

## PERBANDINGAN PERFORMA ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN IMPLEMENTASI PROGRAM KIP-KULIAH

**Ni Putu Ana Rainita<sup>1</sup>, I Md Dendi Maysanjaya<sup>2</sup>, Gede Surya Mahendra<sup>3</sup>**

<sup>1,2,3</sup> Fakultas Teknik dan Kejuruan Universitas Pendidikan Ganesha

E-mail: [ana.rainita@undiksha.ac.id](mailto:ana.rainita@undiksha.ac.id) , [dendi.ms@undiksha.ac.id](mailto:dendi.ms@undiksha.ac.id) , [gmahendra@undiksha.ac.id](mailto:gmahendra@undiksha.ac.id)

---

### Informasi Naskah:

Diterima:  
07 November 2025

Diterbitkan:  
24 Desember 2025

**Abstract:** The selection of a suitable classification model is important in text-based sentiment analysis, especially in conditions of unbalanced data distribution. Naive Bayes and Support Vector Machine (SVM) are two algorithms that are often used in classification, but the comparison of their performance on unbalanced data still needs to be further reviewed. This study aims to compare the performance of the two algorithms in classifying public sentiment towards the Indonesia Smart Card (KIP) Lecture Program. The implementation of the KIP Lecture Program still faces challenges in the accuracy of aid distribution. This situation raises discussions and various controversies among the public, especially on the X platform. The data used were 1,644 tweets, with a distribution of negative sentiment of 1,392 tweets and positive tweets of 252. To overcome the imbalance of data class distribution, the Synthetic Minority Oversampling Technique (SMOTE) method is used. Based on the evaluation results, before SMOTE was applied, SVM obtained 92% accuracy and 91% precision, 77% recall, while Naive Bayes obtained 79% accuracy, 68% precision, and 78% recall. After the application of SMOTE, SVM performance significantly improved with accuracy, precision, and recall reaching 99%, while Naive Bayes improved to 95% on all metrics. These results show that although SVM excels in higher accuracy, Naive Bayes shows a more stable performance on the data neither after nor after the balancing process is performed.

**Keyword:** KIP Kuliah, Sentiment Analysis, Naive Bayes, Support Vector Machine, SMOTE

**Abstrak:** Pemilihan model klasifikasi yang sesuai menjadi hal yang penting dalam analisis sentimen berbasis teks, terutama pada kondisi distribusi data yang tidak seimbang. Naive Bayes dan Support Vector Machine (SVM) merupakan dua algoritma yang kerap digunakan dalam klasifikasi, namun perbandingan performa keduanya pada data tidak seimbang masih perlu ditinjau lebih lanjut. Penelitian ini bertujuan untuk membandingkan performa kedua algoritma dalam mengklasifikasikan sentimen masyarakat terhadap Program Kartu Indonesia Pintar (KIP) Kuliah. Implementasi Program KIP Kuliah masih menghadapi tantangan dalam ketepatan penyaluran bantuan. Situasi tersebut memunculkan diskusi dan berbagai kontroversi di kalangan masyarakat khususnya pada platform X. Data yang digunakan berjumlah 1.644 tweet dengan distribusi sentimen negatif 1.392 tweet dan tweet positif sebanyak 252. Untuk mengatasi ketidakseimbangan distribusi kelas data, digunakan metode Synthetic Minority Oversampling Technique (SMOTE). Berdasarkan hasil evaluasi, sebelum SMOTE diterapkan, SVM memperoleh akurasi sebesar 92% dan presisi 91%, recall 77%, sedangkan Naive Bayes memperoleh akurasi 79%, presisi 68%, dan recall 78%. Setelah SMOTE diterapkan, SVM memperoleh akurasi, presisi, dan recall mencapai 99%, sementara Naive Bayes mencapai 95% pada semua metrik. Hasil ini

menunjukkan bahwa meskipun SVM unggul secara akurasi, Naive Bayes menunjukkan performa yang lebih stabil pada data yang tidak maupun setelah proses penyeimbangan dilakukan.

**Kata Kunci:** KIP Kuliah, Analisis Sentimen, Naive Bayes, Support Vector Machine, SMOTE

## PENDAHULUAN

Kartu Indonesia Pintar (KIP) Kuliah merupakan pengembangan dari program Bidikmisi yang telah dilaksanakan sejak 2010, sebagai bagian dari Program Indonesia Pintar (PIP) di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. Pada tahun 2021, program ini menjamin biaya pendidikan dan bantuan hidup lebih dari 150.000 mahasiswa di seluruh PTN dan PTS melalui berbagai jalur masuk (Negara, 2024) KIP Kuliah bertujuan mendorong mobilitas sosial dan peningkatan potensi ekonomi mahasiswa dari keluarga kurang mampu.

Meski memberikan dampak positif (Ekaptiningrum, 2024), implementasi program ini masih menghadapi masalah dalam ketepatan sasaran penerima. Beberapa kasus mencuat di media sosial, khususnya platform X, memperlihatkan penerima yang tidak mencerminkan kondisi ekonomi yang lemah (Puspapertiwi & Nugroho, 2024) Laporan Badan Pemeriksa Keuangan (BPK) juga menyatakan bahwa dana PIP senilai Rp2,86 triliun yang disalurkan kepada 5.364.986 siswa tidak tepat sasaran (Yuliantri P, 2021). Salah satu contoh kasus terjadi di sebuah universitas di Semarang, di mana bantuan KIP diberikan kepada mahasiswa yang memamerkan gaya hidup mewah di media sosial (Adi Negara, 2024)

Permasalahan tersebut menunjukkan perlunya evaluasi melalui pemahaman persepsi masyarakat. Salah satu pendekatan yang digunakan adalah analisis sentimen berbasis *text mining*. Algoritma yang umum digunakan meliputi Naive Bayes, Support Vector Machine (SVM), Decision Tree, dan K-Nearest Neighbor (KNN). Penelitian Pramudita dkk. (2024) menggunakan Naive Bayes dan memperoleh akurasi 84,99%, namun merekomendasikan penggunaan algoritma lain dan jumlah data yang lebih

besar untuk hasil yang lebih komprehensif. Sementara itu, Amelia & Sarimole (2024) menggunakan SVM dan mencapai akurasi 86,27%, tetapi mengalami bias pada kelas positif akibat ketidakseimbangan data.

Ketidakseimbangan data merupakan kendala umum dalam analisis sentimen, karena model cenderung bias terhadap kelas mayoritas (Nugroho & Rilvani, 2023). Sebelum melakukan penyeimbangan data akan dianalisis terlebih dahulu distribusi masing-masing kelas. Pada penelitian ini menerapkan teknik SMOTE. Teknik SMOTE dipilih karena teknik ini mampu menghasilkan sampel sintetis baru berdasarkan kedekatan antar data minoritas. Teknik ini juga membantu meningkatkan representasi kelas minoritas tanpa memperbesar risiko *overfitting* seperti pada teknik Random Oversampling, serta lebih stabil dibandingkan metode ADASYN. SMOTE juga sesuai untuk data berdimensi tinggi seperti TF-IDF, sehingga membantu model mengidentifikasi pola sentimen secara lebih adil.

Penelitian oleh Iskandar & Nataliani (2021) membandingkan algoritma Naive Bayes, SVM, dan KNN dalam analisis sentimen komentar YouTube dan menemukan bahwa SVM unggul dalam akurasi (96,43%). SVM dikenal mampu menangani data berdimensi tinggi dengan baik dan lebih tahan terhadap *overfitting*, sementara Naive Bayes unggul dalam efisiensi dan kecepatan (Ilmawan & Mude, 2020).

Melihat hal tersebut, penting untuk membandingkan algoritma Naive Bayes dan SVM karena keduanya memiliki pendekatan yang berbeda dalam menangani ketidakseimbangan data. Naive Bayes unggul dalam efisiensi dan kemudahan implementasi. Metode ini bekerja dengan menghitung probabilitas

kemunculan kata. Di sisi lain, SVM lebih unggul dalam menangani data yang kompleks dan berdimensi tinggi.

Kedua metode ini dapat memberikan hasil yang lebih optimal terutama jika dikombinasikan dengan teknik resampling SMOTE, yang berfungsi menyeimbangkan distribusi data dengan menambahkan data sintetis pada kelas minoritas. Karena data opini di media sosial yang terkadang subjektif, tidak seimbang, dan dinamis, pemilihan metode yang tepat menjadi sangat penting. Oleh karena itu, membandingkan performa kedua algoritma ini menjadi langkah penting untuk mengetahui metode mana yang menghasilkan performa yang efektif dalam analisis sentimen yang akurat, adil, dan dapat mendukung evaluasi kebijakan seperti KIP Kuliah.

Berdasarkan latar belakang yang telah dipaparkan sebelumnya maka dirancang sebuah penelitian yang berjudul, Komparasi algoritma Naive Bayes dan SVM dalam analisis sentimen terhadap implementasi program KIP Kuliah.

Mengingat bahwa setiap algoritma memiliki keunggulan dan karakteristik yang berbeda, maka perbandingan antara Naive Bayes dan SVM menjadi langkah yang krusial. Pemilihan metode klasifikasi yang tepat sangat menentukan tingkat akurasi, kemampuan model dalam melakukan generalisasi, serta efektivitasnya dalam mengelola ketidakseimbangan data yang kerap ditemukan dalam opini masyarakat di media sosial.

Penelitian ini tidak hanya berfokus pada perbandingan kinerja kedua algoritma dalam analisis sentimen terkait program KIP Kuliah, tetapi juga bertujuan menghasilkan rekomendasi algoritma yang paling optimal untuk digunakan dalam evaluasi kebijakan berbasis opini publik. Dengan demikian, hasil penelitian ini dapat memberikan kontribusi sebagai dasar pertimbangan bagi pengambil kebijakan dalam meningkatkan ketepatan dan keadilan implementasi program di masa mendatang.

## TINJUAN PUSTAKA

Beberapa penelitian sebelumnya menunjukkan bahwa algoritma *Support Vector Machine* dan *Naive Bayes* banyak digunakan dalam analisis sentimen. Penelitian oleh Hakim dkk. (2021) pada ulasan myIndiHome menunjukkan bahwa SVM dengan linear kernel memiliki akurasi lebih tinggi mencapai 86,54% dibandingkan *Naive Bayes* sebesar 84,69%. Hasil serupa juga diperoleh pada penelitian oleh Hashfi dkk. (2022) pada analisis sentimen pelanggan Indihome melalui Twitter bahwa metode SVM memiliki performa lebih baik dengan akurasi 84%, dibandingkan *Naive Bayes* yang memperoleh akurasi 82%.

Selanjutnya penelitian Khaira dkk. (2023) menganalisis opini masyarakat terhadap bantuan kuota internet dan menunjukkan bahwa SVM lebih unggul dengan akurasi 80% dibandingkan *Naive Bayes* sebesar 64%. Penelitian oleh Pramudita dkk. (2024) menggunakan algoritma *Naive Bayes* dan pendekatan CRISP-DM untuk menganalisis sentimen terhadap Program KIP-K, dengan hasil akurasi 84,99%. Sementara itu, Amelia & Sarimole (2024) menggunakan SVM dalam analisis sentimen KIP Kuliah dan memperoleh akurasi 86,27%, meskipun ditemukan bias terhadap sentimen negatif.

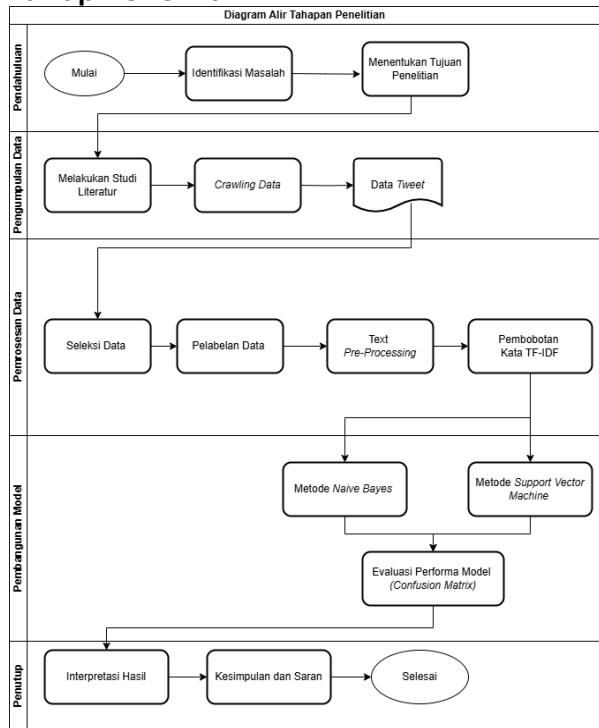
Penelitian ini berbeda dari studi sebelumnya karena memfokuskan pada sentimen terhadap program KIP Kuliah di media sosial X, dengan periode pengambilan data dari Februari 2020-September 2024 dengan menggunakan kata kunci seperti "KIP-K", "KIP Kuliah", serta tagar #KIPK dan #KIPKuliah. Penelitian ini membandingkan performa SVM dan *Naive Bayes* menggunakan fitur TF-IDF, serta mengevaluasi dampak ketidakseimbangan data melalui penerapan teknik *resampling*.

Kontribusi penelitian ini adalah memberikan wawasan mendalam mengenai kinerja dua algoritma klasifikasi yang berbeda saat diterapkan pada dataset tweet terkait Program Kartu Indonesia Pintar Kuliah. Hasil penelitian ini dapat digunakan sebagai panduan untuk

mengembangkan model analisis sentimen yang lebih efektif.

## METODOLOGI PENELITIAN

### Tahap Penelitian



**Gambar 1.** Tahapan Penelitian

Berikut merupakan penjelasan tahapan penelitian pada Gambar 1.

#### a. Pendahuluan

Tahapan pendahuluan merupakan tahapan awalan dari penelitian ini yang di dalamnya bertujuan untuk mengidentifikasi masalah yang akan diteliti, sekaligus juga mendefinisikan tujuan dilaksanakannya penelitian.

#### b. Pengumpulan Data

Tahapan yang kedua adalah pengumpulan data. Pada tahapan ini bertujuan untuk mendapatkan informasi dan data yang bisa membantu mendukung penelitian. Pada penelitian ini ada dua buah cara yang digunakan dalam memperoleh data dan informasi, yang pertama melakukan studi literatur terhadap berbagai sumber, seperti artikel, jurnal, dan lainnya yang bertujuan untuk menambah informasi mengenai teori penelitian yang mendukung dan berkaitan. Tahapan yang kedua adalah melakukan *crawling* data pada sosial

media X untuk memperoleh dataset tweet dengan topik KIP Kuliah berdasarkan beberapa kata kunci tertentu. Metode pengumpulan data melalui media sosial X dengan bantuan tool yaitu *Tweet Harvest*. Data yang diambil berfokus pada topik KIP Kuliah dengan menggunakan kata kunci seperti "KIP-K", "KIP Kuliah", serta hashtag #KIP K dan #KIPKuliah, dalam rentang dari 21 Februari 2020 hingga 30 September 2024.

#### c. Pemrosesan Data

Tahapan ketiga merupakan proses pemrosesan data yang bertujuan untuk mempersiapkan data yang telah dikumpulkan agar menjadi lebih terstruktur. Data yang semula tidak terstruktur akan diolah sehingga lebih mudah dan optimal untuk dianalisis. Tahapan ini dimulai dengan menyeleksi data yang telah diperoleh. Proses seleksi tweet diawali dengan membersihkan data yang duplikat, dan menghapus tweet yang tidak berbahasa Indonesia untuk menjaga konsistensi analisis teks. Kemudian menghapus data tweet yang berasal dari sumber non-individu seperti akun pemerintahan, portal berita, akun institusi, atau akun promosi/iklan. Data tweet yang sudah terseleksi kemudian dilanjutkan ke proses pelabelan data. Pada tahap ini, data akan diklasifikasikan ke dalam dua kategori sentimen, yaitu positif dan negatif. Pelabelan ini dilakukan secara manual oleh dua orang pakar di bidang bahasa Indonesia guna memastikan akurasi klasifikasi.

Setelah pelabelan selesai, dilanjutkan ke tahap *preprocessing* yang terdiri dari beberapa langkah, yaitu *case folding* untuk mengubah semua huruf menjadi huruf kecil, *cleaning* data untuk menyamakan bentuk kata dan menghapus karakter tidak penting, *stopword removal* untuk menghapus kata-kata umum yang tidak memiliki makna penting, *tokenizing* untuk memecah teks menjadi kata-kata atau token, normalisasi untuk mengubah kata tidak baku atau slang menjadi bentuk baku, serta *stemming* untuk mengubah kata ke

bentuk dasarnya. Tujuan dari *preprocessing* ini adalah untuk membersihkan data dari kesalahan atau kekurangan sehingga data menjadi lebih berkualitas dan mudah dianalisis.

Setelah itu, diterapkan algoritma TF-IDF untuk mengekstraksi fitur dari teks tweet, sehingga data teks dapat diubah menjadi bentuk numerik yang dapat digunakan dalam proses analisis lebih lanjut. Untuk mencari nilai TF-IDF, dapat menggunakan persamaan (1-3):

$$tf = ft, d \quad (1)$$

$$idfd = \log\left(\frac{N}{df}\right) \quad (2)$$

$$w_{(t,d)} = tf \times idfd \quad (3)$$

Berdasarkan rumus perhitungan nilai TF-IDF, *term frequency* (TF) merepresentasikan frekuensi kemunculan suatu term ( $ft, d$ ) dalam sebuah dokumen ( $d$ ). Sementara itu, *inverse document frequency* (IDF) dihitung dengan mempertimbangkan jumlah total dokumen ( $N$ ) dan jumlah dokumen yang mengandung term tertentu ( $df$ ). Hasil akhir dari perhitungan ini adalah bobot term ( $w_{(t,d)}$ ) pada dokumen  $d$ , yang menunjukkan seberapa penting term tersebut dalam konteks dokumen tersebut.

#### d. Pembangunan model

Tahapan keempat setelah data diproses *preprocessing* dan pembobotan maka selanjutnya melakukan analisis berdasarkan algoritma *Naive Bayes* dan *Support Vector Machine*. Pada tahapan ini ada dua metode yang akan dilakukan. Pertama model akan dibangun menggunakan metode *Naive Bayes* dan metode *Support Vector Machine*. Setelah itu dilakukan proses validasi menggunakan *10-fold cross validation*, selanjutnya untuk mengevaluasi kinerja algoritma yang dilakukan selama masa pengujian menggunakan *Confusion Matrix* dengan menggunakan tiga metriks yaitu akurasi, presisi, dan recallnya.

#### e. Penutup

Tahapan yang terakhir adalah menyajikan hasil dari hasil komparasi dari kedua algoritma yang digunakan serta membuat kesimpulan mengenai proses penelitian yang sudah dilakukan.

### HASIL PENELITIAN DAN PEMBAHASAN

#### Hasil

Pada tahapan ini akan dilakukan implementasi tahapan penelitian, mulai dari pengumpulan data tweets sampai evaluasi model menggunakan *confusion matrix*.

##### 1. Pengumpulan Data

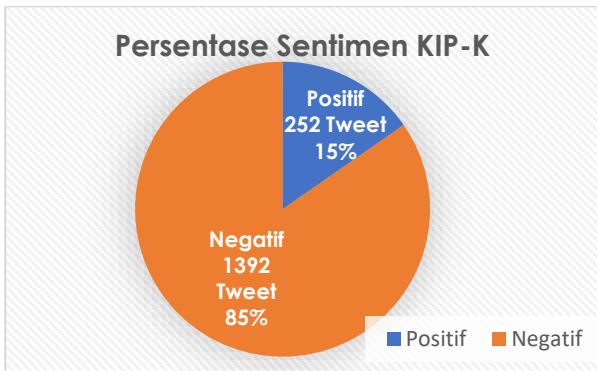
Pengumpulan data dilakukan dengan mengumpulkan data tweets pada media sosial X terkait dengan topik KIP-Kuliah dengan menggunakan kata kunci dan rentang waktu yang telah ditentukan sebelumnya. Berdasarkan tahapan pengumpulan data, maka didapatkan tweets terkait topik KIP-Kuliah dengan jumlah 5.110 tweets.

##### 2. Seleksi Data

Data yang diperoleh dari hasil pengumpulan tweets berjumlah 5.110 tweets. Ada beberapa tahapan yang dilakukan dalam proses seleksi data yaitu dengan menghapus tweet duplikat, menyaring tweet yang tergolong spam seperti promosi, konten berita, atau konten yang tidak bermakna. Setelah seluruh proses seleksi data dilakukan, tersisa 1.644 tweet yang digunakan dalam analisis.

##### 3. Pelabelan Data

Selanjutnya dilakukan proses pelabelan dengan menggunakan dua kelas label yaitu positif dan negatif pada masing-masing tweet. Proses pelabelan ini dibantu oleh dua orang ahli dalam bidang bahasa Indonesia. Dari total 1.644 tweets, didapatkan sentimen positif berjumlah 252 tweets, sentimen negatif berjumlah 1.392 tweets. Perserbaran masing-masing sentimen dapat dilihat pada Gambar 2.



**Gambar 2.** Jumlah Masing-Masing Kelas Sentimen

#### 4. Text-Preprocessing

Proses text-preprocessing terdiri dari case folding, cleaning, tokenize, normalize, stopwords removal, stemming.

##### a. Case folding

Hasil proses case folding dapat dilihat pada Tabel 1.

**Tabel 1.** Hasil Case folding

Sebelum	Kecewa KIPK emg membantu, tapi bnyk yg dapet ga layak !!!
Sesudah	kecewa kipk emg membantu, tapi bnyk yg dapet ga layak !!!

##### b. Cleaning

Hasil proses cleaning dapat dilihat pada Tabel 2.

**Tabel 2.** Hasil Cleaning

Sebelum	kecewa kipk emg membantu, tapi bnyk yg dapet ga layak !!!
Sesudah	kecewa kipk emg bantu tapi bnyk yg dapet ga layak

##### c. Tokenize

Hasil proses tokenize dapat dilihat pada Tabel 3.

**Tabel 3.** Hasil Tokenize

Sebelum	kecewa kipk emg bantu tapi bnyk yg dapet ga layak
Sesudah	["kecewa", "kipk", "emg", "bantu", "tapi", "bnyk", "yg", "dapet", "ga", "layak"]

##### d. Normalize

Hasil proses normalize dapat dilihat pada Tabel 4.

**Tabel 4.** Hasil Normalize

Sebelum	["kecewa", "kipk", "emg", "membantu", "tapi", "bnyk", "yg", "dapet", "ga", "layak"]
Sesudah	["kecewa", "kipk", "memang", "bantu", "tapi", "banyak", "yang", "dapat", "tidak", "layak"]

##### e. Stopwords Removal

Hasil proses stopwords removal dapat dilihat pada Tabel 5.

**Tabel 5.** Hasil Stopwords Removal

Sebelum	["kecewa", "kipk", "memang", "membantu", "tapi", "banyak", "yang", "dapat", "tidak", "layak"]
Sesudah	["kecewa", "kipk", "bantu", "banyak", "dapat", "tidak", "layak"]

##### f. Stemming

Hasil proses stemming dapat dilihat pada Tabel 6.

**Tabel 6. Hasil Steeming**

Sebelum	["kecewa", "kipk", "membantu", "banyak", "dapat", "tidak", "layak"]
Sesudah	["kecewa", "kipk", "bantu", "banyak", "dapat", "tidak", "layak"]

### 5. Ekstraksi TF-IDF

Pada tahapan ekstraksi fitur, proses yang dilakukan perhitungan bobot untuk setiap term atau kata dari data yang sudah melewati tahapan preprocessing. Hasil dari perhitungan ini akan membentuk sebuah matriks dalam bentuk numerik. Tabel 7 merupakan nilai IDF dari term yang memiliki frekuensi kemunculan terbanyak pada dataset.

**Tabel 7. Nilai IDF Top 10 Term**

No	Term	IDF
1	Banget	2.659292
2	Beasiswa	2.155520
3	Iya	2.711764
4	KIP	2.404081
5	KIPK	1.289102
6	Kuliah	2.600361
7	Orang	2.362863
8	Salah	2.072216
9	Sasar	2.081137
10	Terima	2.006901

### 6. Model Klasifikasi

Untuk mengukur dan membandingkan performa masing-masing metode dalam analisis sentimen, dilakukan empat skenario pengujian dengan metode yang berbeda dengan menggunakan teknik 10-fold cross validation. Setiap skenario mengombinasikan algoritma klasifikasi dengan atau tanpa penerapan teknik penyeimbangan data menggunakan SMOTE.

Berikut merupakan rincian dari masing-masing skenario pengujian dapat dilihat pada Tabel 8.

**Tabel 8. Skenario Pengujian Metode Klasifikasi dan SMOTE**

Skenario	Metode
Skenario 1	Naive Bayes
Skenario 2	Support Vector Machine
Skenario 3	Naive Bayes +SMOTE
Skenario 4	Support Vector Machine + SMOTE

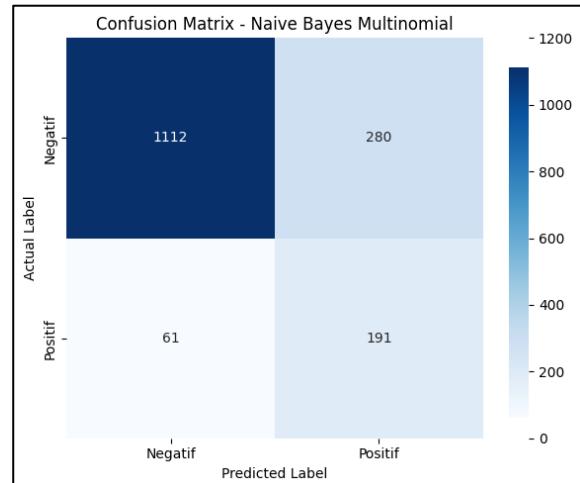
### Skenario 1 (Model Naive Bayes)

Gambar 3 merupakan hasil performa model skenario pertama, yaitu menggunakan metode Naive Bayes.

Classification Report:				
	precision	recall	f1-score	support
Negatif	0.95	0.80	0.87	1392
Positif	0.41	0.76	0.53	252
accuracy			0.79	1644
macro avg	0.68	0.78	0.70	1644
weighted avg	0.86	0.79	0.82	1644

**Gambar 3. Classification Report Model Naive Bayes**

Berdasarkan hasil pengujian model yang dilakukan, skenario pertama menghasilkan nilai akurasi 79%. Berikut merupakan visualisasi confusion matrix pada skenario pertama.

**Gambar 4. Confusion Matrix Naive Bayes**

Berdasarkan visualisasi confusion matrix algoritma Naive Bayes pada Gambar 4, menunjukkan bahwa jumlah tweet yang *true positif* yaitu 191 tweet, 1.112 tweet yang *true negatif*. Sementara itu, terdapat 280 tweet yang diklasifikasikan sebagai *False Positive*, dan 61 tweet *False Negative*.

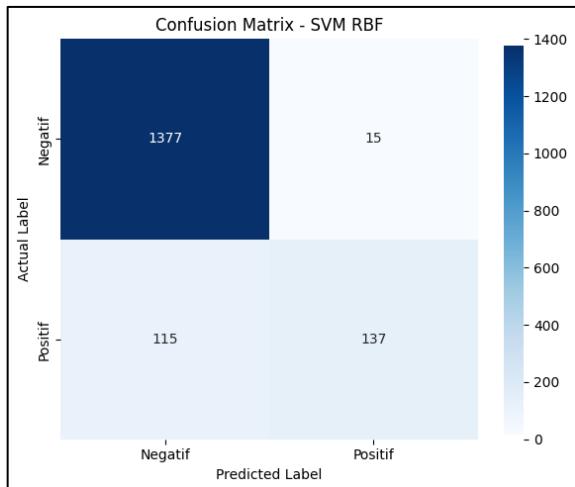
## Skenario 2 (Model Support Vector Machine)

Selanjutnya untuk pengujian model pada skenario kedua menggunakan model SVM mendapatkan nilai akurasi 92% yang ditunjukkan pada Gambar 5.

Classification Report untuk Seluruh Fold:				
	precision	recall	f1-score	support
0	0.92	0.99	0.95	1392
1	0.90	0.54	0.68	252
accuracy			0.92	1644
macro avg	0.91	0.77	0.82	1644
weighted avg	0.92	0.92	0.91	1644

Gambar 5. Classification Report Model SVM

Berikut merupakan hasil perhitungan dan visualisasi untuk confusion matrix model SVM yang ditunjukkan pada Gambar 6.

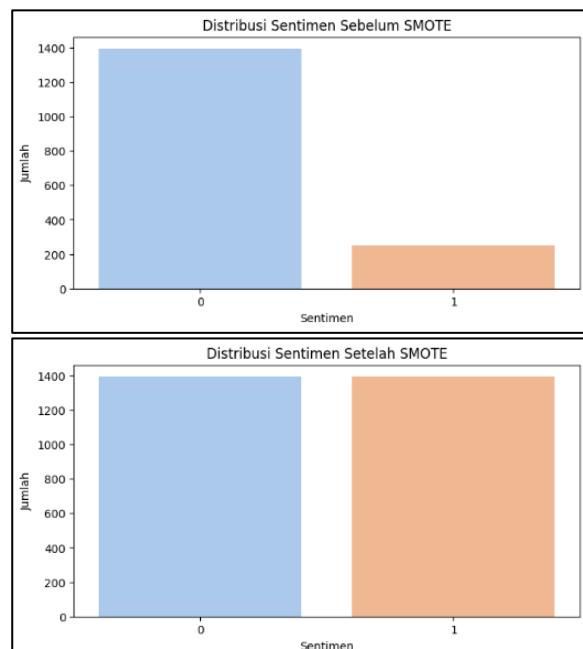


Gambar 6. Confusion Matrix SVM

Berdasarkan visualisasi confusion matrix algoritma SVM pada Gambar 6, menunjukkan bahwa jumlah tweet yang *true positif* yaitu 137 tweet, 1.377 tweet yang *true negatif*. Sementara itu, terdapat 115 tweet yang diklasifikasikan sebagai *False Positive*, dan 15 tweet *False Negative*.

## Skenario 3 (Model Naive Bayes + SMOTE)

Pada skenario model ketiga dan keempat dilakukan proses *resampling* karena distribusi data sentimen yang tidak seimbang. Teknik *resampling* yang digunakan adalah *oversampling* dengan metode SMOTE. Gambar 7 menunjukkan persebaran data sebelum dan sesudah SMOTE.



Gambar 7. Perbandingan Hasil Resampling

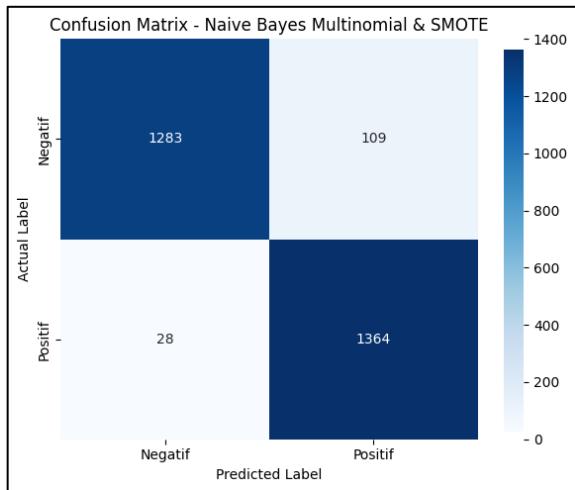
Berdasarkan Gambar 7, setelah dilakukannya proses *oversampling* menggunakan SMOTE, masing-masing jumlah sentimen menjadi seimbang, yaitu dengan jumlah 1.392 pada tiap sentimen. Sehingga saat ini dataset berjumlah 2.784. Setelah distribusi data seimbang, maka berikut merupakan hasil pelatihan model pada skenario ketiga menggunakan model Naive Bayes.

Classification Report:				
	precision	recall	f1-score	support
Negatif	0.98	0.92	0.95	1392
Positif	0.93	0.98	0.95	1392
accuracy			0.95	2784
macro avg	0.95	0.95	0.95	2784
weighted avg	0.95	0.95	0.95	2784

Gambar 8. Classification Report Model Naive Bayes + SMOTE

Berdasarkan Gambar 8, hasil pengujian model Naive Bayes dengan

SMOTE mendapatkan nilai akurasi 95%. Berikut merupakan hasil perhitungan nilai confusion matrix yang ditunjukkan pada Gambar 8.



**Gambar 9.** Confusion Matrix Naive Bayes & SMOTE

Berdasarkan visualisasi confusion matrix algoritma Naive Bayes dengan menerapkan SMOTE pada Gambar 9, menunjukkan bahwa jumlah tweet yang true positif yaitu 1.364 tweet, 1.238 tweet yang true negatif. Sementara itu, terdapat 28 tweet yang diklasifikasikan sebagai False Positive, dan 109 tweet False Negative.

#### Skenario 4 (Model Support Vector Machine + SMOTE)

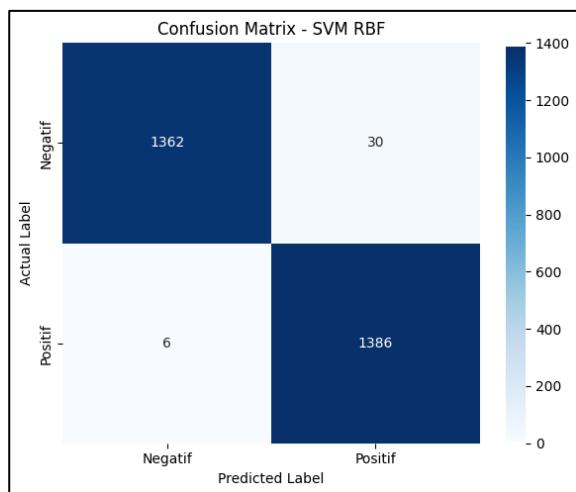
Pengujian model pada skenario keempat menggunakan model SVM dengan SMOTE menghasilkan nilai akurasi mencapai 99% yang ditunjukkan pada classification report pada Gambar 10.

Classification Report untuk Seluruh Fold:				
	precision	recall	f1-score	support
0	1.00	0.98	0.99	1392
1	0.98	1.00	0.99	1392
accuracy			0.99	2784
macro avg	0.99	0.99	0.99	2784
weighted avg	0.99	0.99	0.99	2784

**Gambar 10.** Classification Report SVM & SMOTE

Berikut merupakan hasil perhitungan dan visualisasi untuk

confusion matrix model SVM yang ditunjukkan pada Gambar 11.



**Gambar 11.** Confusion Matrix SVM & SMOTE

Berdasarkan visualisasi confusion matrix algoritma Support Vector Machine dengan menerapkan SMOTE pada Gambar 11, menunjukkan bahwa jumlah tweet yang true positif yaitu 1.386 tweet, 1.362 tweet yang true negatif. Sementara itu, terdapat 8 tweet yang diklasifikasikan sebagai False Positive, dan 30 tweet False Negative.

Perbandingan performa pada keempat model skenario analisis sentimen pada empat skenario analisis dapat dilihat pada Tabel 9.

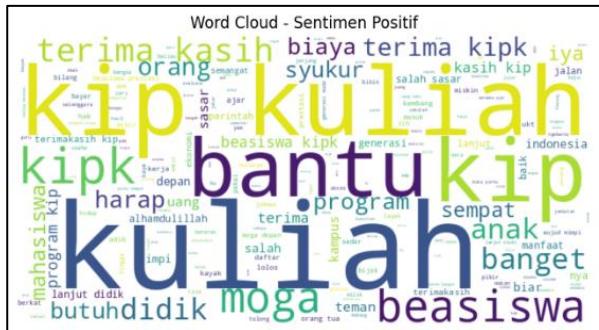
**Tabel 9.** Perbandingan Performa Model Empat Skenario

Skenario	Akurasi	Presisi	Recall
Skenario 1	79%	68%	78%
Skenario 2	92%	91%	77%
<b>Skenario 3</b>	<b>95%</b>	<b>95%</b>	<b>95%</b>
Skenario 4	99%	99%	99%

Berdasarkan Tabel 9 skenario 3 menggunakan metode Naive Bayes dan SMOTE memiliki performa yang lebih stabil dibandingkan dengan 3 skenario sebelumnya.

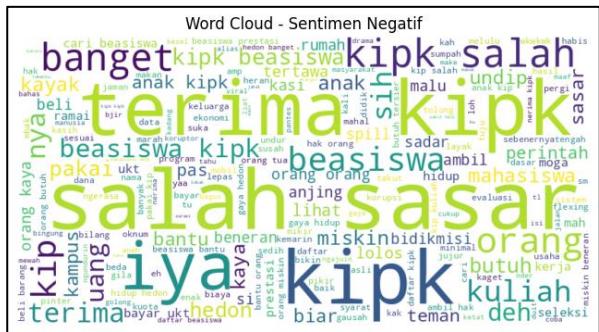
Untuk memahami lebih lanjut mengenai persebaran kata dalam masing-masing kategori sentimen, dilakukan visualisasi kata-kata yang paling sering muncul pada sentimen positif dan negatif. Visualisasi ini bertujuan untuk mengidentifikasi kata-kata dominan yang dapat mencerminkan topik utama yang

sering dibahas oleh masyarakat terkait program KIP-K.



**Gambar 12.** Word cloud pada Sentimen Positif

Gambar 12 menunjukkan visualisasi word cloud pada sentimen positif. Kata yang paling sering muncul pada sentimen positif adalah "kip", "kuliah", "bantu", "moga", "terima", "kasih", "syukur", "beasiswa".



**Gambar 13.** Word Cloud Sentimen Negatif

Gambar 13 menunjukkan visualisasi word cloud pada sentimen negatif. Kata yang sering muncul pada sentimen negatif adalah “salah”, “sasar”, “kipk”, “beasiswa”, “terima”, “uang”.

Secara keseluruhan, respon masyarakat terhadap program KIP-Kuliah menunjukkan bahwa kelas sentimen negatif memiliki jumlah yang paling banyak yakni 1.392 tweet, hal ini berarti respon masyarakat cenderung bersikap negatif terhadap program tersebut.

## Pembahasan

Berdasarkan hasil pengujian empat skenario yang dilakukan untuk mengevaluasi performa algoritma Naive Bayes dan Support Vector Machine dalam analisis sentimen Program Kartu Indonesia Pintar Kuliah (KIP-K), baik tanpa maupun dengan penerapan SMOTE sebagai solusi

untuk mengatasi ketidakseimbangan data, diperoleh sejumlah temuan penting. Pada skenario pertama, algoritma Naive Bayes tanpa SMOTE menunjukkan performa yang cukup baik dengan akurasi sebesar 79%, rata-rata presisi 68%, dan rata-rata *recall* 78%. Namun, terdapat ketimpangan antara presisi dan *recall* pada masing-masing kelas, di mana model lebih baik dalam mengidentifikasi tweet negatif dengan presisi mencapai 95%, sementara presisi untuk kelas positif hanya 41%. Hal ini menunjukkan adanya bias terhadap kelas negatif, di mana model cenderung mengklasifikasikan tweet positif sebagai negatif.

Pada skenario kedua, algoritma SVM menunjukkan performa yang lebih baik dibandingkan Naive Bayes, dengan akurasi 92%, presisi 91%, dan recall 77%. Namun, meskipun presisi untuk kelas positif mencapai 90%, recall-nya masih rendah, yaitu 54%, sementara recall untuk kelas negatif sangat tinggi, mencapai 99%. Ini juga menandakan adanya bias terhadap kelas negatif akibat distribusi data yang tidak seimbang, di mana tweet dengan sentimen negatif mendominasi. Oleh karena itu, pada skenario ketiga dan keempat diterapkan teknik oversampling menggunakan SMOTE untuk mengatasi ketidakseimbangan tersebut. Teknik ini dipilih karena data yang digunakan dalam penelitian ini tergolong kecil yaitu di bawah ratusan ribu sehingga pendekatan oversampling dinilai lebih efektif dibandingkan undersampling, serta tetap menjaga informasi penting dari data asli.

Setelah diterapkan SMOTE pada skenario ketiga, performa algoritma Naive Bayes meningkat secara signifikan. Akurasi, presisi dan *recall* naik menjadi 95%. Ini menunjukkan bahwa SMOTE berhasil meningkatkan performa model secara menyeluruh, terutama dalam mengenali kelas minoritas, yaitu *tweet* dengan sentimen positif. Hasil ini sejalan dengan penelitian sebelumnya (Handoko & Aditya, 2025), yang menyatakan bahwa SMOTE efektif dalam meningkatkan performa klasifikasi pada data tidak seimbang. Skenario keempat juga

menunjukkan bahwa SVM mengalami peningkatan performa setelah SMOTE diterapkan. Akurasi meningkat menjadi 99%, presisi menjadi 99%, dan *recall* menjadi 99%, dengan rata-rata peningkatan seluruh metrik evaluasi sebesar 12,3%.

Secara keseluruhan, penerapan SMOTE berhasil meningkatkan performa kedua algoritma dan mengurangi bias terhadap kelas mayoritas. Naive Bayes menunjukkan performa yang lebih stabil dan seimbang dibandingkan SVM, terutama setelah diterapkan SMOTE. Hal ini menunjukkan keunggulan Naive Bayes dalam menangani klasifikasi pada data kecil dengan distribusi tidak seimbang, serta memiliki struktur yang lebih sederhana dan efisien secara komputasi dibandingkan SVM.

Dari hasil analisis sentimen terhadap Program KIP-K, ditemukan bahwa mayoritas tweet diklasifikasikan sebagai sentimen negatif. Hal ini menunjukkan tingginya ketidakpuasan masyarakat terhadap pelaksanaan program tersebut. Temuan ini diperkuat oleh analisis TF-IDF yang menunjukkan kata-kata seperti "beasiswa", "kip", "kuliah", "orang", "salah", "sasar", dan "terima" memiliki bobot tinggi, serta kemunculan kata-kata tersebut pada word frequency sentimen negatif. Hal ini mencerminkan adanya kritik masyarakat terhadap ketidaktepatan sasaran dan gaya hidup tidak mencerminkan kondisi ekonomi dari sebagian penerima bantuan.

Distribusi data juga menunjukkan ketimpangan antara jumlah tweet negatif dan positif, dengan 1.392 tweet bersentimen negatif dan hanya 252 tweet positif. Ini mengindikasikan bahwa opini publik di platform X cenderung negatif terhadap Program KIP-K. Berdasarkan temuan ini, perlu dilakukan evaluasi menyeluruh terhadap pelaksanaan program, termasuk perbaikan dalam proses seleksi penerima melalui validasi data yang lebih ketat, survei lapangan, serta koordinasi dengan instansi terkait.

Monitoring berkala dan keterlibatan institusi pendidikan juga penting untuk memastikan bantuan tepat sasaran dan tidak disalahgunakan. Selain itu, transparansi dan edukasi publik mengenai program, kriteria penerima, serta etika penerima bantuan perlu ditingkatkan guna membangun kepercayaan masyarakat terhadap Program KIP-K.

## KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan metode SMOTE secara signifikan meningkatkan performa algoritma Naive Bayes dan Support Vector Machine (SVM) dalam analisis sentimen terhadap program Kartu Indonesia Pintar Kuliah (KIP-K). Setelah SMOTE diterapkan, akurasi, presisi, dan *recall* Naive Bayes meningkat menjadi 95%, sementara SVM mencapai 99% pada seluruh metrik evaluasi. Meskipun SVM menunjukkan hasil evaluasi yang lebih tinggi, Naive Bayes dinilai lebih stabil dan konsisten. Mayoritas tweet dalam data diklasifikasikan sebagai sentimen negatif, yang mencerminkan ketidakpuasan masyarakat terhadap pelaksanaan program KIP-K. Oleh karena itu, perlu evaluasi terhadap sistem seleksi dan pemantauan program. Penelitian selanjutnya disarankan untuk mengeksplorasi metode deep learning dan teknik oversampling lainnya seperti ADASYN guna memperoleh hasil yang lebih optimal.

## DAFTAR PUSTAKA

- Amelia, I., & Sarimole, F. M. (2024). Analisis Sentimen Tanggapan Pengguna Media Sosial X Terhadap Program Beasiswa KIP-Kuliah dengan Menggunakan Algoritma Support Vector Machine (SVM). Dalam *Jurnal Indonesia: Manajemen Informatika dan Komunikasi (JIMIK)* (Vol. 5, Nomor 3). <https://journal.stmiki.ac.id>
- Ekaptiningrum, K. (2024, Agustus). Kisah Johar Penerima Beasiswa KIP-K Lulus Cumlaude di FEB UGM, Jadi Sarjana Pertama di Keluarga - FEB UGM. <https://feb.ugm.ac.id/id/berita/4765-kisah-johar-penerima-beasiswa-kip-k>

Hakim, S. N., Putra, A. J., & Khasanah, A. U. (2021). Sentiment Analysis on Myindihome User Reviews Using Support Vector Machine and Naïve Bayes Classifier Method. *International Journal of Industrial Optimization*, 2(2), 141.  
<https://doi.org/10.12928/ijio.v2i2.4449>

Handoko, C. B., & Aditya, C. S. K. (2025). Penerapan Teknik SMOTE Dalam Mengatasi Imbalance Data Penyakit Diabetes Menggunakan Algoritma ANN. *Smart Comp: Jurnalnya Orang Pintar Komputer*, 14(1).  
<https://doi.org/10.30591/smartcomp.v14i1.7045>

Hashfi, F., Sugiarto, D., & Mardianto, I. (2022). Sentiment Analysis of An Internet Provider Company Based on Twitter Using Support Vector Machine and Naïve Bayes Method. *Ultimatics: Jurnal Teknik Informatika*, 14(1), 1-6.  
<https://doi.org/10.31937/ti.v14i1.2384>

Ilmawan, L. B., & Mude, M. A. (2020). Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store. *ILKOM Jurnal Ilmiah*, 12(2), 154–161.  
[https://doi.org/10.33096/ilkom.v12i2.597.154-161](https://doi.org/10.33096/ilkom.v12i2.597)

Iskandar, J. W., & Nataliani, Y. (2021). Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(6), 1120–1126.  
<https://doi.org/10.29207/resti.v5i6.3588>

Khaira, U., Aryani, R., & Hardian, R. W. (2023). Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Kebijakan Kemdikbudristek Mengenai Kuota Internet Selama Covid-19. *Jurnal PROCESSOR*, 18(2).

<https://doi.org/10.33998/processor.2023.18.2.897>

Negara, R. A. (2024, Mei 4). Sudah Viral, Mahasiswa Penerima Bantuan Keluarga Miskin KIP Kuliah Sewa Pengacara Ancam Penjarakan yang Bongkar Identitas - Kilat. <https://www.kilat.com/nasional/84412585023/sudah-viral-mahasiswa-penerima-bantuan-keluarga-miskin-kip-kuliah-sewa-pengacara-ancam-penjarakan-yang-bongkar-identitas>

Pramudita, D., Akbar, Y., & Wahyudi, T. (2024). Analisis Sentimen Terhadap Program Kartu Indonesia Pintar Kuliah Pada Media Sosial X Menggunakan Algoritma Naive Bayes. 4, 1420–1430.  
<https://doi.org/10.57152/malcom.v4i4.1565>

Puspapertiwi, E. R., & Nugroho, R. S. (2024, Mei 1). Ramai soal Mahasiswa Undip Penerima KIP Kuliah Bergaya Hidup Mewah, Mundur Usai Diungkap Warganet Halaman all - Kompas.com. <https://www.kompas.com/tren/read/2024/05/01/204500465/ramai-soal-mahasiswa-undip-penerima-kip-kuliah-bergaya-hidup-mewah-mundur?page=all>

Yuliantri P, F. (2021, Juli 19). Hilangnya Hak Anak dalam Sengkarut Program Indonesia Pintar. <https://wartapemeriksa.bpk.go.id/?p=26572>