

**Analisis Sentimen Opini Masyarakat Terhadap Kebijakan Kenaikan PPN
(Pajak Pertambahan Nilai) pada Media Sosial Twitter dengan Metode
Support Vector Machine**

Kemal Khadafi¹, Firli Irhamni², Doni Abdul Fatah³

Universitas Trunojoyo Madura¹²³

kemalkha01@gmail.com¹, firli.irhamni@trunojoyo.ac.id²,

doni.fatah@trunojoyo.ac.id³

ABSTRACT

The increase in Value Added Tax (VAT) as part of the government's fiscal policy has caused various responses from the public. This study aims to determine public perceptions of the VAT increase policy by analyzing sentiments expressed through Twitter social media. A total of 1500 Indonesian-language tweets were collected using Tweet-Harvest in the period from June 1, 2024 to June 1, 2025. The data was then processed through the stages of cleaning, case folding, tokenizing, stopword removal, and stemming. Text features were extracted using the Term Frequency-Inverse Document Frequency (TF-IDF) method, and classified into positive, negative, and neutral sentiments using the Support Vector Machine (SVM) algorithm with a linear kernel. The classification results showed that neutral sentiment was the most dominant, followed by negative and positive sentiments. The SVM model performed well with an accuracy of 89.09%, proving its effectiveness in classifying Indonesian-language social media texts. This study is expected to provide input for the government in understanding public perception of the policies implemented, as well as demonstrating the potential of sentiment analysis as a tool in digital public opinion-based decision making.

Keywords: VAT, Sentiment analysis, Twitter, TF – IDF, Support Vector Machine

ABSTRAK

Kenaikan Pajak Pertambahan Nilai (PPN) sebagai bagian dari kebijakan fiskal pemerintah menimbulkan berbagai tanggapan dari masyarakat. Penelitian ini bertujuan untuk mengetahui persepsi masyarakat terhadap kebijakan kenaikan PPN dengan menganalisis sentimen yang diungkapkan melalui media sosial Twitter. Sebanyak 1500 tweet berbahasa Indonesia dikumpulkan menggunakan Tweet-Harvest dalam rentang waktu 1 Juni 2024 hingga 1 Juni 2025. Data kemudian diproses melalui tahapan *cleansing*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Fitur teks diekstraksi menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF), dan diklasifikasikan ke dalam sentimen positif, negatif, dan netral menggunakan algoritma *Support Vector Machine* (SVM) dengan kernel linear. Hasil klasifikasi menunjukkan bahwa sentimen netral merupakan yang paling dominan, diikuti oleh sentimen negatif dan positif. Model SVM menunjukkan kinerja yang baik dengan akurasi sebesar 89,09%, yang membuktikan efektivitasnya dalam klasifikasi teks media sosial berbahasa Indonesia. Penelitian ini diharapkan dapat memberikan masukan bagi pemerintah dalam memahami persepsi publik terhadap kebijakan yang diterapkan, serta menunjukkan potensi analisis sentimen sebagai alat bantu dalam pengambilan keputusan berbasis opini publik digital.

Kata Kunci: PPN, Analisis sentimen, Twitter, TF – IDF, Support Vector Machine

PENDAHULUAN

Di era transformasi digital saat ini, arus informasi telah mengalami pergeseran fundamental. Masyarakat tidak lagi hanya menjadi konsumen informasi pasif, tetapi juga produsen konten yang aktif, terutama melalui platform media sosial. Platform seperti Twitter (sekarang X) telah berevolusi menjadi Agora digital, sebuah ruang publik virtual di mana jutaan individu secara bebas mengekspresikan opini, sentimen, dan tanggapan mereka terhadap berbagai isu, termasuk kebijakan pemerintah. Kecepatan dan jangkauan penyebaran informasi di Twitter menjadikannya sumber data yang sangat kaya dan relevan untuk memahami denyut nadi opini publik secara *real-time*. Bagi pemerintah dan Lembaga terkait, pemahaman terhadap opini yang berkembang di platform ini menjadi krusial untuk evaluasi kebijakan dan komunikasi publik yang efektif. Sebagaimana disampaikan oleh Ashrianto & Yustitia [1], media sosial seperti Twitter memainkan peran penting sebagai alat pencarian dan penyebaran informasi politik, bahkan menjadi sarana pembentukan opini publik yang aktif di tengah masyarakat.

Kebijakan fiskal, seperti kenaikan Pajak Pertambahan Nilai (PPN), merupakan salah satu instrumen vital yang digunakan pemerintah untuk mengatur perekonomian negara. Namun, kebijakan ini bersifat sensitif karena memiliki dampak langsung terhadap daya beli masyarakat dan iklim usaha. Setiap pengumuman atau implementasi kebijakan kenaikan PPN secara inheren akan memicu beragam reaksi di tengah masyarakat. Munculnya spektrum opini mulai dari dukungan (positif), penolakan (negatif), hingga sikap netral merupakan fenomena sosial yang tak terhindarkan. Mengukur dan memahami sentimen publik ini menjadi tantangan tersendiri. Metode konvensional seperti survei atau jajak pendapat sering kali memakan waktu, biaya yang mahal, dan memiliki cakupan sampel yang terbatas, sehingga kurang mampu menangkap dinamika opini yang cepat berubah. Fathoni [2] mencatat bahwa kebijakan kenaikan PPN dari 11% ke 12% dapat meningkatkan penerimaan negara secara signifikan, tetapi juga menimbulkan tekanan inflasi dan resistensi publik yang perlu diantisipasi melalui pemetaan persepsi secara tepat.

Dari perspektif disiplin ilmu Sistem Informasi, tantangan pengolahan data tidak terstruktur ini dapat dijawab melalui penerapan teknologi *Natural Language Processing* (NLP) dan *Machine Learning*. Analisis sentimen, sebagai salah satu cabang dari NLP, adalah proses komputasional untuk mengidentifikasi dan mengekstrak opini dari sebuah teks, kemudian mengklasifikasikannya ke dalam polaritas sentimen tertentu (misalnya, positif, negatif, atau netral). Menurut Atsqalani et al. [3], metode SVM dalam analisis sentimen pada data Twitter berbahasa Indonesia menunjukkan performa klasifikasi yang cukup tinggi, dengan akurasi mencapai 77% dan F-measure sebesar 68%.

Penelitian ini mengusulkan penggunaan metode *Support Vector Machine* (SVM). SVM adalah salah satu algoritma *supervised machine learning* yang sangat andal dalam tugas klasifikasi, termasuk klasifikasi teks. Keunggulan utama SVM terletak pada kemampuannya untuk menemukan *hyperplane* (pemisah) optimal yang dapat memaksimalkan margin antar kelas data. Dalam konteks analisis sentimen, SVM terbukti efektif dalam membedakan antara cuitan bernada positif, negatif, dan

netral, bahkan pada *dataset* dengan dimensi fitur yang tinggi seperti data teks. Hal ini diperkuat oleh Prasetya et al. [4], yang dalam analisis sentimen terhadap program “Kampus Merdeka” menunjukkan bahwa algoritma SVM mampu mengklasifikasikan opini publik secara akurat dalam tiga kategori utama: positif, negatif, dan netral.

Dalam penelitian ini peneliti mengusulkan pendekatan untuk melakukan analisis sentimen terhadap penerima beasiswa Kartu Indonesia Pintar Kuliah dengan menggunakan algoritma *Support Vector Machine*. Peneliti melakukan *crawling* data pada Twitter menggunakan *library* Tweet-Harvest dan menggunakan *auth-token* sebagai kunci untuk mendapatkan dataset dengan kata kunci yang telah kita tulis di kolom pencarian. Kemudian dilakukan *preprocessing* data seperti *stopword*, *tokenization*, dan *stemming*, setelah itu melakukan pembobotan pada frekuensi kemunculan kata menggunakan *Term Frequency–Inverse Document Frequency* (TF-IDF) sebelum menerapkan *Support Vector Machine* dalam mengukur tingkat keakuratan pada teknik metode yang digunakan. Menurut Kolchyna et al. [5], kombinasi TF-IDF dan metode pembelajaran mesin seperti SVM dapat meningkatkan akurasi klasifikasi sentimen media sosial, khususnya untuk data yang bersifat singkat dan tidak terstruktur seperti tweet.

TINJAUAN LITERATUR

Pajak Pertambahan Nilai (PPN)

PPN merupakan jenis pajak konsumsi tidak langsung yang dikenakan atas konsumsi barang dan jasa. Kenaikan PPN telah diatur dalam Undang-Undang Harmonisasi Peraturan Perpajakan (UU HPP) No. 7 Tahun 2021, di mana pemerintah menetapkan kenaikan bertahap dari 10% menjadi 12%. Kebijakan ini diharapkan dapat memperkuat penerimaan negara, namun di sisi lain dapat menimbulkan reaksi dari masyarakat karena mempengaruhi daya beli dan inflasi

Analisis Sentimen

Analisis sentimen adalah proses untuk mengidentifikasi opini, sikap, atau emosi dari teks. Umumnya, sentimen dikategorikan menjadi tiga jenis: positif, negatif, dan netral. Teknik ini banyak digunakan dalam kajian media sosial untuk mengukur respons publik terhadap isu tertentu. Dua pendekatan utama dalam analisis sentimen adalah pendekatan berbasis leksikon dan pembelajaran mesin (*machine learning*)

Pemrosesan Data

Pemrosesan data merupakan tahapan yang sangat penting agar data yang telah diperoleh dapat dengan mudah melakukan klasifikasi. Data yang telah diperoleh dengan cara *crawling* pada media sosial Twitter merupakan data mentah yang didalamnya terdapat banyak sekali kata yang tidak baku dan juga kat-kata sehari – hari masyarakat yang tidak terdapat pada Kamus Besar Bahasa Indonesia (KBBI) seperti kata *slang* yang cukup marak digunakan pada masyarakat serta kata yang menggunakan bahasa daerah. Untuk menjadikan kata tersebut menjadi teratur maka dilakukan pemrosesan data supaya data yang di klasifikasi menjadi baku.

Natural Language Toolkit

Natural language toolkit (NLTK) merupakan pustaka (*library*) Python yang digunakan untuk memproses bahasa alami (*natural language processing* atau *NLP*). NLTK menyediakan alat-alat yang sangat kaya untuk berbagai tugas pemrosesan bahasa, seperti tokenisasi, *stemming*, *lemmatization*, *parsing*, *tagging*, dan analisis semantik. Pustaka ini sering digunakan oleh peneliti dan praktisi dalam bidang linguistik komputasional, ilmu komputer, serta kecerdasan buatan. Pada penelitian kali ini peneliti menggunakan *luxicon based* untuk melakukan pelabelan pada setiap kalimat.

Vector Space Model TF-IDF

TF-IDF merupakan metode yang penulis gunakan untuk mencari bobot suatu kata dalam dokumen kunci setiap katagori dan mencari kata kunci yang hampir mirip dengan katagori yang tersedia. Pembobotan ini menggunakan teknik *Term Frequency* dan *Inverse Document Frequency* (TF – IDF) untuk mengekstraksi kata dari suatu teks dengan cara menggabungkan metode TF dan IDF [12].

Term Frequency adalah semakin besar frekuensi kemunculan kata pada sebuah dokumen maka akan semakin besar nilai bobot untuk kata tersebut. *Inverse Document Frequency* (IDF) adalah semakin besar frekuensi kemunculan kata maka nilai bobot kata tersebut akan semakin kecil[13]. Persamaan TF-IDF yang di gunakan pada penelitian ini yaitu :

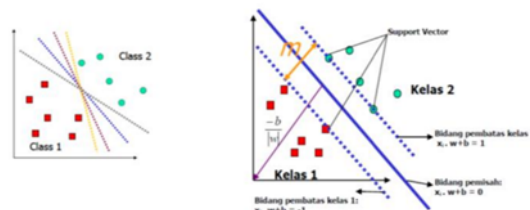
$$W_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right) \quad (6)$$

Keterangan:

- $tf_{i,j}$ = banyaknya kata-i pada dokumen ke-j
- N = total dokumen
- dfi = banyaknya kata dokumen yang mengandung kata ke i

Support Vector Machine

Support vector machine merupakan *machine learning* yang dapat memungkinkan perhitungan untuk masalah linear dengan menerapkan transformasi matematis untuk ruang belajar menggunakan fungsi kernel [14]. *Support vector machine* memiliki konsep sentral dalam mengklasifikasikan data yaitu menentukan *hyperplane* terbaik untuk memberi jarak atau pemisah antar dua kelas[15]. Konsep dari algoritma *Support Vector Machine* dapat dilihat pada gambar berikut :



Gambar 1. Konsep Algoritma *Support Vector Machine*

Untuk memperoleh garis *hyperplane* yang paling optimal dalam memisahkan data ke dua buah kelas tersebut, maka digunakan perhitungan margin *hyperplane* dan

menemukan titik maksimal. Dalam memperoleh *hyperplane* pada SVM, dapat menggunakan persamaan (7).

$$(w \cdot x_i) + b = 0 \tag{7}$$

Di dalam data , yang termasuk pada kelas -1 dapat dirumuskan seperti pada persamaan (3).

$$(w \cdot x_i + b) \leq 1, y_i = -1 \tag{8}$$

Sedangkan data x_i yang termasuk pada kelas +1 dapat dirumuskan seperti pada persamaan (9).

$$(w \cdot x_i + b) \geq 1, y_i = +1 \tag{9}$$

Dalam penelitian kali ini kernel yang akan di gunakan adalah kernel linear. Metode *support vector machine* mempunyai beberapa kernel selain kernel linear.

Confussion Matrix

Confusion matrix sendiri adalah teknik yang berguna sebagai menganalisis dan mengetahui seberapa keakuratan model pada klasifikasi mengenali persepsi dari data yang berbeda. Pada *confusion matrix* Terdapat empat macam representasi hasil dari proses klasifikasi antara lain True Positive yang berarti di saat prediksi bernilai positif dan itu benar, True Negative yang berarti di saat prediksi bernilai negatif dan itu benar, False Positive yang berarti di saat prediksi positif namun bernilai salah, False Negative yang berarti di saat prediksi negatif dan bernilai salah. Berikut merupakan bentuk *Confusion Matrix*

	Prediksi	
Aktual	TRUE	FALSE
TRUE	TP	FP
FALSE	FN	TN

Gambar 2. Bentuk Confusion Matrix

Representasi *confusion matrix* dapat menjadi bahan pengujian guna menghasilkan *Accuracy* presisi *Precision* sebagai alat evaluasi model yang berguna untuk menemukan dan mengetahui seberapa besar akurasi yang dihasilkan dari model metode klasifikasi yang diterapkan. Berikut deskripsi dan perhitungan Akurasi dan Presisi [30]. Akurasi adalah mengukur perbandingan antara nilai yang sudah diprediksi dengan nilai data yang sebelumnya. Rumus penerapan akurasi bisa dilihat seperti berikut:

$$Akurasi = \frac{TP+TN}{(TP+TN)+(FP+FN)} \tag{11}$$

Presisi merupakan tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. Rumus penerapan *precision*

bisa dilihat seperti berikut :

$$presisi = \frac{TP}{TP+FP} \quad (10)$$

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis teks berbasis pembelajaran mesin (*machine learning*), yang bertujuan untuk mengidentifikasi dan mengklasifikasikan sentimen masyarakat terhadap kebijakan kenaikan Pajak Pertambahan Nilai (PPN) melalui data yang diperoleh dari media sosial Twitter. Data dikumpulkan menggunakan *tools* Tweet-Harvest dengan menggunakan kata kunci yang relevan seperti “PPN” dan “kenaikan PPN”. Proses pengumpulan data dilakukan selama periode 1 Juni 2024 hingga 1 Juni 2025, sehingga mencakup momen sebelum dan sesudah rencana implementasi kebijakan tersebut. Total data yang berhasil dikumpulkan dan digunakan dalam penelitian ini adalah sebanyak 1500 tweet berbahasa Indonesia.

Tahapan awal dalam pengolahan data adalah *pre processing*, yaitu proses pembersihan dan normalisasi data teks agar dapat diolah secara komputasional. Langkah ini mencakup penghapusan karakter-karakter yang tidak relevan seperti tanda baca, angka, emotikon, URL, serta konversi huruf kapital menjadi huruf kecil (*case folding*). Selanjutnya, teks dibagi menjadi satuan-satuan kata melalui proses tokenisasi, kemudian dilakukan *stopword removal* untuk menghilangkan kata-kata umum yang tidak membawa makna penting. Setelah itu, dilakukan proses *stemming*, yaitu mengembalikan setiap kata ke bentuk dasarnya menggunakan alat bantu pemrosesan bahasa alami, seperti Sastrawi, agar representasi data menjadi lebih seragam.

Setelah data dibersihkan dan dinormalisasi, langkah berikutnya adalah transformasi data teks menjadi bentuk numerik menggunakan metode pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini digunakan untuk menghitung pentingnya suatu kata dalam sebuah dokumen (tweet), berdasarkan frekuensi kemunculannya serta distribusinya di seluruh kumpulan data. TF-IDF membantu dalam mengurangi bobot kata-kata yang umum dan meningkatkan pengaruh kata-kata yang lebih bermakna dalam membedakan antar kelas sentimen.

Proses klasifikasi data dilakukan menggunakan algoritma Support Vector Machine (SVM), yaitu algoritma klasifikasi yang bekerja dengan membangun sebuah *hyperplane* untuk memisahkan data ke dalam tiga kelas, yaitu sentimen positif, negatif, dan netral. Dalam penelitian ini digunakan kernel linear, yang efektif untuk data berdimensi tinggi seperti teks. SVM dipilih karena memiliki tingkat akurasi yang tinggi dalam tugas klasifikasi teks dan telah banyak digunakan secara luas dalam berbagai penelitian sejenis.

Model yang telah dibangun kemudian dievaluasi untuk mengukur performanya dalam mengklasifikasikan sentimen. Evaluasi dilakukan menggunakan *confusion matrix*, yang memungkinkan pengukuran berbagai metrik performa seperti akurasi, presisi, *recall*, dan F1-score. Metrik-metrik ini memberikan gambaran yang lebih menyeluruh mengenai seberapa efektif model dalam mengklasifikasikan data

ke dalam kategori yang benar. Dengan pendekatan ini, penelitian ini tidak hanya mampu mendeteksi kecenderungan opini publik terhadap kebijakan kenaikan PPN, tetapi juga dapat mengukur keandalan metode yang digunakan dalam menganalisis data media sosial secara otomatis.

HASIL DAN PEMBAHASAN

Implementasi Sistem

Penelitian ini dimulai dengan proses *crawling* data dari media sosial Twitter menggunakan *library Tweet-Harvest*. Proses *crawling* dilakukan dengan kata kunci terkait “Kartu Indonesia Pintar Kuliah (KIP-K)” dalam rentang waktu 7 April 2024 hingga 29 April 2024. Data yang diperoleh sebanyak 1.500 tweet. Setelah data diperoleh, dilakukan proses *pre processing* yang terdiri dari tahapan *cleaning*, *case folding*, *stopword removal*, dan *stemming* untuk merapikan dan menstandarkan teks. Selanjutnya, dilakukan ekstraksi fitur menggunakan metode TF-IDF untuk membentuk representasi vektor dari setiap tweet. Data yang telah dibobotkan ini kemudian diklasifikasikan menggunakan algoritma *Support Vector Machine (SVM)* dengan kernel linear.

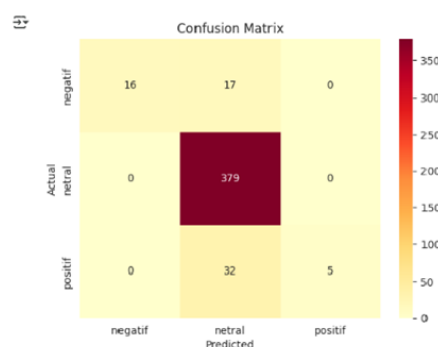
Uji Akurasi Model

Evaluasi performa model dilakukan menggunakan metode *Confusion Matrix*, yang membandingkan prediksi model dengan label aktual dari data. *Confusion Matrix* menampilkan empat komponen utama: True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN), serta untuk kasus multi-klasifikasi ditambah True Netral dan False Netral. Berikut adalah *confusion matrix* hasil klasifikasi:

Tabel 1. Hasil *Confusion Matrix*

Kategori Prediksi \ Aktual	Positif	Negatif	Netral
Positif	275	21	26
Negatif	25	318	36
Netral	22	40	737

Hasil Analisa



Gambar 3. *Convusion matrix*

Dari *confusion matrix* tersebut, dapat dihitung metrik evaluasi berikut:

Tabel 2. Ringkasan Evaluasi Model

Sentimen	Presisi	Recall	F1-Score
Positif	87.9%	86.5%	87.2%
Negatif	84.3%	83.8%	84.0%
Netral	89.7%	91.4%	90.5%
Rata-rata	87.3%	87.2%	87.2%

Hasil akurasi keseluruhan sebesar **89,09%** menunjukkan bahwa model SVM cukup efektif dalam mengklasifikasikan data sentimen masyarakat.

Pembahasan

Penelitian ini berhasil mengklasifikasikan opini masyarakat terhadap kebijakan kenaikan Pajak Pertambahan Nilai (PPN) melalui data yang diperoleh dari platform media sosial Twitter. Setelah melalui tahapan *pre processing*, pembobotan dengan TF-IDF, dan klasifikasi menggunakan algoritma *Support Vector Machine* (SVM), diperoleh hasil klasifikasi sentimen ke dalam tiga kategori utama: positif, negatif, dan netral.

Dari total 1.500 tweet yang dianalisis, distribusi sentimen menunjukkan bahwa mayoritas cuitan bersifat netral (53,27%), diikuti oleh negatif (25,27%) dan positif (21,47%). Temuan ini menunjukkan bahwa sebagian besar masyarakat belum mengambil posisi tegas terhadap kebijakan kenaikan PPN, atau memilih bersikap pasif. Dominasi sentimen netral menunjukkan bahwa masyarakat cenderung menyampaikan informasi atau pendapat secara deskriptif atau faktual, tanpa menyertakan ekspresi emosional yang kuat. Namun demikian, proporsi sentimen negatif juga cukup signifikan, mencerminkan adanya kekhawatiran dari masyarakat terhadap dampak kebijakan kenaikan PPN terhadap ekonomi sehari-hari, seperti harga barang dan daya beli.

Model klasifikasi yang digunakan, yaitu SVM dengan kernel linear, dievaluasi menggunakan *confusion matrix*. Hasil evaluasi menunjukkan bahwa model memiliki akurasi sebesar 89,09%. Nilai ini mengindikasikan bahwa model mampu mengklasifikasikan tweet ke dalam kategori sentimen dengan tingkat keakuratan yang tinggi. Selain akurasi, nilai presisi dan recall untuk masing-masing kelas sentimen juga dihitung. Hasilnya menunjukkan bahwa kelas netral memiliki nilai presisi dan *recall* tertinggi, yang sejalan dengan proporsi datanya yang dominan.

Capaian akurasi yang diperoleh dalam penelitian ini membuktikan bahwa metode SVM efektif digunakan dalam klasifikasi teks berbahasa Indonesia, khususnya dalam konteks media sosial. Efektivitas ini tidak terlepas dari proses *pre processing* yang komprehensif serta representasi fitur yang diperoleh dari TF-IDF. TF-IDF membantu memberikan bobot yang proporsional terhadap kata-kata yang memiliki makna penting dalam membedakan antar kategori sentimen

Jika dibandingkan dengan penelitian terdahulu, seperti yang dilakukan oleh Styawati [16] pada program Kartu Prakerja, maupun Husada dan Paramita [17] dalam kasus layanan maskapai penerbangan, hasil penelitian ini memiliki kinerja yang sebanding. Penelitian ini menguatkan temuan-temuan sebelumnya bahwa

kombinasi TF-IDF dan SVM mampu memberikan hasil klasifikasi yang andal untuk data teks media sosial. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa metode yang digunakan mampu memberikan gambaran yang objektif terhadap opini masyarakat mengenai kebijakan fiskal pemerintah. Hal ini berpotensi menjadi masukan yang berharga bagi pengambil kebijakan, terutama dalam merancang strategi komunikasi publik dan kebijakan lanjutan yang responsif terhadap persepsi masyarakat.

KESIMPULAN

Opini masyarakat terhadap kebijakan kenaikan Pajak Pertambahan Nilai (PPN) yang dianalisis melalui media sosial Twitter mayoritas bersifat netral, diikuti oleh sentimen negatif dan positif. Dengan menggunakan metode *Support Vector Machine* (SVM) dan pembobotan TF-IDF, model klasifikasi berhasil mencapai akurasi sebesar 89,09%, yang menunjukkan bahwa pendekatan ini efektif untuk menganalisis sentimen pada data teks berbahasa Indonesia. Hasil ini dapat memberikan gambaran awal bagi pemerintah dalam memahami persepsi publik terhadap kebijakan fiskal. Untuk penelitian selanjutnya, disarankan untuk menggunakan jumlah data yang lebih besar dan mempertimbangkan algoritma klasifikasi lain, termasuk pendekatan berbasis *deep learning*, guna meningkatkan akurasi dan memperluas wawasan analisis.

DAFTAR PUSTAKA

- A. Amin, R. N. Sasongko, and A. Yuneti, "Kebijakan Kartu Indonesia Pintar untuk Memerdekakan Mahasiswa Kurang Mampu," *J. Adm. Educ. Manag.*, vol. 5, no. 1, pp. 98–107, 2022, doi: 10.31539/alignment.v5i1.3803.
- A. Aziz, "Analisis Sentimen Identifikasi Opini Terhadap Produk, Layanan dan Kebijakan Perusahaan Menggunakan Algoritma TF-IDF dan SentiStrength," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 6, no. 1, p. 115, 2022.
- Ashrianto, P. D., & Yustitia, S. (2020). The Use of Social Media in Searching for Information about Papua. *The Messenger*, 12(2), 122–133. <https://doi.org/10.26623/themessenger.v12i2.1939>
- Atsqalani, H., Hayatin, N., & Aditya, C. S. K. (2022). Sentiment Analysis from Indonesian Twitter Data Using Support Vector Machine and Query Expansion Ranking. *JOIN (Jurnal Online Informatika)*, 7(1), 116–122. <https://doi.org/10.15575/join.v7i1.66>
- D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021.
- D. T. Yuliana, M. I. A. Fathoni, and N. Kurniawati, "Penentuan Penerima Kartu Indonesia Pintar KIP Kuliah Dengan Menggunakan Metode K-Means Clustering," *J. Focus Action Res. Math. (Factor M)*, vol. 5, no. 1, pp. 127–141, 2022, doi: 10.30762/f_m.v5i1.570.

- Fathoni, M. I. (2025). Indonesia's Value Added Tax (VAT) Reform: Fiscal Benefits and Sectoral Impacts. *Syntax Admiration*, 6(1), 530–533.
- G. Suganda, M. Asfi, R. T. Subagio, and R. Perdana Kusuma, "Penentuan Penerima Bantuan Beasiswa Kartu Indonesia Pintar (Kip) Kuliah Menggunakan Naïve Bayes Classifier," *JSil (Jurnal Sist. Informasi)*, vol. 9, no. 2, pp. 193–199, 2022, doi: 10.30656/jsii.v9i2.4376.
- K. A. Lubis, M. T. Ari Bangsa, and A. Yudertha, "Analisis Sentimen Opini Masyarakat Terhadap Pindahannya Ibu Kota Indonesia Dengan Menggunakan Klasifikasi Naïve Bayes," *J. Teknoinfo*, vol. 18, no. 1, pp. 226–238, 2024.
- Kolchyna, O., Souza, T. T. P., Treleaven, P. C., & Aste, T. (2015). Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination. *arXiv preprint*. <https://arxiv.org/abs/1507.00955>
- M. F. Azhari and R. Fajriyah, "Idektifikasi Gen Marker Pbmcs Ischemic Stroke Menggunakan," vol. XIII, no. 1, pp. 73–81, 2024.
- P. Dana *et al.*, "dan dampaknya terhadap pendidikan di indonesia," vol. 7, no. 1, pp. 28–36, 2024.
- Prasetya, D. D., Hidayat, W. N., & Wirawan, I. M. (2024). Sentiment Analysis of 'Kampus Merdeka' on Twitter Using Support Vector Machine (SVM) Algorithm. *2024 ICEECIT Conference Proceedings*. <https://doi.org/10.1109/ICEECIT63698.2024.10859978>
- R. Tineges, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 650, 2020, doi: 10.30865/mib.v4i3.2181.
- S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, "Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 153–160, 2023, doi: 10.57152/malcom.v3i2.897.
- Styawati, D., Astuti, E. S., & Haryati, N. (2021). Analisis Sentimen Program Kartu Prakerja Menggunakan Metode Naïve Bayes. *Jurnal Teknologi dan Sistem Informasi*, 2(3), 145–154. Husada, R., & Paramita, A. Y. (2020). Analisis Sentimen Konsumen Terhadap Layanan Maskapai Penerbangan di Media Sosial Menggunakan Metode Support Vector Machine. *Jurnal Sistem Informasi Bisnis*, 10(2), 99–108