Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329

Analisis Sentimen Kebijakan Penempatan Dana 200 Triliun Bank BUMN Menggunakan Algoritma Support Vector Machine

Taufik Ramlan Alfiansyah, Audy Abdillah Hidayat, Alfarezi Hidayat Pratama, Agil Aqshol Mahenda, Muhammad Rafly, Fuad Nur Hasan*

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Bina Sarana Informatika, DKI Jakarta, Indonesia
Email: [1]taupikramlan3590@gmail.com, [2]audyhidayat31@gmail.com, [3]alfarezihidayat4@gmail.com,
[4]agilaqsholmahenda@gmail.com, [5]muhammadrfy919@gmail.com, [6,*]fuad.fnu@bsi.ac.id
Email Korespondensi: fuad.fnu@bsi.ac.id
Received: **29/10/2025**, Revised: **13/11/2025**, Accepted: **29/11/2025**, Published: **30/11/2025**

Abstrak−Kebijakan penempatan dana sebesar Rp200 triliun di bank-bank milik negara menciptakan berbagai tanggapan dari masyarakat yang bisa dilihat melalui komentar di YouTube. Penelitian ini bertujuan untuk memahami bagaimana sentimen masyarakat terhadap kebijakan itu dan juga menguji seberapa baik algoritma Support Vector Machine (SVM) bisa mengkategorikan pendapat publik secara otomatis dan terukur. Data penelitian diperoleh dari 6.100 komentar YouTube pada enam saluran berita nasional dikumpulkan dengan cara mengambil data dari internet dan diproses melalui beberapa langkah seperti membersihkan teks, menormalkan, memecah kata, menghapus kata yang tidak penting, serta merangkum. Setiap komentar diberikan bobot menggunakan metode TF-IDF, lalu dikategorikan dengan SVM kernel linear dengan proporsi data pelatihan dan pengujian 80:20 agar evaluasinya seimbang. Analisis menunjukkan bahwa 54,9% komentar bernada negatif dan 45,1% bernada positif. Model ini menghasilkan akurasi sebanyak 76,7% dan macro F1-score 0,760, yang berarti kinerjanya stabil dan dapat diandalkan. Komentar negatif banyak membahas masalah transparansi, tanggung jawab, dan risiko keuangan, sedangkan komentar positif lebih fokus pada peningkatan likuiditas, kelancaran pendanaan, dan stabilitas ekonomi nasional. Hasil penelitian ini menunjukkan bahwa cara SVM bisa dipakai untuk memahami pendapat masyarakat dan bisa menjadi dasar untuk sistem pemantauan pandangan publik terhadap kebijakan ekonomi pemerintah yang akan berlanjut dan bisa menyesuaikan di masa depan.

Kata Kunci: Analisis Sentimen; Kebijakan Publik; Support Vector Machine; Dana; Bank BUMN

Abstract−The policy of placing Rp 200 trillion in state-owned banks has generated various public responses, as seen through YouTube comments. This study aims to understand public sentiment toward the policy and to test how well the Support Vector Machine (SVM) algorithm can categorize public opinion automatically and measurably. The research data was obtained from 6,100 YouTube comments on six national news channels collected by scraping data from the internet and processing it through several steps, including text cleaning, normalization, word splitting, removing unnecessary words, and summarizing. Each comment was weighted using the TF-IDF method and then categorized using a linear kernel SVM with an 80:20 training and testing data ratio for a balanced evaluation. The analysis showed that 54.9% of the comments were negative and 45.1% were positive. This model achieved an accuracy of 76.7% and a macro F1-score of 0.760, indicating stable and reliable performance. Negative comments focused primarily on transparency, accountability, and financial risk, while positive comments focused more on increasing liquidity, smooth funding, and national economic stability. The results of this study indicate that the SVM method can be used to understand public opinion and can be the basis for a system for monitoring public views on government economic policies that will continue and can be adjusted in the future.

Keywords: Sentiment Analysis; Public Policy; Support Vector Machine; Dana; Bank BUMN

1. PENDAHULUAN

Kebijakan pemerintah menempatkan dana sekitar Rp200 triliun ke bank-bank BUMN memicu perdebatan publik yang cukup sengit. Di satu sisi, kebijakan ini diharapkan bisa memperkuat likuiditas perbankan dan mendukung penyaluran kredit untuk pemulihan ekonomi. Di sisi lain, muncul kekhawatiran mengenai transparansi dalam penempatan dana, pengelolaan yang baik, serta risiko fiskal yang mungkin muncul. Percakapan dan opini terkait kebijakan ini kini banyak terjadi di ruang digital, sehingga platform seperti YouTube menjadi saluran penting untuk menangkap pendapat masyarakat secara langsung. Penelitian terbaru menunjukkan bahwa komentar dari warganet di YouTube bisa merepresentasikan respons spontan yang relevan untuk menganalisis isu kebijakan publik. Namun, data ini perlu diperlakukan dengan teliti karena cenderung singkat, tidak formal, dan terkadang berisi informasi yang tidak jelas.

Dalam konteks analisis opini warganet, YouTube menawarkan komentar dan metrik interaksi yang dapat diolah untuk menilai polarisasi sentimen. Menurut Ahmad Roshid dkk [1] Pemrosesan komentar YouTube yang dipadu SVM ber-kernel linear mampu mencapai akurasi 81%, sehingga layak dipakai untuk memetakan sikap publik terhadap suatu isu populer. Dalam konteks bahasa dan pra-pemrosesan, komentar warganet berbahasa Indonesia sering memuat kata tidak baku/gaul yang perlu dinormalisasi sebelum analisis sentimen. Penelitian oleh Ahmad Fikri Iskandar dkk [2] menegaskan bahwa modifikasi fonem vokal efektif mengembalikan kata tidak baku ke bentuk dasar mencapai presisi sekitar 90% sehingga meningkatkan kualitas masukan untuk model klasifikasi.

Kualitas stemming sendiri berdampak nyata pada temu balik/IR dan pipeline analitik. Menurut Pardede dan Darmawan [3] Stemming Sastrawi memberi presisi terbaik (70,3%) sedangkan Arifin–Setiono paling efisien dari sisi waktu, menunjukkan pemilihan stemmer perlu menimbang akurasi dan biaya komputasi. Di level tugas yang lebih terstruktur, menurut Mustakim dan Priyanta [4] Analisis sentimen berbasis aspek pada ulasan KAI Access memperlihatkan dominasi sentimen negatif—khususnya aspek *errors*—dan SVM dengan *hyperparameter tuning*

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329

memberi kinerja terbaik (akurasi ~91,6%; F1 ~75,6%). Temuan ini relevan untuk memetakan isu dominan dan memilih algoritma yang stabil.

Untuk implementasi praktis dan pelaporan hasil, penelitian oleh Muhammad Taufiq Hidayat dkk [5] menunjukkan SVM linier yang dipadukan TF-IDF dapat diintegrasikan ke aplikasi Streamlit dengan kinerja solid (akurasi 83%; presisi 85%; recall 83%; F1 81%), sehingga proses prediksi, *confusion matrix*, dan visualisasi (word cloud/plot) dapat diakses interaktif oleh pemangku kepentingan. Di ranah isu pemerintahan, menurut Pradana dkk [6] SVM dengan kernel linier terbukti menjadi model terbaik dibanding Naive Bayes dan KNN saat mengklasifikasi opini kinerja pemerintah, dengan akurasi ±85% disertai presisi dan F1-score tertinggi. Temuan ini memperkuat argumen bahwa SVM andal untuk teks pendek dan bising di media sosial.

Pada layanan publik, menurut Damanhuri dan Husein [7] Pemilihan model perlu dievaluasi lewat metrik akurasipresisi-recall dan visualisasi kata kunci; mereka juga menunjukkan ulasan aplikasi Access by KAI didominasi sentimen negatif, menegaskan pentingnya pengukuran performa yang komprehensif. Di aspek alur kerja, penelitian oleh Mufriz dkk [8] menekankan pipeline umum analisis sentimen YouTube: pembagian data, pembobotan TF-IDF, lalu klasifikasi SVM (kernel linear, C=1) untuk memisahkan kelas sentimen, yang pada kasus politik menghasilkan akurasi di kisaran 85–87%.

Khusus pada teks pendek di YouTube, menurut Maulana Malik dkk [9] SVM tetap kompetitif dengan akurasi sekitar 63-65% pada komentar pertandingan Timnas U-23 performa ini menggambarkan tantangan realis data ringkas dan informal, namun sekaligus menegaskan daya guna SVM. Dari sisi kualitas fitur, menurut Suryanto [10] SVM yang dipasangkan TF-IDF dapat mencapai akurasi 85% dan mengungguli metode term presence yang menarik, performanya tetap terjaga bahkan saat data tidak seimbang.

Dalam studi komparatif berskala lebih luas, menurut Bahri dkk [11] Model IndoBERT mencatat performa keseluruhan terbaik namun di jajaran machine learning klasik, SVM tetap yang terkuat dengan akurasi ±89%, sebuah bukti posisi SVM yang solid untuk baseline kuat. Pada isu kemanusiaan yang ramai diperbincangkan, penelitian oleh Hidayat [12] Menunjukkan SVM efektif mengorganisir komentar YouTube yang tak terstruktur, dengan akurasi 77% dan recall 100%; ini menegaskan relevansi SVM untuk wacana publik yang dinamis.

Literatur bertema sosial-politik juga menggarisbawahi rancangan proses yang transparan. Penelitian oleh Bimantoro dkk [13] merumuskan praktik baik berupa kombinasi pra-pemrosesan, penyeimbangan data (SMOTE), dan SVM untuk membaca persepsi publik; studi mereka menunjukkan akurasi tinggi serta konsistensi metrik utama. Sebagai rujukan kontekstual tambahan, menurut Tane dkk Dedi Darwis penerapan SVM pada opini politik di Twitter pernah mencapai akurasi sekitar 91,5%, menandakan bahwa saat fitur dan parameter disetel tepat, SVM dapat sangat akurat pada topik kebijakan dan politik. Terakhir, karakter komentar singkat di YouTube perlu dicermati terkait panjang kalimat. Menurut Pambudi dan Suprapto [14] , performa SVM+TF-IDF relatif tidak sensitif terhadap panjang kalimat (stabil dan cepat), sementara kombinasi berbasis *embedding* (mis. CNN+Word2Vec) peka terhadap variasi panjang implikasinya, SVM+TF-IDF adalah pilihan baseline yang kuat untuk teks pendek dan tidak baku.

Berangkat dari pendekatan makroprudensial, cara pemerintah menempatkan dana di bank-bank BUMN dapat diartikan sebagai skema pembiayaan untuk pemberian kredit (funding-for-lending/FFL), yaitu cara menyuntikkan dana dengan biaya rendah dengan syarat harus diberikan lagi berupa kredit kepada sektor riil. Penelitian di Indonesia pada masa penerapan PMK 70/2020 menunjukkan bahwa bank BUMN yang terlibat dalam skema ini meningkatkan penyaluran kredit dibandingkan bank yang tidak terlibat, tanpa ada tanda-tanda pengambilan keuntungan secara tidak wajar Dampaknya lebih besar karena kenaikan batas kredit daripada penurunan suku bunga, karena risiko kredit masih tergolong tinggi. Penelitian oleh Naiborhu & Ulfa [15] juga menemukan bahwa kebijakan PMK 70/2020 mendorong bank peserta untuk memberikan kredit lebih banyak daripada bank yang tidak tergabung. Kebijakan ini tidak menyebabkan masalah moral seperti penimbunan dana likuiditas oleh SBN. Selain itu, mekanisme dorongannya bersifat kuantitatif, di mana aturan minimal leverage tiga kali membuat bank melebarkan batas kredit dengan meningkatkan jumlah kredit, bukan terutama dengan menurunkan suku bunga.

Kebijakan alokasi dana sebesar Rp200 triliun ke bank BUMN dianalisis melalui komentar di YouTube untuk memahami bagaimana masyarakat memandang kebijakan fiskal tersebut, termasuk isu-isu seperti kepercayaan, akuntabilitas, dan dampaknya terhadap perekonomian. Dalam penelitian ini, pendapat warganet diproses melalui tahapan text mining lalu diklasifikasikan menggunakan algoritma Support Vector Machine (SVM) yang berfungsi memisahkan dua kategori sentimen positif dan negatif agar hasil penilaian bisa diukur secara kuantitatif dan objektif. Hasil dari penelitian ini diharapkan bisa menjadi masukan bagi pemerintah dan pihak terkait dalam menyusun komunikasi kebijakan yang lebih tepat sasaran dan responsif.

2. METODOLOGI PENELITIAN

2.1 Rancangan Tahapan penelitian

Penelitian ini menggunakan cara deskriptif kuantitatif dan metode analisis sentimen untuk melihat pandangan masyarakat tentang keputusan penempatan dana Rp200 triliun di bank-bank milik negara. Rancangan penelitian ini terdiri dari beberapa langkah penting, seperti mengumpulkan data, mempersiapkan teks, memberi label pada sentimen, melatih model klasifikasi, dan menilai kinerja model. Data penelitian ini diambil dari 6.100 komentar di YouTube yang membahas

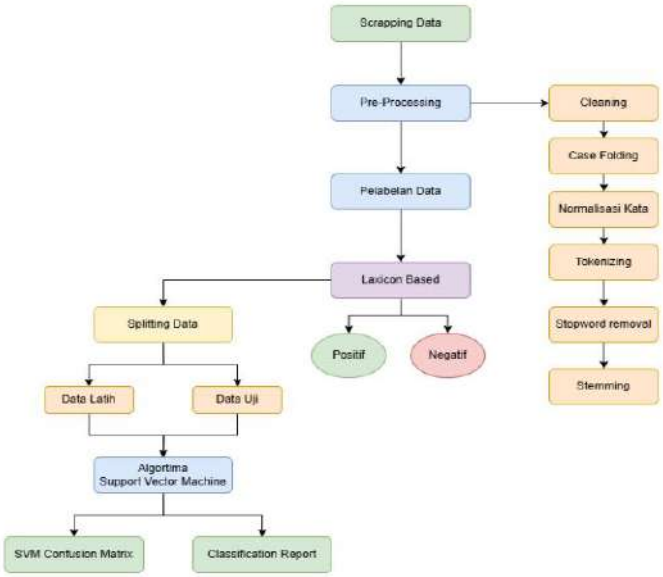
Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 45
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329

kebijakan tertentu di enam saluran berita nasional, yaitu KompasTV, MetroTV, CNBC Indonesia, Official iNews, tvOneNews, dan Kompas.com, dengan menggunakan metode web scraping yang dibuat dengan Python. Data yang dikumpulkan terdiri dari isi komentar dan informasi publik seperti ID video, ID komentar, dan tanggal ketika komentar itu diunggah. Sebelum memberi label, komentar yang ada dipilih dengan membuang komentar yang sama, memastikan semua memakai bahasa Indonesia, dan relevansi dengan topik yang dibicarakan. Seluruh data dikelola dengan cara yang etis tanpa menunjukkan informasi pribadi pengguna dan hanya ditampilkan dalam bentuk agregat.

Setelah semua komentar dibersihkan, setiap komentar akan diberi tanda positif atau negatif untuk analisis sentimen. Langkah ini dilakukan agar dataset yang akan digunakan untuk pelatihan dan pengujian siap. Cara mengumpulkan dan menyiapkan data mengikuti cara umum dalam penelitian analisis sentimen dalam bahasa Indonesia, yaitu dengan mengumpulkan informasi secara teratur, membersihkan atau menormalkan komentar, dan menyusun dataset untuk pelatihan dan pengujian sebelum melakukan pengujian klasifikasi, seperti yang ditunjukkan dalam penelitian Yufridon Charisma Luttu [16].

Dataset yang sudah disiapkan kemudian dibagi menjadi 80 persen untuk data pelatihan dan 20 persen untuk data pengujian. Selanjutnya, data ini diproses dengan menggunakan algoritma Support Vector Machine atau SVM menggunakan pengukuran fitur Term Frequency–Inverse Document Frequency yang dikenal sebagai TF-IDF. Penilaian kinerja model dilakukan dengan menggunakan confusion matrix dan laporan klasifikasi yang mencakup metrik seperti akurasi, presisi, recall, dan F1-score. Desain ini mengikuti cara umum dalam analisis sentimen teks bahasa Indonesia dengan menggunakan TF-IDF untuk membobot fitur, SVM sebagai algoritme klasifikasi, dan evaluasi melalui confusion matrix serta validasi seperti k-fold, seperti yang ditunjukkan oleh Pradana dan Hayaty [17]. Berikut dapat dilihat pada Gambar 1 yg merupakan Tahapan Rancangan Penelitian.



Gambar 1. Rancangan Penelitian

2.2 Pre - Processing

Sebelum dilakukan pelabellan dan klasifikasi, komentar melalui proses pra-pemrosesan untuk membuat teks lebih rapi, konsisten, dan siap dianalisis. Karena komentar di media sosial memiliki karakter tidak baku seperti singkatan, simbol, tautan, dan elemen non-teks, maka dilakukan beberapa langkah secara berurutan, yaitu cleaning, case folding, normalisasi, tokenization, stopword removal, serta stemming. Proses ini membantu mengurangi gangguan, menyamakan penulisan yang berbeda, dan menghasilkan korpus yang terstruktur untuk proses penlabelan, pembuatan model, dan evaluasi selanjutnya.

a. Cleaning
   Cleaning adalah langkah pertama dalam pra-pemrosesan yang bertujuan mengurangi noise dalam teks agar representasi TF-IDF menjadi valid. Langkah ini juga menghilangkan elemen yang bukan teks, seperti tanda baca berlebih, emoji, simbol, angka, URL, tagar, dan mention [16]
b. Case folding
   Case folding mengubah semua huruf menjadi huruf kecil agar penulisan tetap konsisten dan satu kata tidak terbagi menjadi beberapa token karena perbedaan huruf besar dan kecil. Langkah ini mengurangi variasi yang tidak perlu, sehingga proses pencarian fitur, penghitungan bobot TF-IDF, dan klasifikasi menjadi lebih akurat dan efisien. Penjelasan ini sejalan dengan referensi metodologi yang mendefinisikan case folding sebagai langkah untuk "menyamaratakan teks baik dalam bentuk huruf kecil atau huruf besar" agar pemrosesan lanjutan lebih mudah. [16]
c. Normalisasi Kata

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329

Normalisasi adalah proses mengubah kata yang tidak resmi seperti singkatan, bahasa gaul, penulisan salah, atau huruf yang diulang menjadi bentuk yang baku. Tujuannya agar makna tetap konsisten dan frekuensi kata tidak terpecah. Proses ini dilakukan setelah tahap cleaning dan case folding, serta sebelum tahap tokenization. Normalisasi menggunakan kamus KBBI yang telah diperkaya dengan daftar slang dan akronim. Melakukan normalisasi adalah bagian dari proses text preprocessing yang juga mencakup penanganan kata slang dan sangat dianjurkan untuk mempersiapkan data menuju tahap vektorisasi dan klasifikasi sentimen.[18]

d. Tokenization
Tokenisasi membagi teks yang sudah diperbaiki menjadi bagian-bagian kata (token) supaya frekuensi bisa dihitung dan siap untuk langkah selanjutnya seperti stopword removal, stemming, dan ekstraksi fitur. Proses ini dilakukan dengan memisahkan kata berdasarkan batas kata dan menghilangkan tanda baca, sehingga didapatkan rangkaian token yang bersih, konsisten, dan mudah diproses dalam bentuk representasi vektor misalnya TF-IDF serta klasifikasi. Praktik ini sejalan dengan penelitian oleh Fitriyani & Arifin [19] yang mendefinisikan tokenization sebagai "pemecahan kalimat menjadi beberapa token dan sekaligus menghilangkan tanda baca" pada rangkaian *text processing* sebelum langkah *stopword* dan *stemming*.

e. Stopword Removal
Pembersihan kata tidak bermakna menghilangkan kata-kata yang sering muncul dan tidak memberi informasi banyak seperti 'yang', 'di', 'ke', 'ini', 'itu', agar fokus pada kata-kata yang menunjukkan pendapat atau topik. Proses ini dilakukan setelah tokenization dan normalisasi, biasanya menggunakan daftar kata tidak bermakna dalam Bahasa Indonesia, seperti Sastrawi, yang bisa diperluas sesuai dengan kebutuhan korpus. Hasilnya adalah kumpulan kata yang lebih singkat, informatif, dan siap untuk dihitung bobotnya TF-IDF.Menurut Mulyana & Lutfianti stopword merupakan bagian dari rangkaian *preprocessing* dan dilakukan setelah tokenisasi untuk menyingkirkan kata-kata yang tidak memiliki makna seperti "yang", "di", "ke", dan sejenisnya, sebelum lanjut ke tahap berikutnya seperti stemming atau pembobotan.

f. Stemming Data
Stemming mengubah kata yang memiliki imbuhan, seperti awalan, sisipan, akhiran, atau konfiks, menjadi bentuk dasarnya. Tujuannya adalah agar berbagai variasi kata bisa diwakili oleh satu bentuk yang sama. Proses ini membantu mengurangi jumlah dan ketidakteraturan vektor, sehingga metode seperti pembobotan TF-IDF serta proses klasifikasi menjadi lebih stabil. Menurut Pradana & Hayaty, [17] Stemming bertujuan "mengubah kata menjadi kata dasar" dan implementasi yang dipakai adalah Sastrawi Stemmer yang diadaptasi dari Nazief–Andriani, sehingga setiap kata berimbuhan dapat direduksi ke akar katanya sebelum tahap pembobotan TF–IDF dan klasifikasi dilakukan.

2.3 Pelabellan Data (Laxicon Based)

Tahap pengkategorian dilakukan setelah proses pengumpulan dan pembersihan data, dengan cara menggunakan metode klasifikasi berbasis leksikon. Penelitian ini memanfaatkan InSet (Kamus Sentimen Indonesia), yang memiliki sekitar 3.600 kata yang memiliki arti positif dan negatif dalam Bahasa Indonesia yang pas untuk digunakan di teks media sosial. Pelabelan dilakukan setelah tahap crawling dan pembersihan, kemudian semua komentar dikelompokkan menjadi sentimen positif atau negatif untuk analisis selanjutnya. Hal yang sama ditunjukkan pada studi rujukan setelah pengambilan dan penataan data, komentar dilabeli menjadi positif/negatif sebelum pemodelan. [12]

2.4 Splitting Data

Dataset yang sudah berlabel kemudian dibagi menjadi dua subset agar proses pelatihan dan pengujian model berjalan terkontrol serta dapat dievaluasi secara adil. Pada penelitian ini digunakan skema 80% data latih (training set) dan 20% data uji (testing set). Pembagian 80:20 dipilih karena sederhana, banyak dipakai dalam studi analisis sentimen berbahasa Indonesia, dan terbukti memadai untuk mengevaluasi model klasifikasi berbasis SVM setelah tahap prapemrosesan dan pelabelan selesai, sebagaimana praktik pada penelitian sebelumnya yang secara eksplisit membagi data latih–uji dengan perbandingan 80:20 sebelum pemodelan dan evaluasi akurasi–precision–recall–F1 [18].

2.5 Menghitung Term Frequency-Inverse Document Frequency (TF-IDF)

Dalam analisis teks, TF-IDF berfungsi untuk memberikan nilai pada setiap kata, sehingga kata-kata yang muncul sering dalam satu dokumen tetapi jarang terdapat di dokumen lain mendapatkan nilai yang lebih tinggi. Secara prinsip:

a. TF (term frequency): banyaknya kemunculan term t pada dokumen d
b. IDF (inverse document frequency): ukuran "keunikan" term terhadap seluruh korpus; makin sedikit dokumen yang memuat term itu, makin besar nilai IDF
c. TF-IDF adalah hasil perkalian TF × IDF untuk tiap term–dokumen.

penjelasan penggunaan TF-IDF sebagai skema pembobotan untuk klasifikasi sentimen (dengan SVM) dijelaskan jelas pada Pradana & Hayaty [17] Tahap setelah pra-pemrosesan adalah pembobotan TF-IDF, DF dihitung dari log(total dokumen / jumlah dokumen yang memuat term**)**, lalu TF-IDF = TF × IDF bobot ini kemudian menjadi masukan bagi model klasifikasi.

2.6 Algoritma Support Vector Machine

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329

Support Vector Machine adalah algoritma untuk mengklasifikasikan data dengan mencari garis atau hiperbidang yang memiliki jarak terbesar untuk memisahkan dua kelompok. Titik data yang terletak paling dekat dengan garis pembatas disebut support vectors, dan mereka yang menentukan hasil pengambilan keputusan. Dalam teks yang memiliki dimensi tinggi dan cukup jarang, kernel linier biasanya paling efektif. Tingkat ketatnya batas diatur oleh parameter C, yang merupakan kompromi antara memperbesar jarak margin dan mengurangi kesalahan klasifikasi.

Setelah dilakukan pemrosesan awal dan pelabelan, kolom teks yang telah dihilangkan kata umumnya digunakan sebagai fitur dasar. Data kemudian dibagi menjadi dua bagian dengan perbandingan 80:20 untuk keperluan pelatihan dan pengujian, dan selanjutnya diubah menjadi vektor menggunakan CountVectorizer. Vektor X_train dan X_test ini adalah input untuk model. Model yang digunakan untuk pelatihan adalah Support Vector Machine dengan kernel linear SVC(kernel='linear', random_state=42, C=1. 0 [nilai default]). Setelah pelatihan selesai pada data latih, prediksi dilakukan pada data uji dan dievaluasi menggunakan confusion matrix dan classification report (akurasi, presisi, recall, dan F1). Skema *vectorizer → SVM linear → evaluasi confusion matrix*. ini sejalan dengan rujukan metodologi yang menilai kinerja SVM linear pada teks dan melaporkan metrik melalui confusion matrix dan precision/recall/F1 [18].

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Data penelitian dikumpulkan dari kolom komentar di YouTube, khususnya pada kanal berita yang membahas kebijakan penempatan dana sebesar Rp200 triliun di bank BUMN. Setiap komentar yang diambil dilengkapi dengan informasi penulis seperti nama akun, isi komentar, dan waktu posting. Proses ini dilakukan dengan cara scrapping menggunakan bahasa pemrograman Python. Hanya komentar yang bisa diakses secara publik yang digunakan, kemudian diproses untuk menghilangkan duplikat dan memastikan konten tetap relevan dengan topik yang dibahas. Berikut Pada Tabel 1 merupakan Pengumpulan Data

Tabel 1. Pengumpulan data

Author	Comment	Published at
bambanghari5842	Warganet kampung blg SE7 Gaaas pool Pak Menkeu	2025-10-07
NdohaNdoha	Ekonom UGM malu mau mengatakan hal itu. 😊	2025-10-03
kiranaza3827	Dri 200 triliun, Brp lama IDN sdh Dapat merasa	2025-10-02
NeRaka-t1u	Uang sitaan kasus korupsi, kasus narkoba, dan ...	2025-09-28
wachidinsudirohusodho2490	Ada puluhan trilyun digelontorkan ke bank bumn...	2025-09-22

Berdasarkan Tabel 1, kolom author digunakan untuk mengidentifikasi akun tanpa menampilkan data pribadi, comment menjadi sumber utama teks yang digunakan untuk analisis sentimen, dan published_at membantu memahami konteks waktu munculnya opini. Susunan ini sudah cukup baik untuk melacak kronologi percakapan sebelum masuk ke tahap pra-pemrosesan dan pelabelan.

3.2 Hasil Pre-processing

Setelah mengumpulkan data komentar dari YouTube, setiap teks diproses secara bertahap membersihkan, mengubah huruf menjadi huruf kecil, normalisasi, memecah menjadi kata-kata, menghilangkan kata-kata umum, dan mengurangi kata-kata menjadi bentuk dasarnya agar lebih rapi, konsisten, dan siap digunakan untuk memberi bobot dapat dilihat Pada Tabel 2 .

Tabel 2. Hasil Pre-Processing

Tahapan	Comment
Data Komentar	Ada puluhan trilyun digelontorkan ke bank bumn
Cleaning	ada puluhan trilyun digelontorkan ke bank bumn
Case Folding	ada puluhan trilyun digelontorkan ke bank bumn
Normalisasi	ada puluhan triliun digelontorkan ke bank bumn
Tokenization	[ada, puluhan, triliun, digelontorkan, ke, bank, bumn]
Stopword Removal	[puluhan, triliun, digelontorkan, bank, bumn]
Stemming	[puluh, triliun, gelontor, bank, bumn]

Berdasarka Tabel 2 di atas, teks mentah diubah menjadi rangkaian lema yang lebih pendek dan bermakna. Proses cleaning mengurangi noise, case folding menyamakan penggunaan huruf kapital, normalisasi memperbaiki kesalahan penulisan; tokenisasi membagi kalimat, stopword removal menghilangkan kata yang sering muncul, dan stemming mengembalikan kata ke bentuk dasarnya. Hasil akhir ini digunakan dalam pembobotan, seperti TF-IDF, dan pelatihan model SVM.

3.3 Hasil menggunakan Laxicon Based

Setelah tahap pemrosesan awal, setiap komentar mendapatkan label sentimen menggunakan cara yang berbasis kamus dengan InSet (Kamus Sentimen Bahasa Indonesia, dua kategori: positif dan negatif). Langkahnya adalah setiap kata

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329

ditelusuri dalam kamus polaritas; kata yang positif diberi nilai +1 dan kata negatif diberi nilai -1. Kata-kata penyangkal seperti tidak atau bukan akan membalik nilai polaritas kata berikutnya (mengambil 1-2 kata setelahnya). Nilai kalimat dihitung dengan mengurangkan jumlah nilai positif dan negative, label diberikan: jika nilai > 0 maka kategori positif, jika nilai < 0 maka kategori negatif). Untuk nilai hasil labeling data dapat dilihat pada Tabel 2.

Tabel 2. Hasil Labelling Data

Author	Comment	Sentimen
bambanghari5842	warganet kampung bilang se gaaas pool pak menkeu	Positif
NdohaNdoha	ekonom ugm malu mau mengatakan hal itu	Negatif
kiranaza3827	dari triliun berapa lama idn sudah dapat merasa	Positif
NeRaka-t1u	uang sitaan kasus korupsi kasus narkoba dan	Negatif
wachidinsudirohusodho2490	ada puluhan triliun digelontorkan ke bank bumn	Negatif

Berdasarkan Tabel 2 di atas, beberapa komentar dari akun-akun tersebut sudah dikategorikan ke dalam dua kelompok sentimen. Hasil dari pengklasifikasian ini digunakan di tahap berikutnya, yaitu membagi data latih dan data uji, memberi bobot dengan TF-IDF, serta melatih dan mengevaluasi model SVM.

3.4 Hasil Term Frequency-Inverse Document Frequency (TF-IDF)

Setelah proses pra-pemrosesan selesai, metode pembobotan Term Frequency–Inverse Document Frequency (TF-IDF) digunakan untuk menyoroti kata-kata yang sering muncul dalam satu kelas namun tidak banyak muncul di seluruh dokumen. Pembobotan dilakukan pada kolom hasil penghapusan stopword (unigram min_df=2), lalu dihitung rata-rata per kelas agar dapat menemukan kata-kata yang paling mewakili setiap kelas sentimen. Berikut Hasil nya dapat dilihat pada Tabel 4.

Tabel 4. Hasil TF-IDF

Term	Mean TF-IDF
Bank	0,0273
Rakyat	0,0264
Uang	0,0224
Kredit	0,0212
Ekonomi	0,0196
Menkeu	0,0169
Indonesia	0,0151

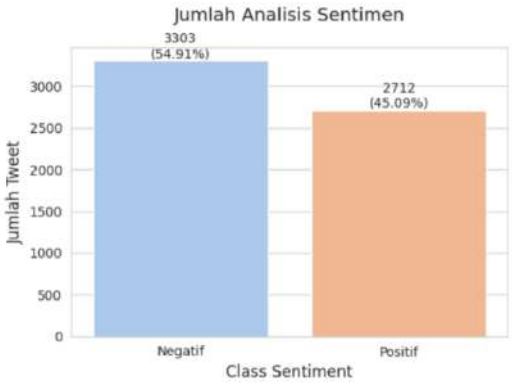
Berdasarkan tabel di atas, kata dengan nilai TF-IDF tertinggi adalah "bank" (0,0273), kemudian "rakyat" (0,0264), "uang" (0,0224), "kredit" (0,0212), "ekonomi" (0,0196), lalu "menkeu" (0,0169) dan "indonesia" (0,0151) pola ini menunjukkan bahwa pembicaraan publik mengenai alokasi dana 200 triliun terpusat pada sektor perbankan dan mekanisme pembiayaan, yaitu cara aliran uang berupa kredit yang mendukung pertumbuhan ekonomi serta fokus pada penerima manfaat (rakyat) dengan penunjuk ke sosok pengambil kebijakan (menkeu) dalam konteks nasional. Secara teknis, nilai TF-IDF yang tinggi menunjukkan kata-kata yang lebih khas, sehingga memiliki potensi besar sebagai fitur utama dalam model sentimen (SVM). Dengan demikian, penggabungan kata-kata seperti bank, uang, kredit, ekonomi, serta rakyat/menkeu memberikan konteks yang jelas sekaligus petunjuk leksikal penting untuk memahami hasil klasifikasi dan menyusun komunikasi kebijakan.

3.5 Hasil Analisis

Dari 6.015 komentar YouTube yang dihimpun dari enam kanal berita nasional, setelah proses pembersihan dan standardisasi teks, seluruh data berhasil dikelompokkan menjadi dua kelas sentimen. Distribusi menunjukkan 3.303 komentar bernada negatif 54,9% dan 2.712 komentar bernada positif 45,1%. Panjang komentar pasca pemrosesan relatif ringkas rata-rata sekitar 11 token median 7 selaras dengan karakter teks di media sosial yang singkat dan langsung.

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329



Gambar 2. Hasil Analisis

Berdasarkan Gambar 2, sentimen negatif terlihat sedikit lebih besar, menunjukkan bahwa kekhawatiran masyarakat masih cukup kuat, umumnya terkait dengan pengelolaan, transparansi, dan risiko. Meski demikian, proporsi sentimen positif juga cukup besar sehingga dukungan terhadap tujuan kebijakan tetap signifikan. Keseimbangan yang hampir seimbang ini menunjukkan bahwa pandangan masyarakat terbagi, tetapi tidak terlalu ekstrem kritik muncul ketika ada pertanyaan terhadap akuntabilitas, sedangkan apresiasi muncul ketika kebijakan tersebut dinilai mampu meningkatkan likuiditas, penyaluran kredit, dan pemulihan aktivitas ekonomi. Secara praktis, komunikasi kebijakan perlu fokus pada mekanisme penyaluran, ukuran keberhasilan, serta dampaknya terhadap sektor riil dan UMKM, sekaligus menampilkan bukti-bukti capaian yang benar-benar terverifikasi untuk memperkuat persepsi positif tanpa mengabaikan kebutuhan akan pengawasan.

3.6 Evaluasi Hasil Algoritma Support Vector Machine

Setelah menyelesaikan pra-pemrosesan dan pelabelan, dataset dipartisi menjadi 80% yang ditunjuk untuk tujuan pelatihan dan 20% dialokasikan untuk tujuan pengujian. Kemudian dinilai melalui empat metrik utama Precision, Recall, F1 Score, dan Accuracy dan dibantu Confusion Matriks untuk menganalisis jenis kesalahan yang terkait dengan setiap kelas. Ringkasan hasil uji pada data Negatif dan Positif.

Tabel 5. *Classification Report for SVM*

	Precision	**Recall**	**F1-Score**	**Support**
Negatif	0.764	0.842	0.801	666.000
Positif	0.772	0.673	0.719	529.000
Accuracy	0.767	0.767	0.767	0.767
Macro avg	0.768	0.758	0.760	1195.000
Weighted avg	0.768	0.767	0.765	1195.000

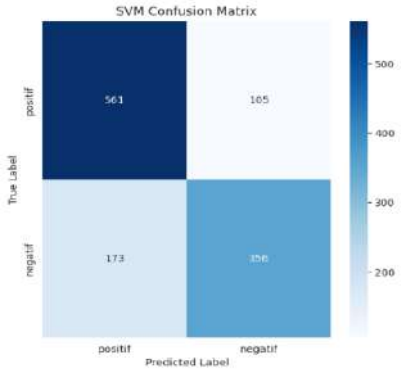
Berdasarkan Tabel 5, model SVM memiliki presisi yang seimbang untuk kedua kelas (Negatif 0,764; Positif 0,772), tetapi recall tidak seimbang Negatif 0,842 lebih tinggi daripada Positif 0,673 yang menunjukkan model lebih mampu mendeteksi kritik dan sebagian komentar positif masih terklasifikasikan sebagai negatif. Sesuai dengan hal tersebut, F1 Negatif (0,801) lebih baik dibandingkan F1 Positif (0,719). Dengan akurasi sebesar 0,767, performa model secara keseluruhan cukup stabil, seperti tercermin dalam macro-F1 0,760 dan weighted-F1 0,765. Perbedaan antara kedua nilai tersebut tidak terlalu besar karena ketidakseimbangan kelas masih cukup moderat, Temuan ini menunjukkan bahwa perlu ada peningkatan pada recall kelas Positif misalnya dengan memperluas kosakata yang bersifat apresiatif, menggunakan n-gram frasa, atau menyesuaikan parameter tanpa mengorbankan presisi yang sudah cukup baik.

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329

3.7 Confusion Matrix

Untuk memperdalam pemahaman atas kinerja klasifikasi, digunakan confusion matrix sebagai alat evaluasi. Matriks ini memperlihatkan perbandingan antara label sebenarnya dan label hasil prediksi, sehingga jenis kesalahan klasifikasi dapat diidentifikasi secara jelas.

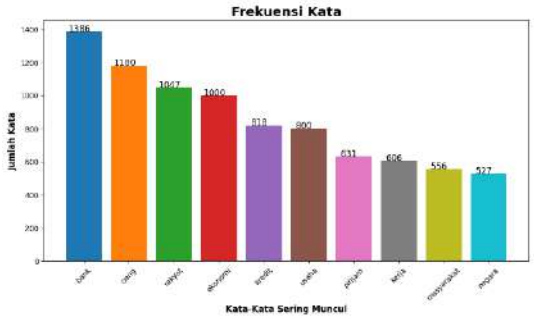


Gambar 3. *Confusion Matrix*

Berdasarkan Gambar 3, sebagian besar prediksi model SVM berada di diagonal utama, yaitu positif ke positif sebanyak 561 dan negatif ke negatif sebanyak 356, yang menunjukkan bahwa sebagian besar komentar berhasil diklasifikasikan sesuai dengan polaritas aslinya. Kesalahan terbesar terjadi pada kategori negatif ke positif sebanyak 173 dan positif ke negatif sebanyak 105. Pola ini menunjukkan bahwa model lebih sensitif terhadap kritik, sehingga sebagian komentar yang sebenarnya mendukung justru terbaca negatif. Sementara itu, beberapa kritik yang bersifat halus atau menggunakan ironi masih terbaca sebagai komentar positif. Untuk mengurangi kesalahan tersebut, dianjurkan untuk menambahkan n-gram frasa (1–2), memperluas kamus ekspresi apresiatif dan kritik, serta menyetel parameter SVM tanpa mengorbankan tingkat presisi.

3.8 Visualisasi

Berikut pada Gambar 4 merupakan Gambar Frekuensi Kata. Pada kumpulan komentar yang sudah diolah sebelumnya, frekuensi menunjukkan sepuluh kata kunci utama bank, uang, rakyat, ekonomi, kredit, usaha, pinjam, kerja, masyarakat, dan negara.



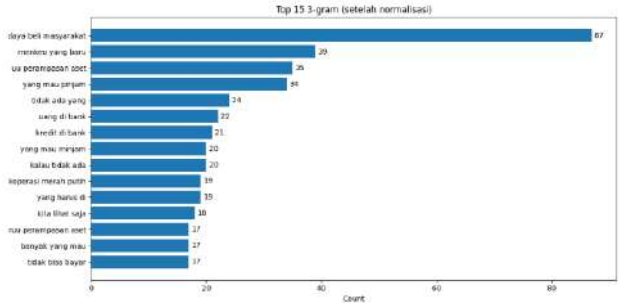
Gambar 4. Frekuensi Kata

Banyaknya kata-kata tersebut menunjukkan bahwa orang lebih banyak berbicara tentang tiga kelompok tema lembaga dan pengelolaan (bank, negara), alat dan cara keuangan (uang, kredit, pinjam), dan orang yang terkena pengaruh kebijakan dalam aspek sosial dan ekonomi (rakyat/masyarakat, usaha/kerja). Secara keseluruhan, gabungan kata "bank–uang–ekonomi–kredit–usaha" menunjukkan bahwa orang sangat peduli dengan likuiditas dan pemberian kredit pada dunia usaha, sedangkan adanya kata "rakyat/masyarakat" menunjukkan perhatian terhadap keadilan dalam manfaat kebijakan. Penting untuk dicatat bahwa banyaknya frekuensi tidak selalu menunjukkan kualitas yang tinggi dalam klasifikasi jadi hasil ini berfungsi sebagai gambaran awal tema yang kemudian akan dilengkapi dengan bobot TF-IDF dan n-gram (misalnya bunga kredit, dana BUMN, bantu usaha) untuk menangkap frasa dengan nuansa yang berbeda. Dengan demikian, analisis frekuensi ini memberikan landasan untuk memahami isu-isu yang paling sering dibahas, sambil memberikan arahan untuk mengeksplorasi fitur-fitur yang lebih informatif di tahap pemodelan sentimen.

Selanjutnya Pada Gambar 5 merupakan Gambar **Diagram 3-Gram**. Untuk melihat kata yang sering muncul dalam percakapan, dibuatlah visualisasi 3-gram yang menunjukkan frasa tiga kata yang paling sering dipakai setelah proses penyesuaian. Pendekatan ini menonjolkan konteks frasal yang lebih ringkas namun informatif.

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

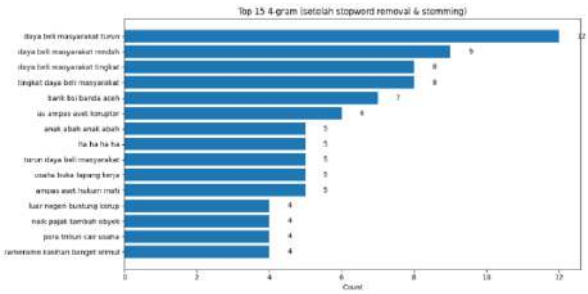
Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329



Gambar 5. Diagram 3-gram

Berdasarkan gambar di atas, rangkaian 3-gram menunjukkan enam tema utama. Yang pertama adalah daya beli dan konsumsi rumah tangga—frasa "daya beli masyarakat" yang sering muncul menggambarkan perhatian terhadap dampak kebijakan terhadap kemampuan belanja masyarakat. Kedua adalah arah kebijakan dan tokoh-tokoh institusional munculnya istilah "menteri yang baru" serta penjelasan tentang "UU/RUU perampasan aset" menunjukkan pembahasan mengenai pergantian wewenang keuangan serta isu hukum dan pengendalian korupsi. Ketiga adalah peran lembaga keuangan dan akses pembiayaan kata-kata seperti "uang di bank", "kredit di bank", dan "yang ingin meminjam" menghubungkan kebijakan dengan ketersediaan dana bagi usaha atau masyarakat. Keempat adalah hambatan dan keterbatasan ungkapan seperti "tidak ada yang/kalau tidak ada/tidak bisa bayar" mencerminkan kekurangan dalam bantuan, syarat yang tidak terpenuhi, atau masalah pembayaran yang menyebabkan reaksi negatif. Kelima adalah entitas atau inisiatif tertentu contohnya "koperasi merah putih" yang menunjukkan referensi terhadap program atau pihak-pihak tertentu saat pengumpulan data. Terakhir adalah frasa fungsional atau retoris seperti "kita lihat saja" dan "yang harus di" yang lebih berperan sebagai penghubung atau penanda sikap hati-hati, sehingga memilikiperan yang rendah dalam hal sentimen namun membantu memberikan konteks dalam sikap menunggu dalam pembicaraan.

Selanjutnya pada Gambar 6 menunjukkan Diagram 4-Gram**.** Untuk memahami konteks dari frasa yang lebih panjang, ditampilkan diagram 4-gram yang berisi 15 frasa terdiri dari empat kata yang paling banyak muncul setelah stopword removal dan stemming.



Gambar 6. Diagram 4-gram

Berdasarkan gambar di atas, rangkaian 4-gram menunjukkan tiga tema utama yang dibahas. Yang pertama adalah daya beli dan konsumsi, yang terlihat dari frasa seperti “daya beli masyarakat turun/rendah/tingkat”, yang menunjukkan perhatian masyarakat terhadap penurunan kemampuan belanjanya hal ini selaras dengan temuan 3-gram yang menunjukkan bahwa dampak dari kebijakan belum sepenuhnya terasa di sisi konsumsi. Kedua adalah tata kelola dan hukum, terlihat dari munculnya frasa seperti “uu ampas aset koruptor”, yang menghubungkan pembicaraan tentang kebijakan dengan penerapan hukum dan pengelolaan aset negara. Ketiga adalah pembiayaan usaha dan pasar kerja, yang terlihat dari frasa seperti “usaha buka lapang kerja” dan “pro triliun cair usaha”, yang menunjukkan hubungan antara kebijakan penempatan dana dengan penciptaan pekerjaan dan penyaluran kredit bagi pelaku usaha. Secara keseluruhan, pola 4-gram membantu memperjelas hubungan antara daya beli, tata kelola, dan pembiayaan usaha/kerja sebagai topik utama yang membentuk sentimen publik.

Pada Gambar 7 menunjukkan Gambar Wordcloud. Yang dimana Visualisasi wordcloud dibuat dari kumpulan kata yang sudah dibersihkan dari kata-kata umum dan diubah sehingga bentuknya lebih sederhana, jadi ukuran setiap kata menunjukkan seberapa sering kata itu muncul dalam komentar.

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 52
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Computing and Informatics Research
ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329



Gambar 7. Wordcloud

Pada gambar terlihat kata-kata yang paling dominan, antara lain "bank, rakyat, uang, usaha, kredit, ekonomi, masyarakat, kerja, bunga, orang, pajak, pinjam, negara atau menkeu, Indonesia". Pola ini menegaskan fokus wacana pada ranah perbankan–pembiayaan (bank, kredit, bunga, pinjam), dampak ke aktivitas ekonomi mikro usaha, kerja/pekerjaan, UMKM, serta persepsi mengenai publik sebagai pemilik kepentingan rakyat/masyarakat, pajak. Kehadiran term seperti "pajak" dan "menkeu/negara" menunjukkan tautan isu ke kebijakan fiskal dan peran pemerintah. Perlu digarisbawahi, wordcloud menampilkan intensitas topik, bukan polaritas karena itu pembacaannya dilengkapi hasil 4-gram tema daya beli, tata kelola, lapangan kerja dan evaluasi model agar makna substantif apakah dukungan atau kritik dapat ditafsirkan dengan tepat.

Hasil dari tes menunjukkan bahwa cara menggunakan SVM dengan kernel linear dapat mengelompokkan pendapat orang di internet dengan baik di komentar YouTube tentang kebijakan penempatan dana 200 triliun. Tingkat akurasinya mencapai 76,7% dan macro-F1-nya 0,760. Presisi untuk kedua kategori cukup seimbang sekitar 0,77, tetapi ada ketidakseimbangan dalam sensitivitas recall untuk Negatif 0,842 lebih tinggi dibandingkan recall untuk Positif 0,673. Confusion Matrix menunjukkan bahwa kesalahan terbesar terjadi karena salah mengklasifikasikan dari Positif menjadi Negatif, ini menunjukkan bahwa model lebih peka terhadap kritik yang jelas dibandingkan dengan pujian yang biasanya dinyatakan dengan singkat, tidak langsung, atau bersyarat**.**

Temuan kata-kata ini menguatkan pola yang ada. Frekuensi distribusi menunjukkan bahwa kata-kata seperti bank, uang, rakyat, ekonomi, kredit, usaha, pinjam, kerja, masyarakat, dan negara adalah yang paling dominan. Di sisi lain, visualisasi n-gram menegaskan kelompok tentang daya beli (contohnya "daya beli masyarakat turun/rendah/tingkat") serta hubungannya dengan akses pembiayaan dan pekerjaan. Pola kata ini menunjukkan cara berpikir masyarakat pengalokasian dana dianggap efektif jika benar-benar digunakan untuk kredit dan mendukung konsumsi keluarga sebaliknya, cerita tentang penurunan daya beli dan masalah pengelolaan membuat sentimen negatif yang mudah dipahami oleh model.

Dari segi cara kerja, gabungan langkah-langkah persiapan yang menyeluruh seperti cleaning data, case folding, normalisasi kata, tokenizing, stopword removal, dan merangkum kata—dengan sistem encoding yang menggunakan metode bag-of-words terbukti cukup bagus dalam menciptakan ruang pemisah di area fitur yang sangat banyak dimensinya. Namun, ada beberapa alasan yang menjelaskan mengapa kinerjanya masih terbatas. Pertama, label yang dibuat berdasarkan pengejaan mungkin belum bisa sepenuhnya mencakup konteks praktis (seperti ironi, sarkasme, atau dukungan yang bersyarat), sehingga beberapa komentar yang positif malah dianggap netral atau negatif. Kedua, sedikitnya variasi antara kelas (Negatif lebih banyak dibanding Positif) dan komentar yang pendek bisa membuat sinyal yang jelas pada kelas Positif menjadi sulit ditemukan. Ketiga, penggunaan unigram dari metode bag-of-words bisa mengurangi kekuatan representasi dari pola yang lebih panjang.

Arah penguatan model tetap sesuai dengan kerangka utama yang terdiri dari empat langkah. Langkah pertama adalah pengayaan sinyal positif, yakni memperluas kosakata yang menyatakan kepuasan dan mengatur ulang contoh komentar positif agar beragam bentuk ekspresi dukungan termasuk yang tidak langsung terwakili dengan baik. Langkah kedua adalah eksperimen fitur, yakni membandingkan representasi TF-IDF dengan rentang n-gram yang lebih informatif, seperti 1–2 atau 1–3, serta menyesuaikan parameter min_df dan max_df untuk mengurangi fitur yang terlalu langka atau terlalu umum. Langkah ketiga adalah penyesuaian hiperparameter C pada SVM agar bisa menyeimbangkan batas pengklasifikasian dan kesalahan, sehingga kesalahan klasifikasi dari Positif ke Negatif berkurang tanpa mengurangi tingkat presisi. Langkah keempat adalah analisis koefisien model untuk menemukan kata atau frasa yang paling berpengaruh pada setiap kelas hasilnya dapat digunakan untuk memperbaiki kosakata sekaligus menjadi dasar interpretasi dalam merancang komunikasi kebijakan.

Secara keseluruhan, hasil ini menempatkan SVM linear sebagai dasar yang kuat, presisi yang tetap stabil, sensitivitas tinggi terhadap kritik, dan konsistensi dalam tema antara fitur kata dan hasil model. Dengan sedikit perubahan pada bagian fitur, label, dan penyetelan parameter, peningkatan recall positif dan skor macro-F1 bisa tercapai, sehingga gambaran tentang pandangan publik terhadap kebijakan menjadi lebih seimbang dan mewakili.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa pandangan masyarakat tentang kebijakan penempatan dana sebesar Rp200 triliun di bank-bank BUMN memiliki perbedaan yang jelas. Sebagian besar, yaitu 54,9% dari pendapat yang ada, bersifat negatif

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

ISSN 2808-375X (Media Online)
Vol 5, No 1, November 2025
Hal 44-54
DOI 10.47065/comforch.v5i1.2329

sementara yang positif hanya 45,1%. Komentar yang negatif biasanya menyoroti masalah seperti kurangnya transparansi, akuntabilitas, dan risiko keuangan. Sementara itu, komentar yang positif lebih fokus pada meningkatnya likuiditas, kemudahan dalam mendapatkan pembiayaan, dan kestabilan ekonomi negara. Hasil dari penelitian ini menunjukkan bahwa kepercayaan masyarakat terhadap kebijakan keuangan sangat dipengaruhi oleh seberapa jelas cara penyaluran dana dan seberapa terbuka pengawasan yang ada. Dari segi teknis, algoritma Support Vector Machine (SVM) yang menggunakan pembobotan TF-IDF memberikan hasil akurasi sebesar 76,7% dan macro F1-score sebesar 0,760. Ini menunjukkan bahwa model ini bekerja dengan baik dan bisa diandalkan dalam mengklasifikasikan pendapat publik. Model ini lebih peka terhadap komentar yang bersifat kritik dibandingkan dengan komen positif yang dinyatakan dengan lembut atau bersyarat. Jadi, untuk meningkatkan akurasi, kita bisa menambah kategori sentimen netral, memperbanyak kosakata positif dan sarkastik, serta mencoba variasi n-gram (1–2 atau 1–3) dan mengatur hiperparameter SVM supaya keseimbangan antara keakuratan dan ingatan semakin baik. Di masa yang akan datang, cara menggunakan word embedding seperti Word2Vec, FastText, atau BERT bisa membantu kita memahami arti dari teks yang rumit dengan lebih baik. Di sisi kebijakan, pemerintah dan lembaga keuangan harus meningkatkan cara mereka berkomunikasi dengan masyarakat yang jelas dan berdasarkan data, terutama saat menjelaskan indikator keberhasilan dan cara mengawasi dana. Tindakan ini sangat penting untuk membangun kepercayaan masyarakat, memperkuat pandangan positif, dan memastikan bahwa penerapan kebijakan ekonomi berjalan dengan baik dan terus-menerus.

REFERENCES

[1] F. Saefulloh, “Analisis Sentimen Masyarakat Terhadap E-Commerce Pada Media Sosial Twitter Menggunakan Metode Support Vector Machine (Svm) (Studi Kasus Shopee Dan Tokopedia),” *S1 thesis, Univ. Muhammadiyah Purwokerto*, vol. 8, no. 2502, pp. 79–88, 2023.

[2] A. F. Iskandar, E. Utami, W. Hidayat, A. P. Budi, and A. D. Hartanto, “Modifikasi Fonem Vokal Pada Stemming Kata Tidak Baku,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 1, pp. 35–42, 2023, doi: 10.25126/jtiik.2023105028.

[3] D. D. Jasman Pardede, “Perbandingan Algoritma Stemming Porter, Sastrawi, Idris, Dan Arifin & Setiono Pada Dokumen Teks Bahasa Indonesia,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 1, pp. 69–76, 2025, doi: 10.25126/jtiik.2025128860.

[4] H. Mustakim and S. Priyanta, “Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 16, no. 2, p. 113, 2022, doi: 10.22146/ijccs.68903.

[5] M. T. Hidayat, M. Arifin, and S. Muzid, “Prediction Sentiment Analysis Grab Reviews using SVM Linear Based Streamlit,” *Indones. J. Comput. Cybern. Syst.*, vol. 19, no. 2, pp. 1–12, 2025, doi: 10.22146/ijccs.104924.

[6] P. Simposium, N. Multidisiplin, and U. M. Tangerang, “Analisis Sentimen Kinerja Pemerintahan Menggunakan Algoritma,” *Pros. Simp. Nas. Multidisiplin*, vol. 4, pp. 114–121, 2022.

[7] R. Damanhuri and V. A. Husein, “Analisis Sentimen pada Ulasan Aplikasi Access by KAI Berbahasa Indonesia Menggunakan Word-Embedding dan Classical Machine Learning,” *J. Masy. Inform.*, vol. 15, no. 2, pp. 97–106, 2024, doi: 10.14710/jmasif.15.2.62383.

[8] Muhammad Fadwa Mufriz, Deden Witarsyah, and Riska Yanu Fa’rifah, “Analisis Sentimen Pada Komentar Youtube Untuk Mengetahui Pandangan Masyarakat Kepada Calon Presiden Indonesia 2024 Menggunakan Algoritma Support Vector Machine,” *e-Proceeding Eng.*, vol. 11, no. 4, pp. 4257–4263, 2024.

[9] M. Malik, D. A. Pangestu, and M. R. Pribadi, “AICOMS Applied Information Technology and Computer Science Analisis Sentimen Hasil Pertandingan Sepakbola Timnas Indo- nesia di Piala Asia U-23 pada Platform Youtube menggunakan Algoritma Suport Vector Machine (SVM),” vol. 3, no. 1, pp. 38–45, 2024.

[10] S. Suryanto and W. Andriyani, “Sentiment Analysis of X Platform on Viral ‘Fufufafa’ Account Issue in Indonesia Using SVM,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 19, no. 1, p. 95, 2025, doi: 10.22146/ijccs.104158.

[11] C. A. Bahri and L. H. Suadaa, “Aspect-Based Sentiment Analysis in Bromo Tengger Semeru National Park Indonesia Based on Google Maps User Reviews,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 17, no. 1, p. 79, 2023, doi: 10.22146/ijccs.77354.

[12] H. Hidayat, F. Santoso, and L. F. Lidimillah, “Analisis Sentimen Pengguna YouTube Tentang Rohingya Menggunakan Algoritma SVM (Support Vector Machine),” *G-Tech J. Teknol. Terap.*, vol. 8, no. 3, pp. 1729–1738, 2024, doi: 10.33379/gtech.v8i3.4497.

[13] T. Arya Bimantoro, Y. Rahmawati, Y. Findawati, and M. A. Rosid, “Sentiment Analysis of YouTube Comments on the East Java Grant Case Using the Support Vector Machine (SVM) Algorithm. [Analisis Sentimen Komentar Youtube Terhadap Kasus Dana Hibah Jawa Timur Menggunakan Algoritma Support Vector Machine (SVM)],” pp. 1–13.

[14] A. Pambudi and S. Suprapto, “Effect of Sentence Length in Sentiment Analysis Using Support Vector Machine and Convolutional Neural Network Method,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 1, p. 21, 2021, doi: 10.22146/ijccs.61627.

[15] E. D. Naiborhu and D. Ulfa, “The lending implication of a funding for lending scheme policy during COVID-19 pandemic: The case of Indonesia Banks,” *Econ. Anal. Policy*, vol. 78, pp. 1059–1069, 2023, doi: 10.1016/j.eap.2023.04.025.

[16] S. A. S. Mola, Y. C. Luttu, and D. N. Rumlaklak, “Perbandingan Metode Machine Learning dalam Analisis Sentimen Komentar Pengguna Aplikasi InDriver pada Dataset Tidak Seimbang,” *J. Sist. Inf. Bisnis*, vol. 14, no. 3, pp. 247–255, 2024, doi: 10.21456/vol14iss3pp247-255.

[17] A. W. Pradana and M. Hayaty, “The Effect of Stemming and Removal of Stopwords on the Accuracy of Sentiment Analysis on Indonesian-language Texts,” *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, no. 3, pp. 375–380, 2019, doi: 10.22219/kinetik.v4i4.912.

[18] D. Triully Prasetyo and Atiqah Meutia Hilda, “Metode Support Vector Machine Untuk Analisis Sentimen Aplikasi Threads di Google Play Store,” *Indones. J. Comput. Sci.*, vol. 13, no. 5, pp. 76–83, 2024, doi: 10.33022/ijcs.v13i5.4446.

[19] Fitriyani and Arifin Toni, “Penerapan_Word_N-Gram_Untuk_Sentiment_Analysis_Rev,” *J. Sist. Informas (Sist.*, vol. 9, no. September, pp. 610–621, 2020.

Copyright © 2025 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.