

ANALISIS SENTIMEN KENAIKAN PBB BERBASIS TIKTOK MENGUNAKAN SUPPORT VECTOR MACHINE

Erhan Perdana¹, Anggun Mentari², Uci Azhari Chaniago³, Juniardo Purba⁴, Agus Perdana Windarto⁵

Program Studi Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar

Jl. Jenderal Sudirman Blok A No. 1/2/3, Siantar Barat, Kota Pematang Siantar, Sumatera Utara

erhanperdana9@gmail.com¹, mentarianggun3@gmail.com², ucichaniago18@gmail.com³

juniardopurba22@gmail.com⁴, agus.perdana@amiktunasbangsa.ac.id⁵

Abstrak

Kebijakan kenaikan Pajak Bumi dan Bangunan Perdesaan dan Perkotaan (PBB-P2) memicu beragam respons masyarakat yang diekspresikan secara luas melalui media sosial. TikTok, sebagai platform dengan tingkat partisipasi tinggi dan gaya komunikasi yang spontan, menjadi ruang penting untuk mengamati opini publik terhadap kebijakan tersebut. Penelitian ini mengkaji pola sentimen masyarakat terhadap kebijakan kenaikan PBB-P2 berdasarkan komentar pengguna TikTok dengan pendekatan klasifikasi berbasis Support Vector Machine (SVM). Analisis dilakukan dengan membandingkan dua konfigurasi kernel, yaitu Linear dan Polynomial, untuk melihat perbedaan karakteristik hasil klasifikasi. Data diperoleh melalui proses pengumpulan komentar secara otomatis, diikuti dengan tahapan pembersihan teks, pembentukan fitur, dan pelabelan sentimen. Evaluasi model dilakukan menggunakan confusion matrix dan metrik klasifikasi untuk menilai distribusi prediksi sentimen. Hasil analisis menunjukkan adanya perbedaan pola kinerja antara kedua kernel, di mana kernel Linear cenderung memberikan hasil yang lebih stabil pada data dengan distribusi sentimen yang tidak seimbang. Temuan ini mengindikasikan bahwa kesesuaian antara karakteristik data teks dan konfigurasi model berperan penting dalam analisis sentimen kebijakan publik berbasis media sosial.

Kata Kunci : Analisis Sentimen, PBB-P2, Support Vector Machine, TikTok, Kebijakan Publik

Abstract

The increase in Land and Building Tax for Rural and Urban Areas (PBB-P2) has generated diverse public reactions that are widely expressed through social media platforms. TikTok, characterized by high user engagement and informal communication patterns, provides a relevant medium for observing public sentiment toward fiscal policies. This study examines public sentiment regarding the PBB-P2 increase by applying a Support Vector Machine (SVM)-based classification approach. Two kernel configurations, Linear and Polynomial, are compared to identify differences in classification behavior when handling social media text data. User comments were collected automatically and processed through text cleaning, feature extraction, and sentiment labeling stages prior to model training. Model performance was analyzed using confusion matrix-based evaluation to observe prediction patterns across sentiment classes. The findings indicate noticeable differences in classification stability between the two kernel types, with the Linear kernel showing more consistent behavior when applied to imbalanced sentiment distributions. These results suggest that selecting a kernel aligned with the characteristics of textual data is an important consideration in social media-based sentiment analysis of public policy issues.

Keyword : Sentiment Analysis, PBB-P2, Support Vector Machine, TikTok, Public Policy

PENDAHULUAN

Pajak Bumi dan Bangunan Perdesaan dan Perkotaan (PBB-P2) merupakan salah satu jenis pajak daerah yang dikenakan atas kepemilikan dan pemanfaatan tanah serta bangunan. PBB-P2 memiliki karakteristik khusus karena kewajibannya dirasakan langsung oleh masyarakat setiap tahun dan berkaitan dengan kondisi ekonomi rumah tangga. Oleh karena itu, kebijakan terkait tarif PBB-P2 berpotensi memengaruhi persepsi masyarakat terhadap beban pajak dan keadilan fiskal (Agustin et al., 2024).

Pada periode 2024–2025, kebijakan kenaikan PBB-P2 di sejumlah daerah di Indonesia memicu perhatian dan respons publik yang cukup luas. Beberapa daerah tercatat menerapkan kenaikan tarif dengan persentase tinggi, seperti Kabupaten Pati yang sempat menaikkan PBB sekitar 250% (BBC

News Indonesia, 2025), Kota Cirebon hingga 1000% (Kompas, 2025a), serta Kabupaten Jombang dan Kabupaten Semarang dengan kenaikan sekitar 400% (Kompas, 2025b; Metro TV News, 2025). Kenaikan tersebut menimbulkan berbagai reaksi masyarakat, mulai dari keluhan hingga penolakan, yang banyak disampaikan melalui media sosial. Kondisi ini menunjukkan bahwa kebijakan PBB-P2 tidak hanya berdampak pada penerimaan daerah, tetapi juga memunculkan persoalan sosial yang sensitif.

Sebagian besar penelitian mengenai PBB-P2 selama ini masih berfokus pada aspek administratif dan fiskal, seperti konsep pajak, dasar hukum, serta kontribusinya terhadap Pendapatan Asli Daerah (PAD). Agustin et al. (2024) menekankan pentingnya pemahaman regulasi dan mekanisme PBB-P2, sedangkan Pratiwi et al. (2020) menunjukkan bahwa efektivitas penerimaan PBB-P2 sangat dipengaruhi oleh tingkat kepatuhan wajib pajak. Namun, penelitian-penelitian tersebut belum secara langsung mengkaji bagaimana masyarakat merespons kebijakan kenaikan tarif PBB-P2 dari sisi persepsi dan opini publik.

Seiring meningkatnya penggunaan media sosial, analisis sentimen mulai digunakan untuk mengkaji respons masyarakat terhadap kebijakan pemerintah. Penelitian terkait kebijakan kenaikan pajak, seperti Pajak Pertambahan Nilai (PPN), menunjukkan bahwa sentimen publik di media sosial cenderung didominasi oleh pandangan negatif (Siagian & Painem, 2024 ; Hasan & Bimby, 2025). Selain itu, beberapa studi membuktikan bahwa algoritma Support Vector Machine (SVM) memiliki kinerja yang stabil dan akurat dalam mengklasifikasikan opini publik terhadap kebijakan pemerintah (Romadhona & Isnain, 2024).

Meskipun demikian, penelitian yang secara khusus menganalisis sentimen publik terhadap kebijakan kenaikan PBB-P2 masih terbatas. Sebagian besar studi sebelumnya menggunakan data dari platform Twitter/X, sementara pemanfaatan komentar pada platform TikTok sebagai sumber data opini publik belum banyak dilakukan. Padahal, TikTok memiliki tingkat partisipasi pengguna yang tinggi dan mencerminkan respons masyarakat yang lebih spontan dan informal (Setiawan & Alafia, 2025). Berdasarkan kondisi tersebut, Penelitian ini mengkaji respons sentimen masyarakat terhadap kebijakan kenaikan PBB-P2 di Indonesia dengan memanfaatkan algoritma Support Vector Machine (SVM) melalui perbandingan kernel Linear dan Polynomial. Analisis yang dilakukan diarahkan untuk mengidentifikasi kecenderungan opini publik yang berkembang di media sosial serta relevansinya sebagai dasar evaluasi kebijakan PBB-P2 di tingkat pemerintah daerah.

PENELITIAN RELEVAN

Penelitian mengenai Pajak Bumi dan Bangunan (PBB) umumnya menempatkan kebijakan ini dalam kerangka fiskal daerah dengan penekanan pada aspek regulasi, tata kelola pemungutan, serta kontribusinya terhadap Pendapatan Asli Daerah (PAD). Studi yang dilakukan oleh Agustin et al. (2024) dan Pratiwi et al. (2020) menekankan bahwa pelaksanaan PBB sangat berkaitan dengan tingkat pemahaman regulasi serta kepatuhan wajib pajak. Kajian-kajian tersebut memberikan dasar konseptual yang kuat mengenai peran PBB dalam sistem keuangan daerah, namun belum secara spesifik membahas respons masyarakat terhadap kebijakan kenaikan tarif PBB dari sudut pandang opini publik.

Seiring meningkatnya penggunaan media sosial sebagai ruang diskursus publik, pendekatan analisis sentimen mulai dimanfaatkan untuk mengkaji persepsi masyarakat terhadap berbagai kebijakan perpajakan dan kebijakan publik. Penelitian yang dilakukan oleh Siagian & Painem (2024) serta Hasan & Bimby (2025) menganalisis sentimen masyarakat terhadap kebijakan kenaikan Pajak Pertambahan Nilai (PPN) dan menemukan kecenderungan respons negatif dalam percakapan publik. Dalam penelitian tersebut, algoritma Support Vector Machine (SVM) digunakan sebagai salah satu metode klasifikasi teks. Temuan ini selaras dengan penelitian Romadhona & Isnain (2024) dan Styawati et al. (2021) serta Sriani et al. (2023) yang menunjukkan bahwa SVM dapat diterapkan untuk mengklasifikasikan opini masyarakat terhadap kebijakan pemerintah pada berbagai konteks isu.

Meskipun penerapan SVM dalam analisis sentimen kebijakan publik telah banyak dilakukan, penelitian yang secara khusus menyoroti sentimen masyarakat terhadap kebijakan kenaikan Pajak Bumi dan Bangunan Perdesaan dan Perkotaan (PBB-P2) masih relatif terbatas. Selain itu, sebagian besar penelitian terdahulu memanfaatkan data dari platform Twitter/X, sementara pemanfaatan komentar pada platform TikTok sebagai sumber data opini publik belum banyak dikaji. Berdasarkan kondisi tersebut, penelitian ini melanjutkan dan mengembangkan kajian sebelumnya dengan menganalisis sentimen publik terhadap kebijakan kenaikan PBB-P2 berbasis data TikTok serta membandingkan kinerja kernel Linear dan Polynomial pada algoritma SVM untuk menyesuaikan model klasifikasi dengan karakteristik data yang digunakan.

METODE PENELITIAN

Tujuan penelitian ini adalah untuk menganalisis data yang diperoleh dari media sosial melalui penerapan metode crawling serta pengolahan data secara terstruktur dan sistematis. Tahapan penelitian yang dilakukan meliputi beberapa proses utama sebagaimana dijelaskan pada bagian berikut.

Akuisisi Data Komentar Media Sosial

Data yang digunakan dalam penelitian ini diperoleh dari komentar pengguna pada platform TikTok. Pemilihan TikTok didasarkan pada karakteristiknya sebagai media sosial dengan tingkat partisipasi pengguna yang tinggi serta kecenderungan penyampaian opini secara spontan terhadap isu-isu publik. Dalam konteks kebijakan fiskal, khususnya kenaikan Pajak Bumi dan Bangunan (PBB) tahun 2025, kolom komentar TikTok menjadi sumber data yang relevan untuk menangkap respons masyarakat secara langsung dan aktual.

Akuisisi data dilakukan secara otomatis menggunakan pendekatan web crawling dengan memanfaatkan bahasa pemrograman Python pada lingkungan Google Colaboratory. Proses pengambilan data difokuskan pada komentar dari unggahan yang mengandung topik kenaikan PBB, dengan pemanfaatan kata kunci yang berkaitan langsung dengan isu tersebut, seperti “kenaikan PBB”, “PBB naik”, dan “pajak bumi dan bangunan”. Rentang waktu pengumpulan data berlangsung dari 6 Agustus hingga 5 Oktober 2025 dan menghasilkan total 5.949 komentar. Seluruh data disimpan dalam format .csv untuk keperluan pemrosesan lanjutan, tanpa menyertakan atribut identitas pengguna guna menjaga aspek etika penelitian dan perlindungan privasi.

Praprocessing dan Labelling Data

Data komentar yang diperoleh dari platform TikTok diproses terlebih dahulu melalui tahapan *preprocessing* sebelum digunakan pada proses pelatihan model klasifikasi. Tahap ini bertujuan untuk mengurangi noise dan meningkatkan kualitas representasi teks. Proses *preprocessing* mencakup penghapusan elemen yang tidak berkontribusi terhadap makna sentimen, seperti tautan, *emoji*, *mention*, *hashtag*, angka, serta simbol dan tanda baca yang tidak memiliki nilai semantik (Susanto & Suparwito, 2023). Setelah itu, dilakukan penyeragaman huruf melalui proses *lowercasing* untuk menyatukan variasi penulisan kata. Selanjutnya, kata-kata umum yang tidak memiliki pengaruh signifikan terhadap sentimen dihilangkan menggunakan daftar *stopword* bahasa Indonesia pada pustaka NLTK yang telah disesuaikan dengan konteks media sosial, termasuk istilah informal yang sering muncul dalam komentar. Tahap akhir *preprocessing* dilakukan dengan menerapkan proses *stemming* menggunakan pustaka Sastrawi untuk mengembalikan kata ke bentuk dasarnya, sehingga menghasilkan teks yang lebih ringkas dan konsisten serta meminimalkan variasi kosakata (Suryanto & Andriyani, 2025).

Setelah proses *preprocessing* selesai, dilakukan pelabelan sentimen secara otomatis menggunakan pendekatan *lexicon-based*. Pada tahap ini, polaritas sentimen ditentukan berdasarkan akumulasi skor kata positif dan negatif yang terdapat dalam setiap komentar. Sentimen diklasifikasikan ke dalam dua kelas, yaitu positif (label 1) untuk komentar yang bersifat mendukung atau memberikan penilaian positif terhadap kebijakan kenaikan Pajak Bumi dan Bangunan (PBB), serta negatif (label 0) untuk komentar yang mengandung kritik atau penolakan terhadap kebijakan tersebut. Pelabelan berbasis *lexicon* dalam penelitian ini digunakan sebagai *proxy label* untuk merepresentasikan

kecenderungan polaritas sentimen, sehingga tidak dimaksudkan untuk mencerminkan kebenaran absolut sentimen secara semantik. Oleh karena itu, hasil evaluasi model klasifikasi merefleksikan tingkat konsistensi algoritma Support Vector Machine (SVM) dalam mempelajari pola sentimen yang dibentuk oleh skema pelabelan *lexicon*, bukan akurasi terhadap label manual. Dataset yang telah melalui tahap *preprocessing* dan pelabelan sentimen selanjutnya digunakan dalam proses pembentukan fitur dan pelatihan model klasifikasi berbasis Support Vector Machine (SVM).

Implementasi Klasifikasi Berbasis Support Vector Machine

Proses klasifikasi sentimen pada penelitian ini dilakukan menggunakan algoritma Support Vector Machine (SVM) sebagai pendekatan pembelajaran terawasi untuk memisahkan komentar TikTok ke dalam dua kelas sentimen, yakni positif dan negatif. Penetapan label numerik digunakan semata-mata sebagai representasi kelas dalam proses komputasi dan tidak menunjukkan hubungan hierarkis antar sentimen. Algoritma SVM dipilih karena kemampuannya dalam membangun fungsi pemisah yang efektif pada data teks yang memiliki jumlah fitur besar dan sparsitas tinggi.

Sebelum model dilatih, data teks yang telah melalui tahap pembersihan selanjutnya direpresentasikan ke dalam bentuk numerik menggunakan skema pembobotan Term Frequency–Inverse Document Frequency (TF-IDF). Pendekatan ini digunakan untuk menangkap karakteristik distribusi kata dalam korpus dengan memberikan bobot yang proporsional terhadap tingkat kemunculan dan kekhasan suatu istilah. Dataset hasil ekstraksi fitur kemudian dipisahkan ke dalam data latih dan data uji dengan proporsi 80% dan 20% guna mengevaluasi kemampuan generalisasi model pada data yang tidak terlibat dalam proses pelatihan.

Evaluasi kinerja model dilakukan melalui perbandingan dua fungsi kernel yang berbeda, yaitu kernel linear dan kernel polynomial. Kernel linear diterapkan untuk memodelkan pemisahan kelas dalam ruang fitur asli, sedangkan kernel polynomial digunakan untuk mengakomodasi kemungkinan hubungan non-linear antar fitur melalui transformasi ruang. Perbandingan ini bertujuan untuk mengidentifikasi konfigurasi kernel yang paling sesuai dengan karakteristik pola bahasa pada komentar publik terkait kebijakan kenaikan Pajak Bumi dan Bangunan (PBB).

Parameter Pengujian

Evaluasi performa model klasifikasi dilakukan untuk menilai kemampuan algoritma Support Vector Machine (SVM) dalam mengklasifikasikan sentimen publik terhadap kebijakan kenaikan Pajak Bumi dan Bangunan (PBB). Evaluasi didasarkan pada *confusion matrix* yang menggambarkan perbandingan antara hasil prediksi model dan label sentimen pada data uji. Empat metrik evaluasi yang digunakan dalam penelitian ini meliputi *precision*, *recall*, *accuracy*, dan *F1-score*, sebagaimana umum diterapkan dalam penelitian analisis sentimen berbasis *machine learning* (Arsi et al., 2021 ; Indrayuni & Acmad Nurhadi, 2025). Metrik *precision* digunakan untuk mengukur tingkat ketepatan prediksi sentimen positif yang dihasilkan oleh model. Nilai *precision* menunjukkan sejauh mana komentar yang diprediksi sebagai sentimen positif benar-benar termasuk ke dalam kategori positif. Metrik ini penting untuk menghindari kesalahan klasifikasi berupa prediksi positif yang tidak relevan, terutama pada konteks kebijakan publik yang sensitif terhadap interpretasi opini masyarakat.

Metrik *recall* digunakan untuk mengukur kemampuan model dalam mengenali seluruh data yang berlabel positif. Nilai *recall* merepresentasikan proporsi komentar positif yang berhasil teridentifikasi dengan benar oleh model (Oktaviyanti et al., 2025). Metrik ini relevan untuk menilai sejauh mana model mampu menangkap sentimen dukungan atau penerimaan masyarakat terhadap kebijakan kenaikan PBB tanpa mengabaikan komentar positif yang ada. Selanjutnya, *accuracy* digunakan untuk mengukur tingkat ketepatan prediksi model secara keseluruhan, baik untuk sentimen positif maupun negatif. Nilai *accuracy* memberikan gambaran umum mengenai proporsi prediksi yang benar dibandingkan dengan keseluruhan data uji (Sihombing et al., 2023). Meskipun demikian, pada kondisi distribusi kelas yang tidak seimbang, metrik ini tidak digunakan sebagai satu-satunya indikator performa model.

Untuk melengkapi evaluasi, digunakan metrik *F1-score* yang merupakan ukuran keseimbangan antara *precision* dan *recall*. Metrik ini digunakan untuk memberikan penilaian yang lebih representatif terhadap performa model, khususnya ketika terdapat ketimpangan distribusi kelas sentimen. Dengan demikian, *F1-score* menjadi indikator penting dalam menilai kestabilan dan keandalan model klasifikasi SVM yang digunakan dalam penelitian ini.

HASIL DAN PEMBAHASAN

Dataset yang Digunakan

Dataset yang digunakan dalam penelitian ini berisi komentar pengguna TikTok yang menanggapi isu kenaikan Pajak Bumi dan Bangunan (PBB) di berbagai daerah di Indonesia. Data ini mencerminkan persepsi publik secara alami terhadap kebijakan fiskal yang menimbulkan beragam reaksi, mulai dari penolakan hingga dukungan. Secara keseluruhan, terkumpul sebanyak 5.949 entri komentar yang diambil pada periode 6 Agustus hingga 5 Oktober 2025, yaitu saat isu kenaikan PBB menjadi perbincangan luas di media sosial.

Komentar yang dikumpulkan berbentuk teks pendek dalam bahasa Indonesia dengan ragam bahasa informal khas media sosial. Beberapa komentar mengandung opini negatif yang menyoroti beban ekonomi akibat kenaikan pajak, sementara lainnya mengekspresikan pandangan positif dengan menekankan pentingnya peningkatan pendapatan daerah untuk Pembangunan. Dataset ini selanjutnya digunakan untuk analisis sentimen guna mengidentifikasi kecenderungan opini publik terhadap kebijakan kenaikan PBB.

Tabel 1. Sampel Dataset

No	Komentar
1	brutal banget 250% anying 🤔
2	250% njirr, apa ga melarat rakyat klo gitu? 😊
3	UMR gak sebanding dengan pajak yg mencekik sedang kan bupatinya tidur nyenyak
4	Viralin banggg, emak gue juga kget pas bayar pajak kemarin 😭😭
5	sy org pati,,sy yg cuma punya rumah dan tanah yg kecil mah santai aja, itu mungkin yg teriak2yg punya rumah gede tanah berhektar2

Text Preprocessing dan Labelling Sentimen

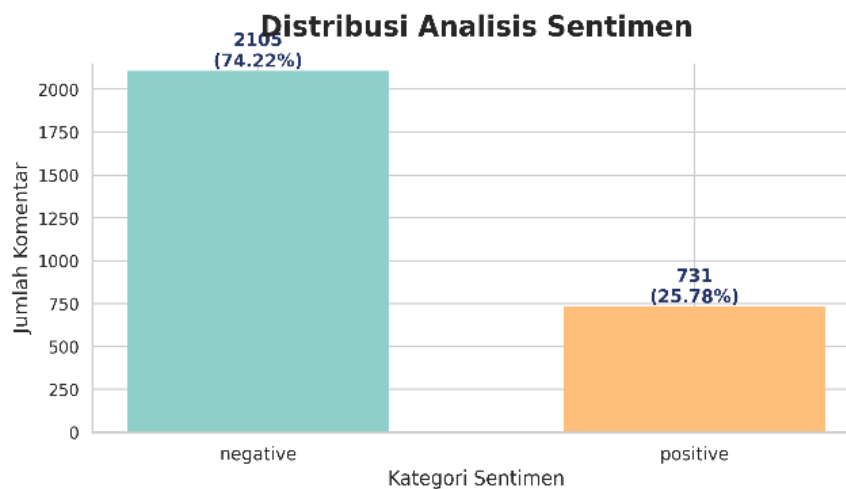
Tahap prapemrosesan teks dilakukan untuk menyiapkan data komentar hasil crawling agar layak digunakan dalam proses klasifikasi sentimen. Pada tahap ini, komentar dibersihkan dari berbagai unsur non-linguistik yang berpotensi mengganggu analisis, termasuk emoji, tanda baca, tautan, serta kata-kata umum yang tidak memiliki kontribusi semantik. Selanjutnya, seluruh teks dinormalisasi ke dalam bentuk huruf kecil dan diproses melalui stemming menggunakan pustaka Sastrawi guna menyatukan variasi morfologis kata ke dalam bentuk dasarnya. Setelah representasi teks terbentuk secara lebih konsisten, pelabelan sentimen dilakukan secara otomatis dengan pendekatan berbasis leksikon, di mana setiap komentar diklasifikasikan ke dalam kategori positif atau negatif berdasarkan kecenderungan polaritas kosakata yang muncul. Label numerik yang digunakan berfungsi sebagai penanda kelas dalam proses pembelajaran model dan tidak menunjukkan hubungan tingkatan antar sentimen.

Tabel 2. Simpel Hasil Praprocessing

Text Praprocessing	Sentiment
brutal an**	0
Anj** melarat rakyat	0
umr banding pajak cekik bupati tidur nyenyak	0

Text Praprocessing	Sentiment
viralin banggg emak kget bayar pajak kemarin	0
orang pati rumah tanah santai teriak rumah gede tanah hektar	1

Setelah proses *preprocessing* dan pelabelan selesai, jumlah data yang dapat digunakan berkurang menjadi 2.836 komentar dari total awal 5.949 data mentah. Pengurangan data terjadi akibat penghapusan komentar duplikat, *spam*, serta teks kosong yang teridentifikasi selama proses pembersihan. Distribusi hasil analisis sentimen ditunjukkan pada Gambar 1, di mana terdapat 2.105 komentar (74,22%) dengan sentimen negatif dan 731 komentar (25,78%) dengan sentimen positif.

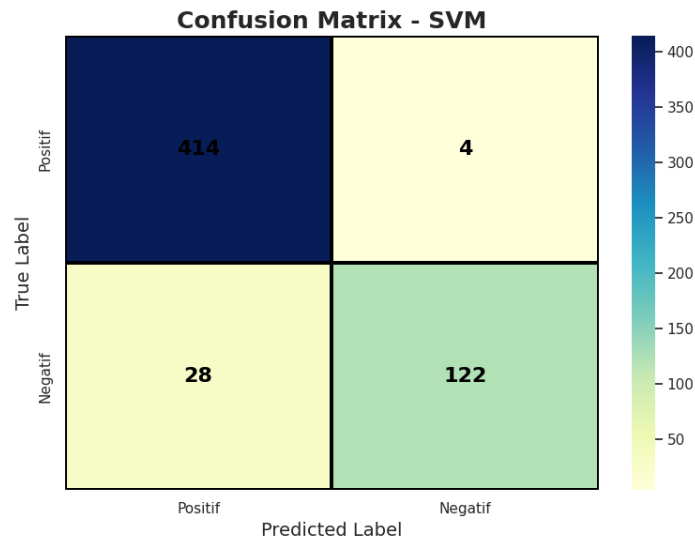


Gambar 1. Distribusi Sentimen

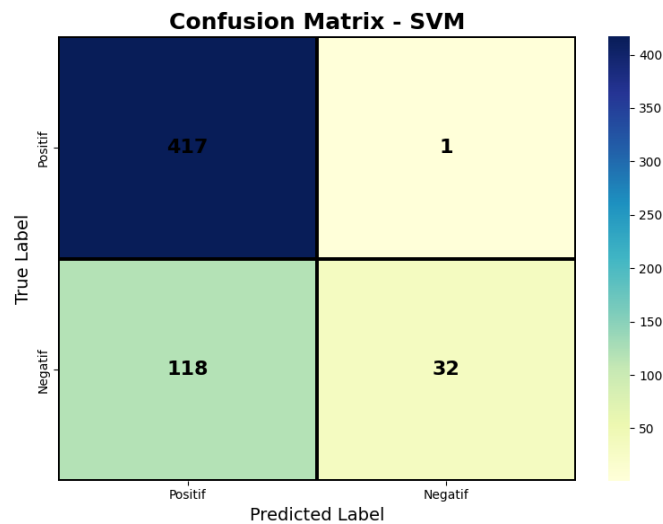
Gambar tersebut menunjukkan ketimpangan distribusi sentimen, di mana komentar dengan sentimen negatif mendominasi perbincangan publik di TikTok. Dataset yang telah melalui tahap *preprocessing* dan pelabelan ini selanjutnya digunakan dalam proses pembentukan fitur dan klasifikasi menggunakan algoritma *Support Vector Machine* (SVM).

Hasil Klasifikasi

Pengujian model dilakukan untuk menilai hasil klasifikasi sentimen yang dihasilkan oleh algoritma *Support Vector Machine* (SVM) dengan pendekatan kernel yang berbeda. Pada bagian ini, kinerja dua varian SVM, yaitu kernel Linear dan kernel Polynomial, dianalisis melalui penyajian confusion matrix. Proses evaluasi menggunakan dataset yang terdiri atas 2.268 data latih dan 568 data uji, yang sebelumnya telah diproses dan diberi label sentimen. Analisis difokuskan pada perbandingan distribusi prediksi benar dan salah dari masing-masing model dalam mengklasifikasikan komentar masyarakat terkait kebijakan kenaikan Pajak Bumi dan Bangunan (PBB). Hasil visualisasi confusion matrix untuk kedua model disajikan pada Gambar 2 dan Gambar 3 sebagai dasar interpretasi performa klasifikasi.



Gambar 2. Confusion Matrix SVM Kernel Linear



Gambar 3. Confusion Matrix SVM Kernal Polynomial

Berdasarkan Gambar 2, hasil klasifikasi menggunakan SVM dengan kernel Linear menunjukkan bahwa model mampu mengklasifikasikan 414 data sentimen positif (*True Positive*) dan 122 data sentimen negatif (*True Negative*) secara tepat. Meskipun demikian, masih terdapat 28 kesalahan klasifikasi *False Positive*, yaitu komentar negatif yang diprediksi sebagai positif, serta 4 kesalahan *False Negative*, yaitu komentar positif yang salah diklasifikasikan sebagai negatif. Pola kesalahan ini menunjukkan bahwa kernel Linear memiliki kemampuan pemisahan kelas yang relatif seimbang dengan tingkat kesalahan yang rendah pada kedua kelas sentimen.

Selanjutnya, Gambar 3 memperlihatkan hasil klasifikasi menggunakan SVM dengan kernel Polynomial. Model ini berhasil mengidentifikasi 417 data sentimen positif dan 32 data sentimen negatif secara benar. Namun, jumlah kesalahan *False Positive* yang dihasilkan cukup tinggi, yaitu sebanyak 118 data, meskipun kesalahan *False Negative* relatif kecil, yakni hanya 1 data. Kondisi ini mengindikasikan bahwa kernel Polynomial cenderung memberikan prediksi positif secara lebih agresif, sehingga berdampak pada rendahnya kemampuan model dalam mengenali sentimen negatif.

Berdasarkan perbandingan kedua *confusion matrix*, SVM dengan kernel Linear menghasilkan jumlah total prediksi benar yang lebih tinggi ($TP + TN = 536$) dibandingkan dengan kernel Polynomial ($TP + TN = 449$). Selain itu, kernel Linear menunjukkan distribusi kesalahan yang

lebih seimbang antara kelas positif dan negatif, sehingga dinilai lebih stabil dalam mengklasifikasikan sentimen publik. Sebaliknya, meskipun kernel Polynomial mampu mengenali sebagian besar sentimen positif, peningkatan jumlah *False Positive* menunjukkan adanya kecenderungan bias prediksi yang dapat menurunkan keandalan model secara keseluruhan.

Parameter Pengujian

Evaluasi kinerja model dilakukan untuk menilai kemampuan klasifikasi sentimen dari dua model Support Vector Machine (SVM), yaitu kernel Linear dan kernel Polynomial. Pengukuran performa didasarkan pada *confusion matrix* hasil pengujian terhadap 568 data uji. Metrik yang digunakan meliputi *accuracy*, *precision*, *recall*, dan *F1-score*. Untuk mengakomodasi ketidakseimbangan distribusi kelas sentimen, nilai *precision*, *recall*, dan *F1-score* dilaporkan menggunakan *weighted average*.

Tabel 3. Perbandingan Kinerja Model

Metrik	SVM Kernel Linear	SVM Kernel Polynomial
Accuracy	0.9437 (94.37%)	0.7905 (79.05%)
Precision (weighted)	0.94	0.83
Recall (weighted)	0.94	0.79
F1-Score (weighted)	0.94	0.74
Jumlah Data Uji	568	568

Berdasarkan Tabel 3, SVM dengan kernel Linear menunjukkan performa yang lebih baik dan konsisten pada seluruh metrik evaluasi. Nilai *accuracy* yang tinggi serta keseimbangan antara *precision* dan *recall* mengindikasikan bahwa model mampu mengklasifikasikan sentimen secara stabil pada data uji yang tidak seimbang.

Sebaliknya, SVM dengan kernel Polynomial menghasilkan performa yang lebih rendah, terutama pada nilai *recall* dan *F1-score*. Kondisi ini menunjukkan bahwa meskipun model masih mampu mengenali kelas mayoritas dengan baik, peningkatan kesalahan klasifikasi pada kelas minoritas menyebabkan penurunan keseimbangan performa secara keseluruhan. Dengan demikian, kernel Linear dinilai lebih sesuai untuk analisis sentimen opini publik dengan karakteristik teks yang beragam dan distribusi kelas yang tidak seimbang.

Pembahasan

Hasil evaluasi memperlihatkan adanya perbedaan kinerja antara dua konfigurasi kernel yang digunakan pada algoritma Support Vector Machine (SVM). Pada data komentar terkait kebijakan kenaikan Pajak Bumi dan Bangunan (PBB-P2), kernel Linear menunjukkan hasil klasifikasi yang lebih stabil dibandingkan kernel Polynomial, sebagaimana tercermin dari pencapaian metrik evaluasi yang relatif lebih seimbang. Temuan ini menunjukkan bahwa efektivitas model tidak hanya ditentukan oleh algoritma yang digunakan, tetapi juga oleh kesesuaian jenis kernel dengan karakteristik data teks yang dianalisis. Dengan demikian, pemilihan kernel menjadi faktor penting dalam membangun model klasifikasi sentimen yang mampu beradaptasi dengan pola bahasa komentar publik.

Secara teknis, perbedaan performa ini dapat dikaitkan dengan sifat data teks komentar TikTok yang digunakan dalam penelitian. Komentar publik pada isu kenaikan PBB-P2 cenderung singkat, bersifat langsung, dan memiliki pola leksikal yang relatif seragam, terutama pada komentar bernada negatif. Dalam kondisi seperti ini, kernel Linear mampu membentuk batas pemisah kelas secara efektif tanpa memerlukan transformasi non-linear yang kompleks. Sebaliknya, kernel Polynomial yang memproyeksikan data ke ruang berdimensi lebih tinggi menunjukkan kecenderungan menghasilkan kesalahan klasifikasi yang lebih besar, khususnya pada kelas negatif, yang mengindikasikan adanya sensitivitas berlebih terhadap variasi kecil pada data teks.

Temuan ini sejalan dengan hasil penelitian Romadhona & Isnain (2024) yang menganalisis sentimen masyarakat terhadap kebijakan kenaikan pajak hiburan menggunakan SVM dan melaporkan bahwa kernel Linear memberikan hasil klasifikasi yang lebih stabil dibandingkan kernel non-linear. Kesamaan hasil ini menunjukkan bahwa pada isu kebijakan fiskal dengan arah sentimen yang relatif jelas, pendekatan linear sering kali lebih sesuai. Selain itu, penelitian Siagian & Painem (2024) yang mengkaji respons masyarakat terhadap rencana kenaikan PPN juga menemukan dominasi sentimen negatif dengan pola bahasa yang homogen, sehingga klasifikasi sentimen dapat dilakukan secara efektif tanpa model yang terlalu kompleks.

Namun demikian, hasil penelitian ini berbeda dengan temuan Sriani et al. (2023) yang melaporkan bahwa kernel non-linear memberikan performa lebih baik pada analisis sentimen kebijakan subsidi kendaraan listrik. Perbedaan tersebut dapat dipahami karena konteks isu subsidi cenderung memunculkan opini yang lebih beragam dan ambivalen, sehingga memerlukan pemodelan non-linear untuk menangkap pola sentimen yang kompleks. Perbandingan ini menunjukkan bahwa kinerja kernel SVM sangat dipengaruhi oleh keragaman bahasa, struktur teks, serta distribusi sentimen pada data yang dianalisis.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa SVM dengan kernel Linear merupakan pilihan yang lebih sesuai untuk analisis sentimen opini publik pada komentar media sosial yang bersifat singkat dan memiliki kecenderungan pola bahasa yang seragam, seperti pada platform TikTok. Temuan ini tidak dimaksudkan untuk menyatakan superioritas mutlak suatu kernel, melainkan menunjukkan bahwa kesesuaian antara karakteristik data dan metode klasifikasi menjadi faktor penting dalam memperoleh hasil yang optimal. Oleh karena itu, pemilihan kernel dalam analisis sentimen perlu mempertimbangkan konteks isu dan sifat linguistik data yang digunakan.

SIMPULAN

Penelitian ini menunjukkan bahwa algoritma Support Vector Machine (SVM) dengan kernel Linear memiliki kinerja yang lebih konsisten dibandingkan kernel Polynomial dalam mengklasifikasikan sentimen publik terhadap kebijakan kenaikan Pajak Bumi dan Bangunan (PBB-P2) berdasarkan komentar media sosial TikTok. Hasil evaluasi menunjukkan bahwa kernel Linear mampu memberikan keseimbangan yang lebih baik antara akurasi dan stabilitas klasifikasi pada data teks yang memiliki distribusi sentimen tidak seimbang dan pola bahasa yang relatif seragam. Meskipun demikian, penelitian ini masih memiliki keterbatasan pada penggunaan pelabelan sentimen berbasis lexicon dan cakupan data yang terbatas pada satu platform media sosial. Oleh karena itu, penelitian selanjutnya disarankan untuk menggabungkan pelabelan otomatis dengan validasi manual serta memperluas sumber data dan pendekatan klasifikasi guna memperoleh pemahaman yang lebih komprehensif terhadap dinamika sentimen publik pada isu kebijakan fiskal.

DAFTAR PUSTAKA

- Agustin, P. D., Dinanty, D., & Nasution, J. H. (2024). Pajak Bumi Dan Bangunan (PBB). *MUSYTARI : Neraca Manajemen, Ekonomi*, 5(5). <https://doi.org/10.8734/mnmae.v1i2.359>
- Arsi, P., Wahyudi, R., & Waluyo, R. (2021). Optimasi SVM Berbasis PSO pada Analisis Sentimen Wacana Pindah Ibu Kota Indonesia. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(2), 231–237. <https://doi.org/10.29207/resti.v5i2.2698>
- BBC News Indonesia. (2025). *Bupati Pati batalkan PBB 250%, warga tetap menuntut dia lengser*. BBC News Indonesia. <https://www.bbc.com/indonesia/articles/cp89dvz1z65o>
- Hasan, M. A., & Bimby, N. P. (2025). Analisis Sentimen Publik Terhadap Kenaikan Pajak PPN di Indonesia Tahun 2024 Menggunakan Algoritma Machine Learning. *JURNAL FASILKOM*, 15(1), 179–184.
- Indrayuni, E., & Acmad Nurhadi. (2025). Analisis Sentimen Aplikasi Tiktok Shop Seller Center Menggunakan Naive Bayes, Svm Dan Logistic Regression. *INTI Nusa Mandiri*, 20(1), 26–34. <https://doi.org/10.33480/inti.v20i1.6851>
- Kompas. (2025a). *PBB Kota Cirebon Naik hingga 1.000 Persen, Warga Berikan Batas Waktu Revisi*. Kompas.Id. <https://www.kompas.id/artikel/pbb-kota-cirebon-hingga-1000-persen-warga-berikan-batas-waktu-revisi>
- Kompas. (2025b). *Polemik PBB Jombang Naik 400 Persen, Pemkab Revisi Perda hingga Diskon BPHTB*. Kompas.Id. <https://www.kompas.id/artikel/polemik-pbb-jombang-revisi-perda-hingga-diskon-bphtb>

- Metro TV News. (2025). *Warga Kabupaten Semarang Kaget, Bayar PBB Naik 400 Persen*. MetroTVNews.Com. <https://www.metrotvnews.com/read/NA0CE540-warga-kabupaten-semarang-kaget-bayar-pbb-naik-400-persen>
- Oktaviyanti, A., Firmansyah, & Kadafi, A. R. (2025). Analisis Sentimen Ulasan Pelecehan Di Media Sosial Tiktok Menggunakan Svm Dan Pengelompokan Pola Kejadian Dengan K-Means Clustering. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(3), 1938–1945. <https://doi.org/10.23960/jitet.v13i3S1.8149>
- Pratiwi, H., Muhaimin, & Rayyani, W. O. (2020). Kontribusi Pajak Bumi Dan Bangunan (Pbb) Dalam Meningkatkan Penerimaan Pajak Daerah. *Amnesty: Jurnal Riset Perpajakan*, 3, 24–32.
- Romadhona, W., & Isnain, A. R. (2024). Analisis Sentimen Pengguna Media Sosial Terhadap Kebijakan Kenaikan Pajak Hiburan Menggunakan Metode Svm (Support Vector Machine). *JIPi (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 9(4), 2185–2195. <https://doi.org/10.29100/jipi.v9i4.5603>
- Setiawan, K., & Alafia, E. (2025). Analisis Sentimen Komentar TikTok terhadap Kebijakan Larangan Wisuda Sekolah oleh Gubernur Jawa Barat Menggunakan Algoritma Naive Bayes. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi (JIMIK)*, 6(3), 1834–1841. <https://doi.org/10.63447/jimik.v6i3.1615>
- Siagian, J. K., & Painem. (2024). Analisis Sentimen Masyarakat Indonesia Terhadap Rencana Kenaikan Ppn Menjadi 12% Di Media Sosial X Dengan Metode Naïve Bayes. *SENAFTI : Seminar Nasional Mahasiswa Fakultas Teknologi Informasi*, 3(2), 779–786.
- Sihombing, E., Halimi Dar, M., & Nasution, F. A. (2023). Comparison Of Machine Learning Algorithms In Public Sentiment Analysis Of TAPERA Policy. *International Journal of Science, Technology & Management*, 5(3), 1089–1098. <http://ijstm.inarah.co.id>
- Sriani, Suhardi, & Gultom, I. F. (2023). Analisis Sentimen Kebijakan Pemberian Subsidi Motor Listrik Menggunakan Metode Support Vector Machine. *JURNAL FASILKOM*, 13, 511–517.
- Styawati, Hendrastuty, N., Isnain, A. R., & Ramadhani, A. Y. (2021). Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine. *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, 6(3), 150–155. <https://doi.org/10.30591/jpit.v6i3.2870>
- Suryanto, & Andriyani, W. (2025). Sentiment Analysis of X Platform on Viral “Fufufafa” Account Issue in Indonesia Using SVM. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 19(1), 95–104. <https://doi.org/10.22146/ijccs.104158>
- Susanto, N. W., & Suparwito, H. (2023). SVM-PSO Algorithm for Tweet Sentiment Analysis #BesokSenin. *Indonesian Journal of Information Systems (IJIS)*, 6(1), 36–37.