Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)
Vol. 9 No. 2, April 2026
P-ISSN 2620-8342
E-ISSN 2621-3052

Analisis Sentimen Terhadap Kebijakan Efisisensi Anggaran Pemerintah Menggunakan Metode Support Vector Machine (SVM) Dan *Vector Space Model* (VSM)

Dwi Fanti Y. Utina[1*], Alter Lasarudin[2], Wahyudin Hasyim[3]
[123]Sistem Informasi, Universitas Muhammadiyah Gorontalo
Jl. Prof. Dr. H. Mansoer Pateda, Desa Pentadio Timur, Kec. Telaga Biru, Kab.Gorontalo, Gorontalo 96181, Indonesia
E-mail: dwifantiutina2@gmail.com

Submitted Date : 13 April 2026 ***Accepted Date : 20 April 2026***

Abstrak - Media sosial saat ini menjadi wadah utama bagi masyarakat dalam menyuarakan aspirasi terhadap kebijakan pemerintah secara terbuka, termasuk isu kebijakan efisiensi anggaran. Penelitian ini bertujuan menganalisis kecenderungan sentimen publik terhadap kebijakan pemerintah menggunakan algoritma *Support Vector Machine* (SVM) yang diintegrasikan dengan *Vector Space Model* (VSM). Data dikumpulkan melalui teknik *crawling* pada *platform* X (Twitter), kemudian diproses melalui tahapan *cleaning*, *case folding*, *tokenizing*, *filtering*, dan *stemming*. Penelitian ini juga menerapkan pendekatan *Lexicon-Based* untuk pelabelan sentimen. Seluruh implementasi sistem dikembangkan menggunakan bahasa *Python* pada *Google Colaboratory*. Hasil pengujian dengan confusion matrix menunjukkan bahwa model mampu mengklasifikasikan data berlabel mandiri dengan akurasi 71%, dan mencapai 96% pada data berlabel *Lexicon-Based Approach*. Hal ini membuktikan integrasi VSM sangat membantu SVM dalam menyajikan representasi fitur dan mengenali pola teks secara akurat. Penelitian ini menyimpulkan bahwa persepsi masyarakat dapat dipetakan secara objektif melalui analisis sentimen, sehingga dapat menjadi bahan evaluasi dan referensi strategis bagi pemerintah dalam mengambil kebijakan di masa depan.
Kata Kunci: Analisis Sentimen, *Support Vector Machine*, Efisiensi Anggaran, X (Twitter), *Vector Space Model*.

Abstract - *Social media is currently the primary platform for the public to openly voice their concerns about government policies, including issues of budget efficiency. This study aims to analyze public sentiment trends toward government policies using the Support Vector Machine (SVM) algorithm integrated with the Vector Space Model (VSM). Data was collected through crawling techniques on the X (Twitter) platform, then processed through cleaning, case folding, tokenizing, filtering, and stemming. This study also applied a Lexicon-Based approach for sentiment labeling. The entire system implementation was developed using Python on Google Collaboratory. Test results using a confusion matrix showed that the model was able to classify self-labeled data with 71% accuracy and achieved 96% accuracy on data labeled with the Lexicon-Based Approach. This demonstrates that the VSM integration significantly assists the SVM in presenting feature representations and accurately recognizing text patterns. This study concludes that public perception can be objectively mapped through sentiment analysis, thus serving as evaluation material and a strategic reference for the government in future policymaking.*
Keywords: Sentiment Analysis, Support Vector Machine, Budget Efficiency, X (Twitter), Vector Space Model.

1. Pendahuluan

Media sosial saat ini menjadi salah satu wadah utama bagi masyarakat dalam menyuarakan pendapat, menyampaikan aspirasi, serta menanggapi berbagai fenomena dan kebijakan pemerintah secara terbuka. Platform seperti X (Twitter), Facebook, dan Instagram memungkinkan opini publik tersebar secara cepat, masif, dan dapat diakses secara real-time. Hal ini menjadikan media sosial sebagai indikator penting dalam membaca persepsi masyarakat terhadap isu-isu nasional yang sedang berlangsung.

Salah satu kebijakan pemerintah yang banyak menjadi bahan perbincangan publik pada tahun 2025 adalah kebijakan efisiensi anggaran dalam APBN 2025. Kebijakan ini ditegaskan melalui Instruksi Presiden Nomor 1 Tahun 2025, yang mengarahkan seluruh kementerian atau lembaga serta pemerintah daerah untuk melakukan pemangkasan belanja non-prioritas, efisiensi operasional, serta pengalihan anggaran menuju sektor-sektor yang dianggap lebih strategis dan berdampak jangka panjang bagi pembangunan nasional. Pemerintah menargetkan penghematan hingga sekitar Rp 306,69 triliun, yang diharapkan menciptakan ruang fiskal untuk mendukung program-program prioritas nasional dan menjaga kestabilan keuangan negara.[1]

Namun, kebijakan ini memunculkan berbagai reaksi dari masyarakat. Sebagian pihak menilai langkah efisiensi anggaran sebagai upaya positif dalam meningkatkan akuntabilitas keuangan negara dan mengurangi pemborosan belanja pemerintah. Di sisi lain, terdapat kekhawatiran mengenai potensi penurunan kualitas

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

layanan publik, perlambatan program pemerintah di daerah, serta dampak terhadap kesejahteraan ASN dan pegawai honorer. Perbedaan pandangan ini semakin tampak melalui diskusi yang marak di media sosial, di mana masyarakat secara aktif menyampaikan dukungan, kritik, maupun opini netral terkait kebijakan tersebut.

Fenomena polarisasi opini publik ini menjadi menarik untuk diteliti secara ilmiah melalui pendekatan analisis sentimen, sebagai metode untuk mengidentifikasi kecenderungan sentimen publik secara objektif berdasarkan bahasa dan kata-kata yang digunakan masyarakat dalam unggahan media sosial. Dengan bantuan teknologi dan metode pembelajaran mesin (*machine learning*), analisis ini mampu melakukan klasifikasi sentimen (positif, negatif, atau netral) dalam skala data besar.

Beberapa algoritma klasifikasi sentimen diantaranya adalah metode berbasis probabilitas seperti *Naïve Bayes* (NB), metode berbasis kedekatan seperti K-*Nearset Neighbors* (KNN), dan metode berbasis struktur keputusan seperti *Decision Tree* (DT), serta berbagai metode lainnya. Dalam penelitian ini, peneliti menggunakan algoritma *Support Vector Machine* (SVM) dan *Vector Space Model* (VSM). Algoritma *Support Vector Machine* (SVM) sebagai metode klasifikasi, yang dikenal efektif untuk pemrosesan teks berdimensi tinggi.[1] Sementara *Vector Space Model* (VSM) dengan pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF) digunakan untuk mengubah data teks menjadi representasi numerik. Kombinasi metode ini telah banyak digunakan dalam penelitian analisis sentimen dan terbukti memberikan hasil klasifikasi yang akurat dan stabil, khususnya pada dataset teks berukuran besar seperti data dari media sosial.

Pemilihan metode *Support Vector Machine* (SVM) sebagai algoritma klasifikasi didasarkan pada keunggulannya yang terbukti efektif dalam memproses data teks berdimensi tinggi, bekerja dengan mencari *hyperplane* pemisah optimal untuk memaksimalkan margin antara kelas sentimen (Positif, Negatif, Netral), sehingga menghasilkan akurasi yang lebih stabil dan tinggi dibandingkan algoritma *Machine Learning* klasik lainnya.[1] VSM dengan pembobotan TF-IDF melengkapi proses ini dengan menyediakan representasi numerik yang diskriminatif sebagai input bagi SVM. Meskipun telah terdapat beberapa penelitian yang membahas analisis sentimen dengan metode SVM dan TF-IDF pada berbagai bidang, seperti ulasan produk, review layanan, dan opini publik secara umum, penelitian yang secara spesifik menganalisis sentimen masyarakat terhadap kebijakan Efisiensi Anggaran Pemerintah 2025 masih terbatas. Penelitian ini diharapkan dapat memberikan kontribusi akademik dalam bidang analisis sentimen berbasis NLP sekaligus memberikan wawasan mengenai persepsi publik terhadap arah kebijakan fiskal pemerintah.

2. Tinjauan Pustaka

2.1. Media Sosial

Media sosial merupakan platform digital yang memungkinkan pengguna untuk berinteraksi, berbagi informasi, serta membentuk jaringan secara online, media sosial telah menjadi sumber informasi yang sangat populer di kalangan masyarakat, terutama generasi muda, serta digunakan sebagai sarana utama dalam mencari informasi. Hal ini menunjukkan bahwa media sosial memiliki peran penting dalam membentuk perilaku pencarian informasi di era digital. [2]

2.2. X (Twitter)

X (Twitter) merupakan media sosial yang digunakan masyarakat untuk menyampaikan opini dan kritik terhadap berbagai isu. Twitter digunakan sebagai sumber data karena memuat berbagai opini publik yang dapat dianalisis menggunakan metode analisis sentimen. Oleh karena itu, Twitter dapat dimanfaatkan sebagai sumber data untuk menganalisis sentimen masyarakat terhadap kebijakan pemerintah. [3]

2.3 Analisis Sentimen

Analisis Sentimen merupakan proses komputasi untuk mengubah data tekstual yang tidak terstruktur (seperti komentar atau ulasan) menjadi data terstruktur. Proses ini meliputi ekstraksi dan pemahaman terhadap opini, emosi, atau penilaian yang disuarakan dalam sebuah kalimat, guna mendapatkan wawasan sentimen secara otomatis. [4]

2.4 Efisiensi Anggaran

Efisiensi Anggaran adalah kebijakan strategis untuk mengoptimalkan penggunaan dana dengan cara meminimalkan pemborosan dan memastikan anggaran yang ada memberikan hasil maksimal. Kebijakan ini bukan sekedar penghematan, tetapi juga mencakup pengalokasian dana secara tepat sasaran.[1]

2.5 Praproses Data Texs (*Text Preprocessing*)

Praproses Teks merupakan tahapan esensial yang menandai awal dari proses Text Mining, di mana fokus utamanya adalah menyeleksi dan membersihkan data dokumen dari elemen yang tidak perlu untuk mengurangi noise. [5]

Praproses merupakan rangkaian tahapan yang dilakukan secara berurutan untuk membersihkan dan menormalkan data, dimulai dari:

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)
Vol. 9 No. 2, April 2026
P-ISSN 2620-8342
E-ISSN 2621-3052

1. Data *Cleaning*, di mana tindakan penghilangan karakter, tanda baca, serta noise lain yang tidak relevan dilakukan untuk membersihkan data.
2. *Case Folding* untuk menstandarisasi semua huruf menjadi huruf kecil (*lowercase*), memastikan konsistensi. Setelah konsistensi tercapai,
3. *Tokenizing* dilakukan untuk memisahkan teks menjadi unit-unit kata yang akan diolah.
4. *Filtering* (*Stopword Removal*) di mana kata-kata sambung yang tidak memiliki makna penting dihilangkan. Terakhir, proses dilanjutkan dengan
5. *Stemming*, yaitu penyederhanaan kata dengan menghilangkan imbuhan agar kata tersebut kembali ke bentuk dasar atau akar katanya.

2.6 *Vector Space Model* Dengan Pembobotan Fitur TF-IDF

Vector Space Model (VSM) digunakan sebagai representasi utama data teks. VSM adalah metode yang fundamental dalam bidang information retrieval untuk menentukan kedekatan tematik antar-teks.[6] Untuk menunjang proses representasi dokumen menjadi vektor pada VSM, penelitian ini memanfaatkan *Term Frequency-Inverse Document Frequency* (TF-IDF) sebagai teknik pembobotan fitur [7]. TF-IDF merupakan teknik yang efektif untuk mengukur tingkat kepentingan suatu kata (*term*) dalam sebuah dokumen dibandingkan dengan seluruh koleksi dokumen. Prinsip kerjanya adalah memberikan bobot tinggi pada kata yang sering muncul di satu dokumen (*Term Frequency* (TF)) namun jarang muncul pada dokumen lain (*Inverse Document Frequency* (IDF)), sehingga menghasilkan fitur yang sangat diskriminatif bagi algoritma klasifikasi. Secara matematis, bobot TF-IDF dihitung sebagai berikut:

1. *Term Frequence* (TF)

$$TF(t,d)\ \frac{\text{jumlah kemunculan kata } t \text{ dalam dokumen } d}{\text{total kata dalam dokumen } d} \qquad (2.1)$$

2. *Inverse Document Frequency* (IDF)

$$IDF(t) = \log \frac{N}{DF(t)} \qquad (2.2)$$

3. *Term Frequency-Inverse Document Frequency* (TF-IDF)

$$TF - IDF(t,d) = TF(t,d)xIDF(t) \qquad (2.3)$$

2.7 *Support Vector Machine*

Algoritma *Support Vector Machine* (SVM) bekerja berdasarkan prinsip penentuan bidang pemisah terbaik (optimal *separating hyperplane*) untuk memisahkan objek-objek data dalam ruang fitur. Meskipun pada mulanya SVM dikembangkan khusus untuk masalah klasifikasi dua kelas, seiring perkembangannya, teknik ini telah diperluas dan diadaptasi secara efektif untuk menangani kasus klasifikasi multi-kelas, seperti yang digunakan dalam analisis sentimen tiga polaritas. [8]

Pemilihan SVM sebagai klasifikator utama didasarkan pada keunggulannya dalam mengatasi data berdimensi tinggi. SVM dikenal efektif dalam menemukan batas pemisah optimal dan menghasilkan **akurasi yang tinggi** dalam klasifikasi sentimen. [9] Kemampuan SVM untuk memetakan data ke dimensi yang lebih besar sangat cocok untuk mengklasifikasikan vektor fitur yang dihasilkan oleh **TF-IDF**, sehingga metode ini terbukti unggul dibandingkan algoritma lain saat menangani data tekstual dengan banyak fitur [10]

Secara matematis, *hyperplane* pemisah didefinisikan melalui persamaan liner sebagai berikut :

$$W \cdot X + b = 0 \qquad (2.4)$$

Setelah model SVM dilatih, proses prediksi atau klasifikasi untuk data baru (X_{baru}) dilakukan menggunakan Fungsi Keputusan (*Decision Function*):

$$f(X_{baru}) = sign\ (W \cdot X + b\ ) \qquad (2.5)$$

Aturan Klasifikasi :

Jika $f(X) > \ 1 : Positif$
Jika $f(X) < -1 : Negatif$
Jika hasil berada di antara -1 dan +1 (atau dekat 0), model memprediksi Netral (atau berada dalam margin)

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)
Vol. 9 No. 2, April 2026
P-ISSN 2620-8342
E-ISSN 2621-3052

2.8 Metrik Evaluasi Kinerja Model

Evaluasi kinerja model klasifikasi multi-kelas, khususnya dalam analisis sentimen, didasarkan pada perhitungan dari Matriks Kebingungan (*Confusion Matrix*) [11]. Matriks ini berfungsi untuk memetakan hasil prediksi model terhadap label sebenarnya dari data uji, sehingga dapat dihitung beberapa metrik sebagai berikut:

1. Akurasi (*Accuracy*): Metrik ini mengukur persentase total prediksi yang berhasil diklasifikasikan dengan tepat (*True Positive* dan *True Negative*) dibandingkan dengan keseluruhan jumlah data. Akurasi memberikan gambaran umum mengenai keseluruhan ketepatan kinerja model.
2. Presisi (*Precision*): Presisi menunjukkan tingkat keandalan model saat memprediksi suatu kelas tertentu. Ukuran ini dihitung dari jumlah prediksi positif yang benar dibagi dengan total seluruh prediksi positif yang dibuat oleh model, merefleksikan ketepatan hasil prediksi.
3. *Recall* (Sensitivitas): Mengukur kemampuan model untuk menangkap semua sampel yang memang termasuk dalam kelas target. *Recall* sangat penting karena mengindikasikan keberhasilan model dalam meminimalisir kesalahan klasifikasi tipe *False Negative* (gagal mendeteksi sentimen yang seharusnya terdeteksi).
4. F1-Score: F1-Score adalah nilai rata-rata harmonis (*harmonic mean*) yang menyatukan nilai *Precision* dan *Recall*. Metrik ini sangat relevan digunakan dalam kasus data yang memiliki ketidakseimbangan kelas (*imbalanced* dataset) karena memberikan tolak ukur tunggal yang seimbang terhadap kinerja model.

Evaluasi dilakukan untuk mengukur sejauh mana model mampu melakukan klasifikasi dengan benar. Perhitungan evaluasi ini didasarkan pada nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

$$Accuary = \frac{TP+TN}{TP+TN+FP+FN} \quad (2.6)$$

$$Precision = \frac{TP}{TP+FP} \quad (2.7)$$

$$Recall = \frac{TP}{TP+FN} \quad (2.8)$$

$$F1 - Scorre = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.9)$$

2.9 Google Colab

Google *Colaboratory* (sering disingkat Colab) merupakan platform cloud yang dikembangkan oleh Google yang menyediakan lingkungan notebook Jupyter untuk menulis dan mengeksekusi kode Python secara penuh melalui browser web, tanpa memerlukan instalasi atau konfigurasi perangkat lunak lokal. Platform ini memanfaatkan infrastruktur cloud Google, menawarkan lingkungan pengembangan yang kuat dan fleksibel, serta memberikan akses gratis ke sumber daya komputasi performa tinggi seperti GPU dan TPU (Tensor Processing Unit) untuk memfasilitasi kebutuhan proyek data *science* dan *machine learning*. [12]

2.10 Python

Python adalah bahasa pemrograman tingkat tinggi yang serbaguna, berorientasi objek, dan dikenal karena sintaksnya yang sederhana serta mudah dibaca (*readable*). Dalam konteks penelitian *Data Science* dan *Machine Learning*, Python menjadi pilihan utama karena ekosistemnya yang kaya akan pustaka (*libraries*) siap pakai, yang sangat krusial untuk mengimplementasikan metodologi kompleks. [13]

3. Meode Penelitian

3.1 Objek Penelitian

Penelitian ini merupakan kumpulan data tekstual dalam bentuk unggahan publik (tweet) yang diekstrak dari platform media sosial X (Twitter). Fokus data diarahkan pada opini dan tanggapan yang disuarakan oleh masyarakat terkait Kebijakan Efisiensi Anggaran Pemerintah yan g diumumkan untuk APBN tahun 2025.

Proses akuisisi data dilakukan secara langsung melalui Twitter API, dengan menggunakan kata kunci utama "Efisiensi Anggaran". Data dikoleksi dalam rentang waktu yang spesifik, yakni sejak 8 Januari 2025 hingga 7 November 2025. Data yang terkumpul kemudian akan diberi label berdasarkan tiga polaritas sentimen, positif, negatif, dan netral sebelum diproses lebih lanjut menggunakan metode *Support Vector Machine* (SVM) untuk klasifikasi.

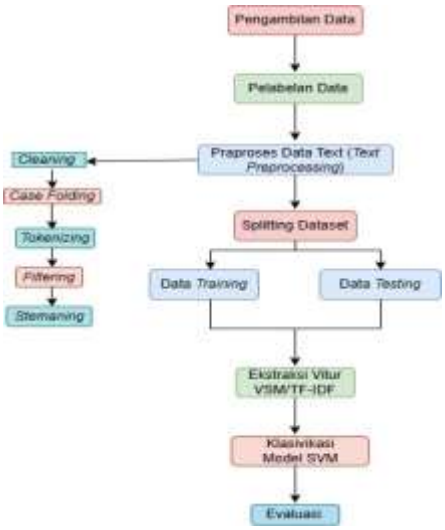
Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Accredited SINTA 5
72/E/KPT/2024

Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)
Vol. 9 No. 2, April 2026
P-ISSN 2620-8342
E-ISSN 2621-3052

3.2 Prosedur Penelitian

Penelitian ini menerapkan prosedur komputasi berbasis sistematisasi text mining untuk mentransformasikan data tekstual menjadi informasi prediktif yang akurat. Landasan prosedural ini berpijak pada metodologi klasifikasi semantik yang dipaparkan oleh [14], di mana sinergi antara V*ector Space Model* dan *Support Vector Machine* menjadi instrumen utama dalam pemrosesan data. Adapun artikulasi dari setiap tahapan kerja tersebut dijabarkan di bawah ini.



Gambar 1. Tahapan Penelitian

4. Hasil Dan Pembahasan

4.1 Pengumpulan Data

Prosedur pengumpulan data dalam penelitian ini dilakukan dengan memanfaatkan platform *Google Colaboratory* melalui teknik *crawling* data pada media sosial X (Twitter). Instrumen pengambilan data difokuskan pada ekstraksi tweet atau cuitan publik yang merepresentasikan opini masyarakat terhadap kata kunci (*search keyword*) spesifik, yakni 'Efisiensi Anggaran'. Adapun periode pengambilan data ditetapkan dalam rentang waktu yang komprehensif, terhitung mulai tanggal 8 Januari 2025 hingga 12 Desember 2025. Dataset yang diperoleh dari proses ini kemudian dihimpun untuk diolah pada tahapan berikutnya.

4.2 Pelabelan Dan Seleksi Data

Dataset yang telah dikumpulkan melalui proses *crawling* menjalani seleksi dan pemberian label. Proses seleksi difokuskan pada pembersihan data, guna mengeliminasi cuitan yang serupa atau duplikat (data *redundant*). Setelah dilakukan proses seleksi dengan menghapus data duplikat, jumlah dataset yang semula berjumlah 1.685 cuitan menyusut menjadi 1.221 cuitan. Setelah dilakukan pelabelan sentimen yang dikategorikan menjadi positif, negatif, dan netral. Dalam penelitian ini, penentuan label pada setiap data dilakukan secara mandiri oleh peneliti dan pelabelan berbasis *Lexicon-Based Approach* untuk menjamin akurasi label sesuai dengan konteks penelitian.

4.3 Pra-pemrosesan Data (*Pre-processing*)

Seluruh cuitan yang telah dikumpulkan perlu melewati tahap pembersihan agar informasi di dalamnya lebih terfokus. Sehingga hanya menyisakan kata kunci penting terkait efisiensi anggaran. Hal ini dilakukan untuk memastikan bahwa data yang masuk ke tahap perhitungan adalah data yang benar-benar berkualitas. Adapun rangkaian langkah untuk merapikan data tersebut meliputi, *cleaning*, *case folding*, *tokenizing*, filtering, *stemming*.

4.4 Data Training Dan Data Testing

Tahap ini mencakup pembagian dataset ke dalam kategori data latih (*training*) dan data uji (*testing*), baik untuk data hasil pelabelan mandiri maupun berbasis *Lexicon-Based Approach*. Dataset yang telah melalui proses *preprocessing* dibagi dengan proporsi 80:20 menggunakan bahasa pemrograman *Python* di lingkungan *Google Colab*. Pembagian secara otomatis ini dilakukan untuk memastikan distribusi data yang objektif sebelum memasuki tahap pemodelan.

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Table 1. Sampel Data Training dan Data Testing

ID	Komentar	Label
D1	hapus pensiun anggota dpr adil efisiensi anggar negara tanda tangan petisi	Positif
D2	gub aceh desperate bupati kabur pusat lepas tangan elit anggar daerah kena efisiensi	Negatif
D3	prabowo jamin efisiensi kelola anggar negara hemat anti korupsi	Positif
D4	tuju benar gaji jabat naik anggar rakyat kena efisiensi	Negatif
D5	indonesia belajar efisiensi anggar	Netral
D6	efisiensi anggar hasil langsung tani keren	Positif
D7	langsung tantrum paksa efisiensi anggar	?

4.5 *Vector Space Model* Dengan Pembobotan TF-IDF

Implementasi *Vector Space Model* (VSM) dalam penelitian ini menggunakan teknik pembobotan TF-IDF untuk mengubah data teks hasil preprocessing menjadi vektor numerik. Tahapan ini dilaksanakan setelah pemisahan data latih (*training*) dan data uji (*testing*) dengan proporsi 80:20. Penggunaan TF-IDF bertujuan untuk memberikan bobot yang lebih tinggi pada istilah kunci yang memiliki signifikansi sentimen dan mengurangi bobot kata-kata umum yang kurang informatif. Transformasi ini sangat penting karena algoritma SVM memerlukan input berupa nilai numerik untuk menentukan posisi koordinat data dalam ruang dimensi tinggi.

Table 2. Hitung TF (*Term Frequency*)

Term	D1	D2	D3	D4	D5	D6
hapus	1	0	0	0	0	0
pensiun	1	0	0	0	0	0
anggota	1	0	0	0	0	0
dpr	1	0	0	0	0	0
adil	1	0	0	0	0	0
efisiensi	1	1	1	1	1	1
anggar	1	1	1	1	1	1
negara	1	0	1	0	0	0
tanda	1	0	0	0	0	0
tangan	1	0	0	0	0	0
petisi	1	0	0	0	0	0
gub	0	1	0	0	0	0
aceh	0	1	0	0	0	0
desperate	0	1	0	0	0	0
bupati	0	1	0	0	0	0
kabur	0	1	0	0	0	0
pusat	0	1	0	0	0	0
lepas	0	1	0	0	0	0
tangan	0	1	0	0	0	0
elit	0	1	0	0	0	0
daerah	0	1	0	0	0	0
kena	0	1	0	1	0	0
prabowo	0	0	1	0	0	0
jamin	0	0	1	0	0	0
kelola	0	0	1	0	0	0
hemat	0	0	1	0	0	0
anti	0	0	1	0	0	0
korupsi	0	0	1	0	0	0
tuju	0	0	0	1	0	0
benar	0	0	0	1	0	0
gaji	0	0	0	1	0	0
jabat	0	0	0	1	0	0
naik	0	0	0	1	0	0
rakyat	0	0	0	1	0	0
indonesia	0	0	0	0	1	0

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)
Vol. 9 No. 2, April 2026
P-ISSN 2620-8342
E-ISSN 2621-3052

Term	D1	D2	D3	D4	D5	D6
belajar	0	0	0	0	1	0
hasil	0	0	0	0	0	1
langsung	0	0	0	0	0	1
tani	0	0	0	0	0	1
keren	0	0	0	0	0	1

Table 3. Hitung IDF (*Inverse Dokument frequency*)

Term	DF	IDF
hapus	1	0.7782
pensiun	1	0.7782
anggota	1	0.7782
dpr	1	0.7782
adil	1	0.7782
efisiensi	6	0.0000
anggar	6	0.0000
negara	2	0.4771
tanda	1	0.7782
tangan	1	0.7782
petisi	1	0.7782
gub	1	0.7782
aceh	1	0.7782
desperate	1	0.7782
bupati	1	0.7782
kabur	1	0.7782
pusat	1	0.7782
lepas	1	0.7782
tangan	1	0.7782
elit	1	0.7782
daerah	1	0.7782
kena	2	0.4771
prabowo	1	0.7782
jamin	1	0.7782
kelola	1	0.7782
hemat	1	0.7782
antik	1	0.7782
korupsi	1	0.7782
tuju	1	0.7782
benar	1	0.7782
gaji	1	0.7782
jabat	1	0.7782
naik	1	0.7782
rakyat	1	0.7782
indonesia	1	0.7782
belajar	1	0.7782
hasil	1	0.7782
langsung	1	0.7782
tani	1	0.7782
keren	1	0.7782

Table 4. Hitung *Term Frequency - Inverse Dokument Frequency*

Term	D1	D2	D3	D4	D5	D6
hapus	0.7782	0	0	0	0	0
pensiun	0.7782	0	0	0	0	0
anggota	0.7782	0	0	0	0	0

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Accredited SINTA 5
72/E/KPT/2024

Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)
Vol. 9 No. 2, April 2026
P-ISSN 2620-8342
E-ISSN 2621-3052

Term	D1	D2	D3	D4	D5	D6
dpr	0.7782	0	0	0	0	0
adil	0.7782	0	0	0	0	0
efisiensi	0	0	0	0	0	0
anggar	0	0	0	0	0	0
negara	0.4771	0	0.4771	0	0	0
tanda	0.7782	0	0	0	0	0
tangan	0.7782	0	0	0	0	0
petisi	0.7782	0	0	0	0	0
gub	0	0.7782	0	0	0	0
aceh	0	0.7782	0	0	0	0
desperate	0	0.7782	0	0	0	0
bupati	0	0.7782	0	0	0	0
kabur	0	0.7782	0	0	0	0
pusat	0	0.7782	0	0	0	0
lepas	0	0.7782	0	0	0	0
tangan	0	0.7782	0	0	0	0
elit	0	0.7782	0	0	0	0
daerah	0	0.7782	0	0	0	0
kena	0	0.4771	0	0.4771	0	0
prabowo	0	0	0.7782	0	0	0
jamin	0	0	0.7782	0	0	0
kelola	0	0	0.7782	0	0	0
hemat	0	0	0.7782	0	0	0
anti	0	0	0.7782	0	0	0
korupsi	0	0	0.7782	0	0	0
tuju	0	0	0	0.7782	0	0
benar	0	0	0	0.7782	0	0
gaji	0	0	0	0.7782	0	0
jabat	0	0	0	0.7782	0	0
naik	0	0	0	0.7782	0	0
rakyat	0	0	0	0.7782	0	0
indonesia	0	0	0	0	0.7782	0
belajar	0	0	0	0	0.7782	0
hasil	0	0	0	0	0	0.7782
langsung	0	0	0	0	0	0.7782
tani	0	0	0	0	0	0.7782

4.6 Klasifikasi *Suport Vector Machine*

Data yang telah melalui serangkaian proses mulai dari pelabelan, seleksi data hingga pembagian data menjadi data training dan data testing untuk dilakukan proses klasifikasi. Metode yang digunakan adalah Support Vector Machine (SVM). SVM merupakan algoritma pembelajaran mesin yang bekerja dengan cara mencari hyperplane atau garis pemisah optimal untuk memisahkan antar kelas sentimen di dalam Vector Space Model. Berdasarkan proses pembagian data tersebut, diperoleh data *training* yang terdiri dari dokumen D1 hingga D6, sedangkan dokumen D7 digunakan sebagai data *testing.*

Table 5. Sampel Data *Testing*

Term	D7(X)
langsung	0.7782
efisiensi	0
anggar	0

$f(X_{baru}) = sign\ (W \cdot X + b\ )$

W = (-0.7, 0.5, 0.4)

b = -02

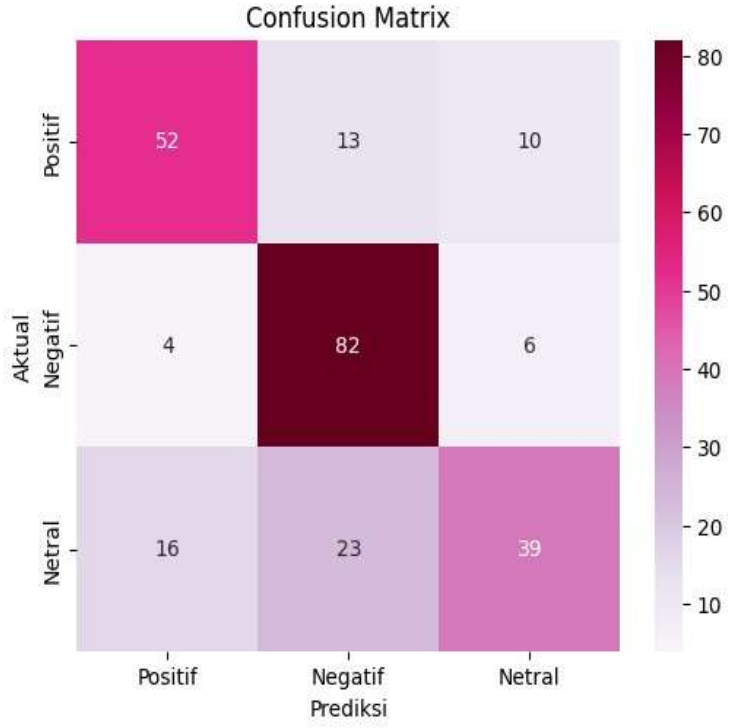
Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)
Vol. 9 No. 2, April 2026
P-ISSN 2620-8342
E-ISSN 2621-3052

$$f(D7) = (-0{,}7 \times 0{,}7782) + (0{,}5 \times 0) + (0.4 \times 0) - 0{,}2$$
$$= -0{,}5447 - 0{,}2 = -0{,}7447$$

4.7 Evaluasi Model

Penilaian ini memanfaatkan *Confusion Matrix* sebagai alat ukur untuk melihat sinkronisasi antara label mandiri dengan label hasil klasifikasi otomatis berbasis *Lexicon-Based Approach*. Melalui matriks ini, diperoleh gambaran menyeluruh mengenai efektivitas model berdasarkan parameter Accuracy, Precision, Recall, dan F1-Score. Confusion matrix digunakan untuk mengevaluasi performa model klasifikasi sentimen dengan menampilkan jumlah prediksi yang benar dan salah pada setiap kelas: Positif, Netral, dan Negatif. Visualisasi ini memudahkan identifikasi kelemahan model, seperti kelas yang sering salah diprediksi, serta memberikan indikasi akurasi model untuk masing-masing sentimen



Gambar 2. Hasil *Confusion Matrix* Pelabelan Mandiri

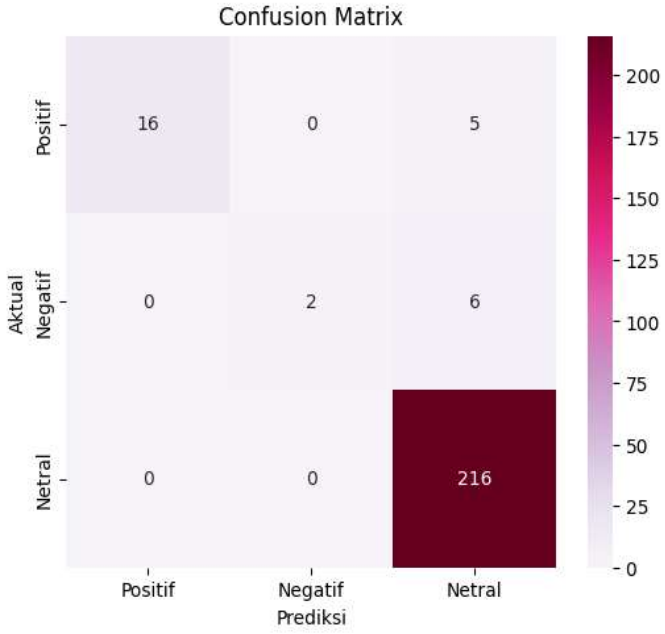
Berdasarkan hasil *confusion matrix* pada pelabelan mandiri, dapat dilihat performa model dalam mengklasifikasikan sentimen menjadi tiga kategori yaitu Positif, Negatif, dan Netral**.** Pada matriks tersebut, baris menunjukkan label aktual**,** sedangkan kolom menunjukkan label hasil prediksi dari model. Nilai pada diagonal matriks menunjukkan jumlah data yang berhasil diprediksi dengan benar oleh model, sedangkan nilai di luar diagonal menunjukkan jumlah kesalahan klasifikasi yang dilakukan oleh model.

1. Kelas Positif
   Pada kelas positif, terdapat 52 data yang berhasil diprediksi dengan benar sebagai Positif. Namun terdapat 13 data Positif yang diprediksi sebagai Negatif dan 10 data Positif yang diprediksi sebagai Netral, yang menunjukkan masih adanya kesalahan klasifikasi pada kelas ini.
2. Kelas Negatif
   Pada kelas Negatif, model menunjukkan performa yang cukup baik dengan 82 data berhasil diprediksi dengan benar sebagai Negatif. Selain itu, terdapat 4 data Negatif yang diprediksi sebagai Positif dan 6 data Negatif yang diprediksi sebagai Netral.
3. Kelas Netral
   Pada kelas Netral, terdapat 39 data yang berhasil diprediksi dengan benar sebagai Netral. Namun masih terdapat 16 data Netral yang diprediksi sebagai Positif dan 23 data Netral yang

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

diprediksisebagai Negatif, yang menunjukkan bahwa kelas Netral merupakan kelas yang paling sering mengalami kesalahan prediksi.

Secara keseluruhan, model mampu mengklasifikasikan sentimen dengan cukup baik terutama pada kelas Negatif, yang memiliki jumlah prediksi benar paling tinggi. Namun demikian, model masih mengalami kesulitan dalam membedakan sentimen Netral, sehingga masih terdapat sejumlah kesalahan klasifikasi pada kelas tersebut.



Gambar 3. Hasil *Confusion Matrix* Pelabelan Berbasis *Lexicon-Based Approach*

Berdasarkan hasil confusion matrix pada pelabelan otomatis berbasis *Lexicon-Based Approach*, dapat dilihat performa model dalam mengklasifikasikan sentimen menjadi tiga kategori yaitu Positif, Negatif, dan Netral. Pada matriks tersebut, baris menunjukkan label aktual, sedangkan kolom menunjukkan label hasil prediksi dari model. Nilai pada diagonal matriks menunjukkan jumlah data yang berhasil diprediksi dengan benar oleh model, sedangkan nilai di luar diagonal menunjukkan jumlah kesalahan klasifikasi.

1. Kelas Positif
   Pada kelas Positif terdapat 16 data yang berhasil diprediksi dengan benar sebagai Positif. Namun terdapat 5 data Positif yang diprediksi sebagai Netral, yang menunjukkan bahwa model masih sering keliru mengklasifikasikan sentimen Positif sebagai Netral.
2. Kelas Negatif
   Pada kelas Negatif terdapat 2 data yang berhasil diprediksi dengan benar sebagai Negatif. Selain itu terdapat 6 data Negatif yang diprediksi sebagai Netral, yang menunjukkan bahwa model masih mengalami kesulitan dalam mengenali sentimen Negatif.
3. Kelas Netral
   Pada kelas Netral model menunjukkan performa yang sangat baik dengan 216 data berhasil diprediksi dengan benar sebagai Netral.

Secara keseluruhan, model menunjukkan performa yang cukup baik terutama pada kelas Netral, yang memiliki jumlah prediksi benar paling tinggi. Namun demikian, model masih mengalami kesulitan dalam membedakan sentimen Positif dan Negatif, karena sebagian data pada kedua kelas tersebut cenderung diprediksi sebagai Netral.

Berikut merupakan hasil evaluasi performa model yang diperoleh dari confusion matrix, yang meliputi nilai precision, recall, F1-score, macro average, dan accuracy untuk mengukur tingkat kinerja model dalam mengklasifikasikan sentimen.

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)
Vol. 9 No. 2, April 2026
P-ISSN 2620-8342
E-ISSN 2621-3052

	Precision	Recall	F-1score	Support
Negatif	0.69	0.89	0.78	92
Netral	0.71	0.50	0.59	78
Positif	0.72	0.69	0.61	75
Accuracy			0.71	245
Macro avg	0.71	0.69	0.69	245

Table 6. Hasil Confusion Matrix Pelabelan Mandiri

	Precision	Recall	F-1score	Support
Negatif	1.00	0.25	0.40	8
Netral	0.95	1.00	0.98	216
Positif	1.00	0.76	0.86	21
Accuracy			0.96	245
Macro avg	0.98	0.67	0.75	245

Table 7. Hasil Confusion Matrix Pelabelan Mandiri

Setelah hasil evaluasi model diperoleh, dilakukan analisis terhadap kinerja model pada setiap kelas sentimen. Analisis ini meliputi nilai *precision*, *recall*, dan *F1-score* pada masing-masing kelas, serta nilai *macro average* dan *accuracy* secara keseluruhan.
Evaluasi model berdasarkan pelabelan mandiri.

1. Kelas Positif
   Berdasarkan grafik, kelas positif menunjukkan performa yang cukup baik. Nilai *precision* sebesar 72% menunjukkan bahwa sebagian besar data yang diprediksi sebagai positif memang benar berasal dari kelas positif. Selain itu, *recall* sebesar 69% menunjukkan bahwa model mampu mengenali sebagian besar data yang benar-benar positif. Kombinasi antara precision dan *recall* tersebut menghasilkan *F1-score* sebesar 71%, yang menunjukkan bahwa model memiliki keseimbangan yang baik dalam mengklasifikasikan data pada kelas positif.
2. Kelas Netral
   Pada kelas netral, model memiliki *precision* sebesar 71%, yang berarti sebagian besar data yang diprediksi sebagai netral benar-benar berasal dari kelas tersebut. Namun, *recall* sebesar 50% menunjukkan bahwa model hanya mampu mengenali setengah dari data yang sebenarnya termasuk ke dalam kelas netral. Akibatnya, *F1-score* yang dihasilkan sebesar 59%, sehingga dapat disimpulkan bahwa model masih kurang optimal dalam mengidentifikasi data pada kelas netral.
3. Kelas Negatif
   Pada kelas negatif, model menunjukkan performa yang cukup stabil. Nilai *precision* sebesar 69% menunjukkan bahwa sebagian besar data yang diprediksi sebagai negatif memang benar berasal dari kelas tersebut. Selain itu, *recall* sebesar 89% menunjukkan bahwa model mampu mengenali hampir seluruh data yang benar-benar negatif. Kombinasi kedua nilai tersebut menghasilkan *F1-score* sebesar 78%, yang menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengklasifikasikan data pada kelas negatif.
4. *Macro Average Dan Accuracy*
   Model dengan pelabelan manual memiliki akurasi 70,61%, yang menunjukkan kemampuan model dalam melakukan prediksi yang benar terhadap sebagian besar data uji. *Macro precision* sebesar 0,71 (71%) dan macro *recall* 0,69 (69%) menunjukkan performa yang relatif seimbang secara rata-rata, meskipun kelas netral memiliki *recall* yang lebih rendah (0,50), menandakan kesulitan model dalam mengenali seluruh instance dari kelas ini. Macro *F1-score* 0,69 (69%) menegaskan kestabilan performa antar kelas. Dengan demikian, pelabelan mandiri menghasilkan performa klasifikasi yang

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

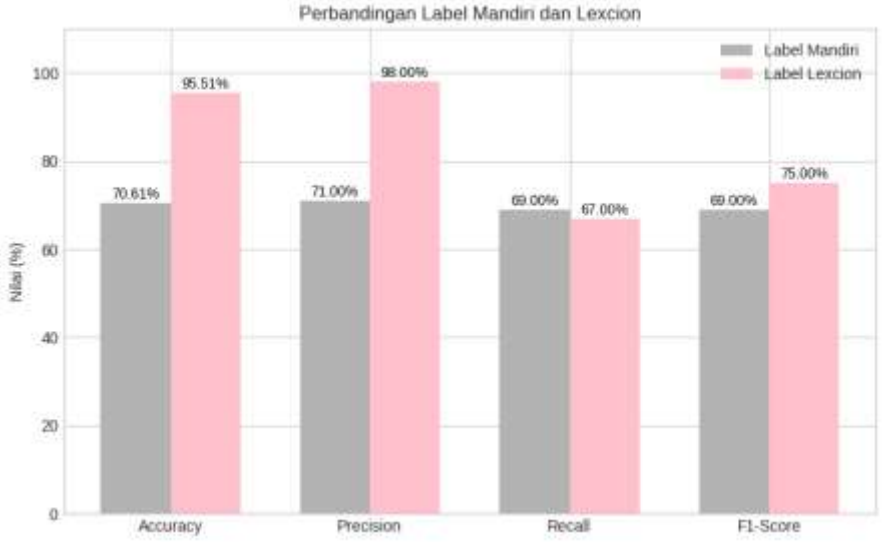
Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)
Vol. 9 No. 2, April 2026
P-ISSN 2620-8342
E-ISSN 2621-3052

lebih merata antar kelas meskipun akurasi keseluruhan sedikit lebih rendah dibanding pelabelan otomatis.

Evaluasi model berdasrkan pelabelan *Lexicon-Based Approach*

1. Kelas Positif
   Pada kelas positif, model memiliki precision sebesar 100%, yang menunjukkan bahwa seluruh data yang diprediksi sebagai positif benar-benar berasal dari kelas positif. Namun, *recall* sebesar 76% menunjukkan bahwa model belum mampu mengenali seluruh data positif yang ada, sehingga masih terdapat sebagian data positif yang tidak terdeteksi. Kombinasi kedua nilai tersebut menghasilkan F1-*score* sebesar 86%, yang menunjukkan bahwa performa model pada kelas positif tergolong cukup baik, meskipun masih perlu peningkatan pada sisi *recall*.
2. Kelas Netral
   Kelas netral menunjukkan performa yang sangat baik. Nilai *precision* sebesar 95% mengindikasikan bahwa hampir seluruh data yang diprediksi sebagai netral memang benar berasal dari kelas tersebut. Selain itu, *recall* sebesar 100% menunjukkan bahwa model mampu mengenali seluruh data yang termasuk dalam kelas netral. Hal ini menghasilkan *F1-score* sebesar 98%, yang menandakan bahwa model memiliki performa yang optimal dalam mengklasifikasikan data netral.
3. Kelas Negatif
   Pada kelas negatif, model menunjukkan performa yang masih rendah. Nilai precision sebesar 100% menunjukkan bahwa semua data yang diprediksi sebagai negatif adalah benar. Namun, *recall* hanya sebesar 25%, yang berarti model hanya mampu mengenali sebagian kecil dari keseluruhan data negatif yang ada. Hal ini menghasilkan *F1-score* sebesar 40%, sehingga dapat disimpulkan bahwa kemampuan model dalam mengklasifikasikan data negatif masih sangat terbatas.
4. *Macro Average* Dan *Acuraccy*
   Secara keseluruhan, model dengan pelabelan berbasis *lexicon-Based* menunjukkan akurasi sebesar 96%, yang berarti sebagian besar data berhasil diklasifikasikan dengan benar. Nilai *macro precision* sebesar 0,98 menunjukkan bahwa model sangat baik dalam menentukan kelas prediksi. Namun, macro *recall* sebesar 0,57 mengindikasikan bahwa model belum mampu mengenali seluruh data pada setiap kelas secara merata, terutama pada kelas negatif yang memiliki nilai *recall* rendah.Akibatnya, macro *F1-score* sebesar 0,75 menunjukkan bahwa performa model masih belum seimbang antar kelas. Dengan demikian, meskipun model sangat baik dalam mengklasifikasikan data pada kelas mayoritas (netral), kinerjanya masih perlu ditingkatkan pada kelas minoritas, khususnya kelas negatif.

Untuk memudahkan pemahaman perbedaan performa model antara pelabelan mandiri dan berbasis *Lexicon-Based Approach*, hasil evaluasi disajikan dalam bentuk perbandingan grafik *macro precision*, *macro recall*, *macro* F1-*score*, dan akurasi pada setiap jenis pelabelan.



Gambar 4. Visualisasi Metrix Pelabelan Mandiri dan *Lexicon-Based*

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Grafik di atas menunjukkan perbandingan performa model antara pelabelan mandiri dan berbasis *lexicon* pada empat metrik utama, yaitu *accuracy*, *precision*, *recall*, dan F1-*score*. Akurasi model pelabelan berbasis *Lexicon-Based Approach* sebesar 96% lebih tinggi dibandingkan pelabelan mandiri sebesar 71% yang menunjukkan bahwa secara keseluruhan model otomatis lebih sering menghasilkan prediksi yang benar. Namun, jika dilihat dari metrik lainnya, performa model menunjukkan perbedaan yang cukup beragam. Nilai *precision* pada pelabelan *lexicon* sebesar 98% lebih tinggi dibandingkan pelabelan mandiri sebesar 71%, yang menunjukkan bahwa model otomatis sangat baik dalam menentukan kelas prediksi dengan tingkat kesalahan yang rendah.Sementara itu, pada metrik *recall*, pelabelan mandiri memiliki nilai sebesar 69%, sedikit lebih tinggi dibandingkan pelabelan *lexicon* sebesar 67%, yang mengindikasikan bahwa model mandiri lebih mampu mengenali seluruh data pada masing-masing kelas.Hal yang sama juga terlihat pada F1-score, di mana pelabelan *lexicon* memperoleh nilai 75%, lebih tinggi dibandingkan pelabelan mandiri sebesar 69%, yang menunjukkan bahwa secara keseimbangan antara precision dan recall, model otomatis memiliki performa yang lebih baik. Dengan demikian, dapat disimpulkan bahwa model berbasis *Lexicon-Based Approach* unggul dalam hal *akurasi*, *precision*, dan *F1-score*, namun masih sedikit tertinggal pada recall dibandingkan model mandiri. Sementara itu, model mandiri menunjukkan kemampuan yang lebih baik dalam mengenali data secara menyeluruh, meskipun tingkat ketepatan prediksinya masih lebih rendah dibandingkan model *Lexicon-Based Approach* .

5. Kesimpulan

Penelitian ini berhasil mengolah sebanyak 1.221 tweet yang telah melewati tahap *preprocessing*. Pada pelabelan mandiri, sentimen negatif mendominasi dengan total 483 tweet, diikuti sentimen positif sebanyak 380 tweet, dan netral 358 tweet. Sebaliknya, pada pelabelan berbasis *Lexicon-Based Approach*, terjadi kecenderungan sistem dalam mengklasifikasikan data ke kategori netral sebanyak 1.010 tweet, sementara positif 157 tweet, dan negatif 54 tweet. Perbedaan kontras ini mengindikasikan adanya variasi subjektivitas serta keterbatasan algoritma dalam menangkap nuansa bahasa manusia yang lebih kompleks. Melalui pengujian dengan confusion matrix, algoritma *Support Vector Machine* terbukti memiliki efektivitas yang baik dalam memetakan sentimen publik. Secara keseluruhan, model mampu mengklasifikasikan data berlabel mandiri dengan tingkat akurasi sebesar 71%, dan mencapai 96 % pada data berlabel berbasis *Lexicon-Based Approach*. Hal ini membuktikan bahwa integrasi *Vector Space Model* mampu menyajikan representasi fitur yang mendukung kinerja SVM dalam mengenali pola teks terkait kebijakan efisiensi anggaran. Bahwa penggunaan *Vector Space Model* sangat membantu SVM dalam mengenali pola teks pada kebijakan pemerintah efisiensi anggaran. Proses klasifikasi dilakukan setelah tahap pembobotan kata menggunakan TF-IDF dengan pembagian rasio data 80:20 (976 data latih dan 245 data uji). Hasil evaluasi menunjukkan bahwa pada pelabelan mandiri, model memiliki performa yang cukup seimbang dengan nilai *precision* sebesar 71%, *recall* sebesar 69%, dan F1-score sebesar 69%. Hal ini menunjukkan bahwa model cukup baik dalam mengenali data serta menjaga keseimbangan antara ketepatan dan cakupan klasifikasi. Sementara itu, pada pelabelan berbasis *Lexicon-Based Approach*, model menunjukkan performa yang lebih tinggi pada beberapa metrik. Model mencapai *precision* sebesar 98%, *recall* sebesar 67%, dan F1-*score* sebesar 75%. Nilai *precision* yang tinggi menunjukkan bahwa model sangat akurat dalam menentukan kelas prediksi, namun nilai *recall* yang lebih rendah mengindikasikan bahwa model belum mampu mengenali seluruh data pada setiap kelas secara menyeluruh. Dengan demikian, dapat disimpulkan bahwa model dengan pelabelan *Lexicon-Based Approach* cenderung lebih selektif dan akurat dalam prediksi, tetapi masih memiliki keterbatasan dalam menjangkau seluruh data, terutama pada kelas tertentu. Sebaliknya, model pelabelan mandiri menunjukkan performa yang lebih seimbang antara *precision* dan *recall*, meskipun tingkat ketepatannya masih di bawah model *Lexicon-Based Approach*.

Daftar Pustaka

[1] F. A.- Zahra, G. Prasiska, T. Anggraini, V. Silvia, And A. Firman, “Efesiensi Anggaran Pemerintahan Presiden Prabowo ; Implementasi Prinsip Transparansi Dan Akuntabilitas Sesuai Pasal 10 Undang-Undang Nomor 30 Tahun 2014 Tentang Administrasi Pemerintahan,” Vol. 10, No. 01, Pp. 610–622, 2025.

[2] Y. Yusfik, K. Nawangwulan, And L. H. Susilo, “Pengaruh Media Sosial Terhadap Pola Pencarian Informasi Kesehatan Pada Generasi Z The Influence Of Social Media On Health Information Search Patterns Among Generation Z,” Vol. 6, No. 2, Pp. 188–198, 2025.

[3] J. Khatib, S. Dalam, T. Methods, U. I. Arsyah, M. Pratiwi, And A. Muhammad, “Twitter Sentiment Analysis Of Public Space Opinions Using Svm And,” Vol. 13, No. 1, Pp. 387–394, 2024.

[4] R. Astuti, “Analisis Sentimen Masyarakat Indonesia Terhadap Sea Games 2023 Di Twitter Dengan Metode Naïve Bayes,” Vol. 7, No. 3, Pp. 728–738, 2023, Doi: 10.52362/Jisamar.V7i3.1150.

[5] P. Arifin, U. Mahdiyah, And A. B. Setiawan, “Performa Machine Learning Terhadap Analisis

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)

Sentimen Pemilu 2024 Melalui Media Digital," Vol. 8, 2024.
[6] N. Hayni, "Perancangan Information Retrieval (Ir) Untuk Pencarian Cerpen Menggunakan Metode Generalized Vector Space Model (Vsm)," Vol. Xv, No. 2, Pp. 197–208, 2022.
[7] O. I. Gifari, M. Adha, I. R. Hendrawan, And F. F. S. Durrand, "Analisis Sentimen Review Film Menggunakan Tf-Idf Dan Support Vector Machine," *J. Inf. Technol.*, Vol. 2, No. 1, Pp. 36–40, 2022.
[8] E. I. Saputra *Et Al.*, "Optimalisasi Algoritma Support Vector Machine Pada Aspect-Based Sentiment Analysis," Vol. 7, No. 1, Pp. 271–278, 2025.
[9] N. Saleh, A. Lasarudin, And W. Hasyim, "Analisis Perbandingan Algoritma Klasifikasi Support Vector Machine , Decision Tree , Naïve Bayes Bencana Banjir Berdasarkan Curah Hujan," 2024.
[10] Robert Antonius, A. R. Zulkarnain, And H. Irsyad, "Pendekatan Tf-Idf, Smote, Dan Svm Dalam Klasifikasi Sentimen Masyarakat Terhadap Pemblokiran Judi Online," *Bul. Ilm. Inform. Teknol.*, Vol. 2, No. 3, Pp. 115–122, 2024, Doi: 10.58369/Biit.V2i3.65.
[11] N. Amini, *Menggunakan Algoritma Klasifikasi Support Vector Machine Dengan Optimasi Particle Swarm Optimization (Studi Kasus : Indonesian Twitter Emotion Dataset) Program Studi Teknik Informatika 2023 M / 1445 H*. 2023.
[12] R. G. Guntara, "Pemanfaatan Google Colab Untuk Aplikasi Pendeteksian Masker Wajah Menggunakan Algoritma Deep Learning Yolov7," *J. Teknol. Dan Sist. Inf. Bisnis*, Vol. 5, No. 1, Pp. 55–60, 2023.
[13] F. A. Rezy *Et Al.*, "Provinsi Dki Jakarta , 2022). Dalam Dua Dekade Terakhir , Berbagai Faktor Seperti Urbanisasi , Migrasi , Wilayah Perkotaan . Ketimpangan Distribusi Penduduk Dapat Mengakibatkan Tekanan Pada Transportasi Ekonomi , Seperti Meningkatnya Pengangguran Dan Kem," Vol. 6, Pp. 50–63, 2025.
[14] K. S. Kumar *Et Al.*, "Sentiment Analysis Of Short Texts Using Svms And Vsms-Based Multiclass Semantic Classification Sentiment Analysis Of Short Texts Using Svms And," *Appl. Artif. Intell.*, Vol. 38, No. 1, 2024, Doi: 10.1080/08839514.2024.2321555.

Copyright © Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)