjadi saat data yang memang benar-benar negatif juga diklasifikasikan dengan benar. Di sisi lain, FP terjadi ketika data yang seharusnya negatif malah salah diklasifikasikan sebagai positif oleh algoritma, sedangkan FN terjadi ketika data yang seharusnya positif justru salah diklasifikasikan sebagai negatif. Setiap kategori ini memberikan *insight* tentang kinerja algoritma klasifikasi, memungkinkan analisis yang lebih mendalam tentang seberapa baik algoritma mampu membedakan antara kelas yang berbeda [23].

## III. HASIL DAN PEMBAHASAN

### A. Dataset

Dataset yang berhasil di dapatkan melalui proses *crawling* menunjukkan bahwa terdapat 1463 sentimen positif dan 1193 sentimen negatif. Karena tidak seimbang, dilakukan optimasi menggunakan metode SMOTE (*Synthetic Minority Over-sampling Technique*) untuk menciptakan keseimbangan antara kelas positif dan negatif. Setelah dilakukan optimasi, jumlah sentimen positif dan negatif menjadi sama, yaitu 1463 untuk masing-masing kelas.

SMOTE melakukan pembuatan sampel sintetis dari kelas minoritas (negatif dalam kasus ini) untuk menyeimbangkan dataset, sehingga meningkatkan kinerja model dalam mengklasifikasikan kedua kelas. Dengan jumlah yang seimbang, model dapat belajar dengan lebih baik dari kedua kelas dan menghasilkan hasil yang lebih baik dalam analisis sentimen.



Gambar 2 Jumlah sentiment sebelum dan sesudah SMOTE

### B. Preprocessing

Berdasarkan hasil *preprocessing* pada Tabel 2 menampilkan teks awal "Best Buy meluncurkan aplikasi Apple Vision Pro untuk pratinjau produk teknologi" telah disiapkan untuk analisis lebih lanjut. Tahapan pertama adalah pembersihan, di mana teks awal telah dihilangkan karakter yang tidak relevan, seperti tautan URL, sehingga sisa teks adalah "Best Buy aplikasi Apple Vision Pro untuk pratinjau produk teknologi". Selanjutnya, teks tersebut diubah menjadi huruf kecil pada tahap pengubahan huruf besar ke kecil, memastikan konsistensi dalam analisis. Kemudian, teks dibagi menjadi kata-kata individual pada tahap *tokenize*, menghasilkan daftar kata yang lebih terstruktur seperti ['best', 'buy', 'aplikasi', 'apple', 'vision', 'pro', 'untuk', 'pratinjau', 'produk', 'teknologi']. Pada tahap *filtering*, kata-kata yang tidak penting atau umum telah dihapus, meninggalkan daftar kata yang lebih relevan: ['best', 'buy', 'aplikasi', 'apple', 'vision', 'pro', 'pratinjau', 'produk', 'teknologi']. Akhirnya, kata-kata yang tersisa dipangkas ke bentuk dasarnya melalui *stemming*, mengurangi variasi kata yang memiliki akar yang sama menjadi "best buy aplikasi apple vision pro pratinjau produk teknologi". Selanjutnya pada Tabel 3 menampilkan hasil *labeling*. Dengan langkah-langkah ini, teks telah disiapkan untuk analisis lebih lanjut dan relevansi dalam representasi teks yang digunakan dalam model analisis.

TABEL II.  
HASIL PREPROCESSING

Tahapan	Hasil <i>Preprocessing</i>
<i>Dataset</i>	Best Buy meluncurkan aplikasi Apple Vision Pro untuk pratinjau produk teknologi <a href="https://t.co/q6x7l26gprpr">https://t.co/q6x7l26gprpr</a>
<i>Cleansing</i>	Best Buy aplikasi Apple Vision Pro untuk pratinjau produk teknologi
<i>Case Folding</i>	best buy aplikasi apple vision pro untuk pratinjau produk teknologi
<i>Tokenize</i>	['best', 'buy', 'aplikasi', 'apple', 'vision', 'pro', 'untuk', 'pratinjau', 'produk', 'teknologi']
<i>Filtering</i>	['best', 'buy', 'aplikasi', 'apple', 'vision', 'pro', 'pratinjau', 'produk', 'teknologi']
<i>Stemming</i>	best buy aplikasi apple vision pro pratinjau produk teknologi

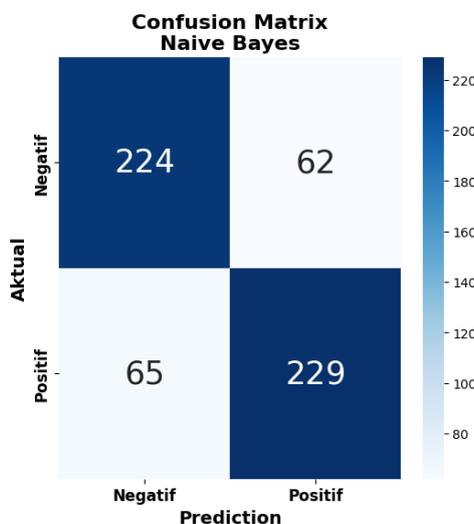
TABEL III.  
HASIL LABELING

<i>Tweets</i>	Hasil <i>Labeling</i>
best buy aplikasi apple vision pro pratinjau produk teknologi	Positif
Pengguna Apple Vision Pro mendapatkan sakit kepala penyakit terminal kanker	Negatif

### C. Tahap pengujian

Setelah tahap *preprocessing* data, yang meliputi normalisasi fitur untuk memastikan skala yang seragam dan mencegah dominasi fitur dengan skala besar, tahap pengujian dilaksanakan. Normalisasi data fitur sebelum pemodelan adalah langkah penting untuk meningkatkan kinerja model, karena memastikan bahwa semua fitur berkontribusi secara setara dalam proses pembelajaran. Pada tahap ini, teknik SMOTE (*Synthetic Minority Over-sampling Technique*) digunakan untuk menghasilkan data sintetik yang lebih seimbang dan representatif. Selanjutnya, evaluasi dilakukan dengan menguji dua algoritma klasifikasi, *Naïve Bayes* dan *Support Vector Machine* (SVM). Kinerja kedua model diukur menggunakan metrik evaluasi seperti confusion matrix, akurasi, presisi, recall, dan F1-score, yang memungkinkan penilaian menyeluruh mengenai efektivitas masing-masing model dalam mengklasifikasikan data dengan akurat dan efektif.

#### 1) *Naïve Bayes*



Gambar 3. Hasil *Confusion Matrix* SVM

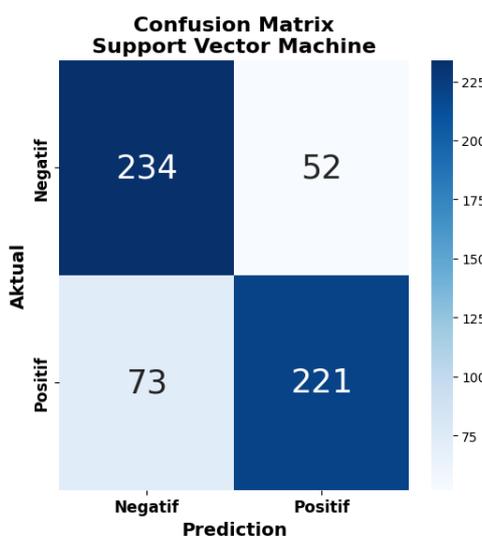
Dalam tahap evaluasi kinerja model klasifikasi *Naive Bayes* pada Gambar 3 di atas, perhatian utama tertuju pada *confusion matrix* yang mencerminkan distribusi prediksi yang benar dan yang salah. Dalam kasus ini, *confusion matrix* akan berisi empat elemen utama: *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN). *True Positive* (TP) sebesar 229 mengindikasikan jumlah data positif yang diprediksi dengan benar oleh model, sementara *False Positive* (FP) sebanyak 65 menunjukkan jumlah data negatif yang salah diprediksi sebagai positif. Sebaliknya, *False Negative* (FN) menunjukkan jumlah data positif yang salah diprediksi sebagai negatif sebanyak 62, dan *True Negative* (TN) sebesar 224 mencerminkan jumlah data negatif yang diprediksi dengan benar oleh model. Dengan memperhatikan setiap elemen dalam *confusion matrix*, kita dapat menganalisis secara lebih mendalam kinerja model *Naive Bayes*, mengidentifikasi area di mana model mungkin perlu ditingkatkan, serta memahami bagaimana model mengelompokkan data yang diberikan dalam konteks klasifikasi yang bersangkutan.

TABEL III.  
 HASIL CLASSIFICATION REPORT NAÏVE BAYES

<i>Naïve Bayes using SMOTE</i>				
	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
Positif	78%	79 %	78%	78%
Negatif		78 %	78%	78%

Dari hasil *classification report* pada Tabel 3, Setelah penerapan *Synthetic Minority Over-sampling Technique* (SMOTE), model *Naïve Bayes* menunjukkan performa klasifikasi yang seimbang dan konsisten dengan *accuracy* sebesar 78%. Analisis metrik evaluasi menunjukkan bahwa *precision* untuk kelas positif tercatat sebesar 79%, sementara kelas negatif mencapai 78%. *Recall* untuk kedua kelas menunjukkan nilai yang identik, yaitu 78%, mencerminkan bahwa model memiliki kemampuan yang setara dalam mendeteksi kedua kelas tanpa adanya bias signifikan. *F1-score* untuk kedua kelas juga menunjukkan kesamaan, masing-masing sebesar 78%, yang menegaskan bahwa model ini mampu menangani trade-off antara *precision* dan *recall* dengan baik. Hasil ini menunjukkan bahwa *Naïve Bayes*, setelah optimasi menggunakan SMOTE, berhasil mencapai performa yang stabil dan konsisten dalam klasifikasi, menjadikannya sebagai model yang andal untuk tugas-tugas klasifikasi pada data dengan ketidakseimbangan kelas.

## 2) Support Vector Machine (SVM)



Gambar 4. Hasil *Confusion Matrix* SVM

Setelah mengalami proses optimasi melalui SMOTE, model klasifikasi sentimen menggunakan *Support Vector Machine* telah menunjukkan peningkatan yang signifikan. Terlihat dari hasil yang dicapai, dengan jumlah *True Positive* (TP) sebanyak 221, model mampu mengenali ulasan positif secara akurat. Namun, masih ada ruang untuk perbaikan yang perlu diperhatikan, terutama dalam mengurangi jumlah *False Negative* (FN) sebanyak 52, menandakan adanya kecenderungan model untuk salah mengklasifikasikan ulasan positif sebagai negatif. Meski demikian, hasil *False Positive* (FP) sebanyak 73 dan *True Negative* (TN) sebesar 221 mengindikasikan bahwa model jarang membuat kesalahan dalam memprediksi ulasan negatif sebagai positif, walaupun masih ada beberapa kasus yang terjadi. Dengan demikian, analisis *confusion matrix* ini memberikan gambaran yang lebih terperinci tentang kinerja model setelah dilakukan proses optimasi, sekaligus menyoroti area-area yang memerlukan perbaikan guna meningkatkan akurasi dalam prediksi sentimen. Hasil *confusion matrix* dapat dilihat pada Gambar 4 berikut ini.

TABEL IV.  
 HASIL CLASSIFICATION REPORT SVM

<i>Support Vector Machine using SMOTE</i>				
	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
Positif	78%	81 %	75%	78%
Negatif		76 %	82%	79%





*WordCloud* analisis sentimen negatif, terlihat bahwa kata yang paling dominan adalah "Apple", diikuti oleh "vision" dan "pro". Ini menandakan bahwa ketika pengguna mengekspresikan sentimen negatif terhadap produk tersebut, fokus utama mereka tetap pada merek dan produk itu sendiri. Selanjutnya, terdapat kata-kata kunci yang mencolok, seperti "beli" dan "sakit kepala". Keberadaan kata "beli" menunjukkan ketidakpuasan terhadap pengalaman pembelian, mungkin terkait dengan harga atau proses transaksi. Sementara itu, kata "sakit kepala" dapat mengindikasikan bahwa pengguna mengalami masalah atau ketidaknyamanan saat menggunakan produk tersebut, entah itu terkait dengan kinerja, keandalan, atau aspek lainnya. Secara keseluruhan, *WordCloud* ini memberikan gambaran yang jelas tentang aspek-aspek yang menjadi perhatian utama dalam ulasan negatif terkait Apple Vision Pro, yang meliputi pengalaman pembelian dan masalah penggunaan yang mungkin menyebabkan ketidakpuasan atau ketidaknyamanan. Hasil *WordCloud* dapat dilihat pada Gambar 8.

Analisis sentimen tingkat kata berdasarkan *WordCloud* pada Gambar, 6, 7, dan 8) menunjukkan bahwa kata – kata seperti “teknologi”, “fitur”, “nyata”, “keren”, dan “bagus” mendominasi sentimen positif, mencerminkan apresiasi terhadap inovasi dan kualitas fitur Apple Vision Pro. Pengguna umumnya memuji aspek teknis dan pengalaman yang dihadirkan produk ini. Di sisi lain, sentimen negatif didominasi oleh kata – kata seperti “beli” dan “sakit kepala”, yang menunjukkan ketidakpuasan terkait harga tinggi dan pengalaman penggunaan yang kurang nyaman, seperti masalah teknis atau fisik saat menggunakan produk. Pola ini mengindikasikan bahwa meskipun inovasi dan desain Vision Pro dihargai, perhatian serius perlu diberikan pada aspek harga dan kemudahan penggunaan untuk mengatasi sumber utama ketidakpuasan pengguna. Temuan ini penting untuk panduan pengembangan produk dan strategi pemasaran yang lebih efektif.

Dengan demikian, penelitian ini tidak hanya mendukung temuan-temuan sebelumnya tetapi juga memberikan kontribusi baru dengan mengevaluasi secara komprehensif keefektifan algoritma *Naïve Bayes* dan SVM dalam konteks produk teknologi inovatif, serta mengidentifikasi pola-pola sentimen spesifik yang berkaitan dengan aspek harga, inovasi, dan kenyamanan penggunaan. Penelitian ini memberikan konteks lebih lanjut mengenai bagaimana persepsi pengguna terhadap produk teknologi baru dapat dianalisis dan dimengerti, sekaligus menyoroti area yang memerlukan perhatian dalam pengembangan dan pemasaran produk teknologi masa depan.

#### IV. KESIMPULAN

Berdasarkan hasil penelitian menunjukkan bahwa setelah penerapan *Synthetic Minority Over-sampling Technique* (SMOTE), model *Naïve Bayes* mencapai performa klasifikasi yang lebih konsisten dengan *accuracy* sebesar 78%, *precision* 79% untuk kelas positif dan 78% untuk kelas negatif, serta *recall* yang identik sebesar 78% untuk kedua kelas. *F1-score* yang juga konsisten pada 78% menunjukkan stabilitas model dalam menangani *trade-off* antara *precision* dan *recall*. Sementara itu, model *Support Vector Machine* (SVM) juga mencapai *accuracy* 78%, namun menunjukkan variasi yang lebih besar antara *precision* dan *recall* dengan *precision* 81% untuk kelas positif dan 76% untuk kelas negatif, serta *recall* 75% untuk kelas positif dan 82% untuk kelas negatif. Meskipun SVM menunjukkan peningkatan performa yang signifikan setelah optimasi dengan SMOTE, model ini masih memerlukan penyempurnaan dalam mengurangi kesalahan klasifikasi pada ulasan positif. Dengan demikian, *Naïve Bayes* menunjukkan keunggulan dalam hal stabilitas dan konsistensi, menjadikannya pilihan yang lebih tepat untuk klasifikasi data dengan ketidakseimbangan kelas.

#### DAFTAR PUSTAKA

- [1] P. Produk And D. I. Korea, “No Title,” Vol. 2, No. 2, Pp. 1–15.
- [2] E. Fran, “Apple Vision Pro : Comments In Healthcare,” No. June 2023.
- [3] Normalisa, A. Rachmaniar, D. Diana, M. Saefudin, And R. Parulian, “Application Of Computer Vision Detection Of Apples And Oranges Using Python Language,” *J. Inf. Syst. Informatics Comput.*, Vol. 6, No. 2, Pp. 455–466, 2022, Doi: 10.52362/Jisicom.V6i2.946.
- [4] S. Lestari And S. Berliani, “Analisis Sentimen Masyarakat Terhadap Isu Pecat Sri Mulyani Pada Twitter Menggunakan Metode Naive Bayes Dan Support Vector Machine,” Vol. 5, No. 3, Pp. 951–960, 2024.
- [5] W. Francis, “Analisis Sentimen Masyarakat Terhadap Isu Penundaan Pemilu 2024 Pada Twitter Dengan Metode Naive Bayes Dan Support Vector Machine,” *J. Sains Dan Teknol.*, Vol. 5, No. 3, Pp. 890–899, 2023, [Online]. Available: [Http://Ejournal.Sisfokomtek.Org/Index.Php/Saintek/Article/View/1359](http://ejournal.sisfokomtek.org/index.php/saintek/article/view/1359)
- [6] D. N. Novianti, D. F. Shiddieq, F. F. Roji, And W. Susilawati, “Komparasi Algoritma Support Vector Machine Dan Naïve Bayes Untuk Analisis Sentimen Pada Metaverse,” *Malcom Indones. J. Mach. Learn. Comput. Sci.*, Vol. 4, No. 1, Pp. 231–239, 2024, Doi: 10.57152/Malcom.V4i1.1061.
- [7] J. Homepage, W. Ningsih, B. Alfianda, And D. Wulandari, “Malcom: Indonesian Journal Of Machine Learning And Computer Science Comparison Of Naive Bayes And Svm Algorithms In Twitter Sentiment Analysis On Electric Car Use In Indonesia Perbandingan Algoritma Svm Dan Naïve

- Bayes Dalam Analisis Sentimen Twitter Pada,” Vol. 4, No. 2, Pp. 556–562, 2024.
- [8] F. Matheos Sarimole And K. Kudrat, “Analisis Sentimen Terhadap Aplikasi Satu Sehat Pada Twitter Menggunakan Algoritma Naive Bayes Dan Support Vector Machine,” *J. Sains Dan Teknol.*, Vol. 5, No. 3, Pp. 783–790, 2024, Doi: 10.55338/Saintek.V5i3.2702.
- [9] A. Nurian And T. N. Padilah, “Disdukcapil Karawang Menggunakan Naive,” Vol. 12, No. 2, 2024.
- [10] A. R. Raharja, A. Pramudianto, And Y. Muchsam, “Penerapan Algoritma Decision Tree Dalam Klasifikasi Data ‘ Framingham ’ Untuk Menunjukkan Risiko Seseorang Terkena Penyakit Jantung Dalam 10 Tahun Mendatang”.
- [11] D. Darwis, N. Siskawati, And Z. Abidin, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional,” *J. Tekno Kompak*, Vol. 15, No. 1, P. 131, 2021, Doi: 10.33365/Jtk.V15i1.744.
- [12] D. Oktavia, Y. R. Ramadahan, And M. Minarto, “Analisis Sentimen Terhadap Penerapan Sistem E-Tilang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm),” *Klik Kaji. Ilm. Inform. Dan Komput.*, Vol. 4, No. 1, Pp. 407–417, 2023, Doi: 10.30865/Klik.V4i1.1040.
- [13] T. Rivanie, R. Pebrianto, T. Hidayat, A. Bayhaq, W. Gata, And H. B. Novitasari, “Analisis Sentimen Terhadap Kinerja Menteri Kesehatan Indonesia Selama Pandemi Covid-19,” *J. Inform.*, Vol. 21, No. 1, Pp. 1–13, 2021, Doi: 10.30873/Ji.V21i1.2864.
- [14] Y. Pratama, D. T. Murdiansyah, And K. M. Lhaksana, “Analisis Sentimen Kendaraan Listrik Pada Media Sosial Twitter Menggunakan Algoritma Logistic Regression Dan Principal Component Analysis,” *J. Media Inform. Budidarma*, Vol. 7, No. 1, Pp. 529–535, 2023, Doi: 10.30865/Mib.V7i1.5575.
- [15] P. K. Sari And R. R. Suryono, “Komparasi Algoritma Support Vector Machine Dan Random Forest Untuk Analisis Sentimen Metaverse,” *J. Mnemon.*, Vol. 7, No. 1, Pp. 31–39, 2024, Doi: 10.36040/Mnemonic.V7i1.8977.
- [16] R. Nurul Ikhsani And F. Fauzi Abdulloh, “Optimasi Svm Dan Decision Tree Menggunakan Smote Untuk Mengklasifikasi Sentimen Masyarakat Mengenai Pinjaman Online,” *J. Media Inform. Budidarma*, Vol. 7, No. 4, Pp. 1667–1677, 2023, Doi: 10.30865/Mib.V7i4.6809.
- [17] P. Widyantara *Et Al.*, “Analisis Sentimen Pada Teks Berbahasa Bali Menggunakan Metode Multinomial Naive Bayes Dengan Tf-Idf Dan Bow,” *Jnatia*, Vol. 2, No. 1, Pp. 37–46, 2023.
- [18] Muhammad Daffa Al Fahreza, Ardytha Luthfiarta, Muhammad Rafid, And Michael Indrawan, “Analisis Sentimen: Pengaruh Jam Kerja Terhadap Kesehatan Mental Generasi Z,” *J. Appl. Comput. Sci. Technol.*, Vol. 5, No. 1, Pp. 16–25, 2024, Doi: 10.52158/Jacost.V5i1.715.
- [19] Rayuwati, Husna Gemasih, And Irma Nizar, “Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid,” *Jural Ris. Rumpun Ilmu Tek.*, Vol. 1, No. 1, Pp. 38–46, 2022, Doi: 10.55606/Jurritek.V1i1.127.
- [20] R. Alfarez And V. Purwayoga, “Penerapan Naive Bayes Untuk Prediksi Customer Churn ( Studi Kasus : Pt Hutchison 3 Indonesia ),” Vol. 05, No. 02, Pp. 301–307, 2024.
- [21] H. C. Husada And A. S. Paramita, “Analisis Sentimen Pada Maskapai Penerbangan Di Platform Twitter Menggunakan Algoritma Support Vector Machine (Svm),” *Teknika*, Vol. 10, No. 1, Pp. 18–26, 2021, Doi: 10.34148/Teknika.V10i1.311.
- [22] Y. A. Singgalen, “Analisis Sentimen Wisatawan Terhadap Kualitas Layanan Hotel Dan Resort Di Lombok Menggunakan Servqual Dan Crisp-Dm,” *Build. Informatics, Technol. Sci.*, Vol. 4, No. 4, Pp. 1870–1882, 2023, Doi: 10.47065/Bits.V4i4.3199.
- [23] A. S. Rahayu, A. Fauzi, And R. Rahmat, “Komparasi Algoritma Naive Bayes Dan Support Vector Machine (Svm) Pada Analisis Sentimen Spotify,” *J. Sist. Komput. Dan Inform.*, Vol. 4, No. 2, P. 349, 2022, Doi: 10.30865/Json.V4i2.5398.