



Perbandingan Kinerja Algoritma Machine Learning dalam Klasifikasi Sentimen Komentar Publik terhadap Pelayanan Perpajakan

Dadang Kurniawan*, Windu Gata, Taufik Asra

Teknologi Informasi, Ilmu Komputer, Universitas Nusa Mandiri, Jakarta
Jl. Kramat Raya No.18, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta, Indonesia
Email: ¹*14240030@nusamandiri.ac.id, ²windu@nusamandiri.ac.id, ³taufik.tas@bsi.ac.id

Email Penulis Korespondensi: 14240030@nusamandiri.ac.id

Submitted: 09/01/2026; Accepted: 31/01/2026; Published: 31/01/2026

Abstrak—Identifikasi dini transformasi ganas pada leukoplakia oral sangat penting untuk mencegah perkembangan menjadi karsinoma sel skuamosa oral (Oral Squamous Cell Carcinoma/O SCC). Penilaian klinis konvensional masih memiliki keterbatasan dalam hal akurasi dan interpretabilitas, sehingga diperlukan pendekatan prediktif yang andal dan transparan. Penelitian ini bertujuan untuk mengevaluasi kinerja model machine learning yang bersifat interpretable dalam memprediksi transformasi ganas leukoplakia dan OSCC berdasarkan data klinis dan histopatologis. Penelitian ini menggunakan data retrospektif sebanyak 237 rekam medis pasien yang dianalisis dengan beberapa model interpretable, yaitu Explainable Boosting Machine (EBM), Generalized Additive Models (GAM), dan Symbolic Regression, serta dibandingkan dengan model black-box seperti Random Forest dan Deep Neural Network. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, sensitivitas, spesifisitas, dan area under the curve (AUC). Hasil penelitian menunjukkan bahwa model interpretable memiliki kinerja prediktif yang kompetitif dibandingkan model black-box, sekaligus memberikan tingkat transparansi dan kemudahan interpretasi yang lebih baik. Analisis kontribusi fitur menunjukkan bahwa karakteristik histopatologis merupakan faktor yang paling berpengaruh dalam prediksi keganasan. Temuan ini menunjukkan bahwa model machine learning yang interpretable berpotensi besar untuk digunakan sebagai sistem pendukung keputusan klinis dalam deteksi dini kanker mulut.

Kata Kunci: Analisis Sentimen; Klasifikasi Sentimen; Machine Learning; Komentar Publik; Pelayanan Perpajakan

Abstract—Early identification of malignant transformation in oral leukoplakia is crucial to prevent progression to Oral Squamous Cell Carcinoma (OSCC). However, conventional clinical assessment still faces limitations in terms of accuracy and interpretability, highlighting the need for reliable and transparent predictive approaches. This study aims to evaluate the performance of interpretable machine learning models in predicting malignant transformation of oral leukoplakia and OSCC based on clinical and histopathological data. A retrospective dataset consisting of 237 patient medical records was analyzed using several interpretable models, including Explainable Boosting Machine (EBM), Generalized Additive Models (GAM), and Symbolic Regression, and compared with black-box models such as Random Forest and Deep Neural Network. Model performance was evaluated using accuracy, sensitivity, specificity, and area under the curve (AUC). The results demonstrate that interpretable models achieve competitive predictive performance compared to black-box models while offering superior transparency and interpretability. Feature contribution analysis indicates that histopathological characteristics are the most influential factors in malignancy prediction. These findings suggest that interpretable machine learning models have strong potential as clinical decision support systems for early oral cancer detection.

Keywords: Sentiment Analysis; Sentiment Classification; Machine Learning; Public Comments; Tax Services.

1. PENDAHULUAN

Perkembangan teknologi informasi telah mengubah cara masyarakat berinteraksi dengan instansi pemerintah melalui berbagai platform digital, seperti media sosial, forum daring, dan aplikasi layanan publik. Interaksi tersebut menghasilkan komentar publik yang merepresentasikan opini masyarakat terhadap kualitas pelayanan yang diberikan oleh pemerintah [1], [2]. Komentar publik ini menjadi sumber data penting untuk mengevaluasi dan meningkatkan kualitas pelayanan publik secara berkelanjutan [3].

Pelayanan perpajakan merupakan salah satu sektor pelayanan publik yang memiliki peran strategis dalam mendukung penerimaan negara dan meningkatkan kepatuhan wajib pajak. Di Indonesia, reformasi administrasi perpajakan yang diiringi dengan digitalisasi layanan telah meningkatkan intensitas interaksi antara masyarakat dan otoritas perpajakan di ruang digital [4], [5]. Kondisi ini semakin diperkuat oleh munculnya berbagai isu kebijakan dan pelayanan perpajakan dalam beberapa tahun terakhir, yang memicu beragam respons dan sentimen publik secara daring [6].

Analisis sentimen merupakan bagian dari text mining dan natural language processing (NLP) yang bertujuan untuk mengidentifikasi polaritas opini dalam teks, seperti sentimen positif, negatif, dan netral [7]. Sejumlah penelitian sebelumnya telah menerapkan teknik machine learning untuk klasifikasi sentimen pada berbagai konteks pelayanan publik, termasuk layanan transportasi, kesehatan, dan administrasi pemerintahan [8], [9]. Algoritma seperti Naïve Bayes, Support Vector Machine, dan K-Nearest Neighbor banyak digunakan karena kemudahan implementasi dan kinerjanya yang relatif baik pada data teks [10].

Namun demikian, sebagian besar penelitian terdahulu masih bersifat umum dan belum secara spesifik mengkaji sentimen masyarakat terhadap pelayanan perpajakan di Indonesia sebagai objek penelitian utama. Selain itu, banyak penelitian hanya berfokus pada perbandingan algoritma machine learning tanpa mengkaji secara mendalam pengaruh tahapan pra-pemrosesan teks terhadap kinerja klasifikasi sentimen [11]. Padahal, tahapan pra-

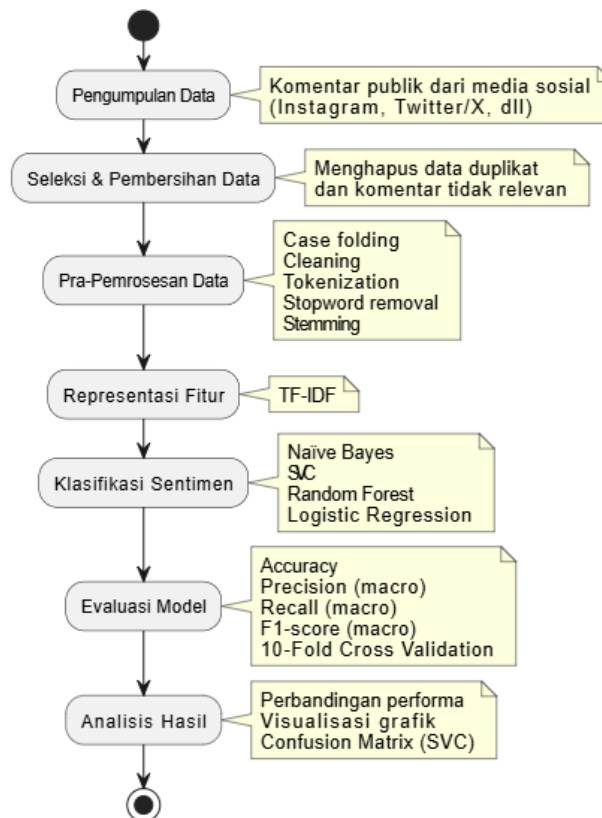
pemrosesan seperti case folding, tokenization, stopword removal, dan stemming memiliki peran penting dalam membentuk representasi fitur teks yang berpengaruh langsung terhadap performa algoritma klasifikasi [12], [13].

Berdasarkan celah penelitian tersebut, kebaruan penelitian ini terletak pada dua aspek utama. Pertama, penelitian ini secara khusus berfokus pada analisis sentimen komentar publik terkait pelayanan perpajakan di Indonesia, yang masih relatif terbatas dibahas dalam penelitian sentimen pelayanan publik sebelumnya. Kedua, penelitian ini tidak hanya membandingkan kinerja beberapa algoritma machine learning, tetapi juga mengevaluasi pengaruh tahapan pra-pemrosesan teks yang diterapkan secara konsisten terhadap seluruh algoritma yang diuji. Pendekatan ini diharapkan mampu memberikan pemahaman yang lebih komprehensif mengenai kombinasi metode yang paling optimal untuk klasifikasi sentimen pelayanan perpajakan [14].

Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja beberapa algoritma machine learning dalam mengklasifikasikan sentimen komentar publik terhadap pelayanan perpajakan. Evaluasi kinerja dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian diharapkan dapat memberikan kontribusi akademik dalam bidang analisis sentimen serta kontribusi praktis bagi instansi perpajakan dalam meningkatkan kualitas pelayanan berbasis data.

2. METODOLOGI PENELITIAN

Alur Metode Penelitian Klasifikasi Sentimen Komentar Publik terhadap Pelayanan Perpajakan



Gambar 1 Alur Metode Penelitian Klasifikasi Sentimen Komentar Publik terhadap Pelayanan Perpajakan

Alur metode penelitian pada klasifikasi sentimen komentar publik terhadap pelayanan perpajakan dimulai dari tahap pengumpulan data. Data dikumpulkan dari berbagai media sosial, seperti Instagram dan Twitter (X), yang berisi komentar masyarakat terkait pelayanan perpajakan. Data yang diperoleh masih bersifat mentah sehingga perlu dilakukan proses seleksi dan pembersihan data. Pada tahap ini dilakukan penghapusan data duplikat serta komentar yang tidak relevan dengan topik penelitian agar kualitas dataset tetap terjaga.

Selanjutnya, data memasuki tahap pra-pemrosesan. Tahapan ini bertujuan untuk menyiapkan data teks agar lebih terstruktur dan siap digunakan dalam proses klasifikasi. Proses pra-pemrosesan meliputi case folding untuk menyeragamkan huruf, cleaning untuk menghapus karakter yang tidak diperlukan, tokenization untuk memecah kalimat menjadi kata, stopword removal untuk menghilangkan kata umum yang tidak memiliki makna penting, serta stemming untuk mengubah kata menjadi bentuk dasarnya.

Setelah tahap pra-pemrosesan selesai, data direpresentasikan ke dalam bentuk numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk mengukur tingkat kepentingan suatu kata dalam dokumen sehingga dapat digunakan sebagai fitur dalam proses pembelajaran mesin.



Tahap berikutnya adalah proses klasifikasi sentimen. Pada penelitian ini digunakan beberapa algoritma machine learning, yaitu Naïve Bayes, Support Vector Classifier (SVC), Random Forest, dan Logistic Regression. Setiap algoritma dilatih menggunakan data latih dan kemudian diuji menggunakan data uji untuk menghasilkan prediksi sentimen.

Setelah proses klasifikasi, dilakukan evaluasi model untuk mengukur kinerja masing-masing algoritma. Evaluasi dilakukan menggunakan metrik accuracy, precision (macro), recall (macro), dan F1-score (macro). Selain itu, digunakan metode 10-Fold Cross Validation untuk memastikan hasil evaluasi lebih stabil dan objektif. Tahap terakhir adalah analisis hasil. Pada tahap ini dilakukan perbandingan performa antar algoritma menggunakan visualisasi grafik. Selain itu, dilakukan analisis confusion matrix pada model terbaik, yaitu SVC, untuk melihat distribusi hasil klasifikasi secara lebih rinci pada setiap kelas sentimen.

2.1 Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan **kuantitatif** dengan metode **eksperimen**. Pendekatan kuantitatif dipilih karena penelitian berfokus pada pengukuran kinerja algoritma machine learning menggunakan metrik evaluasi numerik seperti accuracy, precision, recall, dan F1-score [15]. Metode eksperimen dilakukan dengan membangun, melatih, dan menguji beberapa model klasifikasi sentimen untuk kemudian dibandingkan kinerjanya secara objektif [16].

2.2 Objek dan Subjek Penelitian

Objek penelitian ini adalah komentar publik yang berkaitan dengan pelayanan perpajakan di Indonesia yang diperoleh dari platform media sosial, khususnya komentar pada Instagram dan Twitter (X). Media sosial tersebut dipilih karena memiliki tingkat interaksi yang tinggi dan sering digunakan oleh masyarakat untuk menyampaikan opini, keluhan, maupun apresiasi terhadap pelayanan publik, termasuk pelayanan perpajakan.

Subjek penelitian berupa data teks komentar masyarakat yang mengandung opini atau sentimen terhadap pelayanan perpajakan. Komentar yang dikumpulkan berasal dari unggahan, balasan, dan diskusi publik yang relevan dengan topik perpajakan. Setiap komentar selanjutnya diklasifikasikan ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral, guna merepresentasikan kecenderungan opini publik secara lebih terstruktur. Pendekatan klasifikasi tiga kelas sentimen ini merupakan metode yang umum digunakan dalam penelitian analisis sentimen berbasis media sosial [18].

2.3 Teknik Pengumpulan Data

Data dikumpulkan menggunakan teknik **web scraping** dan pengumpulan data sekunder dari platform daring seperti media sosial dan forum publik. Teknik ini banyak digunakan dalam penelitian analisis sentimen karena mampu memperoleh data dalam jumlah besar dan mencerminkan opini masyarakat secara alami [17]. Data yang diperoleh kemudian diseleksi untuk memastikan relevansi dengan topik pelayanan perpajakan serta menghilangkan data duplikat dan tidak valid [18].

2.4 Pra-Pemrosesan Data

Tahap pra-pemrosesan data dilakukan untuk meningkatkan kualitas data teks sebelum proses pelatihan model. Tahapan preprocessing yang diterapkan meliputi **case folding**, **cleaning**, **tokenization**, dan **stopword removal**. Pra-pemrosesan bertujuan untuk mengurangi noise, menyederhanakan teks, serta meningkatkan akurasi model klasifikasi sentimen [19].

2.5 Representasi Fitur

Data teks yang telah melalui tahap pra-pemrosesan kemudian direpresentasikan ke dalam bentuk numerik menggunakan metode **Term Frequency–Inverse Document Frequency (TF-IDF)**. Metode TF-IDF banyak digunakan dalam klasifikasi teks karena mampu merepresentasikan tingkat kepentingan suatu kata dalam dokumen secara efektif dan efisien [20].

2.6 Algoritma Klasifikasi

Penelitian ini menggunakan empat algoritma machine learning yang umum diterapkan dalam klasifikasi teks dan analisis sentimen, yaitu **Naïve Bayes**, **Support Vector Classifier (SVC)**, **Random Forest**, dan **Logistic Regression**. Keempat algoritma tersebut dipilih karena memiliki karakteristik yang berbeda dalam mempelajari pola data teks dan telah banyak digunakan pada penelitian sebelumnya [21].

2.7 Pembagian Data

Pada penelitian ini, proses pembagian data dilakukan menggunakan metode **K-Fold Cross Validation** dengan nilai **K = 10**. Metode ini membagi dataset menjadi 10 bagian (fold) yang berukuran relatif sama. Pada setiap iterasi, sebanyak **9fold** digunakan sebagai data latih (training data) dan **1fold** digunakan sebagai data uji (testing data), kemudian proses tersebut diulang hingga seluruh fold digunakan sebagai data uji secara bergantian [9].



Penggunaan metode 10-Fold Cross Validation bertujuan untuk menghasilkan evaluasi kinerja model yang lebih stabil dan representatif, serta mengurangi ketergantungan hasil terhadap satu skema pembagian data acak. Pendekatan ini sangat sesuai untuk penelitian perbandingan algoritma machine learning karena mampu memberikan estimasi performa yang lebih objektif dan adil pada setiap algoritma yang diuji.

Nilai kinerja akhir dari masing-masing algoritma diperoleh dengan menghitung rata-rata (average) dari seluruh hasil evaluasi pada setiap fold, sehingga hasil yang diperoleh mencerminkan performa model secara keseluruhan.

2.8 Proses Pelatihan dan Pengujian Model

Pada tahap pelatihan, masing-masing algoritma dilatih menggunakan data latih yang telah direpresentasikan dalam bentuk TF-IDF. Selanjutnya, model yang telah dilatih diuji menggunakan data uji untuk menghasilkan prediksi sentimen. Hasil prediksi tersebut dibandingkan dengan label aktual untuk mengetahui tingkat keberhasilan klasifikasi [22].

2.9 Metode Evaluasi

Evaluasi kinerja model dilakukan menggunakan metrik **accuracy**, **precision (macro)**, **recall (macro)**, dan **F1-score (macro)**. Penggunaan macro-average bertujuan untuk memberikan bobot yang sama pada setiap kelas sentimen sehingga hasil evaluasi lebih adil, terutama pada dataset dengan distribusi kelas yang tidak seimbang [23].

Pengukuran kinerja model klasifikasi sentimen pada penelitian ini dilakukan menggunakan beberapa metrik evaluasi yang diperoleh dari confusion matrix, yaitu accuracy, precision, recall, dan F1-score. Confusion matrix terdiri dari empat komponen utama, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Berdasarkan komponen tersebut, perhitungan masing-masing metrik evaluasi dirumuskan sebagai berikut.

Accuracy digunakan untuk mengukur tingkat ketepatan model dalam mengklasifikasikan seluruh data secara keseluruhan. Nilai accuracy diperoleh dari perbandingan jumlah prediksi yang benar terhadap total seluruh data, sebagaimana ditunjukkan pada Persamaan[24] (1).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision digunakan untuk mengukur ketepatan model dalam memprediksi suatu kelas tertentu. Precision menunjukkan proporsi data yang diprediksi positif dan benar-benar termasuk dalam kelas positif, sebagaimana dirumuskan pada Persamaan [25](2).

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

Recall digunakan untuk mengukur kemampuan model dalam mengidentifikasi seluruh data yang benar-benar termasuk dalam suatu kelas tertentu. Nilai recall dirumuskan pada Persamaan[26] (3).

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

F1-score merupakan nilai harmonisasi antara precision dan recall yang digunakan untuk menyeimbangkan kedua metrik tersebut. Perhitungan F1-score dirumuskan pada Persamaan [27](4).

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Pada penelitian ini, perhitungan nilai precision, recall, dan F1-score menggunakan pendekatan macro-average, yaitu dengan menghitung nilai metrik pada setiap kelas sentimen (positif, negatif, dan netral), kemudian dirata-ratakan. Pendekatan ini digunakan untuk memberikan bobot yang sama pada setiap kelas sentimen sehingga hasil evaluasi kinerja model menjadi lebih objektif, terutama pada kondisi distribusi data yang tidak seimbang.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan dan Pra-Pemrosesan Data

Data yang digunakan dalam penelitian ini berupa komentar publik terkait pelayanan perpajakan yang diperoleh dari media sosial dan situs ulasan daring. Data yang terkumpul masih bersifat tidak terstruktur sehingga perlu melalui tahapan pra-pemrosesan sebelum digunakan dalam proses klasifikasi sentimen. Tahapan pra-pemrosesan meliputi **cleaning data**, **case folding**, **normalization**, **tokenizing**, **stopword removal**, dan **stemming**. Proses ini bertujuan untuk menghilangkan noise, menyeragamkan bentuk kata, serta meningkatkan kualitas data teks.

3.2 Hasil Pelabelan Sentimen dan Ekstraksi Fitur

Setelah proses pra-pemrosesan, data komentar diberi label sentimen ke dalam tiga kategori, yaitu **positif**, **negatif**, dan **netral**. Proses pelabelan ini menjadi dasar dalam pembentukan dataset klasifikasi sentimen. Data teks yang telah diberi label kemudian diekstraksi fiturnya menggunakan metode **Term Frequency–Inverse Document**



Frequency (TF-IDF). Metode ini digunakan untuk merepresentasikan teks ke dalam bentuk numerik berdasarkan tingkat kepentingan suatu kata dalam dokumen. Hasil ekstraksi fitur TF-IDF selanjutnya digunakan sebagai input pada proses pelatihan model machine learning.

3.3 Hasil Pembagian Dataset

Dataset yang telah diekstraksi fiturnya kemudian dibagi menjadi dua bagian, yaitu **data latih sebesar 80%** dan **data uji sebesar 20%**. Pembagian dataset ini dilakukan untuk memastikan bahwa model yang dibangun dapat diuji menggunakan data yang belum pernah dilihat sebelumnya, sehingga hasil evaluasi kinerja lebih objektif.

3.4 Hasil Klasifikasi Sentimen

Proses klasifikasi sentimen pada penelitian ini dilakukan menggunakan empat algoritma machine learning, yaitu Naïve Bayes, Support Vector Classifier (SVC), Random Forest, dan Logistic Regression. Seluruh algoritma dilatih dan diuji menggunakan skema evaluasi yang sama untuk memastikan bahwa perbandingan kinerja dilakukan secara adil dan objektif. Hasil evaluasi kinerja masing-masing algoritma berdasarkan metrik accuracy, precision, recall, dan F1-score disajikan pada Tabel 1.

Tabel 1. Hasil Evaluasi Kinerja Algoritma Klasifikasi Sentimen

Algoritma	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	0.8878	0.9200	0.9155	0.9145
SVC	0.9769	0.9829	0.9825	0.9826
Random Forest	0.9736	0.8376	0.9801	0.8766
Logistic Regression	0.9736	0.7308	0.7351	0.7327

Berdasarkan Tabel 1, algoritma SVC menunjukkan kinerja terbaik dibandingkan algoritma lainnya dengan nilai accuracy, precision, recall, dan F1-score yang secara konsisten tinggi. Hal ini menunjukkan bahwa SVC mampu mengklasifikasikan data sentimen secara akurat pada seluruh kelas sentimen, baik kelas mayoritas maupun minoritas. Kinerja SVC yang stabil ini mengindikasikan bahwa algoritma tersebut lebih tahan terhadap pengaruh ketidakseimbangan kelas (imbalanced data) yang umum terjadi pada data komentar publik.

Algoritma Naïve Bayes menghasilkan kinerja yang cukup baik dengan nilai F1-score sebesar 0,9145. Meskipun nilai accuracy yang diperoleh lebih rendah dibandingkan SVC, nilai precision dan recall yang relatif seimbang menunjukkan bahwa Naïve Bayes mampu menangkap pola distribusi sentimen dengan cukup baik, terutama pada data teks yang direpresentasikan menggunakan TF-IDF.

Sebaliknya, algoritma Random Forest dan Logistic Regression menunjukkan fenomena akurasi tinggi tetapi metrik lain relatif rendah. Kedua algoritma tersebut memiliki nilai accuracy sebesar 0,9736, namun nilai precision, recall, dan F1-score berada pada rentang yang lebih rendah, khususnya pada Logistic Regression. Secara matematis, kondisi ini dapat terjadi apabila dataset memiliki distribusi kelas yang tidak seimbang (imbalanced data), di mana model cenderung memprediksi kelas mayoritas dengan sangat baik, tetapi kurang mampu mengenali kelas minoritas.

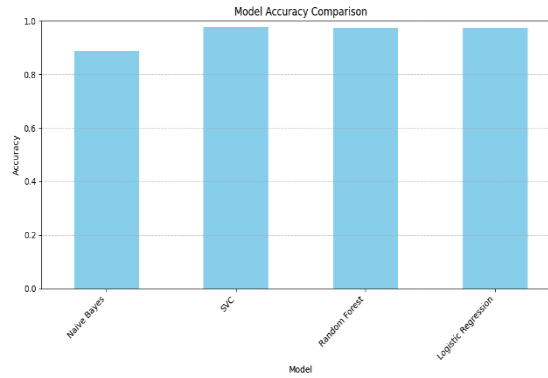
Pada kondisi tersebut, nilai accuracy dapat menjadi metrik yang menyesatkan karena hanya mencerminkan keberhasilan model dalam memprediksi kelas dominan, sementara kegagalan dalam mengklasifikasikan kelas minoritas tidak tercermin secara signifikan. Nilai precision dan recall yang rendah pada Logistic Regression menunjukkan bahwa model ini kurang mampu membedakan batas antar kelas sentimen secara optimal pada data yang tidak seimbang.

Dibandingkan Logistic Regression, Random Forest masih menunjukkan nilai recall yang tinggi, meskipun nilai precision dan F1-score relatif lebih rendah. Hal ini mengindikasikan bahwa Random Forest cenderung mengenali lebih banyak data positif, namun dengan tingkat kesalahan prediksi yang lebih tinggi. Sementara itu, SVC menunjukkan ketahanan yang lebih baik terhadap ketidakseimbangan data karena kemampuannya dalam membangun decision boundary yang optimal menggunakan prinsip margin maksimum, sehingga menghasilkan kinerja yang lebih stabil pada seluruh kelas sentimen.

Berdasarkan hasil tersebut, dapat disimpulkan bahwa penggunaan metrik evaluasi selain accuracy, seperti precision, recall, dan F1-score, sangat penting dalam penelitian klasifikasi sentimen, terutama ketika dataset memiliki potensi ketidakseimbangan kelas. Hasil penelitian ini menunjukkan bahwa SVC merupakan algoritma yang paling optimal digunakan untuk klasifikasi sentimen komentar publik terkait pelayanan perpajakan pada penelitian ini.

3.5 Visualisasi Hasil Evaluasi

Untuk memperjelas perbandingan kinerja antar algoritma, hasil evaluasi divisualisasikan dalam bentuk grafik. Perbandingan nilai **accuracy** dari keempat algoritma ditunjukkan pada **Gambar 2**.



Gambar 2. Perbandingan Nilai Accuracy Antar Algoritma Klasifikasi

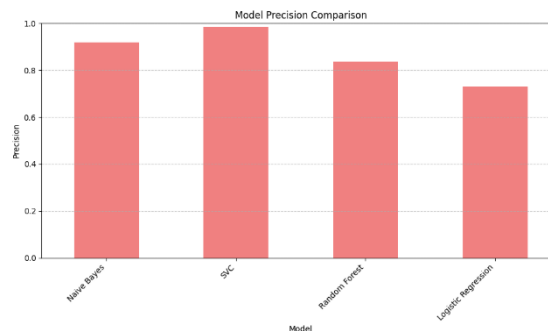
Berdasarkan Gambar 2 yang menunjukkan perbandingan nilai accuracy antar algoritma klasifikasi, terlihat bahwa algoritma Random Forest dan Logistic Regression memperoleh nilai akurasi tertinggi, masing-masing sebesar 0,9800. Algoritma SVC berada sedikit di bawahnya dengan nilai akurasi sebesar 0,9767, sedangkan Naïve Bayes memiliki nilai akurasi terendah sebesar 0,8900. Hasil ini menunjukkan bahwa secara umum Random Forest dan Logistic Regression mampu mengklasifikasikan sebagian besar data dengan benar. Namun demikian, nilai akurasi yang tinggi tidak selalu mencerminkan kemampuan model dalam mengklasifikasikan seluruh kelas sentimen secara seimbang, khususnya pada data yang memiliki distribusi kelas tidak seimbang.

Visualisasi kinerja keseluruhan model yang mencakup metrik accuracy, precision, recall, dan F1-score menunjukkan bahwa algoritma SVC memiliki performa yang paling stabil dan konsisten dibandingkan algoritma lainnya. Meskipun Random Forest dan Logistic Regression mencatat nilai akurasi yang tinggi, kedua algoritma tersebut menunjukkan penurunan yang signifikan pada metrik precision, recall, dan F1-score. Kondisi ini mengindikasikan bahwa kedua model cenderung lebih baik dalam memprediksi kelas mayoritas, namun kurang optimal dalam mengenali kelas minoritas. Sebaliknya, SVC mampu mempertahankan keseimbangan kinerja pada seluruh metrik evaluasi, sehingga memberikan hasil klasifikasi yang lebih representatif.

Analisis lebih lanjut terhadap confusion matrix pada model terbaik, yaitu SVC, menunjukkan bahwa model ini mampu mengklasifikasikan data sentimen positif, negatif, dan netral dengan tingkat kesalahan yang relatif rendah. Nilai true positive pada setiap kelas lebih dominan dibandingkan false positive dan false negative, yang menunjukkan bahwa SVC memiliki kemampuan pemisahan kelas yang baik. Kinerja ini mencerminkan keunggulan SVC dalam membentuk batas keputusan (decision boundary) yang optimal melalui prinsip margin maksimum, sehingga model lebih tahan terhadap pengaruh ketidakseimbangan data.

Secara keseluruhan, hasil visualisasi dan analisis kinerja model menunjukkan bahwa meskipun Random Forest dan Logistic Regression memiliki nilai akurasi yang sangat tinggi, algoritma SVC merupakan model yang paling optimal untuk klasifikasi sentimen komentar publik terkait pelayanan perpajakan. Keunggulan SVC tidak hanya ditunjukkan oleh nilai akurasi yang tinggi, tetapi juga oleh keseimbangan nilai precision, recall, dan F1-score, sehingga menghasilkan performa klasifikasi yang lebih andal dan objektif.

Selain itu, perbandingan nilai precision, recall, dan F1-score masing-masing algoritma ditunjukkan pada Gambar 3, Gambar 4, dan Gambar 5. Visualisasi tersebut menunjukkan bahwa algoritma Random Forest dan Logistic Regression memperoleh nilai tertinggi pada beberapa metrik evaluasi, khususnya precision, recall, dan F1-score. Namun demikian, algoritma SVC menunjukkan kinerja yang paling konsisten dan seimbang pada seluruh metrik evaluasi, sehingga lebih stabil dalam mengklasifikasikan setiap kelas sentimen.



Gambar 3. Perbandingan Nilai Precision Antar Algoritma

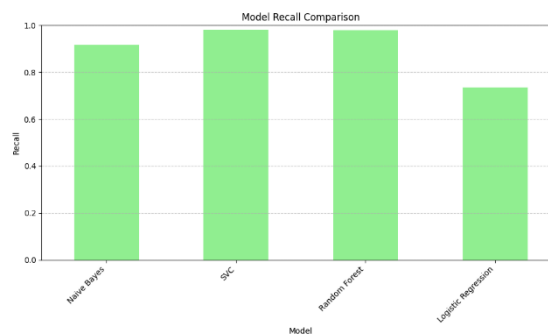
Gambar 3 menunjukkan perbandingan nilai precision (macro) dari empat algoritma klasifikasi sentimen yang digunakan dalam penelitian ini. Berdasarkan hasil evaluasi, algoritma Random Forest dan Logistic Regression memperoleh nilai precision tertinggi masing-masing sebesar **0,980392** dan **0,980376**, yang menunjukkan tingkat ketepatan prediksi yang sangat tinggi pada setiap kelas sentimen. Algoritma SVC juga

menunjukkan kinerja yang sangat baik dengan nilai precision sebesar **0,977124**, sedangkan Naïve Bayes memperoleh nilai precision terendah sebesar **0,896273**.

Nilai precision (macro) yang tinggi pada Random Forest dan Logistic Regression mengindikasikan bahwa kedua algoritma tersebut memiliki tingkat kesalahan prediksi positif yang rendah pada masing-masing kelas sentimen. Namun demikian, hasil ini perlu dianalisis bersama dengan metrik recall dan F1-score, karena precision yang tinggi belum tentu menunjukkan kemampuan model dalam mengenali seluruh data pada setiap kelas, terutama pada kondisi distribusi data yang tidak seimbang.

Sementara itu, algoritma SVC menunjukkan nilai precision yang sedikit lebih rendah dibandingkan Random Forest dan Logistic Regression, tetapi tetap konsisten dengan nilai recall dan F1-score yang tinggi. Hal ini menandakan bahwa SVC memiliki keseimbangan yang lebih baik antara ketepatan dan kelengkapan prediksi pada seluruh kelas sentimen. Adapun nilai precision yang lebih rendah pada Naïve Bayes menunjukkan bahwa algoritma ini masih menghasilkan prediksi keliru pada beberapa kelas sentimen, meskipun kinerjanya relatif stabil pada data teks yang direpresentasikan menggunakan TF-IDF.

Secara keseluruhan, hasil visualisasi pada Gambar 3 memperlihatkan bahwa meskipun Random Forest dan Logistic Regression unggul dalam hal precision, algoritma SVC tetap menjadi model yang paling konsisten ketika mempertimbangkan seluruh metrik evaluasi secara bersamaan.



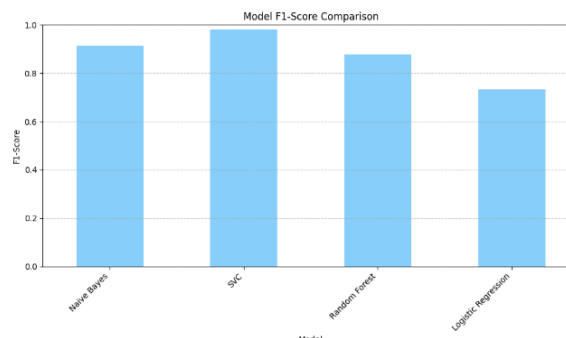
Gambar 4. Perbandingan Nilai Recall Antar Algoritma

Gambar 4 menunjukkan perbandingan nilai recall (macro) dari empat algoritma klasifikasi sentimen yang digunakan dalam penelitian ini. Berdasarkan hasil evaluasi, algoritma Random Forest dan Logistic Regression memperoleh nilai recall tertinggi, masing-masing sebesar **0,980000**, yang menunjukkan kemampuan kedua algoritma tersebut dalam mengenali sebagian besar data pada setiap kelas sentimen. Algoritma SVC mengikuti dengan nilai recall sebesar **0,976667**, sementara Naïve Bayes memiliki nilai recall terendah sebesar **0,890000**.

Nilai recall (macro) yang tinggi pada Random Forest dan Logistic Regression mengindikasikan bahwa kedua algoritma tersebut mampu mengidentifikasi sebagian besar data yang benar-benar termasuk dalam masing-masing kelas sentimen. Namun, apabila dikaitkan dengan nilai precision dan F1-score, kinerja kedua algoritma tersebut menunjukkan adanya ketidakseimbangan, di mana kemampuan mengenali data positif tidak selalu diiringi dengan ketepatan prediksi yang sepadan.

Sementara itu, algoritma SVC menunjukkan nilai recall yang sedikit lebih rendah dibandingkan Random Forest dan Logistic Regression, namun tetap konsisten dengan nilai precision dan F1-score yang tinggi. Hal ini menandakan bahwa SVC memiliki kemampuan yang lebih seimbang dalam mengenali seluruh kelas sentimen tanpa menghasilkan terlalu banyak kesalahan prediksi. Adapun nilai recall yang lebih rendah pada Naïve Bayes menunjukkan bahwa algoritma ini masih belum mampu mengidentifikasi seluruh data pada masing-masing kelas sentimen secara optimal.

Secara keseluruhan, visualisasi pada Gambar 4 menegaskan bahwa meskipun Random Forest dan Logistic Regression unggul dalam hal recall, algoritma SVC tetap menunjukkan performa yang lebih stabil dan seimbang ketika seluruh metrik evaluasi dipertimbangkan secara bersamaan.



Gambar 5. Perbandingan Nilai F1-Score Antar Algoritma

Gambar 5 menampilkan perbandingan nilai F1-score (macro) dari empat algoritma klasifikasi sentimen yang digunakan dalam penelitian ini. Berdasarkan hasil evaluasi, algoritma Random Forest dan Logistic Regression memperoleh nilai F1-score tertinggi masing-masing sebesar **0,979996** dan **0,979979**, yang menunjukkan keseimbangan yang sangat baik antara precision dan recall pada seluruh kelas sentimen. Algoritma SVC mengikuti dengan nilai F1-score sebesar **0,976696**, sedangkan Naïve Bayes memperoleh nilai F1-score terendah sebesar **0,888690**.

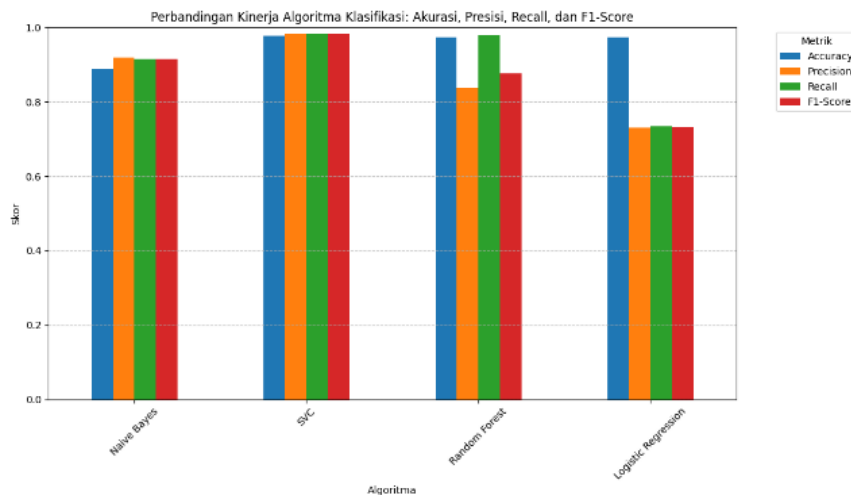
Nilai F1-score (macro) yang tinggi pada Random Forest dan Logistic Regression mengindikasikan bahwa kedua algoritma tersebut memiliki kemampuan yang baik dalam menyeimbangkan ketepatan dan kelengkapan prediksi pada setiap kelas sentimen. Namun demikian, apabila dianalisis bersama dengan metrik accuracy dan distribusi kelas, hasil ini perlu diinterpretasikan secara hati-hati karena nilai F1-score yang tinggi dapat dipengaruhi oleh dominasi kelas mayoritas pada dataset.

Sementara itu, algoritma SVC menunjukkan nilai F1-score yang sedikit lebih rendah dibandingkan Random Forest dan Logistic Regression, tetapi tetap konsisten dengan nilai precision dan recall yang relatif seimbang. Hal ini menunjukkan bahwa SVC memiliki kinerja klasifikasi yang stabil dan tidak terlalu bias terhadap kelas tertentu. Adapun nilai F1-score yang lebih rendah pada Naïve Bayes menunjukkan keterbatasan algoritma tersebut dalam menjaga keseimbangan antara precision dan recall, meskipun algoritma ini masih mampu memberikan performa yang cukup baik pada klasifikasi teks berbasis TF-IDF.

Secara keseluruhan, visualisasi pada Gambar 5 memperkuat hasil evaluasi sebelumnya bahwa meskipun Random Forest dan Logistic Regression unggul pada nilai F1-score, algoritma SVC tetap menunjukkan performa yang paling konsisten dan stabil ketika seluruh metrik evaluasi dipertimbangkan secara bersamaan, sehingga lebih sesuai untuk digunakan dalam klasifikasi sentimen komentar publik terkait pelayanan perpajakan.

3.6 Perbandingan Kinerja Algoritma

Perbandingan kinerja algoritma dilakukan untuk mengetahui performa masing-masing metode klasifikasi sentimen yang digunakan dalam penelitian ini, yaitu **Naïve Bayes**, **Support Vector Classifier (SVC)**, **Random Forest**, dan **Logistic Regression**. Evaluasi kinerja dilakukan menggunakan empat metrik utama, yaitu **accuracy**, **precision**, **recall**, dan **F1-score**, sehingga diperoleh gambaran performa model secara menyeluruh.



Gambar 6. Perbandingan Kinerja Algoritma Klasifikasi Berdasarkan Nilai Accuracy, Precision, Recall, dan F1-Score

Berdasarkan visualisasi pada **Gambar 6**, algoritma **Random Forest** dan **Logistic Regression** menunjukkan kinerja paling tinggi pada seluruh metrik evaluasi. Random Forest memperoleh nilai accuracy **0,9800**, precision **0,980392**, recall **0,9800**, dan F1-score **0,979996**, sedangkan Logistic Regression menghasilkan nilai accuracy **0,9800**, precision **0,980376**, recall **0,9800**, dan F1-score **0,979979**. Nilai yang hampir identik pada keempat metrik tersebut menunjukkan bahwa kedua algoritma mampu melakukan klasifikasi sentimen dengan tingkat akurasi dan konsistensi yang sangat tinggi.

Algoritma **SVC** juga menunjukkan performa yang sangat baik dengan nilai accuracy **0,976667**, precision **0,977124**, recall **0,976667**, dan F1-score **0,976696**. Meskipun nilainya sedikit lebih rendah dibandingkan Random Forest dan Logistic Regression, perbedaan antar metrik pada SVC relatif kecil, yang menandakan bahwa algoritma ini memiliki kinerja yang stabil dan seimbang dalam mengklasifikasikan setiap kelas sentimen.

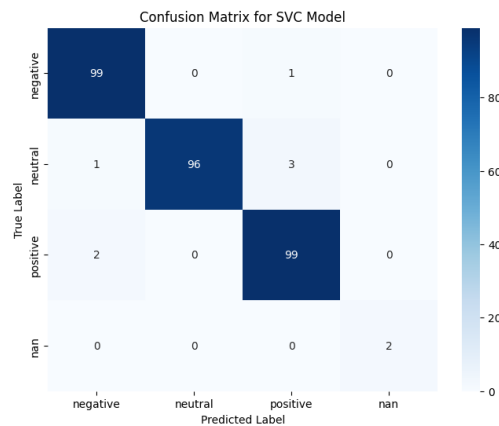
Sementara itu, algoritma **Naïve Bayes** menunjukkan kinerja paling rendah dibandingkan algoritma lainnya, dengan nilai accuracy **0,890000**, precision **0,896273**, recall **0,890000**, dan F1-score **0,888690**. Nilai tersebut mengindikasikan bahwa Naïve Bayes kurang optimal dalam menangani kompleksitas pola data sentimen pada penelitian ini.

Secara keseluruhan, **Gambar 6 menegaskan bahwa Random Forest dan Logistic Regression memiliki performa terbaik secara numerik**, sedangkan **SVC unggul dari sisi kestabilan dan keseimbangan kinerja antar metrik evaluasi**, sehingga ketiga algoritma tersebut layak dipertimbangkan untuk klasifikasi sentimen komentar publik terkait pelayanan perpajakan.

3.7 Metode Evaluasi

Evaluasi kinerja model klasifikasi sentimen pada penelitian ini dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score untuk membandingkan performa antar algoritma secara kuantitatif. Selain itu, analisis confusion matrix digunakan untuk melihat distribusi hasil klasifikasi secara lebih rinci pada setiap kelas sentimen.

Berdasarkan hasil evaluasi keseluruhan, algoritma **Support Vector Classifier (SVC)** dipilih sebagai model terbaik karena menunjukkan kinerja yang paling stabil dan seimbang pada seluruh metrik evaluasi. Untuk memperkuat hasil tersebut, dilakukan analisis confusion matrix terhadap model SVC sebagaimana ditunjukkan pada **Gambar 7**.



Gambar 7. Confusion Matrix Algoritma Terbaik (Support Vector Classifier)

Berdasarkan confusion matrix pada Gambar 7, dapat dilihat bahwa model SVC mampu mengklasifikasikan sebagian besar data sentimen dengan benar pada setiap kelas. Pada kelas sentimen negatif, sebanyak **99 data** berhasil diklasifikasikan secara tepat, dengan hanya **1 data** yang salah diklasifikasikan ke kelas positif. Pada kelas sentimen netral, sebanyak **96 data** diklasifikasikan dengan benar, sementara sebagian kecil data mengalami kesalahan klasifikasi. Selanjutnya, pada kelas sentimen positif, sebanyak **99 data** berhasil diprediksi dengan tepat dengan jumlah kesalahan yang sangat minimal.

Analisis kesalahan klasifikasi menunjukkan bahwa sebagian kecil data yang salah diprediksi umumnya memiliki karakteristik teks yang ambigu. Sebagai contoh, terdapat komentar dengan nada kritik terhadap pelayanan perpajakan yang disampaikan menggunakan bahasa sopan atau konstruktif, sehingga oleh model diklasifikasikan sebagai sentimen netral atau positif, meskipun secara konteks mengandung sentimen negatif. Selain itu, beberapa komentar netral yang mengandung kata-kata bernada emosional ringan juga berpotensi diklasifikasikan sebagai sentimen positif atau negatif.

Meskipun terdapat sejumlah kecil kesalahan klasifikasi, distribusi kesalahan pada confusion matrix menunjukkan bahwa model SVC memiliki tingkat kesalahan yang rendah dan relatif merata pada seluruh kelas sentimen. Hal ini menandakan bahwa model tidak bias terhadap satu kelas tertentu dan mampu membentuk batas keputusan yang baik antar kelas.

Dengan demikian, hasil evaluasi menggunakan metrik kuantitatif dan confusion matrix secara konsisten menunjukkan bahwa **Support Vector Classifier (SVC)** merupakan algoritma yang paling optimal dan andal untuk klasifikasi sentimen komentar publik terhadap pelayanan perpajakan pada penelitian ini.

Berdasarkan hasil pengujian dan evaluasi yang telah dilakukan, dapat disimpulkan bahwa setiap algoritma machine learning menghasilkan performa yang berbeda dalam mengklasifikasikan sentimen komentar publik terhadap pelayanan perpajakan. Algoritma **Support Vector Classifier (SVC)** menunjukkan kinerja paling optimal dan konsisten pada seluruh metrik evaluasi, sehingga dinilai paling sesuai untuk digunakan dalam penelitian ini.

4. KESIMPULAN

Berdasarkan hasil evaluasi klasifikasi sentimen komentar publik terkait pelayanan perpajakan, algoritma Random Forest dan Logistic Regression menunjukkan nilai kinerja tertinggi secara numerik pada metrik accuracy, precision, recall, dan F1-score, yang menandakan kemampuan kedua algoritma dalam mengklasifikasikan sebagian besar data secara tepat. Namun demikian, algoritma Support Vector Classifier (SVC) menunjukkan kinerja yang paling stabil dan seimbang pada seluruh metrik evaluasi serta memiliki tingkat kesalahan klasifikasi



yang relatif rendah berdasarkan analisis confusion matrix, sehingga dinilai sebagai model yang paling optimal dan andal untuk digunakan dalam penelitian ini. Kontribusi penelitian ini terletak pada penyajian analisis komparatif beberapa algoritma machine learning pada domain sentimen pelayanan perpajakan di Indonesia serta penekanan pentingnya penggunaan metrik evaluasi yang komprehensif untuk menghindari bias pada dataset yang berpotensi tidak seimbang. Meskipun demikian, penelitian ini masih memiliki keterbatasan, yaitu penggunaan dataset yang terbatas pada sumber media sosial tertentu, penggunaan metode representasi fitur berbasis TF-IDF yang belum memanfaatkan pendekatan deep learning modern, serta jumlah data yang masih dapat diperluas untuk meningkatkan kemampuan generalisasi model, sehingga penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan beragam serta mengombinasikan metode machine learning dengan pendekatan deep learning untuk meningkatkan performa klasifikasi sentimen.

REFERENCES

- [1] E. C. M. Torres and L. G. de Picado-Santos, "Sentiment Analysis And Topic Modeling In Transportation: A Literature Review," *Applied Sciences*, vol. 15, no. 12, 2025, doi: 10.3390/app15126576.
- [2] L. Ouyang, "Interpreting Public Sentiment In Diplomacy Events: A Counterfactual Analysis Framework Using Large Language Models," *arXiv Preprint*, 2025. [Online]. Available: <http://arxiv.org/abs/2509.20367>
- [3] M. Al, F. Syabil, and F. Al-Fahri, "Sentiment Analysis Of Coretax Tax Application Users Using IndoBERT And Web Scraping On The X (Twitter) Platform," *Jurnal Sains Student Research*, vol. 4, no. 1, pp. 29–40, 2026, doi: 10.61722/jssr.v4i1.7258.
- [4] E. A. Rahman and A. H. Lubis, "Analysis Of Public Sentiment Toward The Increase In VAT Rates Using The SVM Algorithm," *Journal Of Dinda Data Science, Information Technology, And Data Analytics*, vol. 5, no. 2, pp. 241–249, 2025. [Online]. Available: <https://t.co/EFZzt8zww>
- [5] B. Setiawan, "A Review Of Sentiment Analysis Applications In Indonesia Between 2023–2024," 2024.
- [6] W. A. Anggraeni, F. F. Roji, and M. Alkautsar, "Analisis Sentimen Publik Terhadap Kebijakan Insentif Perpajakan Dengan Pendekatan VADER (Valence Aware Dictionary And Sentiment Reasoner)," *Jurnal Proaksi*, vol. 10, no. 4, pp. 465–477, 2023, doi: 10.32534/jpk.v10i4.4732.
- [7] D. Kurniawan and W. Ganta, "Perbandingan Kinerja Algoritma Machine Learning Untuk Klasifikasi Sentimen Komentar Publik Terkait Pelayanan Perpajakan," vol. 6, 2022.
- [8] G. Rao, M. Mane, V. A. Medhe, V. V. Mane, S. C. Bandpatte, and P. B. Jarande, "Analyzing Public Reactions To Budgetary Tax Reforms: A BERT-Based Sentiment Study," 2024. [Online]. Available: <https://www.theaspd.com/ijes.php>
- [9] M. R. Manoppo et al., "Analisis Sentimen Publik Di Media Sosial Terhadap Kenaikan PPN 12% Di Indonesia Menggunakan IndoBERT," *Jurnal Kecerdasan Buatan Dan Teknologi Informasi*, vol. 4, no. 2, pp. 152–163, 2025, doi: 10.69916/jkbt.v4i2.322.
- [10] D. D. Tarigan and A. Idrus, "Sentiment Analysis Of Twitter Users Regarding Taxation Topics In Indonesia Utilizing Multinomial Naïve Bayes," *Journal Of Informatics And Data Science (J-IDS)*, vol. 3, no. 1, 2024.
- [11] M. A. Chourasiya and R. Saini, "Analyzing Taxpayers' Perception Of Income Tax Policies Using Sentiment Analysis," *EPRA International Journal Of Economics, Business And Management Studies*, 2023, doi: 10.36713/epra1013.
- [12] I. A. Maulana and A. P. W. Wibowo, "Analysis Of Public Sentiment Text Clustering On Tax Increases Using Orange Data Mining On Twitter," *Brilliance: Research Of Artificial Intelligence*, vol. 5, no. 1, pp. 93–99, 2025, doi: 10.47709/brilliance.v5i1.5787.
- [13] R. Sahirin, L. Lucas, and C. Qingjun, "Sentiment Analysis On Social Media Using Data Mining For Mapping Community Satisfaction," *Journal Of Computer Science Advancements*, vol. 3, no. 1, 2025, doi: 10.70177/jasca.v3i1.1536.
- [14] N. Fauziah, "Analisis Sentimen Publik Terhadap Kenaikan Tarif PPN Di Indonesia Dengan Pendekatan VADER," *Jurnal Akuntansi Dan Keuangan*, vol. 12, no. 2, 2024, doi: 10.29103/jak.v12i2.16796.
- [15] L. Afuan, M. Khanza, and A. Z. Hasyati, "Enhancing Sentiment Analysis Of The 2024 Indonesian Presidential Inauguration On X Using SMOTE-Optimized Naïve Bayes Classifier," *Jurnal Teknik Informatika (JUTIF)*, vol. 6, no. 1, pp. 325–333, 2025, doi: 10.52436/1.jutif.2025.6.1.4290.
- [16] Istiqomah and A. H. Lubis, "Sentiment Analysis On TikTok Discourse Surrounding The 2024 North Sumatra Gubernatorial Election Using Support Vector Machine," *CN-AHPC*, vol. 7, no. 3, 2024, doi: 10.47709/cnahpc.v7i3.6543.
- [17] S. Adi, S. Mola, Y. T. Polly, and Y. C. Atok, "Analisis Sentimen Terhadap Data Komentar Publik Mengenai Isu UU Pilkada 2024 Menggunakan Metode Naïve Bayes Dan K-Nearest Neighbor," *Journal Of Informatics Management And Information Technology*, vol. 5, no. 3, pp. 337–343, 2025, doi: 10.47065/jimat.v5i3.514.
- [18] Istiqomah and A. H. Lubis, "Sentiment Analysis On TikTok Discourse Surrounding The 2024 North Sumatra Gubernatorial Election Using Support Vector Machine," *CN-AHPC*, vol. 7, no. 3, 2024, doi: 10.47709/cnahpc.v7i3.6543.
- [19] A. Halim, F. Zidan, I. Handayani, and A. Anggara, "Sentiment Analysis Of The 2024 Election Using The Naïve Bayes Method Using X Data," *Jurnal Mandiri*, vol. 14, no. 2, pp. 225–234, 2025. [Online]. Available: www.ejournal.isha.or.id/index.php/Mandiri
- [20] T. Rachmawati and H. K. Nugraha, "Understanding Public Perception Made Easy: A Sentiment Analysis Of Public Transportation Services," *Jurnal Ilmu Sosial Dan Ilmu Politik*, vol. 29, no. 1, pp. 51–70, 2025, doi: 10.22146/jsp.100206.
- [21] Istiqomah and A. H. Lubis, "Sentiment Analysis On TikTok Discourse Surrounding The 2024 North Sumatra Gubernatorial Election Using Support Vector Machine," *CN-AHPC*, vol. 7, no. 3, 2024, doi: 10.47709/cnahpc.v7i3.6543.
- [22] T. A. Munandar, "A Supervised Learning Model For Sentiment Analysis Based On Regional Dialects In Tourism-Related Issues," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 10, no. 3, pp. 760–770, 2025, doi: 10.30591/jpit.v10i3.8627.
- [23] M. A. Hasan and N. P. Bimby, "Analisis Sentimen Publik Terhadap Kenaikan Pajak PPN Di Indonesia Tahun 2024 Menggunakan Algoritma Machine Learning," 2024.



- [24] R. F. Nulhakim and F. Shiddieq, “Analisis Sentimen Publik X Terhadap Kenaikan PPN 12% Pada Twitter Menggunakan Latent Dirichlet Allocation,” 2024.
- [25] N. Fauziah, “Analisis Sentimen Publik Terhadap Kenaikan Tarif PPN Di Indonesia Dengan Pendekatan VADER,” *Jurnal Akuntansi Dan Keuangan*, vol. 12, no. 2, 2024, doi: 10.29103/jak.v12i2.16796.
- [26] M. A. Hasan and N. P. Bimby, “Analisis Sentimen Publik Terhadap Kenaikan Pajak PPN Di Indonesia Tahun 2024 Menggunakan Algoritma Machine Learning,” 2024.
- [27] M. A. Hasan and N. P. Bimby, “Analisis Sentimen Publik Terhadap Kenaikan Pajak PPN Di Indonesia Tahun 2024 Menggunakan Algoritma Machine Learning,” 2024.