

Analisis Sentimen Masyarakat Twitter Terhadap Kebijakan Efisiensi Anggaran Kementerian Menggunakan SVM

Muhammad Arif Kurniawan^{1*}, Samsul Makin², Angger Styo Yuniarti³, Andi Rukmana⁴,

^{1,2,3,4}Dosen Tetap, Universitas Insan Pembangunan
awan.insanpembangunan@gmail.com^{1*}, samsulmakin25@gmail.com²,
anggeruniarti06@gmail.com³, rukmana.andy@gmail.com⁴

Abstrak

Analisis sentimen terhadap kebijakan efisiensi anggaran kementerian sangat penting untuk memahami respons publik terhadap kebijakan pemerintah. Penelitian ini menggunakan metode **Support Vector Machine (SVM)** untuk mengklasifikasikan sentimen positif dan negatif dari **1.418 tweet** yang dikumpulkan melalui proses *crawling* menggunakan **Twitter API v2** pada periode **10–22 Februari**. Tahapan pengolahan teks yang dilakukan meliputi *case folding*, *cleaning*, *tokenizing*, *stopword removal*, *stemming*, serta pembobotan menggunakan metode **Term Frequency–Inverse Document Frequency (TF-IDF)**. Hasil analisis menunjukkan bahwa **sentimen negatif lebih dominan** dibandingkan sentimen positif, yang mencerminkan adanya kritik dan ketidakpuasan publik terhadap kebijakan tersebut. Model SVM dievaluasi menggunakan **k-fold cross-validation** dengan nilai k antara 2 hingga 10, dan memperoleh akurasi terbaik sebesar **94,76%** pada **10-fold validation**. Evaluasi menggunakan **confusion matrix** menghasilkan nilai **precision sebesar 92,85%**, **recall sebesar 91,32%**, serta **AUC sebesar 0,972**, yang menunjukkan kinerja model yang sangat baik dalam klasifikasi sentimen. Temuan ini menunjukkan bahwa model SVM efektif dalam menganalisis sentimen publik terhadap kebijakan pemerintah dan dapat dikembangkan lebih lanjut dengan memperkaya fitur serta membandingkannya dengan algoritma lain guna meningkatkan akurasi prediksi.

Kata kunci: Analisis Sentimen, Support Vector Machine, Efisiensi Anggaran, Twitter, Pengolahan Teks, Kebijakan Publik.

PENDAHULUAN

Dalam beberapa tahun terakhir, kebijakan efisiensi anggaran yang diterapkan oleh berbagai kementerian di Indonesia menjadi topik yang menarik untuk diperbincangkan di tengah masyarakat. Pemerintah sering kali melakukan penyesuaian anggaran untuk meningkatkan efektivitas penggunaan dana, mengurangi pemborosan, dan mengalokasikan sumber daya secara lebih optimal. Kebijakan semacam ini kerap kali mendapat respons beragam dari masyarakat, baik dalam bentuk dukungan maupun kritik. Di era digital saat ini, salah satu platform yang sering digunakan untuk menyampaikan pendapat terkait kebijakan pemerintah adalah media sosial, terutama twitter (wati dan ernawati, 2021).

Twitter telah menjadi media diskusi publik umum yang dinamis, sehingga masyarakat bisa dengan cepat merespons kebijakan pemerintah melalui cuitan yang bersifat opini, kritik, maupun apresiasi (Alhaq

dan Mustopa, 2021). Dengan sifatnya yang terbuka dan *real-time*, twitter menyediakan data berharga bagi pemerintah, akademisi, dan peneliti untuk memahami bagaimana suatu kebijakan diterima oleh publik. Sentimen masyarakat terhadap kebijakan efisiensi anggaran kementerian dapat memberikan wawasan mengenai sejauh mana kebijakan tersebut dipahami dan diterima oleh publik.

Analisis sentimen menjadi parameter pendekatan yang dapat digunakan untuk mendalami opini publik secara luas. Analisis ini memungkinkan identifikasi pola sentimen negatif, positif, maupun netral dari sekumpulan teks yang diperoleh dari media sosial (Widowati dan Sadikin, 2020). Hal ini membantu dalam memahami reaksi publik tanpa harus membaca satu per satu komentar yang berjumlah sangat besar. Dalam penelitian ini, peneliti memilih metode

Support Vector Machine (SVM) yang digunakan untuk melakukan pengelompokan sentimen. SVM merupakan salah satu algoritma *machine learning* yang dikenal efektif dalam mengklasifikasikan data berbasis teks, termasuk dalam analisis sentimen (Handayani dan Zufria, 2023). Algoritma ini bekerja dengan mencari hiperplane optimal yang dapat memisahkan berbagai kelas sentimen dalam ruang vektor tinggi (Samsir, dkk. 2021). Keunggulan SVM terletak pada kemampuan menangani data teks yang kompleks dan memberikan hasil pengelompokan yang akurat, menjadikannya pilihan yang tepat dalam studi analisis sentimen berbasis media sosial (Utami dan Erfina, 2021).

Penerapan analisis sentimen dalam konteks kebijakan publik, khususnya efisiensi anggaran kementerian, bisa memberikan pemahaman yang lebih mendalam mengenai bagaimana kebijakan tersebut sesuai dengan penerimaan oleh masyarakat (Pamungkas dkk., 2021). Pemahaman terhadap hal tersebut membantu pengambil kebijakan dalam mengevaluasi dampak kebijakan yang diterapkan serta merancang strategi komunikasi yang lebih efektif. Dengan memahami pola sentimen masyarakat, pemerintah dapat mengidentifikasi aspek yang perlu diperbaiki serta merancang pendekatan yang lebih responsif terhadap kebutuhan publik. Selain itu, penggunaan media sosial sebagai sumber data dalam penelitian kebijakan publik semakin relevan di era digital ini. Masyarakat kini lebih aktif menyuarakan pendapatnya melalui platform daring, sehingga informasi yang terkandung dalam media sosial dapat menjadi refleksi dari persepsi publik secara luas. Dengan meningkatnya volume data yang tersedia, metode analisis berbasis kecerdasan buatan seperti SVM menjadi semakin penting dalam menggali wawasan dari data yang besar dan tidak terstruktur (Noviana dan Rasal, 2023).

Beberapa penelitian terdahulu telah menggunakan analisis sentimen berbasis twitter dengan algoritma SVM untuk memahami opini publik terhadap berbagai isu kebijakan. Misalnya, penelitian sebelumnya menganalisis sentimen masyarakat terhadap penyedia jasa penerbangan di *platform* twitter menggunakan SVM dan menunjukkan bahwa algoritma ini mampu memberikan akurasi tinggi dalam mengklasifikasikan opini publik (Husada dan

Paramita, 2021). Selain itu, penelitian lain juga memanfaatkan SVM dalam mengkaji sentimen mengenai kebijakan pemindahan ibu kota, di mana hasilnya menunjukkan bahwa SVM memiliki hasil performa yang tepat dalam mengolah teks tidak terstruktur dari twitter dan membedakan sentimen positif, negatif, serta netral dengan presisi yang tinggi (Arsi dan Waluyo, 2021).

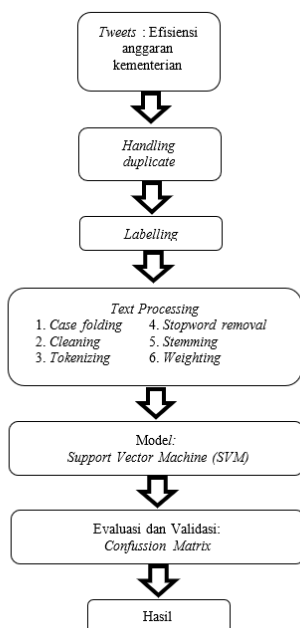
Namun, sebagian besar penelitian sebelumnya masih berfokus pada kebijakan ekonomi makro atau kesehatan, sementara penelitian terkait sentimen masyarakat terhadap kebijakan efisiensi anggaran kementerian masih terbatas. Beberapa penelitian menggunakan metode klasifikasi lain seperti *Naïve Bayes* atau *Random Forest*, tetapi hasilnya cenderung kurang akurat dibandingkan SVM dalam menangani data teks yang kompleks (Fikri dkk., 2020). Sehingga, penelitian ini bertujuan untuk mengisi kesenjangan tersebut dengan menerapkan analisis sentimen berbasis SVM untuk mengeksplorasi opini masyarakat di twitter mengenai kebijakan efisiensi anggaran kementerian. Meskipun menggunakan jumlah dataset yang terbatas, algoritma Support Vector Machine terbukti mampu menghasilkan performa yang baik dalam klasifikasi sentimen biner (positif dan negatif) dengan memanfaatkan pemisahan bidang (*hyperplane*) yang optimal antara kedua kelas tersebut (Syahputra dkk., 2022).

Penelitian ini bertujuan untuk mengembangkan pendekatan analisis sentimen berbasis SVM guna mengklasifikasikan opini masyarakat terkait kebijakan efisiensi anggaran kementerian. selanjutnya, penelitian ini juga memiliki tujuan untuk mengevaluasi efektivitas SVM dibandingkan dengan metode klasifikasi lainnya yang umum digunakan dalam analisis sentimen. Hasil dari penelitian yang dilakukan diharapkan bisa menjadi referensi bagi pemerintah dalam memahami pola reaksi publik serta implikasi dari kebijakan efisiensi anggaran yang diterapkan. Dengan mengetahui respons masyarakat secara lebih mendalam, pemerintah dapat menyusun strategi komunikasi kebijakan yang lebih efektif serta mengantisipasi potensi resistensi dari publik (Darwis dkk., 2020).

Selain itu, penelitian yang dilakukan juga dapat memberikan kontribusi bagi akademisi dalam mengembangkan model analisis sentimen yang lebih tepat dan aplikatif dalam konteks kebijakan publik. Dengan demikian, pendekatan berbasis data dalam evaluasi kebijakan dapat semakin dioptimalkan untuk bisa mendukung proses pengambilan kebijakan yang lebih tepat sasaran dan responsif terhadap aspirasi masyarakat.

METODOLOGI PENELITIAN

Dalam penelitian ini, akan digunakan metode analisis sentimen berbasis *Support Vector Machine (SVM)* untuk mengklasifikasikan opini masyarakat terhadap kebijakan efisiensi anggaran kementerian yang diungkapkan melalui twitter. Metode yang digunakan melibatkan beberapa tahapan utama, yaitu pengumpulan data (*crawling*), pemrosesan awal (*preprocessing*), pembagian data menjadi *training* dan *testing*, serta validasi dan evaluasi model menggunakan metrik yang sesuai (Fitriyah dkk., 2020). Dengan metode seperti yang terdapat pada Gambar 1, diharapkan dapat diperoleh pemahaman yang lebih mendalam mengenai sentimen publik terhadap kebijakan tersebut serta meningkatkan akurasi dalam klasifikasi sentimen.



Gambar 1. Metode Penelitian Dataset

Dataset yang akan digunakan dalam penelitian ini diperoleh dari median sosial twitter dengan menggunakan beberapa kata kunci terkait kebijakan efisiensi anggaran kementerian, seperti "efisiensi anggaran", "pemotongan anggaran kementerian", dan "anggaran pemerintah". Dataset terdiri dari teks cuitan beserta atribut tambahan seperti *timestamp*, jumlah *retweet*, jumlah *like*, dan *user ID* (yang dianonimkan untuk menjaga privasi) (Supian dkk., 2024).

Crawling Data

Proses dalam pengambilan data dilakukan melalui *crawling* menggunakan Twitter API v2 dengan bantuan pustaka *Python*. *Crawling* dilakukan pada rentang waktu 10–22 Februari 2025, dengan total 1.423 *tweet* yang berhasil dikumpulkan. Data diambil berdasarkan beberapa kata kunci yang sudah ditentukan sebelumnya untuk memastikan relevansi terhadap topik penelitian (Setiawan dkk., 2021). Selain itu, dilakukan penyaringan terhadap cuitan yang mengandung unsur spam atau tidak relevan, seperti promosi produk atau postingan berulang dari akun yang sama (Styawati dkk., 2021).

TABEL 1. Tweet Topik Kebijakan Efisiensi Anggaran Kementerian

	Tweet
1	Gak perlu dipuji, Bare minimum ini dimanapun, gak ada yang spesial, biasa aja. Dimana-mana ketika melakukan efisiensi anggaran, negara yang normal ya melakukan ini Lho, kok di Indonesia gak gini? ya kita gak normal aja. Udah.
2	Efisiensi Anggaran tidak menasar belanja pegawai. Ingat ya kawan jangan terprovokasi hoaks!
3	Efisiensi bisa dilakukan Pemerintah demi prioritas anggaran ke depan #EfisiensiUntukBangsa #EfisiensiCegahKorupsi #EfisiensiUntukRakyat #indonesiagelap
4	Pemerintah Tidak Berpihak Kepada Rakyat: Efisiensi Anggaran Yang Mengambil Hak Rakyat Dan Komunikasi Yang Kasar Kepada Rakyat
5	Anggaran untuk rakyat dipangkas habis hingga pintu perlitasan kereta api tak dijaga lagi atas nama efisiensi. Menteriya sendiri malah bebas foya-foya dengan Dewa 19.
...
1422	Dengan efisiensi anggaran, kita bisa capai stabilitas ekonomi yang kita impikan
1423	Optimisme, dengan efisiensi anggaran ekonomi kita bakal makin tangguh

Preprocessing Data

Tahap preprocessing dilakukan untuk dapat meningkatkan kualitas data sebelum dianalisis lebih lanjut (Nasution dkk., 2019). Langkah-langkah yang dilakukan dalam tahap ini meliputi:

1. *Handling Duplicate*: Menghapus data yang memiliki konten identik untuk menghindari bias dalam analisis.
2. *Labelling*: Menentukan kategori sentimen setiap teks, seperti *negatif*, *positif*, atau *netral*, secara manual atau menggunakan teknik berbasis *lexicon*.
3. *Case Folding*: Mengubah semua teks yang ada menjadi huruf kecil untuk menyamakan format.
4. *Cleaning*: Menghapus beberapa karakter yang tidak diperlukan seperti angka, tanda baca, emoji, serta simbol yang tidak relevan.
5. *Tokenizing*: Memecah atau memisahkan teks menjadi kata-kata individu untuk mempermudah proses analisis lebih lanjut.
6. *Stopword Removal*: Menghapus kata-kata umum yang tidak memiliki makna signifikan dalam analisis (misalnya: "yang", "dan", "di", dll.).
7. *Stemming*: Mengubah beberapa kata menjadi bentuk dasar menggunakan algoritma *stemming* Bahasa Indonesia (misalnya: "mengurangi" menjadi "kurang").
8. *Weighting (TF-IDF)*: Mengonversi teks yang telah dibersihkan ke dalam bentuk vektor fitur menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)* untuk memberi bobot pada kata-kata sesuai tingkat kepentingannya dalam dokumen.

Dengan menerapkan langkah- langkah ini, maka data menjadi lebih terstruktur dan bisa digunakan untuk dianalisis menggunakan model klasifikasi.

Pembagian Data: Training dan Testing

Dataset yang sebelumnya telah diproses dibagi menjadi dua bagian, yaitu data uji (*testing data*) dan data latih (*training data*) (Musthafa dkk., 2025). Rasio pembagian yang digunakan adalah 90:10, dengan porsi 90% data digunakan untuk melatih model (*training*), sedangkan 10% sisanya digunakan untuk menguji tingkat performa model (*testing*) (Safitri dkk., 2023). Pembagian yang dilakukan ini secara acak untuk memastikan bahwa model bisa belajar dari variasi data yang beragam.

Validasi dan Evaluasi Model

Validasi model dilakukan dengan teknik *k-fold validation*, yang membagi dataset menjadi *k* bagian sama besar, di mana setiap iterasi model dilatih dengan *k-1* bagian dan diuji dengan bagian sisanya (Pravina dkk., 2023). Dalam penelitian ini, digunakan *10-fold validation*, di mana proses validasi berlangsung 10 kali dengan kombinasi data training dan testing yang berbeda di setiap iterasi untuk mencegah *overfitting* dan meningkatkan stabilitas akurasi.

Evaluasi terhadap model dilakukan dengan menggunakan *confusion matrix* yang mengukur performa klasifikasi berdasarkan *True Negative (TN)*, *True Positive (TP)*, *False Negative (FN)*, dan *False Positive (FP)* (Novantirani dkk., 2015). Dari matriks yang didapatkan, dihitung metrik evaluasi seperti akurasi, *recall*, *precision*, dan *F1-score* guna menilai ketepatan serta kemampuan model dalam mengklasifikasikan data. Dengan kombinasi validasi dan evaluasi ini, penelitian memastikan bahwa model yang diusulkan memiliki kinerja optimal dalam analisis sentimen.

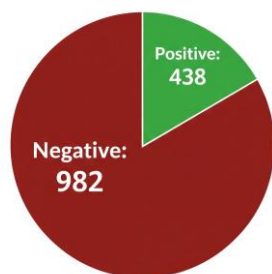
HASIL DAN PEMBAHASAN

Pada tahap awal analisis, dilakukan *handling duplicate* untuk menghapus *tweet* ganda, terutama yang berasal dari *retweet*, sehingga jumlah data berkurang dari 1423 menjadi 1418 sebelum masuk ke tahap selanjutnya. Setelah itu, dilakukan *labeling* data secara manual oleh dua annotator, di mana annotator pertama mengklasifikasikan *tweet* ke dalam sentimen negatif dan positif, sementara annotator kedua melakukan *cross-check* untuk memastikan konsistensi hasil. Sentimen positif diberikan pada *tweet* yang mendukung wacana kebijakan efisiensi anggaran, sedangkan negatif untuk *tweet* yang menolak atau mengkritisi kebijakan tersebut. Hasil akhir pelabelan menampilkan jumlah *tweet* untuk masing-masing kategori sentimen dan disajikan dalam bentuk tabel guna memperlihatkan sampel hasil klasifikasi.

Hasil Text Processing

Proses *text processing* dilakukan untuk dapat memastikan data siap digunakan dalam analisis sentimen. Tahapan

ini mencakup berbagai teknik pemrosesan teks yang bertujuan meningkatkan kualitas data sebelum masuk ke algoritma *Support Vector Machine* (SVM). Setiap langkah dalam *text processing* berpengaruh terhadap hasil akhir analisis sentimen, baik dalam aspek akurasi maupun efektivitas pemodelan. Proses preprocessing seperti case folding dan cleansing bukan sekadar tahap awal, melainkan langkah krusial untuk menyeragamkan fitur kata dan menghapus kebisingan (noise) seperti simbol atau tautan yang tidak memiliki nilai semantik dalam penentuan sentimen masyarakat (Adilah dkk., 2023). Gambar 2 memperlihatkan sentimen data yang akan melalui *text processing*.



Gambar 2. Data sentimen tweet setelah handle duplicates

Case Folding

Seluruh teks yang ada diubah menjadi huruf kecil (*lowercase*) untuk memastikan format teks seragam. Proses ini membantu mengurangi perbedaan akibat perbedaan kapitalisasi kata. Sampel hasil *case folding* dapat terlihat pada Tabel 2.

TABEL 2. CASE FOLDING

Sebelum	Sesudah
Efisiensi anggaran, Danantara, lalu Bank Emas.. Dengan kabinet gemuk berisi orang asal njeplak, gak peka dan nir empati, apa kita bisa percaya nitipin harta kita ke mereka? #IndonesiaGelap #KaburAjaDulu Kawal perjuangan mahasiswa https://cnnindonesia.com/ekonomi/20250218105812-92-1199513/prabowo-kita-sekarang-punya-bank-emas-kita-resmikan-26-februari	efisiensi anggaran, danantara, lalu bank emas.. dengan kabinet gemuk berisi orang asal njeplak, gak peka dan nir empati, apa kita bisa percaya nitipin harta kita ke mereka? #indonesiagelap #kaburajadulu kawal perjuangan mahasiswa https://cnnindonesia.com/ekonomi/20250218105812-92-1199513/prabowo-kita-sekarang-punya-bank-emas-kita-resmikan-26-februari

Cleaning

Teks dibersihkan dari elemen yang tidak relevan, seperti URL, *username* (@mention), *retweet* (RT), karakter HTML, dan *hashtag* (#). Langkah ini bertujuan menghilangkan *noise* yang dapat mengganggu hasil analisis. Sampel hasil *cleaning* dapat terlihat pada Tabel 3.

TABEL 3. CLEANING

Sebelum	Sesudah
efisiensi anggaran, danantara, lalu bank emas.. dengan kabinet gemuk berisi orang asal njeplak, gak peka dan nir empati, apa kita bisa percaya nitipin harta kita ke mereka? #indonesiagelap #kaburajadulu kawal perjuangan mahasiswa https://cnnindonesia.com/ekonomi/20250218105812-92-1199513/prabowo-kita-sekarang-punya-bank-emas-kita-resmikan-26-februari	efisiensi anggaran danantara lalu bank emas dengan kabinet gemuk berisi orang asal njeplak gak peka dan nir empati apa kita bisa percaya nitipin harta kita ke mereka kawal perjuangan mahasiswa

Tokenizing

Kalimat dipecah menjadi kata-kata individu berdasarkan tanda baca dan spasi. Proses ini mempermudah analisis pada tingkat kata per kata. Sampel hasil *tokenizing* dapat terlihat pada Tabel 4.

TABEL 4. TOKENIZING

Sebelum	Sesudah
efisiensi anggaran danantara lalu bank emas dengan kabinet gemuk berisi orang asal njeplak gak peka dan nir empati apa kita bisa percaya nitipin harta kita ke mereka kawal perjuangan mahasiswa	efisiensi, anggaran, danantara, lalu, bank, emas, dengan, kabinet, gemuk, berisi, orang, asal, njeplak, gak, peka, dan, nir, empati, apa, kita, bisa, percaya, nitipin, harta, kita, ke, mereka, kawal, perjuangan, mahasiswa

Stopword Removal

Menghapus beberapa kata-kata umum yang tidak mempunyai makna signifikan dalam analisis, seperti "yang", "ke", "di", "dan", untuk mengurangi jumlah kata yang diproses dan meningkatkan efisiensi algoritma. Sampel hasil *stopword removal* dapat terlihat pada Tabel 5.

TABEL 5. STOPWORD REMOVAL

Sebelum	Sesudah
efisiensi anggaran danantara lalu bank emas dengan kabinet gemuk berisi orang asal njeplak gak peka dan nir empati apa kita bisa percaya nitipin harta kita ke mereka kawal perjuangan mahasiswa	efisiensi, anggaran, danantara, bank, emas, kabinet, gemuk, berisi, orang, asal, njeplak, gak, peka, nir, empati, percaya, nitipin, harta, mereka, kawal, perjuangan, mahasiswa

Stemming

Setiap kata diubah menjadi bentuk dasar menggunakan metode stemming Bahasa Indonesia. Contohnya, kata "mengurangi" diubah menjadi "kurang", sehingga variasi kata dapat diminimalkan dan analisis lebih akurat. Sampel hasil dari *stemming* dapat terlihat pada Tabel 6.

TABEL 6. STEMMING

Sebelum	Sesudah
efisiensi, anggaran, danantara, bank, emas, kabinet, gemuk, berisi, orang, asal, njeplak, gak, peka, nir, empati, percaya, nitipin, harta, mereka, kawal, perjuangan, mahasiswa	efisiensi, anggaran, danantara, bank, emas, kabinet, gemuk, isi, orang, asal, njeplak, gak, peka, nir, empati, percaya, titip, harta, mereka, kawal, juang, mahasiswa

Weighting

Metode *Term Frequency-Inverse Document Frequency* (TF-IDF) digunakan untuk memberi bobot pada kata-kata dalam teks. (Septiani dkk., 2022). Bobot ini menentukan seberapa penting sebuah kata dalam dokumen, sehingga kata-kata dengan relevansi tinggi mendapatkan perhatian lebih besar dalam analisis sentimen. Tabel 7 dan 8.

TABEL 7. PENENTUAN IDF

Token	Kk	Tf D1	Tf D2	Tf D3	Df	d/Df	IDF (log d/Df)
anggaran	6	2	1	0	2	3/2	0.1761
bank	4	1	1	2	3	3/3	0
emas	3	0	2	1	2	3/2	0.1761
kabinet	2	1	0	1	2	3/2	0.1761
mahasiswa	1	1	1	0	2	3/2	0.1761

TABEL 8. PENENTUAN NILAI W (BOBOT SETIAP KATA)

Token	Kk	W D1	W D2	W D3
anggaran	6	0.3522	0.1761	0
bank	4	0	0	0
emas	3	0	0.3522	0.1761
kabinet	2	0.1761	0	0.1761
mahasiswa	1	0.1761	0.1761	0

Eksperimen Model Support Vector Machine (SVM)

Model SVM diterapkan untuk menganalisis sentimen terhadap kebijakan efisiensi anggaran kementerian. Eksperimen dilakukan dengan mengatur parameter C dan Epsilon untuk mendapatkan performa optimal. Hasil terbaik diperoleh dengan kombinasi C = 0,5 dan Epsilon = 0,5, menghasilkan akurasi 94,76%.

Validasi dan Evaluasi

Validasi dilakukan menggunakan *k-fold cross validation* dengan nilai k = 2 hingga k = 10, dan hasil terbaik diperoleh pada *10-fold validation*. Dalam pembagian data, digunakan rasio 9:1, di mana 1.276 data digunakan sebagai *training* dan 142 data digunakan sebagai *testing*. Hasil evaluasi memberikan gambaran bahwa model memiliki akurasi yang mencapai 94,76%, yang menunjukkan bahwa model memiliki tingkat kesalahan yang rendah dalam mengklasifikasikan sentimen.

Precision yang diperoleh adalah 92,85%, yang berarti dari semua prediksi sentimen positif, 92,85% di antaranya benar-benar positif. Sementara itu, *recall* mencapai 91,32%, menunjukkan bahwa model mampu

mengidentifikasi 91,32% dari total data yang benar-benar memiliki sentimen positif. Selain itu, nilai *AUC (Area Under Curve)* sebesar 0,972 menunjukkan bahwa model ini memiliki performa yang sangat baik dalam membedakan antara sentimen negatif dan positif.

Sebagai tambahan, evaluasi menggunakan *ROC Curve* mengonfirmasi bahwa model memiliki tingkat optimasi yang tinggi, dengan nilai *AUC* mendekati 1,0, yang menandakan kemampuan model dalam mengklasifikasikan sentimen secara akurat. Hasil validasi ini menunjukkan bahwa model yang dikembangkan mampu memberikan prediksi yang cukup andal terhadap data uji, sehingga dapat digunakan untuk analisis sentimen lebih lanjut.

KESIMPULAN

Penelitian yang dilakukan ini menunjukkan bahwa analisis sentimen terhadap kebijakan efisiensi anggaran kementerian menggunakan metode *support vector machine* dapat memberikan hasil yang akurat dengan tingkat akurasi 94,76%, *precision* 92,85%, *recall* 91,32%, dan *AUC* 0,972. Proses *text processing*, yang mencakup *case folding*, *cleaning*, *tokenizing*, *stopword removal*, *stemming*, dan *weighting*, berperan cukup signifikan dalam meningkatkan kualitas data sebelum bisa diterapkan pada model. Hasil analisis memberikan gambaran bahwa sentimen negatif muncul lebih dominan dibandingkan dengan sentimen positif, yang mencerminkan adanya ketidakpuasan masyarakat terhadap kebijakan tersebut. Validasi menggunakan *k-fold cross validation* dengan *10-fold validation* memberikan hasil terbaik, menunjukkan bahwa model yang dikembangkan memiliki performa yang optimal dalam mengklasifikasikan sentimen.

Untuk penelitian selanjutnya, disarankan untuk menggunakan *dataset* yang lebih besar dan beragam agar model dapat mengenali lebih banyak variasi sentimen dalam masyarakat. Selain itu, dapat dilakukan eksperimen dengan algoritma lain, seperti *LSTM* atau *BERT*, untuk melihat apakah ada peningkatan akurasi dibandingkan dengan *support vector machine*. Penelitian juga dapat memperkaya

fitur dengan mempertimbangkan aspek emosional dan subjektivitas dalam teks, serta menambahkan analisis temporal untuk melihat bagaimana sentimen berubah seiring waktu. Dengan demikian, hasil analisis dapat memberikan wawasan yang lebih komprehensif bagi pengambil kebijakan dalam memahami persepsi publik terhadap kebijakan yang diterapkan.

DAFTAR PUSTAKA

- Adilah, H. S., & Alit, R. (2023). Analisis Sentimen Masyarakat Twitter Terhadap Kebijakan Pemerintah Dalam Menaikkan Harga Bahan Bakar Minyak Dengan Menggunakan Metode Support Vector Machine. *JINACS: Journal of Informatics and Computer Science*, 5(02), 120-128.
- Alhaq, Z., Mustopa, A., Mulyatun, S., & Santoso, J. D. (2021). Penerapan Metode Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter. *Journal of Information System Management (JOISM)*, 3(1), 16-21.
- Arsi, P., & Waluyo, R. (2021). Analisis sentimen wacana pemindahan ibu kota Indonesia menggunakan algoritma Support Vector Machine (SVM). *J. Teknol. Inf. dan Ilmu Komput*, 8(1), 147.
- Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). Penerapan Algoritma Svm Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. *Jurnal Ilmiah Edutic: Pendidikan dan Informatika*, 7(1), 1-11.
- Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Perbandingan metode naïve bayes dan support vector machine pada analisis sentimen twitter. *SMATIKA Jurnal: STIKI Informatika Jurnal*, 10(02), 71-76.
- Fitriyah, N., Warsito, B., & Di Asih, I. M. (2020). Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (SVM). *Jurnal Gaussian*, 9(3), 376-390.
- Handayani, A., & Zufria, I. (2023). Analisis sentimen terhadap bakal capres ri 2024 di twitter menggunakan algoritma svm. *J. Inf. Syst. Res*, 5(1), 53-63.
- Husada, H. C., & Suryaputra Paramita, A. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM).
- Musthafa, A., Harmini, T., Rafiq, A., & Marantika, N. (2025). Pemanfaatan Machine Learning dalam Menganalisis Sentimen Terhadap Program TAPER A di Platform Digital X: Utilization of Machine Learning in Analyzing Sentiment Towards the TAPER A Program on Digital X Platform. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 5(2), 587-597.
- Nasution, M. R. A., & Hayaty, M. (2019). Perbandingan akurasi dan waktu proses algoritma K-NN dan SVM dalam analisis sentimen twitter. *Jurnal Informatika*, 6(2), 226-235.
- Ningsih, W., Alfianda, B., Rahmaddeni, R., & Wulandari, D. (2024). Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Twitter pada Penggunaan Mobil Listrik di Indonesia: Comparison of Naïve Bayes and SVM Algorithms in Twitter Sentiment Analysis on Electric Car Use in Indonesia. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 556-562.
- Novantirani, A., Sabariah, M. K., & Effendy, V. (2015). Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota dengan Metode Support Vector Machine. *eProceedings of Engineering*, 2(1).

- Noviana, R., & Rasal, I. (2023). Penerapan Algoritma Naive Bayes Dan Svm Untuk Analisis Sentimen Boy Band Bts Pada Media Sosial Twitter. *Jurnal Teknik dan Science*, 2(2), 51-60.
- Pamungkas, B., Purbaya, M. E., & AK, D. J. (2021). Analisis Sentimen Twitter Menggunakan Metode Support Vector Machine (SVM) pada Kasus Benih Lobster 2020. *Journal of Informatics Information System Software Engineering and Applications (INISTA)*, 3(2), 10-20.
- Pravina, A. M., Cholisoddin, I., & Adikara, P. P. (2019). Analisis sentimen tentang opini maskapai penerbangan pada dokumen twitter menggunakan algoritme support vector machine (svm). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3(3), 2789-2797.
- Safitri, T., Umaidah, Y., & Maulana, I. Analisis Sentimen Pengguna Twitter Terhadap BTS Menggunakan Algoritma Support Vector Machine. *Journal of Applied Informatics and Computing (JAIC)*, 7(1), 34-41.
- Samsir, A., Verawardina, U., Edi, F., & Watianthos, R. (2021). Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes. *Jurnal Media Informatika Budidarma*, 5(1), 157-163.
- Septiani, D., & Isabela, I. (2022). Analisis term frequency inverse document frequency (tf-idf) dalam temu kembali informasi pada dokumen teks. *Sistem dan Teknologi Informasi Indonesia (SINTESIA)*, 1(2), 81-88.
- Setiawan, H., Utami, E., & Sudarmawan, S. (2021). analisis sentimen twitter kuliah online pasca covid-19 menggunakan algoritma support vector machine dan naive bayes. *Jurnal Komtika (Komputasi dan Informatika)*, 5(1), 43-51.
- Styawati, S., Hendrastuty, N., & Isnain, A. R. (2021). Analisis sentimen masyarakat terhadap program kartu prakerja pada twitter dengan metode support vector machine. *Jurnal Informatika: Jurnal Pengembangan IT*, 6(3), 150-155.
- Supian, A., Revaldo, B. T., Marhadi, N., Efrizoni, L., & Rahmaddeni, R. (2024). Perbandingan Kinerja Naïve Bayes Dan Svm Pada Analisis Sentimen Twitter Ibukota Nusantara. *Jurnal Ilmiah Informatika*, 12(01), 15-21.
- Syahputra, D. W., Rahayudi, B., & Muflikhah, L. (2022). Analisis Sentimen Twitter terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat menggunakan Metode Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(3), 1067-1072.
- Utami, D. S., & Erfina, A. (2021, September). Analisis Sentimen Pinjaman Online di Twitter Menggunakan Algoritma Support Vector Machine (SVM). In *Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra* (Vol. 1, pp. 299-305).
- Wati, R., & Ernawati, S. (2021). Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python. *Jurnal Teknik Informatika UNIKA Santo Thomas*, 240-247.
- Widowati, T. T., & Sadikin, M. (2020). Analisis Sentimen Twitter terhadap Tokoh Publik dengan Algoritma Naive Bayes dan Support Vector Machine. *Jurnal SIMETRIS*, 11(2), 626-636.