



ANALISIS SENTIMEN ULASAN PENGGUNA CORETAX DI APLIKASI TWITTER MENGGUNAKAN ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE

Riza Fahlapi ^{1*}, Laila Awalia Ramadani ², Dinda Aulia Julianti ³, Aldilah Rhamadani ⁴,
Dhita Amalia ⁵

^{1,2,3,4,5} Program Studi Akuntansi, Fakultas Bisnis dan Manajemen, Universitas Bina Sarana
Informatika

E-mail: riza.rzf@bsi.ac.id^{1*}, 63230944@bsi.ac.id², 63231042@bsi.ac.id³,
63231374@bsi.ac.id⁴, 63231395@bsi.ac.id⁵

ABSTRACT

This study aims to evaluate user sentiment towards the latest tax application, Coretax, and compare the effectiveness of two classification algorithms, namely Support Vector Machine (SVM) and Naive Bayes. The data was collected using web scraping techniques from the X platform (formerly Twitter) and through a series of pre-processing processes, including data cleansing, case folding, normalization, stopwords, stemming, labeling, and visualization. After that, feature weighting was carried out using the Term Frequency-Inverse Document Frequency (TF-IDF) method, as well as feature selection using the SelectKBest and Chi-Square methods.

The results of the evaluation showed that both algorithms performed very well in classifying sentiment. SVM recorded an accuracy of 92%, with precision, recall, and F1-scores of 93%, 94%, and 94%, respectively, for negative sentiment, and 90%, 88%, and 89% for positive sentiment. Meanwhile, the Naive Bayes algorithm also showed similar performance with 92% accuracy, 90% precision, 99% recall, and 94% F1-score for negative sentiment, as well as 97% precision, 80% recall, and 88% F1-score for positive sentiment. These results indicate that both algorithms are quite reliable at recognizing negative sentiments, although there is still room for improvement in detecting positive sentiment.

Keywords: Coretax, Sentiment Analysis, Support Vector Machine, Naive Bayes

ABSTRAK

Penelitian ini bertujuan untuk mengidentifikasi persepsi atau sentimen pengguna terhadap aplikasi perpajakan terbaru, yaitu Coretax, serta membandingkan performa kedua metode klasifikasi, yakni Support Vector Machine (SVM) dan Naive Bayes. Data diperoleh dengan teknik web scraping dari platform X (dahulu dikenal sebagai Twitter), kemudian diolah melalui serangkaian tahap pra-pemrosesan yang mencakup pembersihan teks (cleaning), perubahan huruf menjadi format seragam (case folding), normalisasi (normalized), penghapusan kata-kata umum

Article History

Received: Juni 2025

Reviewed: Juni 2025

Published: Juni 2025

Plagiarism Checker No
235

Prefix DOI :

[10.8734/Koehesi.v1i2.365](https://doi.org/10.8734/Koehesi.v1i2.365)

Copyright : Author

Publish by : Koehesi



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



(stopword), proses stemming (stemming), pelabelan data (labeling), dan visualisasi (visualized). Setelah semua langkah itu, dilakukan pembobotan kata menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) dan pemilihan fitur menggunakan pendekatan SelectKBest serta pengujian Chi-Square.

Dari hasil pengujian, kedua algoritma menunjukkan hasil yang sangat baik dalam mengklasifikasikan data sentimen. Algoritma SVM mencapai akurasi sebesar 92%, dengan nilai precision, recall, dan F1-score masing-masing adalah 93%, 94%, dan 94% untuk sentimen negatif. Sementara itu, untuk sentimen positif, angkanya adalah 90%, 88%, dan 89%. Di sisi lain, algoritma Naive Bayes juga mencatat akurasi sebesar 92%, dengan precision 90%, recall 99%, dan F1-score 94% pada sentimen negatif, serta precision 97%, recall 80%, dan F1-score 88% untuk sentimen positif. Temuan ini menunjukkan bahwa kedua metode efektif dalam mengidentifikasi sentimen negatif, tetapi klasifikasi untuk sentimen positif masih perlu perbaikan dalam akurasinya.

Kata Kunci: Analisis Sentimen, Coretax, Support Vector Machine, Naive Bayes

1. PENDAHULUAN

Transformasi digital dalam layanan publik merupakan inisiatif penting yang dilakukan oleh pemerintah guna meningkatkan efisiensi operasional, keterbukaan informasi, serta kemudahan akses bagi masyarakat luas. Di sektor perpajakan, Direktorat Jenderal Pajak (DJP) Indonesia mendukung proses digitalisasi ini melalui pengembangan aplikasi **Coretax**, yang menjadi bagian integral dari program **Pembaruan Sistem Inti Administrasi Perpajakan (PSIAP)**. Inovasi ini dirancang untuk menyatukan berbagai fungsi layanan perpajakan dalam satu sistem digital yang terpusat dan lebih modern, dengan harapan dapat meningkatkan mutu pelayanan dan memperkuat kepatuhan para wajib pajak.

Meski demikian, keberhasilan sebuah sistem digital tidak semata-mata bergantung pada teknologi yang diimplementasikan, melainkan juga dipengaruhi oleh tanggapan dan kepuasan para penggunanya. Saat ini, masyarakat cenderung menyampaikan pendapat, pengalaman, atau keluhan mereka melalui platform seperti **Twitter**, yang menawarkan ruang publik yang dinamis untuk menyuarakan opini secara langsung. Informasi yang terkandung dalam unggahan pengguna di media sosial dapat dimanfaatkan sebagai sumber data alternatif untuk memahami persepsi masyarakat terhadap layanan seperti **Coretax**.



Untuk mengolah opini-opini tersebut secara terstruktur, pendekatan **analisis sentimen** menjadi metode yang tepat. Analisis sentimen atau *sentiment analysis* merupakan salah satu proses komputasi dengan menganalisis teks digital yang digunakan untuk menentukan apakah kata-kata atau kalimat yang disampaikan memiliki makna atau emosional pesan. Sentimen analisis akan memahami dan mengelompokkan emosi dalam kategori positif, negatif, dan netral. Teknik ini bertujuan untuk mengidentifikasi opini masyarakat ke dalam kategori sentimen tertentu—positif, dan negatif—sehingga dapat memberikan gambaran umum mengenai sikap publik terhadap layanan yang diteliti. Dalam praktiknya, algoritma Naive Bayes dan Support Vector Machine (SVM) lumrah digunakan dalam tugas klasifikasi sentimen, karena keduanya terbukti efektif dalam mengolah data tidak terstruktur. Naive Bayes menawarkan proses klasifikasi yang efisien untuk dataset berukuran besar, sementara SVM memiliki keunggulan dalam membedakan data antar kelas dengan akurasi yang tinggi. Naive Bayes merupakan salah satu algoritma klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes untuk memprediksi kategori atau kelas berdasarkan data yang ada. Prinsip dasar dari Naive Bayes adalah mengasumsikan bahwa setiap fitur dalam dataset bersifat independen satu sama lain, yang dikenal dengan istilah *naive* (naif). Sedangkan Support Vector Machine (SVM) merupakan pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. Svm bekerja dengan mencari Hyperplane optimal yang memisahkan data ke dalam kelas kelas yang berbeda dengan margin terbesar. Keunggulan utama svm adalah kemampuannya menangani data non linear melalui pengguna fungsi kernel, seperti kernel linear, polynomial, Radial Basis Function (RBF), dan Sigmoid.

Penelitian ini bertujuan untuk mengevaluasi persepsi publik terhadap aplikasi Coretax dengan menganalisis review pengguna di platform Twitter. Selain itu, studi ini juga melakukan perbandingan performa antara algoritma Naive Bayes dan Support Vector Machine dalam mengklasifikasikan sentimen yang muncul dari data tersebut. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi nyata terhadap peningkatan kualitas layanan perpajakan berbasis digital, serta memperkaya pemahaman akademik dalam bidang analisis sentimen media sosial.

2. METODE PENELITIAN

Penelitian ini mengadopsi penelitian kuantitatif, yaitu penelitian yang berfokus pada pengumpulan data dalam bentuk angka dan penggunaan teknik analisis statistik untuk



menentukan fakta-fakta objektif. Memanfaatkan data yang berasal dari media sosial, khususnya Twitter, pendekatan kuantitatif dalam penelitian ini digunakan untuk menilai dan mengukur opini publik terhadap aplikasi Coretax

Secara lebih rinci, penelitian ini menganalisis umpan balik pengguna mengenai aplikasi Coretax itu sendiri. Berdasarkan konteksnya, tujuan dari analisis ini adalah untuk mengkategorikan tweet ke dalam dua jenis sentimen: positif dan negatif. Penelitian ini juga membandingkan efektivitas dua algoritma klasifikasi teks yang umum digunakan dalam analisis sentimen, yaitu Naive Bayes dan Support Vector Machine (SVM), untuk melihat bagaimana kinerja masing-masing algoritma dalam mengklasifikasikan opini yang telah disebutkan di atas. Kami telah menyertakan daftar uraian tahapan penelitian yang kami lakukan untuk memudahkan Anda memahami penelitian kami.

2.1 Data Scraping

Data selection adalah proses pemilihan data yang telah dieliminasi dan kemudian diberi label. Metode yang akan digunakan untuk mengumpulkan data adalah teknik scraping Google Colab. Data yang akan digunakan adalah opini publik pengguna Coretax terhadap aplikasi X. Sekitar 684 data akan digunakan selama periode waktu dari Januari 2025 hingga April 2025.

2.2 Pre-Procesing

Pre-Processing merupakan proses transformasi data tidak terstruktur menjadi data terstruktur (Alvis Lowell, 2025). Dalam Pre-Processing terdapat beberapa tahapan, yaitu;

a. Cleaning

Cleaning adalah tahapan yang bertujuan untuk membersihkan atau menghapus atribut yang tidak penting, mengeliminasi data duplikat, dan memilih atribut data yang akan diperlukan. Tahapan ini dilakukan untuk memastikan data siap digunakan.

b. Case Folding

Case Folding dilakukan untuk mengganti kata dalam tulisan menjadi huruf kecil. Hal ini dilakukan untuk menstandarisasikan kata agar mudah untuk digunakan.

c. Normalized

Normalized adalah tahapan mengubah seluruh kata yang tidak baku menjadi baku.

d. Stopwords Removal



Stopwords Removal adalah tahapan menghapus kata-kata umum yang tidak relevan.

e. Tokenized

Tokenized adalah tahapan yang dilakukan untuk memecah atau memisahkan teks menjadi beberapa kata serta menghapus tanda baca yang ada. Tujuannya adalah untuk mempermudah perhitungan frekuensi kemunculan kata.

f. Stemming

Stemming adalah tahapan dimana kata-kata yang telah dipisahkan akan diubah menjadi bentuk dasarnya.

g. Labeling

Labeling adalah tahapan memberi bentuk sentimen pada semua data atau pendapat yang digunakan untuk penelitian ini, bertujuan agar mempermudah memilih data. Data yang ada akan dilabeli dengan dua macam sentimen, yaitu negatif atau positif.

h. Visualisasi

Visualisasi adalah tahapan mengubah visual dari data menjadi gambar sehingga mudah untuk dipahami dan dilihat.

2.3 Transformation

Transformasi data adalah proses mengubah data mentah menjadi format lain sehingga sesuai dengan asumsi yang mendasari analisis berbagai teknik data mining. Teks berikut ini akan diberikan bobot dengan menggunakan teknik TF-IDF (Term Frequency-Inverse Document Frequency). Tujuannya adalah untuk meringkas data yang berbentuk kata menjadi numerik, yang nantinya akan mempermudah proses data mining. Adapun rumus dari TF-IDF Score yaitu sebagai berikut:

$$\text{TF-IDF (t,d)} = \text{TF (t,d)} \times \text{IDF (t)}$$

Keterangan:

- TF(t,d) = frekuensi kata t dalam dokumen d
- IDF(t) = $\log(\text{total dokumen} / \text{jumlah dokumen yang mengandung t})$

Setelah dilakukan ekstraksi dengan TF-IDF, data kemudian akan diseleksi dengan metode Chi-Square. Metode Chi-Square (χ^2) merupakan salah satu teknik statistik non parametrik yang digunakan untuk mengetahui ada tidaknya hubungan atau perbedaan antara dua variabel yang berskala nominal atau ordinal dalam suatu populasi. Tujuan



dari metode ini adalah untuk mengetahui apakah ada perbedaan yang signifikan antara frekuensi yang dihasilkan dari observasi dengan frekuensi yang diprediksi berdasarkan hipotesis nol (H_0).

2.4 Data Mining

Data mining adalah jenis proses klasifikasi yang terdiri dari beberapa langkah yang dirancang untuk menganalisa informasi dan data dari data tekstual atau data sentimental yang digunakan. Dalam proses klasifikasi ini, data akan dibagi menjadi dua kategori: data training dan data testing. Langkah selanjutnya adalah klasifikasi dengan menggunakan dua algoritma: Naive Bayes dan Support Vector Machine. Proses klasifikasi ini akan dilakukan dengan menggunakan bahasa pemrograman Python pada Google Colab.

2.5 Evaluation

Langkah terakhir dari analisis sentimen adalah evaluasi. Tujuan dari evaluasi ini adalah untuk meninjau dan meningkatkan kinerja klasifikasi yang sebelumnya telah diselesaikan dengan menggunakan Confusion Matrix. Langkah terakhir dari analisis sentimental adalah evaluasi. Tujuan dari evaluasi ini adalah untuk menilai dan meningkatkan kinerja dari dua klasifikasi yang telah diselesaikan menggunakan Confusion Matrix. Confusion matrix adalah metode evaluasi yang digunakan untuk menentukan akurasi atau ambang batas kinerja dari suatu proses klasifikasi. Confusion matrix dapat digunakan untuk menganalisis beberapa keefektifan model klasifikasi dalam mengidentifikasi data dari berbagai kelas yang berbeda. Confusion Matrix digunakan untuk menghitung nilai F1-score, akurasi, dan presisi.

1. Accuracy digunakan untuk membandingkan jumlah prediksi yang dibuat oleh model yang akurat dengan jumlah total prediksi yang dibuat, apakah akurat atau tidak. Rumus yang digunakan untuk mengukur accuracy yaitu:

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

Keterangan:

- TP (True Positive): Prediksi positif yang benar
- TN (True Negative): Prediksi negatif yang benar



- FP (False Positive): Prediksi positif yang salah
- FN (False Negative): Prediksi negatif yang salah

2. Precision digunakan untuk meningkatkan akurasi prediksi yang dibuat oleh model. Presisi memberikan prediksi positif yang akurat dan relevan. Rumus yang digunakan untuk mengukur precision yaitu:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

Keterangan:

- TP (True Positive): Prediksi positif yang benar
- FP (False Positive): Prediksi positif yang salah

3. F1-score adalah rata-rata harmonis dari presisi dan recall, yang berarti hanya memberikan nilai yang tinggi jika kedua metrik yang sesuai sama. F1-score digunakan untuk menghitung perbedaan antara recall dan precision dalam model klasifikasi. Rumus yang digunakan untuk mengukur F1-score yaitu:

$$F1 = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$$

Dimana:

- Precision = $\text{TP} / (\text{TP} + \text{FP})$
- Recall = $\text{TP} / (\text{TP} + \text{FN})$

3. HASIL DAN PEMBAHASAN

3.1 Hasil Data Scraping

Hasil penelitian yang kami lakukan, kami menggunakan data opini pengalaman pengguna Coretax yang disuarakan oleh para pengguna di aplikasi X. Data yang ada kemudian dikumpulkan menggunakan metode scraping dengan bahasa pemrograman Python. Dari hasil scraping tersebut, kami berhasil mengumpulkan 684 opini pengguna dari periode Januari 2025 hingga April 2025. Data tersebut kemudian kami simpan dalam format CSV.



Gambar 1 : Hasil Data Scraping

conversation_id_str	created_at	favorite_count	full_text	id_str	image_url	in_reply_to_screen_name	lang	location	quote_count	reply_count
0	1917250184079368314 Tue Apr 29 16:10:39 +0000 2025	0	@kring_pajak kak ini solusinya gimana ya ? Pad...	1917250184079368314	https://pbs.twimg.com/media/Gptw4DobAAE0Z7p.jpg	kring_pajak	in	NaN	0	2
1	1917229842418036834 Tue Apr 29 15:39:15 +0000 2025	0	@Dinaa_dini12 @kring_pajak Samaaa banget abah ...	1917242280999878785	NaN	Dinaa_dini12	in	NaN	0	1
2	1917229842418036834 Tue Apr 29 14:48:49 +0000 2025	2	Selamat malam min @kring_pajak Dari sore sampa...	1917229842418036834	https://pbs.twimg.com/media/GploYWraEAA89Nq.jpg	NaN	in	NaN	0	7
3	1917221396048090598 Tue Apr 29 14:27:12 +0000 2025	0	Kepikiran coretax sangat syulit https://t.co/D...	1917224152127455734	https://pbs.twimg.com/text_video_thumb/19172...	my_cening_j	in	NaN	0	0
4	1917151946303304064 Tue Apr 29 14:02:57 +0000 2025	0	@hrsnotes Sama kena gini juga Coretax parah pa...	1917218046835609644	NaN	hrsnotes	in	NaN	0	0
...
679	1915449011819176004 Thu Apr 24 16:53:28 +0000 2025	0	dengan coretax pekerjaan saya jadi semakin ruma	1915449011819176004	NaN	NaN	in	NaN	0	0

3.2 Hasil Preprocessing Data

Selanjutnya, setelah data dikumpulkan dan diberi label, data tersebut akan melalui berbagai langkah pre-processing untuk menghasilkan hasil yang siap digunakan.

a. Cleaning Data

Tahap Cleaning Data dilakukan untuk menghilangkan atau menyembapkan data dari berbagai elemen yang tidak relevan, seperti karakter tertentu, tanda baca, simbol khusus, emoticon, angka, dan URL.

Sebelum	Sesudah
Coretax itu membantu apa mempersulit?	Coretax itu membantu apa mempersulit

Tabel 1. Hasil Cleaning Data

b. Case Folding

Tahap Case Folding dilakukan agar seluruh kata menjadi data dengan huruf kecil.

Sebelum	Sesudah
Coretax itu membantu apa mempersulit	coretax itu membantu apa mempersulit

Tabel 2. Hasil Case Folding



c. Normalized

Tahap Normalized dilakukan untuk memperbaiki kesalahan ejaan dan penulisan kata.

Sebelum	Sesudah
aku gk bs akses coretax dari pagi ini	aku tidak bisa akses coretax dari pagi ini

Tabel 3. Hasil Normalized

d. Stopwords Removal

Tahap Stopwords Removal digunakan agar kata yang tidak relevan dan kata yang tidak mempengaruhi sentimen tersebut hilang, sehingga kalimat akan menjadi singkat dan jelas.

Sebelum	Sesudah
coretax sempat berjalan baik pagi ini untuk pembuatan faktur namun saat malam ingin melaporkan ppn justru tidak bisa sangat menyulitkan	coretax berjalan pagi pembuatan faktur malam melaporkan ppn menyulitkan

Tabel 4. Hasil Stopwords Removal

e. Tokenized

Tahap Tokenized bertujuan untuk memisahkan kalimat menjadi kata per kata.

Sebelum	Sesudah
coretax berjalan pagi pembuatan faktur malam melaporkan ppn menyulitkan	['coretax', 'berjalan', 'pagi', 'pembuatan', 'faktur', 'malam', 'melaporkan', 'ppn', 'menyulitkan']

Tabel 5. Hasil Tokenized

f. Stemming

Tahap Stemming bertujuan agar kata-kata yang sudah dipisahkan menjadi bentuk aslinya.

Sebelum	Sesudah
---------	---------

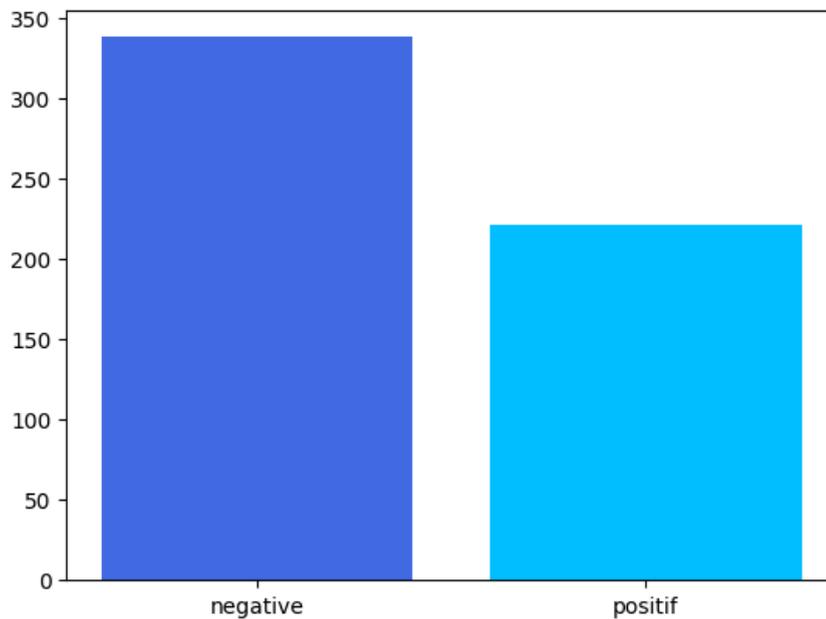


['coretax', 'berjalan', 'pagi', 'pembuatan', 'faktur', 'malam', 'melaporkan', 'ppn', 'menyulitkan']	['coretax', 'jalan', 'pagi', 'buat', 'faktur', 'malam', 'lapor', 'ppn', 'sulit']
---	---

Tabel 6. Hasil Stemming

g. Labeling

Labeling dilakukan untuk memberikan label sentimen. Setiap data akan dilabeli dengan sentimen 'positive', atau 'negative' berdasarkan skor yang diperoleh. Labeling dilakukan pada 553 data yang telah melalui proses pre-processing. Dengan begitu, berikut hasil distribusi sentimen yang telah dilakukan.



Gambar 3. Hasil Distribusi Sentimen Coretax

Sentimen	Jumlah
Negative	338
Positive	221
Total	559

Tabel 7. Hasil Labeling Data

h. Visualisasi

Gambar 5. Hasil TF-IDF

Kemudian setelah melakukan pembobotan dengan TF-IDF, tahap selanjutnya adalah melakukan seleksi fitur menggunakan metode Chi-Square (χ^2 test). Metode Chi-Square berfungsi untuk menilai hubungan antara masing-masing fitur (kata) dengan variabel target (label kelas). Dengan menggunakan metode Select K-Best, sejumlah fitur terbaik dipilih berdasarkan skor Chi-Square tertinggi. Fitur-fitur yang terpilih kemudian disimpan dalam sebuah dictionary bernama k-best feature, yang berisi nama fitur dan indeksinya sesuai dengan representasi TF-IDF. Selanjutnya, data tersebut dibuat dalam bentuk DataFrame agar lebih mudah dianalisis dan digunakan dalam pengembangan model klasifikasi yang lebih optimal.

index	fitur	indeks
0	sulf	604
1	akses	10
2	coretax	120
3	buruk	99
4	tasowa	274
5	mirah	372
6	slars	20
7	kendala	282
8	ganggu	192
9	pagi	434
10	lapor	333
11	pajak	437
12	hasil	213
13	agikasi	38
14	hambat	207
15	beban	63
16	kerja	285
17	rapot	515
18	error	167
19	jalan	248
20	berasmenerus	644
21	lamban	325
22	sabi	2
23	jangkel	261
24	sistem	579

Gambar 6. Hasil Chi-Square Feature

3.4 Hasil Data Mining

Dataset yang telah melalui proses pre-processing dan transformasi akan melalui proses modelling. Sebanyak 20% dari total data akan digunakan menjadi data pelatihan (training), dan 80% dari total data digunakan menjadi data pengujian (testing). Pembagian data dilakukan agar dapat mengevaluasi kinerja model secara akurat dan adil. Setelah itu, langkah selanjutnya adalah menggunakan algoritma Naïve Bayes dan SVM untuk mengklasifikasikan data uji guna memperoleh probabilitas dari sentimen yang terdeteksi.

A. Hasil Klasifikasi SVM



```
classification report:
              precision    recall  f1-score   support

     0         0.93         0.94         0.94         71
     1         0.90         0.88         0.89         41

 accuracy          0.92         112
 macro avg         0.92         0.91         0.91         112
 weighted avg      0.92         0.92         0.92         112
```

Gambar 7. Hasil Klasifikasi SVM

Berdasarkan hasil klasifikasi terhadap 112 data uji menggunakan algoritma SVM, dapat diketahui bahwa 71 sentimen negatif dan 41 sentimen positif. Hasil tersebut juga menunjukkan bahwa akurasi keseluruhan data sebanyak 92%, yang menunjukkan bahwa model yang digunakan mampu mengklasifikasikan 92% data dengan andal.

Pada data dengan sentimen negatif (label 0), presisi model adalah 93%, yang mengindikasikan bahwa keluaran model seakurat mungkin dengan kondisi yang diberikan. Di sisi lain, recall sebesar 94% mengindikasikan bahwa model dapat mengidentifikasi mayoritas data yang memiliki sentimen negatif. Dengan F1-score sebesar 94%, dapat disimpulkan bahwa model memiliki performa yang sangat baik dalam hal akurasi dan kemampuan untuk mendeteksi sinyal negatif. Secara keseluruhan, hasil tersebut menunjukkan bahwa model memiliki performa yang baik ketika menganalisis dan mengklasifikasikan data yang mengandung sentimen negatif.

Sebaliknya, untuk sentimen positif (label 1), model memiliki tingkat presisi 90%, yang mengindikasikan bahwa prediksi model terhadap sentimen positif setidaknya sama tingginya dengan kondisi yang mendasarinya. Recall sebesar 88% mengindikasikan bahwa model dapat secara akurat menganalisis 88% dari semua data yang jelas-jelas positif. Di sisi lain, F1-score sekitar 88% mengindikasikan bahwa performa model dalam mendeteksi dan mengklasifikasikan sentimen positif secara umum sudah baik, meskipun agak lebih buruk daripada sentimen negatif. Hal ini mengindikasikan bahwa model ini sudah cukup efektif dalam menghasilkan emosi positif, meskipun masih banyak yang harus dilakukan untuk meningkatkan kemampuan pendeteksiannya.

B. Hasil Klasifikasi Naive Bayes



```
classification report:
              precision    recall  f1-score   support

     0         0.90         0.99         0.94         71
     1         0.97         0.80         0.88         41

 accuracy         0.92         112
 macro avg         0.93         0.90         0.91         112
 weighted avg         0.92         0.92         0.92         112
```

Gambar 8. Hasil Klasifikasi Naive Bayes

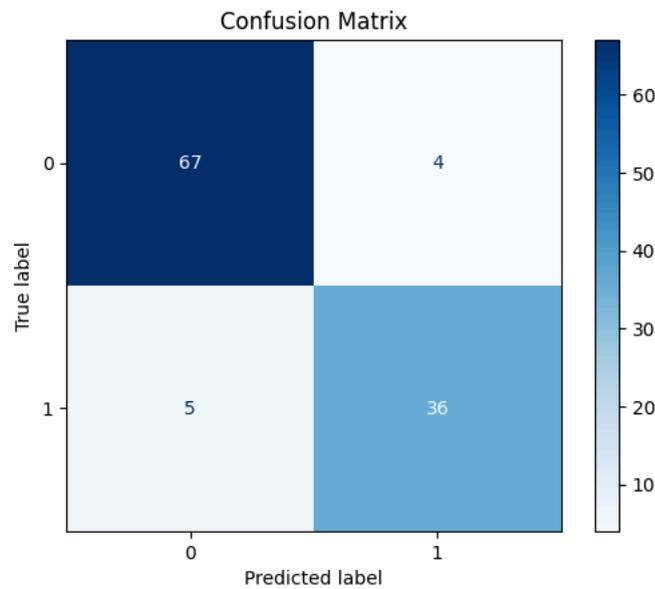
Berdasarkan hasil klasifikasi terhadap 112 data uji menggunakan algoritma Naive Bayes, dapat diketahui bahwa 71 sentimen negatif dan 41 sentimen positif. Hasil tersebut juga menunjukkan bahwa akurasi keseluruhan data sebanyak 92%, yang menunjukkan bahwa model yang digunakan mampu mengklasifikasikan 92% data dengan benar lebih besar daripada hasil klasifikasi menggunakan metode SVM.

Hasil yang diperoleh oleh data dengan sentimen negatif (label 0), nilai precision sebesar 90% untuk sentimen negatif (label 0), yang berarti sebagian besar prediksi negatif yang dihasilkan model sesuai dengan label sebenarnya. Sementara itu, recall pada kelas negatif mencapai 99%, menunjukkan bahwa model mampu mengenali mayoritas data dengan sentimen negatif. F1-score untuk kelas ini sebesar 94%, mencerminkan keseimbangan yang baik antara ketepatan dan cakupan model. Dengan performa tersebut, kami disimpulkan bahwa model sangat andal dalam mendeteksi dan mengklasifikasikan data yang mengandung sentimen negatif.

Sedangkan untuk sentimen positif (label 1), nilai precision sebesar 90% menunjukkan bahwa prediksi positif tergolong akurat, dan recall sebesar 80% menandakan bahwa sebagian besar data dengan label positif berhasil dikenali. F1-score pada kelas ini sebesar 88% mengindikasikan performa model yang cukup stabil. Dengan akurasi keseluruhan sebesar 92%, dapat disimpulkan bahwa model berkemampuan baik dan cukup seimbang dalam menganalisis data dengan sentimen negatif maupun positif.

3.5 Hasil Evaluation

A. Confusion Matrix SVM



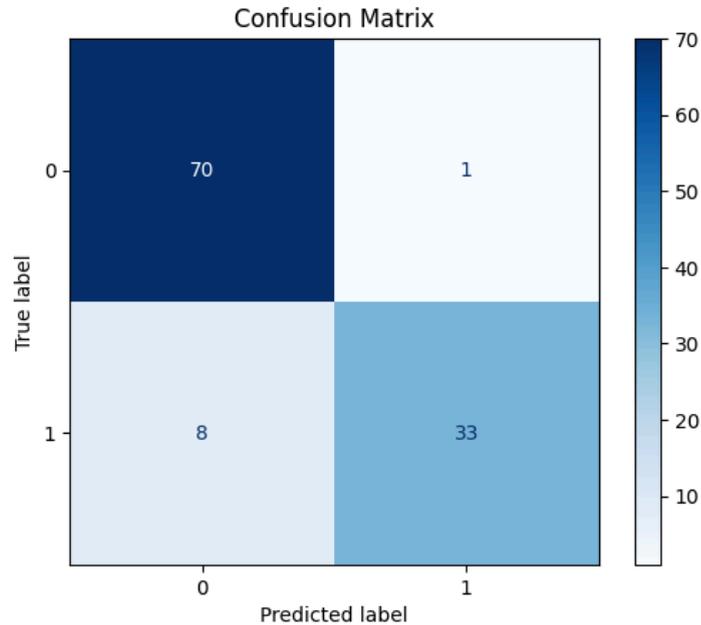
Dari hasil evaluasi yang diperoleh melalui confusion matrix, menunjukkan kinerja yang andal dalam menganalisis dua kategori sentimen, yaitu negatif (label 0) dan positif (label 1), dengan tingkat akurasi keseluruhan sebesar 92%. Model mampu mengidentifikasi dengan benar 71 data berlabel negatif dan 41 data berlabel positif. Dari total 112 data uji, berikut distribusinya.

- True Negative (TN): Dari seluruh data uji sebanyak 112, sejumlah 67 data negatif berhasil dikenali dan diprediksi sesuai label aslinya sebagai negatif
- True Positive (TP) : Dari seluruh data uji sebanyak 112, sejumlah 36 data positif berhasil dikenali dan diprediksi sesuai label aslinya sebagai positif
- False Negative (FN) : Dari seluruh data uji sebanyak 112, sejumlah 5 data positif salah dikenali dan diprediksi sebagai label negatif
- False Positive (FP) : Dari seluruh data uji sebanyak 112, sejumlah 4 data negatif salah dikenali dan diprediksi sebagai label positive

Secara umum, model menunjukkan kinerja klasifikasi yang solid dan proporsional, dengan tingkat akurasi yang tinggi serta nilai presisi dan recall yang konsisten dalam mengenali data positif. Model berhasil mengklasifikasikan sebagian besar data secara tepat, baik untuk kategori negatif maupun positif. Walaupun masih ditemukan sejumlah kecil kesalahan dalam prediksi kedua kelas, jumlahnya tergolong rendah,

sehingga model dapat dianggap cukup handal, meski tetap terbuka peluang untuk peningkatan lebih lanjut.

A. Confusion Matrix Naive Bayes



Gambar 9. Confusion Matrix

Hasil evaluasi yang diperoleh melalui confusion matrix, menunjukkan kinerja yang cukup andal dalam menganalisis dua kategori sentimen, yaitu negatif (label 0) dan positif (label 1), dengan tingkat akurasi keseluruhan sebesar 92%. Model mampu mengidentifikasi dengan benar 71 data berlabel negatif dan 41 data berlabel positif. Dari total 112 data uji, berikut distribusinya.

- True Negative (TN): Dari seluruh data uji sebanyak 112, sejumlah 70 data negatif berhasil dikenali dan diprediksi sesuai label aslinya sebagai negatif
- True Positive (TP) : Dari seluruh data uji sebanyak 112, sejumlah 33 data positif berhasil dikenali dan diprediksi sesuai label aslinya sebagai positif
- False Negative (FN) : Dari seluruh data uji sebanyak 112, sejumlah 8 data positif salah dikenali dan diprediksi sebagai label negatif
- False Positive (FP) : Dari seluruh data uji sebanyak 112, sejumlah 1 data negatif salah dikenali dan diprediksi sebagai label positive

Secara keseluruhan, model menunjukkan kinerja klasifikasi yang baik dengan akurasi tinggi serta ketepatan yang kuat dalam memprediksi data positif. Meski demikian,



kemampuan model dalam mengenali seluruh data positif belum maksimal, terlihat dari hasil recall tidak cukup baik. Kami simpulkan model masih perlu ditingkatkan pada aspek kepekaan terhadap data positif agar tidak melewatkan kasus-kasus yang seharusnya terdeteksi.

KESIMPULAN

Berdasarkan penelitian yang dilakukan dengan metode web scraping terhadap 684 ulasan pengguna aplikasi Coretax di Twitter menunjukkan bahwa sentimen negatif lebih dominan dibandingkan sentimen positif. Dari jumlah tersebut, sebanyak 338 ulasan tergolong negatif, sementara 221 lainnya bersentimen positif. Hasil ini sejalan dengan evaluasi terhadap 112 data uji, di mana 71 ulasan juga dikategorikan sebagai negatif. Banyaknya keluhan dari pengguna, seperti gangguan sistem, hambatan penggunaan, dan masalah teknis yang tercermin melalui kata-kata seperti "kendala", "alami", dan "gangguan", mengindikasikan adanya ketidakpuasan terhadap performa aplikasi. Temuan ini menjadi sinyal penting bagi pengembang untuk melakukan perbaikan guna meningkatkan pelayanan publik dan pengalaman pengguna.

Ketika algoritma Support Vector Machine (SVM) dan Naive Bayes digunakan secara bersamaan, tingkat akurasi mencapai sekitar 92%. Analisis terhadap 112 data uji dan hasil evaluasi menunjukkan bahwa algoritma SVM memberikan hasil yang sangat baik, dengan akurasi 92% dan keandalan untuk sinyal negatif (presisi 93%, recall 94%, dan F1-score 94%) dan sinyal positif (presisi 90%, recall 88%, dan F1-score 89%). Selain itu, algoritma Naive Bayes memberikan hasil klasifikasi yang tinggi dengan akurasi yang sebanding, yaitu 92% dengan rincian untuk sinyal negatif (precision 90%, recall 99%, dan F1-score 94%) dan sinyal positif (precision 97%, recall 80%, dan F1-score 88%).

Berdasarkan temuan penelitian di atas, dapat disimpulkan bahwa kedua model ini cukup efektif dalam menganalisis sentimen negatif, yang juga merupakan jenis sentimen yang paling banyak muncul di dalam data. Dominasi negatif ini disertai dengan munculnya kata-kata seperti kendala, buruk, dan gangguan, yang mengindikasikan adanya berbagai kesulitan teknis dan kesulitan dalam menggunakan aplikasi. Hasil penelitian ini memberikan informasi penting bagi para pengembang aplikasi untuk segera melakukan



pemeliharaan dan perbaikan sistem demi meningkatkan kualitas layanan dan memberikan pengalaman terbaik.

DAFTAR PUSTAKA

- [1] Ane, Tanjea. "A Review of Machine Learning Applications and Their Predictive Solutions in Agriculture." *Asian Journal of Advances in Agricultural Research*, vol. 24, no. 7, 1 July 2024, pp. 80-90,
<https://journalajaar.com/index.php/AJAAR/article/view/525>
- [2] Desiani, Anita, et al. "Penerapan Metode Support Vector Machine Dalam Klasifikasi Bunga Iris." *Indonesian Journal of Applied Informatics*, vol. 7, no. 1, 9 Apr. 2023, pp. 12-12,
<https://repository.unsri.ac.id/view/subjects/art=5Fop.html>
- [3] Elnaz Putri, Hilda Zaqya, and Hisyam Fahmi. "Implementasi Metode Support Vector Machine Pada Klasifikasi Diagnosis Penyakit Hipertensi." *Jurnal Riset Mahasiswa Matematika*, vol. 3, no. 5, 1 June 2024, pp. 241-250,
https://www.researchgate.net/publication/386119942_Implementasi_Metode_Support_Vector_Machine_pada_Klasifikasi_Diagnosis_Penyakit_Hipertensi
- [4] Matemilola, Alice Sikemi, and Salisu Aliyu. "Development of an Enhanced Naive Bayes Algorithm for Fake News Classification." *Science World Journal*, vol. 19, no. 2, 28 July 2024, pp. 512-517,
https://www.researchgate.net/publication/382804999_Development_of_an_enhanced_naive_bayes_algorithm_for_fake_news_classification
- [5] Sri Suryani Prasetyowati, and Yuliant Sibaroni. "Unlocking the Potential of Naive Bayes for Spatio Temporal Classification: A Novel Approach to Feature Expansion." *Journal of Big Data*, vol. 11, no. 1, 5 Aug. 2024
<https://journalofbigdata.springeropen.com/articles/10.1186/s40537-024-00958-x>
- [6] Trujillo, Fabricio, et al. "Artificial Intelligence in Education: A Systematic Literature Review of Machine Learning Approaches in Student Career Prediction." *Journal of Technology and Science Education*, vol. 15, no. 1, 7 Mar. 2025, pp. 162-162
https://www.researchgate.net/publication/389649687_Artificial_intelligence_in_education_A_systematic_literature_review_of_machine_learning_approaches_in_student_career_prediction
- [7] Putra, H. R., & Novayelinda. (2021). Analisis Hubungan Antara Tingkat Pendidikan dan Jenis Pekerjaan dengan Uji Chi-Square di Kabupaten Lima Puluh Kota. *Jurnal Statistika*



Universitas Andalas, 10(2), 55-61

[8] Alexander Pak, Patrick Paroubek. (2010). *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC)*. Retrieved from <https://www.researchgate.net/publication/220746311>

[9] Solihin, F., Awaliyah, S., & Shofa, A. M. A. (2021). Pemanfaatan Twitter Sebagai Media Penyebaran Informasi Oleh Dinas Komunikasi dan Informatika. *Jurnal Pendidikan Ilmu Pengetahuan Sosial (JPIPS)*, 13(1), 52-58.

<http://e-journal.upr.ac.id/index.php/JP-IPS>

[10] Wala, G. N., & Tesalonika, R. (2024). Transformasi Administrasi Perpajakan Melalui Coretax: Analisis Hukum dan Akuntansi. *JKIS*, 2(4), 149-158.

<https://doi.org/10.38035/jkis.v2i4>