

Analisis Sentimen Terhadap Kinerja Wakil Presiden Pada Tahun 2025 Menggunakan Metode Support Vector Machine

Try Fani*, Yusuf Ramadhan Nasution

Fakultas Sains dan Teknologi, Program Studi Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Email: ¹*fanipku01@gmail.com, ²ramadhannst@uinsu.ac.id

Email Penulis Korespondensi : fanipku01@gmail.com

Submitted 02-08-2025; Accepted 19-08-2025; Published 20-08-2025

Abstrak

Penelitian ini menganalisis persepsi publik terhadap kinerja Wakil Presiden Indonesia tahun 2025 dengan memanfaatkan data opini di media sosial X/Twitter. Masalah yang diangkat adalah belum tersedianya gambaran kuantitatif terkini mengenai polaritas sentimen publik pasca-pelantikan, khususnya terhadap Gibran Rakabuming Raka yang memicu pro-kontra. Tujuan penelitian adalah mengklasifikasikan sentimen (positif/negatif) dan mengukur kinerja model klasifikasi pada korpus unggahan warganet. Data berjumlah 898 tweet yang dihimpun menggunakan tagar #wapres, #Gibran, dan #WapresGibran. Alur pengolahan mencakup pembersihan teks (cleaning), penyamaan huruf (case folding), tokenisasi, normalisasi, penghapusan stopword, dan stemming. Representasi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF), sedangkan pemodelan dilakukan dengan Support Vector Machine (SVM). Hasil menunjukkan 647 tweet bernada positif dan 251 bernada negatif, menandakan kecenderungan sentimen yang relatif positif namun tetap menyimpan polaritas. Model SVM mencapai akurasi 80,68%, yang mengindikasikan performa klasifikasi yang andal pada data berfitur tinggi. Temuan ini menyuguhkan potret ringkas opini publik yang dapat menjadi rujukan bagi perumus kebijakan dan strategi komunikasi pemerintah. Kontribusi utama studi ini ialah memberikan bukti empiris berbasis media sosial tentang dinamika sentimen terhadap kinerja Wakil Presiden, sekaligus membuktikan efektivitas kombinasi TF-IDF dan SVM pada tugas analisis sentimen politik kontemporer.

Kata Kunci: Analisis Sentimen; Media Sosial; Support Vector Machine; TF-IDF; Wakil Presiden

Abstract

This study examines public perception of the performance of Indonesia's Vice President in 2025 by utilizing opinion data from social media X/Twitter. The research addresses the lack of up-to-date quantitative insights into public sentiment polarity following the inauguration, particularly regarding Gibran Rakabuming Raka, whose appointment has sparked mixed reactions. The objective of this study is to classify sentiments as positive or negative and to evaluate the performance of the classification model on a corpus of user posts. The dataset consists of 898 tweets collected using the hashtags #wapres, #Gibran, and #WapresGibran. Data processing involved cleaning the text, converting all characters to lowercase (case folding), tokenization, normalization, removal of stopwords, and stemming. Feature representation was carried out using Term Frequency-Inverse Document Frequency (TF-IDF), while modeling was performed with the Support Vector Machine (SVM) algorithm. Results show 647 tweets with positive sentiment and 251 tweets with negative sentiment, indicating a generally positive tendency while maintaining some diversity of opinion. The SVM model achieved an accuracy of 80.68%, demonstrating reliable performance on high-dimensional textual data. These findings provide a concise overview of public opinion that can serve as a reference for policymakers and government communication strategies. The study's main contribution lies in offering empirical evidence from social media on sentiment dynamics toward the Vice President's performance, while also highlighting the effectiveness of combining TF-IDF and SVM in contemporary political sentiment analysis.

Keywords: Sentiment Analysis; Social Media; Support Vector Machine; TF-IDF; Vice President

1. PENDAHULUAN

Kinerja pejabat publik pada posisi strategis seperti wakil presiden berperan penting dalam menjaga legitimasi politik dan stabilitas pemerintahan[1]. Seiring dengan kemajuan teknologi digital, media sosial hadir sebagai ruang utama masyarakat mengekspresikan opini politik secara luas dan real time, sehingga relevan sebagai sumber analisis persepsi publik[2]. Pengangkatan Gibran Rakabuming Raka sebagai Wakil Presiden Indonesia periode 2024–2029 memicu beragam respons, dipengaruhi faktor usia muda, latar belakang sebagai pengusaha dan mantan Wali Kota Surakarta, serta hubungan keluarga dengan Presiden Joko Widodo. Putusan Mahkamah Konstitusi yang memungkinkan pencalonan di luar batas usia minimal memunculkan narasi politik dinasti, menambah sentimen negatif berupa keraguan terhadap kapasitas dan potensi konflik kepentingan[3]. Namun, terdapat pula sentimen positif yang melihat Gibran sebagai simbol regenerasi dan pembaruan pemerintahan. Kondisi ini mencerminkan polarisasi opini publik yang perlu dianalisis secara ilmiah untuk memperoleh gambaran objektif dan komprehensif mengenai persepsi masyarakat terhadap kinerja wakil presiden.

Perbedaan pandangan terhadap kinerja Wakil Presiden Gibran Rakabuming Raka memunculkan polarisasi opini di masyarakat. Pendukung menilai inisiatif seperti program “Lapor Mas Wapres” sebagai langkah inovatif yang mendekatkan wakil presiden dengan publik, sementara kritik menyoroti kurangnya pengalaman dan insiden yang memicu sentimen negatif, seperti kesalahan dalam kunjungan kerja yang viral di media sosial. Cepatnya arus informasi di ruang digital memperbesar dampak isu dan membentuk persepsi publik secara dinamis[4]. Kondisi ini menunjukkan perlunya analisis berbasis data yang mampu menangkap nuansa sentimen secara mendalam dan objektif[5].

Untuk memahami polarisasi opini publik, diperlukan metode yang mampu mengolah data teks secara sistematis dan akurat. Analisis sentimen dengan pembelajaran mesin menjadi solusi efektif untuk mengklasifikasikan opini publik secara otomatis[6]. Penelitian ini menggunakan Support Vector Machine (SVM) dipilih karena kemampuannya yang

unggul dalam mengelola data dengan dimensi tinggi., meminimalkan risiko overfitting, dan memberikan kinerja klasifikasi yang konsisten[7], sehingga dapat menghasilkan gambaran objektif mengenai sentimen masyarakat terhadap kinerja Wakil Presiden Gibran Rakabuming Raka pada tahun 2025.

Analisis sentimen, yang merupakan bagian dari text mining, berfungsi untuk mengkategorikan pandangan, perasaan, dan sikap individu mengenai suatu objek menjadi positif dan negatif [8][9]. Metode ini memanfaatkan kecerdasan buatan, *natural language processing*, dan komputasi linguistik untuk mengenali, mengekstrak, serta menganalisis informasi teks secara terstruktur, sehingga berfungsi dalam menangkap persepsi publik dan memberikan dukungan bagi pengambilan keputusan berbasis data[10]. Analisis sentimen umumnya dilakukan melalui tahapan pengumpulan data, preprocessing untuk membersihkan dan menormalkan teks, ekstraksi fitur, klasifikasi sentimen menggunakan algoritma pembelajaran mesin, serta evaluasi kinerja model dengan metrik seperti akurasi, presisi, recall, dan F1-score guna mendukung berbagai kebutuhan analisis berbasis data[11][12].

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang dirancang untuk menemukan hyperplane dengan margin terbesar guna memisahkan dua kelas secara optimal[13][14]. Metode ini menggunakan fungsi linier dalam ruang fitur berdimensi tinggi dan efektif menangani data dengan banyak variabel, jumlah sampel kecil hingga menengah, serta mampu meminimalkan risiko overfitting. SVM bekerja berdasarkan teori optimasi dan prinsip pembelajaran statistik, dengan memanfaatkan support vector untuk membentuk hyperplane yang memisahkan kelas data secara akurat, bahkan pada kasus dengan margin besar[15]. Support Vector Machine (SVM) adalah algoritma pembelajaran mesin dengan dasar teoritis yang kuat dan kemampuan generalisasi yang baik. Metode ini populer karena mampu menghasilkan klasifikasi yang lebih akurat dibanding banyak algoritma lain, sehingga efektif digunakan dalam berbagai aplikasi seperti analisis sentimen dan pengenalan pola[16].

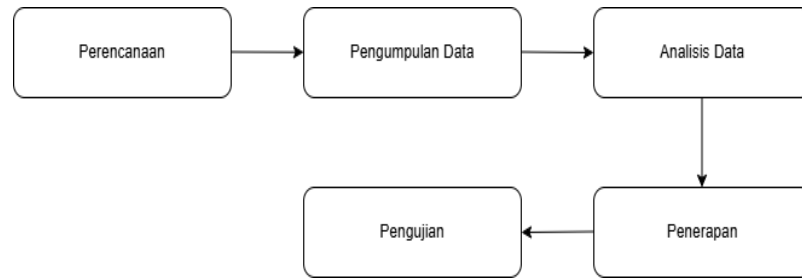
Penelitian Zahra Pruwanti dan Sugiyono berjudul “Pemodelan Text Mining untuk Analisis Sentimen Terhadap Program Makan Siang Gratis di Media Sosial X Menggunakan Algoritma Support Vector Machine (SVM)” menunjukkan bahwa data dari media sosial X dapat dianalisis secara efektif menggunakan SVM. Dengan tingkat akurasi 94% dari 606 data uji, model berhasil memprediksi 497 tanggapan positif dan 37 tanggapan negatif. Hasil ini mengindikasikan bahwa mayoritas masyarakat memiliki persepsi positif terhadap program makan siang gratis, meskipun terdapat sebagian kecil tanggapan negatif. Temuan tersebut memberikan gambaran bahwa program ini diterima dengan baik oleh publik di media sosial[17]. Penelitian lain yang dilakukan oleh Ilka Zufria, Aidil Halim Lubis, dan Siti Septia Febiyaula berjudul “Analisis Sentimen Kepercayaan Masyarakat Terhadap Kepolisian Republik Indonesia Menggunakan Algoritma SVM” menganalisis opini publik di media sosial Twitter menggunakan 801 data, dengan 90% (720 data) sebagai data latih dan 10% (81 data) sebagai data uji. Pelabelan sentimen dilakukan menggunakan kamus lexicon, menghasilkan 444 data positif dan 357 data negatif. Penerapan algoritma Support Vector Machine menghasilkan akurasi 83,95%, precision 85,71%, recall 87,50%, dan F1-score 86,60%, menunjukkan kinerja klasifikasi yang cukup baik dalam mengidentifikasi sentimen masyarakat terhadap kepolisian[18].

Penelitian terdahulu telah memanfaatkan algoritma SVM untuk menganalisis sentimen publik pada berbagai konteks, seperti evaluasi program makan siang gratis maupun pengukuran tingkat kepercayaan masyarakat terhadap institusi kepolisian. Hasil-hasil tersebut menunjukkan bahwa SVM efektif dalam mengolah opini publik berbasis media sosial. Namun, penelitian-penelitian sebelumnya belum secara spesifik menyoroti persepsi masyarakat terhadap kinerja Wakil Presiden, padahal posisi ini memiliki peran strategis dalam sistem pemerintahan Indonesia. Dengan demikian, terdapat celah penelitian (research gap) dalam kajian analisis sentimen politik yang berfokus pada pejabat eksekutif tingkat kedua. Penelitian ini hadir untuk mengisi celah tersebut dengan menerapkan analisis sentimen berbasis SVM terhadap kinerja Wakil Presiden Indonesia tahun 2025, sehingga memberikan kontribusi baru berupa perspektif ilmiah terkait penerimaan publik terhadap wakil kepala pemerintahan melalui pendekatan analisis sentimen media sosial.

Penelitian ini bertujuan mengkaji sentimen masyarakat mengenai kinerja Wakil Presiden Indonesia tahun 2025 dengan memanfaatkan data dari media sosial, serta menggunakan algoritma Support Vector Machine untuk mengelompokkan opini publik menjadi kategori positif dan negatif. Penelitian ini juga bertujuan memberikan gambaran objektif dan terukur mengenai persepsi publik di tengah dinamika politik nasional, sekaligus menyediakan temuan yang dapat dijadikan acuan bagi pengambil kebijakan dan pihak terkait dalam merumuskan strategi komunikasi serta kebijakan publik yang responsif terhadap aspirasi masyarakat.

2. METODOLOGI PENELITIAN

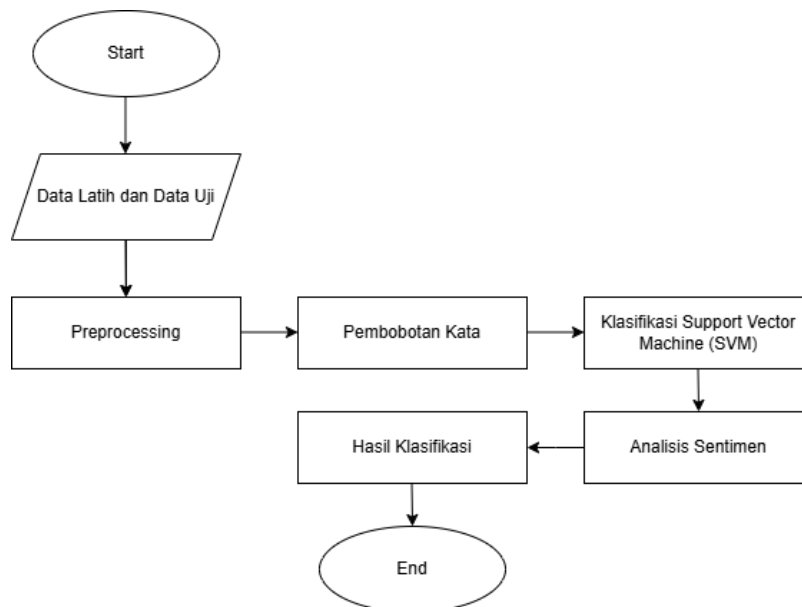
Penelitian ini dilakukan secara kuantitatif untuk memahami persepsi publik terhadap kinerja Wakil Presiden Indonesia tahun 2025. Tahapannya meliputi perencanaan penelitian, pengumpulan data dari media sosial, analisis dan pengolahan data, penerapan algoritma Support Vector Machine untuk klasifikasi sentimen, serta pengujian hasil untuk menghasilkan gambaran objektif mengenai opini masyarakat. Tahap penelitian dapat dilihat pada tabel berikut.



Gambar 1. Tahapan Penelitian

2.1 Perencanaan

Pengangkatan Gibran Rakabuming Raka sebagai Wakil Presiden Indonesia periode 2024-2029 menimbulkan beragam reaksi publik. Penunjukannya yang relatif muda dan berlatar belakang pengusaha serta wali kota Surakarta memicu kontroversi, termasuk isu politik dinasti dan keraguan kapasitas. Di sisi lain, sebagian masyarakat mendukungnya sebagai simbol regenerasi pemerintahan. Mengacu pada permasalahan ini, Studi ini menerapkan metode Support Vector Machine untuk menganalisis sentimen. Berikut adalah alur penerapan metode SVM dalam penelitian ini.



Gambar 2. Alur Metode SVM

2.2 Pengumpulan Data

Penelitian ini mengumpulkan data melalui dua metode. Metode pertama adalah studi literatur, yakni dengan membaca, mencatat, dan mengolah berbagai sumber referensi yang relevan sebagai landasan konseptual serta dasar dalam penyusunan kerangka penelitian[19]. Metode kedua adalah observasi, yaitu mengumpulkan data primer secara langsung dari Twitter dengan memanfaatkan teknik crawling melalui Twitter API [20]. Proses crawling dilakukan menggunakan Google Colab dengan bantuan library Python, yakni socrate, untuk mengambil data berupa teks tweet, nama akun pengguna, serta tanggal publikasi, sehingga informasi yang dikumpulkan lebih terstruktur, akurat, dan siap dianalisis lebih lanjut.

2.3 Analisis Data

Sesudah data terkumpul, tahap selanjutnya adalah melakukan analisis guna mengukur sentimen masyarakat terhadap kinerja Wakil Presiden. Data hasil crawling dari media sosial terlebih dahulu melalui proses preprocessing yang mencakup pembersihan teks, case folding, tokenisasi, penghapusan stopword, serta stemming agar data yang digunakan lebih konsisten dan terstruktur[21]. Selanjutnya, data diolah menggunakan algoritma Support Vector Machine (SVM) dengan kernel linier karena data hasil representasi teks dengan TF-IDF berbentuk vektor berdimensi tinggi dan cenderung terpisahkan secara linier. Proses pelatihan SVM diformulasikan sebagai optimisasi kuadrat dengan melibatkan matriks Hessian untuk menentukan fungsi tujuan dan kendala optimasi. Penyelesaian masalah optimasi ini dilakukan menggunakan metode Sequential Minimal Optimization (SMO) yang bekerja dengan memperbaiki pasangan variabel Lagrange secara bertahap hingga mencapai konvergensi. Hasil akhir dari proses klasifikasi berupa pemisahan data ke dalam kategori sentimen positif dan negatif. Hasil klasifikasi kemudian dievaluasi dan dianalisis untuk mengidentifikasi

pola, tren, serta persepsi publik secara keseluruhan, sehingga penelitian mampu memberikan gambaran objektif mengenai opini masyarakat terhadap kinerja Wakil Presiden.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pada tahap ini, disajikan data hasil crawling dari media sosial Twitter (X) sebanyak 898 tweet. Data dikumpulkan secara otomatis menggunakan kata kunci berupa tagar #wapres, #Gibran, dan #WapresGibran, yang secara langsung berkaitan dengan topik kinerja Wakil Presiden Indonesia. Tweet-tweet tersebut mencerminkan opini dan tanggapan publik yang akan dianalisis lebih lanjut. Karena data yang diperoleh masih berupa data mentah, selanjutnya akan melalui tahapan pemrosesan, termasuk pembersihan (preprocessing), untuk memastikan kualitas dan konsistensi data. Penyajian data pada tahap ini bertujuan memberikan gambaran awal mengenai jenis dan karakteristik opini publik yang menjadi fokus penelitian. Tabel dibawah merupakan representasi data pada penelitian ini.

Tabel 1. Representasi Data

No	created at	full text	Sentiment
1	Wed Jul 16 2025	Wapres Gibran serukan pemimpin ke lapangan #beritanasional #wapres #gibran #pimpinan https://t.co/aH8Yl6BAbJ	Positif
2	Wed Jul 16 2025	(P3N) XXV dan Pendidikan Penyiapan serta Pemantapan Pimpinan Nasional (P4N) LXVIII di Istana Wakil Presiden Jakarta. #wapres #gibran #hilirisasi #kemenyan #LV #gucci #parfum #kuatbacacom #connectingideas	Positif
3	Tue Jul 15 2025	Wakil Presiden (Wapres) Gibran Rakabuming Raka menyentil ibu-ibu yang memakai parfum mahal merek Louis Vuitton (LV) dan Gucci @gibran_rakabuming #gibran #wapres #prabowo #tiktok #berita #news #beritahariini #trending https://t.co/rIaznH851U	Positif
4	Tue Jul 15 2025	Prosesnya itu masih dalam mekanisme yang ada kata Puan di Kompleks Parlemen Senayan Jakarta Selasa (15/7/2025). #Wapres #WapresGibran #Pemakzulan #PuanMaharani #Serambinews https://t.co/G6GWa8s8SW	Positif
5	Tue Jul 15 2025	Bansos dipake judul? Gimana tuh? https://t.co/2bP3h7pHJ5 #gibranrakabumingraka #wapres #judul #judionline #bantuansosial #bansos	Positif
...	...	...	...
896	Tue Jun 24 2025	Wapres #Gibran Ajak #UMKM Mekaar Bondowoso Naik Kelas Lewat Platform Digital https://t.co/tWxsmrprMf	Positif
897	Tue Jun 24 2025	Masa Reses Selesai Pimpinan DPR Segera Bahas Usulan Pemakzulan Gibran Dasco Ungkap Mekanismenya #Gibran #PemakzulanGibran #DPR https://t.co/gchCF3vj13 lewat @tribunkaltim	Positif
898	Tue Jun 24 2025	Masa Reses Selesai Pimpinan DPR Segera Bahas Usulan Pemakzulan Gibran Dasco Ungkap Mekanismenya #Gibran #PemakzulanGibran #DPR https://t.co/gchCF3vQQB lewat @tribunkaltim	Positif

3.2 Analisis Data

3.2.1 Preprocessing Data

Tahap ini merupakan proses preprocessing data, yaitu langkah penting untuk menyiapkan teks agar siap dianalisis dan diklasifikasikan. Proses ini meliputi pembersihan data dari elemen tidak relevan seperti tanda baca, angka, link, dan simbol; case folding untuk menyamakan huruf menjadi kecil; tokenizing untuk memisahkan teks menjadi kata-kata; normalisasi kata tidak baku; penghapusan stopword; dan stemming untuk mengembalikan kata ke bentuk dasarnya. Tahapan preprocessing ini krusial karena kualitas data yang dihasilkan akan mempengaruhi akurasi model dalam mengklasifikasikan sentimen.

Penelitian ini menggunakan 5 data latih dan 2 data uji yang dipilih secara acak dari dataset yang tersedia. Data latih digunakan untuk membangun dan melatih model agar mengenali pola sentimen, sedangkan data uji berfungsi untuk mengevaluasi kemampuan model dalam mengklasifikasikan data baru, sehingga hasil pengujian mencerminkan kondisi nyata.

Tabel 2. Sampel Data

sentimen	Data Latih
Positif	Wapres Gibran serukan pemimpin ke lapangan #beritanasional #wapres #gibran #pimpinan https://t.co/aH8Yl6BAbJ
Negatif	Bansos dipake judul? Gimana tuh? https://t.co/2bP3h7pHJ5 #gibranrakabumingraka #wapres #judul #judionline #bantuansosial #bansos

Negatif	Sarjana Banyak Nganggur Ketersediaan Lapangan Kerja Mengkhawatirkan https://t.co/DXhniVgMlF #LapanganPekerjaan #Sarjana #GibranLapanganPekerjaan #Wapres
Positif	wapres gibran tinjau pelaksanaan ckg di bondowoso jawa timur #beritanasional #wapres #gibran #jawatimur https://t.co/0xfTY4vQMY
Positif	Lapor Mas Wapres Tindak Lanjuti Ribuan Aduan Masyarakat #wapres #rabu #gibran #masyarakat #beritanasional https://t.co/BXkqU3ah3D
Data Uji	
?	Saatnya dua kekuatan besar politik bersatu! Akankah Puan dan Prabowo menciptakan perubahan besar bagi Indonesia? #puanmaharani #tetappuan #mbakpuan #prabowo #prabowosubianto #PDIPerjuangan #pdip #gerindra #jokowi #ketuadpri #dprri #presidenri #wapres #politikindonesia https://t.co/p9tkNHEeKkm
?	Gaji Pensiun Wapres #wapres #marufamin #tunjangan #wapres #marufamin #tunjangan #tunjanganwapres #mivotv #kamugaktahu https://t.co/jlsjf94j8K

Setelah pemilihan 5 data latih dan 2 data uji secara acak dari keseluruhan dataset, tahap berikutnya adalah melakukan preprocessing pada data terpilih. Proses ini bertujuan memastikan data berada dalam kondisi bersih, konsisten, dan siap untuk dianalisis secara komputasional. Tahapan preprocessing mencakup cleaning untuk menghapus elemen tidak relevan, case folding untuk menyamakan format huruf, tokenizing untuk memisahkan teks menjadi kata-kata, normalisasi untuk mengubah kata tidak baku menjadi bentuk standar, penghapusan stopword, serta stemming untuk mengembalikan kata ke bentuk dasarnya agar model dapat mengenali pola sentimen dengan lebih akurat. Hasil akhir preprocessing Dapat dilihat pada tabel berikut.

Tabel 3 Hasil Akhir Preprocessing Data

sentimen	Data Latih
Positif	['wakil', 'presiden', 'gibran', 'pimpin', 'lapang']
Negatif	['bantu', 'sosial', 'pakai', 'judi', 'online']
Negatif	['sarjana', 'nganggur', 'sedia', 'lapang', 'kerja', 'khawatir']
Positif	['wakil', 'presiden', 'gibran', 'tinjau', 'laksana', 'giat', 'bondowoso', 'jawa', 'timur']
Positif	['wakil', 'presiden', 'tindak', 'lanjut', 'adu', 'masyarakat']
Data Uji	
?	['kuat', 'politik', 'satu', 'puan', 'prabowo', 'cipta', 'ubah', 'indonesia']
?	['gaji', 'pensiun', 'wakil', 'presiden']

3.2.2 Pembobotan TF-IDF

Setelah tahap preprocessing selesai, langkah selanjutnya adalah menghitung bobot kata menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency). Proses ini bertujuan untuk menilai tingkat kepentingan setiap kata dalam dokumen relatif terhadap keseluruhan kumpulan data, sehingga fitur yang dihasilkan dapat digunakan secara efektif dalam pelatihan dan pengujian model untuk klasifikasi sentimen. Hasil TF-IDF dapat dilihat pada tabel berikut.

Tabel 4. Nilai TF-IDF

TERM	TF-IDF				
	D1	D2	D3	D4	D5
wakil	1,176	0	0	1,176	1,176
presiden	1,176	0	0	1,176	1,176
gibran	1,301	0	0	1,301	0
pimpin	1,477	0	0	0	0
lapang	1,301	0	1,301	0	0
...	...	...	...	...	...
timur	0	0	0	1,477	0
tindak	0	0	0	0	1,477
lanjut	0	0	0	0	1,477
adu	0	0	0	0	1,477
masyarakat	0	0	0	0	1,477

Langkah selanjutnya dalam representasi data adalah melakukan normalisasi terhadap nilai TF-IDF yang telah dihitung. Normalisasi bertujuan untuk menyamakan skala nilai TF-IDF antar dokumen sehingga setiap kata memiliki kontribusi yang seimbang, mencegah dominasi kata yang muncul dengan frekuensi tinggi. Dengan begitu, setiap vektor dokumen berada dalam skala yang seragam, yang sangat penting untuk menjaga akurasi klasifikasi, khususnya ketika menggunakan algoritma seperti Support Vector Machine (SVM) yang peka terhadap perbedaan skala data. Proses ini

memastikan bahwa model dapat mempelajari pola sentimen secara lebih adil dan konsisten. Hasil normalisasi data dapat dilihat pada tabel berikut.

Tabel 5. Normalisasi Data

No	D1	D2	D3	D4	D5
1	0,1507	0	0	0,1507	0,1507
2	0,1507	0	0	0,1507	0,1507
3	0,1667	0	0	0,1667	0
4	0,1893	0	0	0	0
5	0,1667	0	0,1667	0	0
...	...	...	...	...	...
20	0	0	0	0,1893	0
21	0	0	0	0,1893	0
22	0	0	0	0	0,1893
23	0	0	0	0	0,1893
24	0	0	0	0	0,1893
25	0	0	0	0	0,1893

3.3 Penerapan Algoritma Support Vector Machine

Setelah data melalui tahap pembersihan dan preprocessing, langkah berikutnya adalah klasifikasi menggunakan algoritma Support Vector Machine (SVM), yang dipilih karena kemampuannya memisahkan kelas dengan margin optimal. Dataset dibagi secara acak menjadi 80% data latih untuk membangun model dan 20% data uji untuk mengevaluasi performa. Proses SVM dimulai dengan perhitungan kernel linier untuk menilai kemiripan antar data, membentuk matriks Hessian untuk optimasi, dan dilanjutkan dengan sequential training untuk menemukan parameter terbaik dalam membentuk hyperplane yang memisahkan kelas. Tahapan ini memastikan model dapat mengklasifikasikan data uji dengan akurat.

3.3.1 Kernel Linier

Pada tahap ini, dilakukan perhitungan kernel linier pada data latih untuk mengukur kesamaan antar pasangan data melalui perkalian skalar vektor fitur. Hasil perhitungan ini digunakan untuk membentuk matriks Hessian sebagai dasar optimasi SVM, sehingga model dapat membangun pemisah antar kelas secara matematis dalam ruang fitur berdimensi tinggi. Hasil kernel linier dapat dilihat pada tabel berikut.

Tabel 6. Kernel Linier

No	1	2	3	...	23	24	25
1	0,06815	0,06815	0,05026	...	0,02853	0,02853	0,02853
2	0,06815	0,06815	0,05026	...	0,02853	0,02853	0,02853
3	0,05026	0,05026	0,05560	...	0	0	0
4	0,02853	0,02853	0,03156	...	0	0	0
5	0	0	0	...	0	0	0
...	...	...	...	...	...	...	...
20	0,02853	0,02853	0,03156	...	0	0	0
21	0,02853	0,02853	0,03156	...	0	0	0
22	0,02853	0,02853	0	...	0,03583	0,03583	0,03583
23	0,02853	0,02853	0	...	0,03583	0,03583	0,03583
24	0,02853	0,02853	0	...	0,03583	0,03583	0,03583
25	0,02853	0,02853	0	...	0,03583	0,03583	0,03583

3.3.2 Matriks Hessian

Setelah kernel linier dihitung, langkah selanjutnya adalah membentuk matriks Hessian, yang berisi nilai hubungan antar data latih dalam konteks margin maksimum. Matriks ini diperoleh dengan mengalikan kernel linier dengan label kelas, dan digunakan dalam tahap optimasi untuk menentukan hyperplane terbaik yang memisahkan data secara optimal. Hasil matrik hessian dapat dilihat pada tabel berikut.

Tabel 7. Matriks Hessian

	wakil	presiden	gibran	...	adu	masyarakat
wakil	0,31815	0,31815	0,30026	...	0,27853	0,27853
presiden	0,31815	0,31815	0,30026	...	0,27853	0,27853

gibran	0,30026	0,30026	0,30560	...	0,25	0,25
pimpin	0,27853	0,27853	0,28156	...	0,25	0,25
lapang	0,27513	0,27513	0,27780	...	0,25	0,25
...	...	...	...	...	...	...
timur	0,27853	0,27853	0,28156	...	0,25	0,25
tindak	0,27853	0,27853	0,25	...	0,28583	0,28583
lanjut	0,27853	0,27853	0,25	...	0,28583	0,28583
adu	0,27853	0,27853	0,25	...	0,28583	0,28583
masyarakat	0,27853	0,27853	0,25	...	0,28583	0,28583

3.3.3 Sequential Training SVM

Setelah matriks Hessian terbentuk, tahap berikutnya adalah sequential training, di mana SVM secara iteratif memperbarui parameter model berdasarkan hasil sebelumnya hingga mencapai konvergensi. Proses ini melibatkan pemilihan pasangan data yang melanggar kondisi optimalitas untuk memperbaiki parameter α secara bertahap. Tabel di bawah merupakan hasil dari nilai alpha.

Tabel 8. Nilai Alpha

No.	$\alpha 1$	$\alpha 2$	$\alpha 3$
1	0,1	0,132245	0,223284
2	0,1	0,132245	0,223284
3	0,1	0,133452	0,224571
4	0,1	0,135940	0,227231
5	0,1	0,137426	0,228827
6	0,1	0,135708	0,226983
7	0,1	0,135708	0,226983
8	0,1	0,135708	0,226983
9	0,1	0,135708	0,226983
10	0,1	0,135708	0,226983
11	0,1	0,136024	0,227322
12	0,1	0,136024	0,227322
13	0,1	0,136024	0,227322
14	0,1	0,136024	0,227322
15	0,1	0,136024	0,227322
16	0,1	0,134464	0,225652
17	0,1	0,134464	0,225652
18	0,1	0,134464	0,225652
19	0,1	0,134464	0,225652
20	0,1	0,134464	0,225652
21	0,1	0,134464	0,225652
22	0,1	0,135496	0,226756
23	0,1	0,135496	0,226756
24	0,1	0,135496	0,226756
25	0,1	0,135496	0,226756

Langkah berikutnya adalah menghitung fungsi kernel untuk masing-masing kelas menggunakan nilai α tertinggi dari tahap sebelumnya, yakni 0,228827 untuk kelas positif dan 0,227322 untuk kelas negatif. Perhitungan ini bertujuan menilai kesesuaian antara data latih yang paling berpengaruh dengan data lainnya dalam ruang fitur, sehingga mendukung pembentukan hyperplane dalam klasifikasi. Nilai hyper plane dapat dilihat pada tabel berikut.

Tabel 9. Nilai Kernel Pemisah

Kernel Kelas Positif	Kernel Kelas Negatif
0,352384	0,2312

Setelah nilai kernel untuk kelas positif dan negatif diperoleh, tahap berikutnya adalah menghitung bias, yang berperan penting dalam menentukan posisi akhir hyperplane pemisah dua kelas pada algoritma SVM.

$$b = -\frac{1}{2} [w \cdot x^+ + w \cdot x^-]$$

$$= -\frac{1}{2} [0,352384 + 0,2312]$$

$$= -\frac{1}{2}(0,584)$$

$$= -0,29179$$

3.4 Pengujian

Dalam analisis sentimen terhadap Wakil Presiden Gibran, hasil klasifikasi positif menunjukkan dukungan atau apresiasi terhadap kinerjanya, sedangkan hasil negatif menandakan kritik atau ketidakpuasan. Pendekatan ini memungkinkan sistem secara otomatis mengenali pola opini publik dari kata-kata dalam komentar.

1. Data uji 1

['kuat', 'politik', 'satu', 'puan', 'prabowo', 'cipta', 'ubah', 'indonesia']

$$f(x) = w.x + b$$

$$= \sum \alpha_i y_i K(x_i, x_j) + b$$

$$= (0,0349 + 0,0349 + 0,0349 + 0,0349 + 0,0349 + 0,0349 + 0,0349) + (-0,29179)$$

$$= 0,571330518$$

$$\text{Fungsi Klasifikasi} = \text{sign}(0,571330518)$$

$$= 1$$

2. Data uji 2

['gaji', 'pensiun', 'wakil', 'presiden']

$$f(x) = w.x + b$$

$$= \sum \alpha_i y_i K(x_i, x_j) + b$$

$$= (0,0349 + 0,0349 + 0,0231 + 0,0231) + (-0,29179)$$

$$= 0,40789927$$

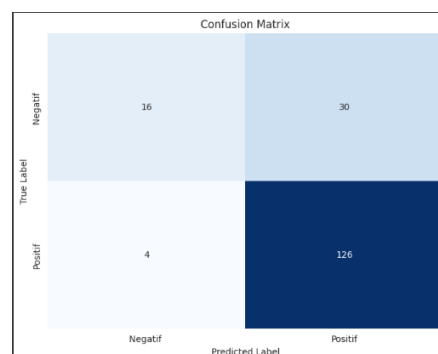
$$\text{Fungsi Klasifikasi} = \text{sign}(0,40789927)$$

$$= 1$$

Hasil pengujian terhadap dua data uji menunjukkan fungsi keputusan $f(x) = 1$, menandakan keduanya diklasifikasikan sebagai sentimen positif. Temuan ini menunjukkan bahwa model SVM mampu mengenali pola kata dalam teks yang menyampaikan dukungan atau apresiasi terhadap Wakil Presiden Gibran, serta efektif dalam mengidentifikasi persepsi masyarakat secara otomatis.

3.5 Evaluasi

Gambar berikut menunjukkan confusion matrix, tabel yang mengevaluasi kinerja model klasifikasi dengan membandingkan prediksi model dan label sebenarnya dalam analisis sentimen terhadap kinerja Wakil Presiden Gibran.



Gambar 3. Confusion Matrix

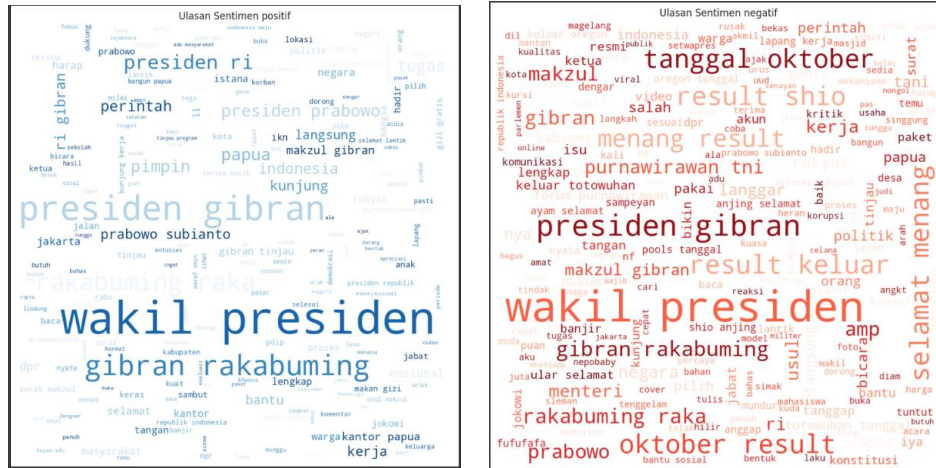
Gambar berikut menampilkan evaluasi kinerja model menggunakan metrik akurasi, presisi, recall, dan F1-score, yang mengukur ketepatan, kemampuan model dalam menemukan data relevan, sensitivitas, serta keseimbangan antara presisi dan recall dalam klasifikasi sentimen terhadap kinerja Wakil Presiden Gibran.

	precision	recall	f1-score	support
Negatif	0.80	0.35	0.48	46
Positif	0.81	0.97	0.88	130
accuracy			0.81	176
macro avg	0.80	0.66	0.68	176
weighted avg	0.81	0.81	0.78	176

Gambar 4. Akurasi Model

Gambar di atas menunjukkan laporan klasifikasi model untuk kelas Negatif dan Positif dengan akurasi keseluruhan 81%. Kelas Negatif memiliki presisi 0,80, recall 0,35, dan F1-score 0,48, menunjukkan kemampuan deteksi yang rendah, sementara kelas Positif memiliki presisi 0,81, recall 0,97, dan F1-score 0,88, menandakan performa baik. Nilai macro F1-score 0,66 dan weighted F1-score 0,78 mencerminkan kecenderungan model lebih akurat pada kelas dengan jumlah data lebih besar.

Gambar berikut menampilkan word cloud untuk sentimen positif dan negatif, yang memvisualisasikan kata-kata paling sering muncul dalam komentar publik. Visualisasi ini membantu mengidentifikasi istilah kunci yang mencerminkan dukungan maupun kritik terhadap kinerja Wakil Presiden Gibran.



Gambar 5. Word Cloud Sentimen Positif dan Negatif

4. KESIMPULAN

Penelitian ini menganalisis 898 tweet dari Twitter yang menggunakan tagar #wapres, #Gibran, dan #WapresGibran, sebagai representasi opini publik terhadap kinerja Wakil Presiden Indonesia pada tahun 2025. Dari jumlah tersebut, ditemukan 647 tweet mengandung sentimen positif dan 251 tweet negatif, yang menunjukkan mayoritas masyarakat cenderung memberikan penilaian positif, meskipun terdapat perbedaan pendapat. Data mentah kemudian melalui tahap preprocessing agar siap dianalisis menggunakan algoritma Support Vector Machine (SVM). Proses preprocessing mencakup cleaning untuk menghapus karakter tidak relevan, case folding untuk menyamakan huruf menjadi kecil, tokenizing untuk memisahkan kata per kata, normalisasi untuk mengubah kata tidak baku menjadi bentuk baku, stopword removal untuk menghapus kata umum yang tidak bermakna signifikan, dan stemming untuk mengembalikan kata ke bentuk dasarnya. Setelah data bersih, representasi fitur dilakukan menggunakan metode TF-IDF untuk menentukan bobot kata dalam dokumen. Tahap selanjutnya adalah klasifikasi menggunakan SVM, yang efektif untuk memisahkan kelas data teks berdimensi tinggi. Dataset dibagi menjadi data latih dan data uji untuk membangun model dan mengukur performanya. Hasil pengujian menunjukkan akurasi keseluruhan model sebesar 81%. Kelas Positif memiliki presisi 0,81, recall 0,97, dan F1-score 0,88, sementara kelas Negatif presisi 0,80, recall 0,35, dan F1-score 0,48. Macro average F1-score tercatat 0,66 dan weighted average 0,78, menunjukkan model lebih efektif dalam mendeteksi kelas dengan jumlah data lebih banyak. Secara keseluruhan, hasil penelitian menegaskan bahwa algoritma SVM mampu mengenali pola kata dalam teks dan mengklasifikasikan opini publik terhadap Wakil Presiden dengan baik, terutama pada sentimen positif, sambil tetap mencerminkan variasi opini masyarakat. Tahapan preprocessing dan representasi fitur menjadi faktor penting yang memengaruhi efektivitas model klasifikasi.

REFERENCES

- [1] M. Farhan and U. N. Huda, "Jurnal Hukum Legalita Analisis Respon Negatif Publik terhadap Putusan MK Nomor," vol. 6, 2024.
- [2] Dinda Aurellia and K. Katimin, "Arah Baru Politik di Era Digital (Perspektif Generasi Milenial di Media Sosial)," *SOSMANIORA J. Ilmu Sos. dan Hum.*, vol. 4, no. 2, pp. 231–236, 2025, doi: 10.55123/sosmaniora.v4i2.5149.
- [3] N. Aykar and Muhammad Arfan Ahwady, "Batas Usia Capres-Cawapres Dalam Keputusan MK No. 90/PUU-XXI/2023 Dan Implikasinya Terhadap Dinasti Politik Perspektif Fikih," *Syariah J. Fiqh Stud.*, vol. 2, no. 2, pp. 21–46, 2024, doi: 10.61570/syariah.v2i2.79.
- [4] S. Suhendra and F. Selly Pratiwi, "Peran Komunikasi Digital dalam Pembentukan Opini Publik: Studi Kasus Media Sosial," *Iapa Proc. Conf.*, p. 293, 2024, doi: 10.30589/proceedings.2024.1059.
- [5] Junaedi, A. Hendra Gunawan, V. Kuswanto, and Jonathan, "Tinjauan Support Vector Machine dalam Text-Mining untuk Analisis Sentimen di Sektor Pariwisata," *bit-Tech*, vol. 7, no. 2, pp. 323–330, 2024, doi: 10.32877/bt.v7i2.1810.
- [6] V. Agustina and A. Herliana, "Analisis Sentimen Publik atas Kebijakan Efisiensi Anggaran 2025 dengan Text Mining dan Natural Language Processing," *J. Media Inform.*, vol. 6, no. 3, pp. 2182–2194, 2025, [Online]. Available: <https://ejournal.sisfokomtek.org/index.php/jumin/article/view/6301>
- [7] J. K. Nainggolan, F. Sinaga, A. M. Sitorus, A. Khairia, and B. A. Wijaya, "Feature Selection using / for T ransductive S upport

- Vector Machine,” *Science (80-.)*, vol. 4, no. 1, pp. 62–70, 2015.
- [8] R. Muhammad Nusantara, “Analisis Sentimen Masyarakat terhadap Pelayanan Bank Central Asia: Text Mining Cuitan Satpam BCA pada Twitter,” *Co-Value J. Ekon. Kop. dan kewirausahaan*, vol. 15, no. 9, 2025, doi: 10.59188/covalue.v15i9.5099.
 - [9] A. K. Ginting, D. Purba, L. M. Sinaga, and M. Sagala, “Analisis Sentimen Masyarakat Twitter terhadap Emas Digital Menggunakan Algoritma Naïve Bayes,” *KAKIFIKOM (Kumpulan Artik. Karya Ilm. Fak. Ilmu Komput.)*, vol. 07, no. 01, pp. 8–14, 2025.
 - [10] A. Fauzi and A. H. Yunial, “Analisis Sentimen Pada Media Sosial Menggunakan Perbandingan Algoritma Data Mining,” *J. Edukasi dan Penelit. Inform.*, vol. 10, no. 2, p. 277, 2024, doi: 10.26418/jp.v10i2.76024.
 - [11] H. Y. Pradana, I. Slamet, and E. Zuhronah, “Analisis Sentimen Kinerja Pemerintahan Menggunakan Algoritma Nbc, Knn, Dan Svm,” *Pros. Simp. Nas. Multidisiplin*, vol. 4, p. 114, 2023, doi: 10.31000/sinamu.v4i1.7869.
 - [12] L. N. Haq and A. Mulyani, “Analisis Sentimen Aplikasi Jakone Mobile Pada Google Play Store Menggunakan Metode Support Vector Machine,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 1, pp. 1–9, 2025, [Online]. Available: https://repository.nusamandiri.ac.id/repo/files/254376/download/file_10.pdf
 - [13] Danis Rifa Nurqotimah, A. Naseh Khudori, and R. Siwi Pradini, “Implementasi Algoritma Support Vector Machine (SVM) Untuk Klasifikasi Penyakit Stroke,” *J. Appl. Comput. Sci. Technol.*, vol. 5, no. 2, p. press, 2024, doi: 10.52158/jacost.v5i2.817.
 - [14] M. N. Ichsan, M. Ayu, D. Widyadara, and U. Mahdiyah, “Pemanfaatan Support Vector Machine dalam Mendeteksi Biji Kopi,” vol. 9, pp. 1024–1033.
 - [15] J. E. Br Sinulingga and H. C. K. Sitorus, “Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF,” *J. Manaj. Inform.*, vol. 14, no. 1, pp. 42–53, 2024, doi: 10.34010/jamika.v14i1.11946.
 - [16] I. S. K. Idris, Y. A. Mustofa, and I. A. Salihi, “Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM),” *Jambura J. Electr. Electron. Eng.*, vol. 5, no. 1, pp. 32–35, 2023, doi: 10.37905/jjee.v5i1.16830.
 - [17] Z. Purwanti and Sugiyono, “Pemodelan Text Mining untuk Analisis Sentimen Terhadap Program Makan Siang Gratis di Media Sosial X Menggunakan Algoritma Support Vector Machine (SVM),” *J. Indones. Manaj. Inform. dan Komun.*, vol. 5, no. 3, pp. 3065–3079, 2024, doi: 10.35870/jimik.v5i3.1001.
 - [18] I. Zufria, A. H. Lubis, S. S. Febiyaula, U. Islam, and N. Sumatera, “Kepolisian Republik Indonesia Menggunakan,” vol. 4307, no. August, pp. 1266–1272, 2024.
 - [19] M. Ridwan, S. AM, B. Ulum, and F. Muhammad, “Pentingnya Penerapan Literature Review pada Penelitian Ilmiah (The Importance Of Application Of Literature Review In Scientific Research),” *J. Masohi*, vol. 2, no. 1, pp. 42–51, 2021, [Online]. Available: <http://journal.fdi.or.id/index.php/jmas/article/view/356>
 - [20] J. Budiarto, “Identifikasi Kebutuhan Masyarakat Nusa Tenggara Barat pada Pandemi Covid-19 di Media Sosial dengan Metode Crawling,” *JTIM J. Teknol. Inf. dan Multimed.*, vol. 2, no. 4, pp. 244–250, 2021, doi: 10.35746/jtim.v2i4.119.
 - [21] R. Rahmadani, A. Rahim, and R. Rudiman, “Analisis Sentimen Ulasan ‘Ojol the Game’ Di Google Play Store Menggunakan Algoritma Naive Bayes Dan Model Ekstraksi Fitur Tf-Idf Untuk Meningkatkan Kualitas Game,” *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4988.