

**ANALISIS SENTIMEN PRESIDEN TERPILIH PRABOWO-GIBRAN
MENGUNAKAN METODE SVM DI MEDIA TWITTER/X**

SKRIPSI



SYARIFUL ABRORI

2021503054

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMY
SITUBONDO**

2025

HALAMAN JUDUL**ANALISIS SENTIMEN PRESIDEN TERPILIH PRABOWO-GIBRAN
MENGUNAKAN METODE SVM DI MEDIA TWITTER/X****SKRIPSI**

Diajukan untuk Memenuhi Salah Satu Persyaratan dalam Menyelesaikan Program Sarjana (S-1) pada Program Studi Teknologi Informasi Fakultas Sains dan

Teknologi Universitas Ibrahimi

Oleh:

SYARIFUL ABRORI

2021503054

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMI
SITUBONDO**

2025

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : **Syariful Abrori**

NPM : 2021503054

Program Studi : Teknologi Informasi S-1

Fakultas : Sains Dan Teknologi

Menyatakan dengan sebenarnya, bahwa tugas akhir/skripsi ini secara keseluruhan adalah hasil penelitian atau karya sendiri, kecuali pada bagian-bagian yang dirujuk sebagai sumber referensi dan disebutkan dalam daftar pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa tugas akhir/skripsi ini hasil plagiasi, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Situbondo, 23 juni 2025

Menyatakan,



Syariful Abrori

PERSETUJUAN PEMBIMBING

Nama : Syariful Abrori
NPM : 2021503054
Judul : Analisis Sentimen Presiden Terpilih Prabowo-Gibran
Menggunakan Metode Svm Di Media Twitter/X



Telah disetujui oleh:

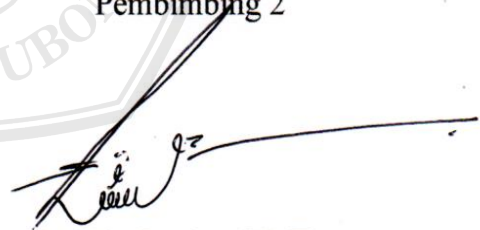
Pembimbing 1



Firman Santoso, M. Kom.

NIDN: 0722129201

Pembimbing 2



Farihin Lazim, M. Tr. t.

NIDN: 0711099201

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN PRESIDEN TERPILIH PRABOWO-GIBRAN
MENGUNAKAN METODE SVM DI MEDIA TWITTER/X**

SYARIFUL ABRORI

2021503054

Telah dipertahankan di depan dewan penguji Sidang/Munaqasah skripsi pada hari Kamis, Tanggal 24 juli 2025. Sebagai salah satu syarat memperoleh gelar Sarjana (S.kom) pada Fakultas Sains dan Teknologi Universitas Ibrahimi.

Tim Penguji,

Ketua Sidang,

Sekretaris Sidang,

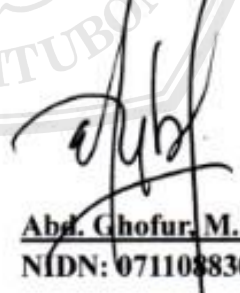



Firman Santoso, M. Kom.
NIDN: 0722129201

Fajro Sabut Panotogo, A.Md. Pi
NIDN: -

Penguji I,

Penguji II,

Zachol Fatah, M. Kom.
NIDN: 0715057801

Abd. Ghofur, M. Kom.
NIDN: 0711088303

Mengetahui
Dekan,



Abd. Ghofur, M. Kom.
NIDN: 0711088303

MOTTO

“NIAT KAYA UNTUK BERDAKWAH LEWAT HARTA”



PERSEMBAHAN

Semoga skripsi ini dapat terselesaikan dengan sukses dan tepat waktu, setelah kita panjatkan segala shalawat kepada Allah SWT dan kepada amal serta doa orang-orang disekitar kita. Oleh karena itu, saya ingin mengucapkan terima kasih kepada Anda dan saudara laki-laki saya dengan perasaan berikut:

1. Allah SWT. Karena terbatas pada mata dan telinga, skripsi ini mungkin selesai dan selesai dalam jangka waktu tertentu.
2. Ayah dan Ibu tercinta terima kasih atas dukungan dan pengorbanannya," doa serta "kasih sayangnya" tak akan pernah terlupakan.
3. Kepada seluruh anggota dan pendukung, saya ingin mengucapkan terima kasih dan penghargaan atas dukungan Anda yang tak tergoyahkan selama perjalanan ini.
4. Sahabat-sahabat yang selalu ada dalam keadaan suka maupun duka.
5. Ucapkan terima kasih kepada sang "kekasih" yang sebesar-besarnya atas dukungan yang tidak pernah berkurang dan dukungannya yang tulus dalam skripsi ini.
6. Semua orang yang tidak dapat saya ajak bicara tentang satu hal pun, Namanya, yang perlahan-lahan menjadi lebih dekat dengan saya sehingga saya dapat mengedit naskah dengan baik.

KATA PENGANTAR

Segala Puji Syukur Peneliti Sampaikan Kepada Allah SWT, Karena Atas Rahmat Dan Hidayah-Nya, Perencanaan, Pelaksanaan Dan Penyelesaian Tugas Akhir/Skripsi Dengan Judul **“ANALISIS SENTIMEN PRESIDEN TERPILIH PRABOWO-GIBRAN MENGGUNAKAN METODE SVM DI MEDIA TWITTER/X”** Sebagai Salah Satu Syarat Penyelesaian Program Sarjana Dapat terselesaikan Dengan Baik Dan Lancar. Kesuksesan Ini Dapat Peneliti Peroleh Karena Dukungan Beberapa Pihak. Peneliti Menyampaikan Terima Kasih Kepada:

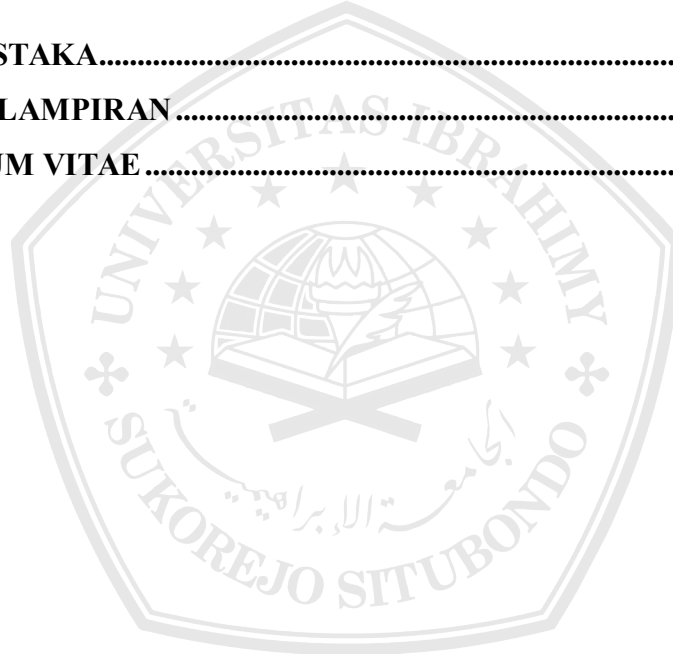
1. KHR. Ach. Azaim Ibrahimi Selaku Pengasuh Pondok Pesantren Salafiyah Syafi'iah Sukorejo Situbondo.
2. KH. Ach. Fadlail, S.H, M.H Selaku Rektor Universitas Ibrahimi Sukorejo Situbondo.
3. Abd. Ghofur, M. Kom Selaku Dekan Fakultas Sains Dan Teknologi Universitas Ibrahimi.
4. Firman Santoso, M. Kom. Selaku Ketua Program Studi Teknologi Informasi.
5. Firman Santoso, M. Kom. Dan Farihin Lazim , M Tr, T. Selaku Dosen Pembimbing I Dan II.
6. Orang Tuaku Bapak Abdul Ghafir Dan Ibu Makkawiyah Serta Keluarga Besar Serta Para Guru-Guru Yang Selalu Mendoakan.
7. Temen – Teman Asrama Sunan Drajat No. 18 Yang Menemani Kami Dalam Suka Maupun Duka.
8. Seluruh Dosen Fakultas Sains Dan Teknologi Yang Telah Memberikan Kami Ilmu Sehingga Sampai Pada Masa Tugas Akhir (Skripsi) Saat Ini. Semoga Semua Amal Baik Yang Telah Diberikan Oleh Bapak/Ibu Kepada Peneliti Mendapat Balasan Yang Sebaik Mungkin Dari Allah SWT, Aamiin.

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN KEASLIAN TULISAN.....	ii
PERSETUJUAN PEMBIMBING.....	iii
HALAMAN PENGESAHAN	iv
MOTTO	v
PERSEMBAHAN	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL	xiii
DAFTAR SEGMENT	xiv
DAFTAR LAMPIRAN	xv
ABSTRAK	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	4
1.4 Batasan Masalah	4
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian.....	5
1.7 Metode Penelitian.....	6
1.7.1 Jenis Penelitian	6
1.7.2 Teknik Pengumpulan Data	6
1.7.3 Metode Pengembangan Sistem.....	7
BAB II TINJAUAN PUSTAKA	12
2.1 Penelitian Terdahulu.....	12
2.2 Landasan Teori	17

2.2.1 Prabowo-Gibran Presiden	17
2.2.2 Analisis Sentimen	18
2.2.3 Twitter (X)	19
2.2.4 Data Mining.....	20
2.2.5 Support Vector Machine.....	21
2.2.6 Python Notebook	23
BAB III ANALISIS DAN PERANCANGAN SISTEM	26
3.1 Analisis Masalah	26
3.2 Alur Proses.....	27
3.2.1 Pengumpulan Data.....	27
3.2.2 Preprocessing	29
3.2.3 Transformation	31
3.2.4 Data Mining.....	32
3.3 Perangkat yang Digunakan	34
3.3.1 Perangkat Keras	34
3.3.2 Perangkat Lunak.....	35
BAB IV PEMBAHASAN	36
4.1 Program Kerja Presiden	36
4.2 Pengumpulan Data <i>Twitter/X</i>	39
4.2.1 Persiapan Scraping	39
4.2.2 Kata Kunci dan Rentang Waktu.....	42
4.2.3 Pengumpulan Data <i>Twitter/X</i>	43
4.3 Preprocessing Data.....	46
4.3.1 Cleaning Text.....	46
4.3.2 Tokenisasi dan Stemming	47
4.3.3 Penghapusan Duplikat Data	49
4.3.4 Labelling Sentimen.....	50
4.3.5 TF-IDF Vectorization	53
4.4 Implementasi SVM	54
4.4.1 Splitting Data	54

4.4.2 Klasifikasi Support Vector Machine	56
4.4.3 Evaluasi Model.....	58
4.5 Pembahasan Hasil	60
4.5.1 Interpretasi Sentimen	60
4.5.2 Evaluasi Kinerja Model.....	63
4.5.3 Pembahasan Hasil	66
BAB V PENUTUP	68
5.1 Kesimpulan.....	68
5.2 Saran	69
DAFTAR PUSTAKA.....	71
LAMPIRAN-LAMPIRAN.....	77
CURRICULUM VITAE	88



DAFTAR GAMBAR

Gambar 1. 1 Metode Knowledge Discovery database	9
Gambar 2. 1 Hasil Rekapitulasi KPU Pilpres 2024	17
Gambar 2. 2 Contoh Aplikasi X	18
Gambar 2. 3 Logo Aplikasi X.....	19
Gambar 2. 4 Gambaran Proses Data Mining	20
Gambar 2. 5 Ilustrsi Algoritma SVM.....	21
Gambar 2. 6 Jupyter Notebook.....	23
Gambar 2. 7 Perhitungan Confusion Matrix.....	25
Gambar 3. 1 Tahapan Pengumpulan Data.....	27
Gambar 3. 2 Tahapan Preprocessing Data	31
Gambar 3. 3 Diagram Alur Analisis Sentimen	34
Gambar 4. 1 Presiden Prabowo dan Gibran.....	36
Gambar 4. 2 Siswa menikmati makan siang gratis	37
Gambar 4. 3 Python.....	40
Gambar 4. 4 Selenium	41
Gambar 4. 5 Code Scrapping Data Twitter	42
Gambar 4. 6 Chrome Mode Debugging.....	42
Gambar 4. 7 Masukkan Kata Kunci Dan Waktu	43
Gambar 4. 8 Nama file yang tersimpan	43
Gambar 4. 9 File CSV	44
Gambar 4. 10 Code.....	45
Gambar 4. 11 Merged Data.....	45
Gambar 4. 12 Head Data Cleanning Text	47
Gambar 4. 13 Hasil Tokenisasi & Stemming	48
Gambar 4. 14 Hasil Penghapusan duplikat data	49
Gambar 4. 15 Labelling Sentimen.....	50
Gambar 4. 16 Diagram Batang Distribusi Sentimen.....	52
Gambar 4. 17 Diagram Distribusi Sentimen.....	53
Gambar 4. 18 TF-IDF vectorization.....	54
Gambar 4. 19 Hasil Splitting Data	55

Gambar 4. 20 Implementasi Algoritma SVM.....	57
Gambar 4. 21 Output Implementasi SVM.....	58
Gambar 4. 22 Akurasi Model	59
Gambar 4. 23 Confusion Matrix.....	60



DAFTAR TABEL

Tabel 1 Hasil Prediksi Model	66
------------------------------------	----



DAFTAR SEGMENT

Segment 1 Chrome mode Debugging	41
---------------------------------------	----



DAFTAR LAMPIRAN

Lampiran 1 Surat Penelitian	77
Lampiran 2 Kartu bimbingan	78
Lampiran 3 Dataset Twitter/X	80
Lampiran 4 Program Scrapping.....	83
Lampiran 5 Program Algoritma.....	85
Lampiran 6 Letter of Acceptance Journal	86
Lampiran 7 Lembar Keaslian Tulisan	87



ABSTRAK

Syariful Abrori.2025 **Analisis Sentimen Presiden Terpilih Prabowo-Gibran Menggunakan Metode Svm Di Media Twitter/X**. Skripsi, Program Studi Teknologi Informasi, Universitas Ibrahimi. Pembimbing: (1) Firman Santoso, M. Kom, (II) Farihin Lazim, M.Tr.T.

Periode 100 hari pertama pemerintahan menjadi momentum krusial untuk menganalisis persepsi masyarakat terhadap kepemimpinan baru. Penelitian ini bertujuan mengimplementasikan Support Vector Machine (SVM) untuk menganalisis sentimen masyarakat Indonesia terhadap 100 hari pertama pemerintahan Prabowo-Gibran berdasarkan data Twitter/X. Metodologi penelitian menggunakan framework Knowledge Discovery in Databases (KDD) yang mencakup pengumpulan data, preprocessing, transformasi, data mining, dan evaluasi. Data dikumpulkan melalui web scraping menggunakan Selenium dari periode 20 Oktober 2024 hingga 28 Januari 2025 berdasarkan enam kata kunci relevan, menghasilkan 5.136 tweet yang kemudian diproses menjadi 4.631 data unik setelah penghapusan duplikat. Preprocessing meliputi text cleaning, tokenisasi, stemming menggunakan Sastrawi, dan pelabelan sentimen dengan pendekatan lexicon-based.

Transformasi data menggunakan TF-IDF dengan pembagian 80:20 untuk training dan testing. Model SVM dengan kernel linear berhasil mencapai akurasi 85% dengan performa terbaik pada klasifikasi sentimen negatif (precision 0.86, recall 0.95, F1-score 0.90) dan sentimen positif (precision 0.83, recall 0.75, F1-score 0.79). Hasil analisis menunjukkan dominasi sentimen negatif yang mencerminkan sikap kritis masyarakat terhadap implementasi kebijakan awal. Penelitian ini memberikan kontribusi dalam pengembangan sistem analisis sentimen berbahasa Indonesia untuk monitoring opini publik terhadap pemerintahan.

Kata Kunci: Analisis Sentimen, Support Vector Machine, Twitter/X, Prabowo-Gibran, Media Sosial

ABSTRACT

Syariful Abrori.2025 **Analisis Sentimen Presiden Terpilih Prabowo-Gibran Menggunakan Metode Svm Di Media Twitter/X**. Skripsi, Program Studi Teknologi Informasi, Universitas Ibrahimy. Pembimbing: (1) Firman Santoso, M. Kom, (II) Farihin Lazim, M.Tr.T.

The first 100 days of government represents a crucial momentum for analyzing public perception toward new leadership. This study aims to implement Support Vector Machine (SVM) to analyze Indonesian public sentiment toward the first 100 days of Prabowo-Gibran administration based on Twitter/X data. The research methodology employs Knowledge Discovery in Databases (KDD) framework encompassing data collection, preprocessing, transformation, data mining, and evaluation. Data was collected through web scraping using Selenium from October 20, 2024, to January 28, 2025, based on six relevant keywords, yielding 5,136 tweets which were processed into 4,631 unique data after duplicate removal. Preprocessing includes text cleaning, tokenization, stemming using Sastrawi, and sentiment labeling with lexicon-based approach.

Data transformation utilized TF-IDF with 80:20 split for training and testing. The SVM model with linear kernel achieved 85% accuracy with best performance on negative sentiment classification (precision 0.86, recall 0.95, F1-score 0.90) and positive sentiment (precision 0.83, recall 0.75, F1-score 0.79). Analysis results show negative sentiment dominance reflecting critical public attitudes toward early policy implementation. This research contributes to developing Indonesian sentiment analysis systems for monitoring public opinion toward government administration.

Keywords: Sentiment Analysis, Support Vector Machine, Twitter/X, Prabowo-Gibran, Social Media

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi diharapkan dapat membantu memberikan kemudahan dan efisiensi dalam kehidupan manusia di zaman ini. Dunia selalu berkembang seiring munculnya permasalahan baru yang tidak pernah terduga oleh generasi sebelumnya[1]. Salah satu perkembangan yang signifikan adalah penggunaan media sosial sebagai sarana komunikasi dan wadah diskusi publik. Oleh karena itu, manusia harus selalu beradaptasi dan berinovasi untuk dapat bertahan hidup. Jejaring sosial seperti *Twitter/X* sekarang menjadi salah satu perangkat komunikasi yang sangat populer di kalangan pengguna dunia maya dan dapat digunakan sebagai media untuk menyampaikan pendapat dan opini publik[2].

Pemilihan *Twitter/X* sebagai fokus penelitian ini didasarkan pada beberapa alasan yang membuatnya lebih unggul dibandingkan platform media sosial lainnya, seperti *TikTok*, *YouTube*, atau *Instagram*[3]. Pertama, *Twitter/X* memiliki karakteristik unik dengan fokus pada teks singkat (*tweet*) yang dibatasi jumlah karakternya. Hal ini memudahkan analisis data teks, terutama ketika menggunakan metode seperti *Support Vector Machine (SVM)*, karena data yang diolah lebih terstruktur dan fokus. Kedua, *Twitter/X* dikenal sebagai platform yang dinamis dan *real-time*. Pengguna dapat dengan cepat merespons isu-isu terkini, sehingga data yang dihasilkan sangat relevan untuk penelitian yang berkaitan dengan opini publik. *Support Vector Machine (SVM)* telah terbukti menjadi salah satu metode yang paling efektif dalam analisis sentimen, khususnya untuk data teks berbahasa

Indonesia. Metode ini memiliki keunggulan dalam kemampuannya untuk menangani data dengan dimensi tinggi dan menghasilkan model klasifikasi yang optimal melalui pencarian *hyperplane* terbaik yang memisahkan kelas-kelas sentimen[4]. SVM telah menunjukkan performa yang konsisten dengan tingkat akurasi mencapai 85-90% pada berbagai studi sebelumnya. Keunggulan lain dari SVM adalah kemampuannya dalam menangani karakteristik unik bahasa Indonesia, seperti variasi dialek dan penggunaan bahasa informal yang umum ditemukan di media sosial[5]. SVM dipilih karena kemampuannya dalam mengklasifikasikan data dengan akurasi tinggi, bahkan ketika data memiliki dimensi yang kompleks. Hyperplane yang dihasilkan oleh SVM memastikan pemisahan optimal antara kelas-kelas sentimen, seperti positif, negatif, atau netral. Selain itu, fleksibilitas SVM dalam menangani bahasa informal dan variasi dialek membuatnya cocok untuk analisis teks dari platform seperti Twitter/X, di mana bahasa yang digunakan seringkali tidak formal dan beragam.

Untuk memahami bagaimana sentimen masyarakat terhadap presiden dan wakil presiden terpilih, diperlukan sebuah analisis yang mendalam terhadap opini publik di media sosial. Hal ini menjadi krusial mengingat dinamika politik pasca pemilihan presiden 2024 yang menimbulkan berbagai reaksi dan polarisasi di masyarakat. Munculnya berbagai narasi dan opini yang berkembang di media sosial dapat mempengaruhi stabilitas sosial dan politik, serta berpotensi menimbulkan kesalahpahaman atau konflik di masyarakat jika tidak dipahami dengan baik. *Twitter/X* menjadi sumber data yang ideal karena sifatnya yang *real-time* dan memungkinkan pengguna untuk mengekspresikan pendapat mereka secara

bebas[6]. Menurut laporan yang dirilis oleh Wearesocial pada Oktober 2023, pengguna internet di Indonesia mencapai 212,9 juta dan pengguna sosial media aktif mencapai 167 juta. Indonesia merupakan 5 negara terbesar pengguna Twitter/X di dunia di bawah negara Brazil dengan angka 24 juta pengguna[7]. Besarnya jumlah pengguna ini menimbulkan tantangan tersendiri dalam menganalisis dan memahami sentimen publik secara akurat, mengingat volume data yang besar dan kompleksitas bahasa yang digunakan di media sosial. Tanpa adanya metode analisis yang sistematis dan terukur, akan sulit bagi pemangku kepentingan untuk memahami aspirasi dan kekhawatiran masyarakat secara komprehensif, yang pada akhirnya dapat mempengaruhi efektivitas pengambilan kebijakan dan komunikasi publik[8].

Berdasarkan hal tersebut, penelitian ini mengusulkan penggunaan SVM untuk menganalisis sentimen masyarakat Indonesia terhadap Prabowo-Gibran di *Twitter/X*. Dengan mengumpulkan *tweet* dalam periode Februari hingga Oktober 2024, penelitian ini bertujuan untuk memberikan gambaran yang komprehensif tentang sentimen publik pasca terpilihnya presiden dan wakil presiden baru. Data yang dikumpulkan akan melalui proses *preprocessing* yang meliputi pembersihan teks, normalisasi, dan *tokenisasi* sebelum dianalisis menggunakan model SVM[9]. Hasil dari penelitian ini diharapkan dapat memberikan wawasan berharga tentang persepsi masyarakat terhadap presiden dan wakil presiden terpilih, sekaligus memvalidasi efektivitas metode SVM dalam analisis sentimen berbahasa Indonesia. Penelitian ini juga dapat menjadi referensi untuk pengembangan sistem

analisis sentimen yang lebih baik di masa depan, khususnya untuk konteks politik dan sosial di Indonesia.

1.2 Identifikasi Masalah

Berdasarkan latar belakang diatas, maka ditemukan beberapa masalah dalam penelitian ini adalah:

1. Besarnya volume dan kompleksitas data *tweet* terkait presiden dan wakil presiden terpilih memerlukan metode analisis yang sistematis dan terukur.
2. Perlunya pengembangan model klasifikasi sentimen yang akurat untuk menganalisis opini publik berbahasa Indonesia di *Twitter/X*.
3. Tantangan dalam menganalisis variasi bahasa dan konteks dalam *tweet* berbahasa Indonesia yang sering menggunakan bahasa informal dan campuran.

1.3 Rumusan Masalah

Berdasarkan identifikasi masalah tersebut, maka dapat dirumuskan masalah sebagai berikut:

1. Bagaimana cara mengolah dan memproses data *tweet* berbahasa Indonesia untuk analisis sentimen menggunakan metode SVM?
2. Bagaimana mengimplementasikan metode SVM untuk mengklasifikasikan sentimen masyarakat terhadap presiden dan wakil presiden terpilih?
3. Bagaimana tingkat akurasi dan performa model SVM dalam mengklasifikasikan sentimen dari data *Twitter/X*?

1.4 Batasan Masalah

Agar penelitian lebih terarah, maka ditentukan batasan masalah sebagai berikut:

1. Dataset yang digunakan adalah *tweet* berbahasa Indonesia dengan periode 20 Oktober 2024 – 28 Januari 2025 kerja 100 hari
2. Analisis sentimen difokuskan pada opini masyarakat terhadap presiden dan wakil presiden terpilih Prabowo-Gibran.
3. Klasifikasi menggunakan metode SVM dengan tiga kategori sentimen: positif, negatif, dan netral.

1.5 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka tujuan dari penelitian ini adalah:

1. Merancang dan mengimplementasikan *preprocessing* data untuk mengolah *tweet* berbahasa Indonesia.
2. Mengembangkan model klasifikasi sentimen menggunakan metode SVM untuk analisis opini publik.
3. Mengukur dan mengevaluasi performa model SVM dalam klasifikasi sentimen dengan metrik akurasi, presisi, *recall*, dan F1-score.

1.6 Manfaat Penelitian

Manfaat dari penelitian ini yaitu:

a. Bagi Peneliti

Penelitian ini memberikan pengalaman yang berharga bagi peneliti untuk menerapkan secara langsung ilmu yang diperoleh selama menempuh pendidikan di perguruan tinggi khususnya dalam bidang analisis sentimen dan *machine learning*.

b. Bagi Akademik

Penelitian ini dapat dijadikan sebagai referensi untuk pengembangan sistem analisis sentimen berbahasa Indonesia dan penerapan metode SVM dalam konteks media sosial.

c. Bagi Masyarakat

Memberikan gambaran objektif tentang sentimen publik terhadap presiden dan wakil presiden terpilih berdasarkan analisis data media sosial.

1.7 Metode Penelitian

1.7.1 Jenis Penelitian

Metode yang digunakan adalah text mining berbasis Natural Language Processing (NLP) dengan algoritma Support Vector Machine (SVM) untuk mengklasifikasikan sentimen tweets ke dalam kategori positif, negatif, atau netral. Penelitian ini akan menghitung distribusi sentimen dan mengevaluasi performa model menggunakan metrik seperti akurasi, presisi, recall, dan F1-score guna memastikan keakuratan klasifikasi yang dilakukan. Dengan pendekatan ini, penelitian bertujuan untuk mengidentifikasi pola dan kecenderungan sentimen masyarakat secara objektif berdasarkan data yang diperoleh[10].

1.7.2 Teknik Pengumpulan Data

Tahapan pada pengumpulan data ini adalah sebagai berikut:

a. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan menggunakan metode web scraping dari platform Twitter/X, bukan melalui API

resmi. Teknik ini memanfaatkan library *Selenium* dan *ChromeDriver* yang secara otomatis mengakses laman pencarian Twitter/X dan mengambil data berupa tweet-tweet publik. Pengumpulan data difokuskan pada tweet berbahasa Indonesia yang membahas topik seputar Prabowo-Gibran dalam periode Oktober 2024 hingga 28 Januari 2025, yang merepresentasikan 100 hari kerja pertama presiden dan wakil presiden terpilih.

b. Scrapping Twitter

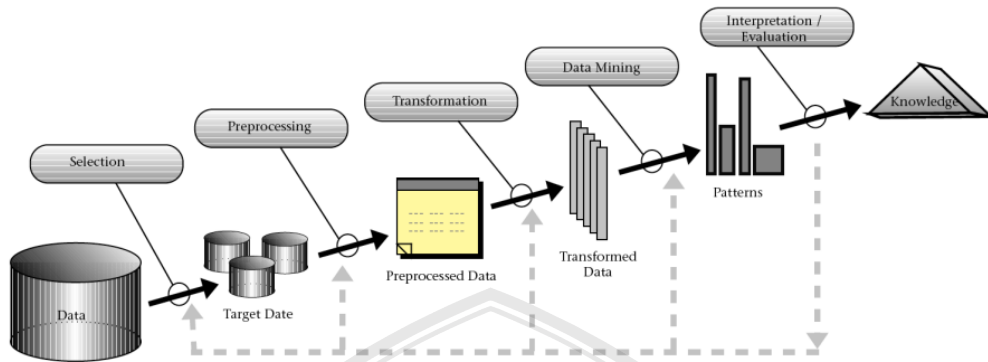
Scraping dilakukan dengan mengakses URL pencarian Twitter/X yang difilter menggunakan kata kunci tertentu seperti "100 hari presiden", "program kerja presiden", "makan siang gratis", dan lainnya. Data yang diambil terdiri dari tanggal posting, nama pengguna, dan isi tweet. File hasil scraping kemudian disimpan dalam format Excel (CSV/XLSX) dan digunakan sebagai bahan mentah untuk proses preprocessing dan analisis selanjutnya.

1.7.3 Metode Pengembangan Sistem

Penelitian ini menggunakan metodologi *Knowledge Discovery in Databases* (KDD) untuk menganalisis sentimen masyarakat terhadap presiden dan wakil presiden terpilih Prabowo-Gibran di media sosial *Twitter/X*. KDD dipilih karena menyediakan kerangka kerja yang sistematis dan terstruktur dalam proses penemuan pengetahuan dari data yang kompleks[11]. Metodologi ini sangat relevan untuk penelitian analisis sentimen di media sosial karena beberapa alasan kritis[12]:

Pertama, KDD memiliki tahapan seleksi data yang kritis sebagai fondasi awal proses. Tahap ini memungkinkan identifikasi dan pemfilteran data relevan dari platform Twitter/X (seperti tweet, hashtag, atau thread) sesuai tujuan analisis, sekaligus mengeliminasi data redundan atau tidak relevan sebelum masuk ke tahap preprocessing. Kedua, tahap preprocessing data yang komprehensif menjadi keunggulan KDD dalam menangani karakteristik unik data Twitter/X seperti noise, bahasa informal, slang, emoji, dan struktur teks tidak teratur. Seleksi data di awal meminimalisir beban komputasi pada tahap ini, karena hanya data terpilih yang menjalani proses pembersihan (cleansing), normalisasi, dan transformasi. Ketiga, KDD menekankan proses iteratif dan evaluasi berkelanjutan, memungkinkan penyempurnaan model secara bertahap (misalnya tuning parameter SVM) untuk mencapai akurasi klasifikasi optimal. Keempat, kerangka terstruktur KDD memfasilitasi integrasi metode machine learning seperti SVM melalui tahapan yang jelas: seleksi → preprocessing → transformasi → mining → evaluasi. Dibandingkan metodologi lain (CRISP-DM/SEMMA), KDD lebih fleksibel dalam preprocessing dan transformasi data,

terutama untuk analisis sentimen berbahasa Indonesia yang kompleks akibat variasi linguistik (dialek, singkatan, campuran bahasa[13].



Gambar 1.1 Metode Knowledge Discovery database

Adapun tahapan Proses KDD sebagai berikut:

1. Tahap Selection

Tahap Selection merupakan tahapan awal dalam proses KDD (Knowledge Discovery in Database) yang bertujuan untuk mengumpulkan dan menyeleksi data yang relevan. Pada penelitian ini, data dikumpulkan dari platform Twitter/X dengan menggunakan web scraping berbasis Selenium, bukan API resmi. Data yang dikumpulkan difokuskan pada tweet berbahasa Indonesia yang membahas isu-isu terkait Presiden dan Wakil Presiden terpilih Prabowo-Gibran selama periode Oktober 2024 – 28 Januari 2025, yang mencerminkan momentum 100 hari awal pemerintahan[14].

2. Preprocessing

Tahap Preprocessing adalah tahapan yang dimana data tweet yang telah dikumpulkan melalui serangkaian proses pembersihan. Proses

ini mencakup penghapusan URL, mention, dan karakter khusus yang tidak relevan, mengubah semua teks menjadi huruf kecil (case folding), memecah kalimat menjadi kata-kata (tokenization), menghapus kata-kata umum yang tidak bermakna (stopword removal), dan mengubah kata-kata menjadi bentuk dasarnya (stemming). Tahap ini sangat krusial untuk memastikan kualitas data yang akan dianalisis[15].

3. Transformation

Pada tahap ini, data yang sudah dibersihkan diubah menjadi format yang bisa diproses oleh algoritma machine learning. Proses ini menggunakan teknik TF-IDF untuk mengubah teks menjadi angka dan melakukan normalisasi fitur. Pelabelan data dilakukan secara otomatis dengan sistem yang telah dirancang untuk mengklasifikasikan data dengan presisi dan efisiensi. Pelabelan otomatis ini menggunakan algoritma machine learning yang telah dilatih sebelumnya, seperti Random Forest, SVM, atau Deep Learning. Sistem ini dapat menganalisis fitur-fitur penting dari teks dan memberikan label yang sesuai berdasarkan pola yang telah dipelajari[16].

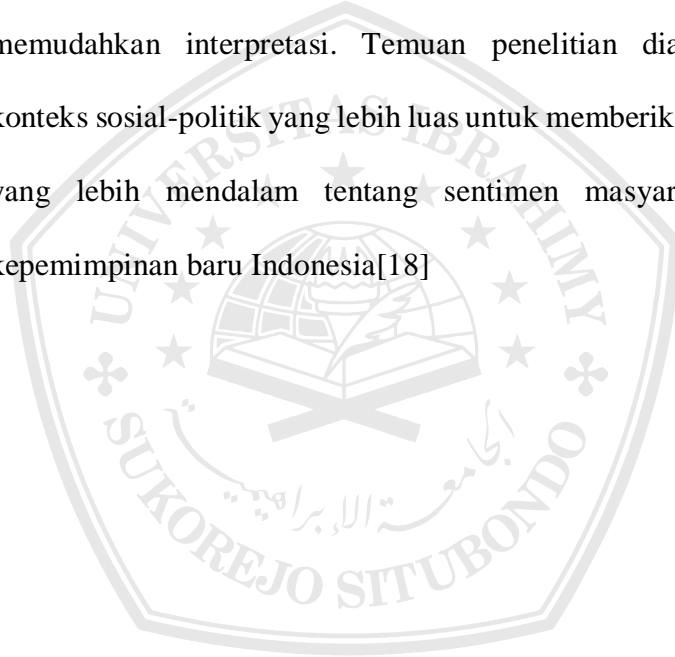
4. Data Mining

Data Mining adalah tahap dimana algoritma Support Vector Machine (SVM) diterapkan untuk mengklasifikasikan sentimen. Pemilihan SVM didasarkan pada kemampuannya dalam menangani data teks

dengan dimensi tinggi dan menghasilkan model klasifikasi yang optimal. [17].

5. Interpretation/Evaluation,

Tahap ini merupakan tahapan akhir dimana hasil klasifikasi dievaluasi menggunakan confusion matrix untuk mengukur tingkat akurasi, presisi, recall, dan F1-score dari model. Hasil analisis kemudian divisualisasikan dalam bentuk grafik dan diagram untuk memudahkan interpretasi. Temuan penelitian dianalisis dalam konteks sosial-politik yang lebih luas untuk memberikan pemahaman yang lebih mendalam tentang sentimen masyarakat terhadap kepemimpinan baru Indonesia[18]



BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

“Analisis Sentimen Terhadap Presiden Terpilih Dimedia Sosial Twitter (X) Menggunakan Algoritma Support Vector Machine”[19]

Penelitian ini dilakukan untuk menganalisis sentimen terhadap presiden terpilih Indonesia dengan menggunakan data dari media sosial Twitter (X) melalui teknik scraping. Data yang dikumpulkan berjumlah 2.000 tweet dan diklasifikasikan ke dalam lima kategori sentimen, yaitu positif, sangat positif, netral, negatif, dan sangat negatif. Pengumpulan data dilakukan dengan kata kunci seperti "Paslon02", "Prabowo", dan "Gibran", kemudian diberi label secara manual untuk memastikan keakuratan klasifikasi sebelum diproses lebih lanjut. Setelah pengumpulan data, dilakukan preprocessing untuk meningkatkan kualitas data dengan menerapkan beberapa teknik, antara lain case folding untuk mengubah semua huruf menjadi kecil, normalisasi untuk mengonversi kata tidak baku ke bentuk standar, stopword removal untuk menghapus kata-kata umum yang tidak memiliki makna signifikan, dan lemmatization untuk mengubah kata ke bentuk dasarnya. Langkah ini bertujuan untuk memastikan konsistensi dan keterbacaan data, serta meminimalisir kesalahan pada proses klasifikasi.

Kemudian, analisis dilakukan menggunakan algoritma Support Vector Machine (SVM), yang dikenal sebagai salah satu metode machine learning yang kuat dalam klasifikasi teks, termasuk analisis sentimen. Untuk mengevaluasi performa model, dilakukan pengujian dengan lima skenario berbeda berdasarkan

jumlah data, yaitu 100, 500, 1.000, 1.500, dan 2.000 tweet. Hasil pengujian menunjukkan bahwa semakin banyak data yang digunakan, semakin meningkat akurasi model, karena model memiliki lebih banyak informasi untuk mempelajari pola sentimen dalam data. Pada pengujian dengan 100 data, akurasi yang diperoleh masih rendah, yaitu 0,40%, dengan presisi 0,08%, dan recall 0,20%, menunjukkan bahwa data yang terbatas tidak cukup untuk membentuk pola klasifikasi yang akurat. Seiring bertambahnya jumlah data, model menunjukkan peningkatan performa yang signifikan. Pengujian dengan 500 data menghasilkan akurasi 0,67%, presisi 0,33%, dan recall 0,24%. Saat jumlah data mencapai 1.000, akurasi meningkat menjadi 0,73%, dengan presisi 0,52%, dan recall 0,29%. Pada 1.500 data, model mencapai akurasi 0,74%, dengan presisi 0,41%, dan recall 0,29%. Pengujian terakhir dengan 2.000 data menghasilkan tingkat akurasi tertinggi sebesar 0,75%, dengan presisi 0,47%, dan recall 0,30%, menunjukkan bahwa model dapat mengklasifikasikan sentimen dengan lebih baik ketika dataset yang digunakan lebih besar.

Penelitian ini berbeda dengan jurnal IJCS karena lebih spesifik menganalisis sentimen terhadap Prabowo-Gibran di Twitter/X selama 100 hari kerja pertama, sedangkan jurnal sebelumnya membahas presiden terpilih secara umum tanpa batasan waktu. Penelitian ini juga menggunakan tiga kategori sentimen (positif, negatif, netral) agar lebih fokus, berbeda dengan jurnal yang menggunakan lima kategori. Selain itu, penelitian ini lebih memperhatikan bahasa informal dan campuran dalam tweet serta menambahkan F1-score sebagai metrik evaluasi

tambahan. Dengan pendekatan ini, penelitian ini diharapkan lebih akurat dalam memahami opini publik terhadap pemerintahan baru.

“Analisis Sentimen Pengguna Aplikasi X Terhadap Program Makan Siang Gratis Dengan Metode Naïve Bayes Classifier”[20]

Berdasarkan jurnal penelitian tersebut, hasil analisis sentimen terhadap program makan siang gratis menggunakan metode Naïve Bayes Classifier menunjukkan temuan yang menarik. Dari total 4.038 data cuitan yang diambil dari aplikasi *Twitter/X* pada periode 14 Februari - 20 Maret 2024, setelah melalui proses pembersihan data (*text preprocessing*) didapatkan 2.909 data yang layak dianalisis. Dalam proses analisis, peneliti melakukan pelabelan manual terhadap 730 data (25% dari total data) sebagai data *training*, yang terdiri dari 409 sentimen positif dan 321 sentimen negatif. Setelah diterapkan algoritma Naïve Bayes Classifier, dari total keseluruhan data ditemukan 1.543 cuitan (53,04%) mengandung sentimen positif dan 1.366 cuitan (44,95%) mengandung sentimen negatif.

Untuk mengukur keakuratan model, peneliti menggunakan metode *K-fold validation* dengan nilai $K=10$. Hasil pengujian menunjukkan tingkat akurasi yang cukup tinggi yaitu 80,31% dengan *margin error* $\pm 5,27\%$. Ini berarti model dapat mengklasifikasikan sentimen dengan tepat pada 80,31% kasus. Lebih detail lagi, untuk sentimen positif menghasilkan nilai presisi 80,69% dan *recall* 79,71%, sedangkan untuk sentimen negatif menghasilkan nilai presisi 79,95% dan *recall* 80,93%. Visualisasi data menggunakan *wordcloud* menunjukkan kata-kata yang paling sering muncul dalam cuitan terkait program makan siang gratis. Sebelum dilakukan analisis, data melalui beberapa tahap *preprocessing* seperti *cleaning*,

tokenize, *transform case*, *stopword removal*, dan *filtering* untuk memastikan kualitas data yang akan dianalisis.

Hasil penelitian ini menunjukkan bahwa mayoritas pengguna aplikasi X memberikan respon positif terhadap program makan siang gratis, meskipun selisihnya tidak terlalu jauh dengan sentimen negatif. Tingkat akurasi yang mencapai 80,31% menunjukkan bahwa metode Naïve Bayes Classifier cukup efektif dalam menganalisis sentimen untuk kasus ini.

“Analisis Sentimen X Terhadap Pemilihan Presiden Indonesia 2024 dengan Metode K-Nearest Neighbor”[21]

Penelitian tentang analisis sentimen telah banyak dilakukan sebelumnya dalam berbagai konteks dan metodologi. Salah satu penelitian terkini yang menjadi acuan penting adalah penelitian yang dilakukan oleh Linggar Nursinggah dkk (2024) dengan judul "Analisis Sentimen Pengguna Aplikasi *twitter/X* Terhadap Program Makan Siang Gratis dengan Metode Naïve Bayes Classifier". Penelitian ini menggunakan pendekatan *machine learning* untuk menganalisis sentimen publik terhadap kebijakan program pemerintah di media sosial *Twitter/X*. Dalam penelitiannya, Nursinggah dkk melakukan pengumpulan dan analisis data secara sistematis dari platform *Twitter/X* selama periode 14 Februari - 20 Maret 2024. Dari total 1.000 *tweet* yang dianalisis, penelitian tersebut berhasil mengklasifikasikan sentimen publik dengan rincian 211 *tweet* bersentimen positif, 175 *tweet* bersentimen negatif, dan 614 *tweet* bersentimen netral. Data ini menunjukkan bahwa mayoritas pengguna *Twitter/X* cenderung memberikan respon netral terhadap program tersebut.

Menggunakan metode Naïve Bayes Classifier, penelitian tersebut mencapai hasil yang cukup menjanjikan dengan tingkat akurasi sebesar 80,31% dan *margin error* $\pm 5,27\%$. Lebih detail, penelitian ini mencapai presisi yang baik dengan nilai 80,69% untuk sentimen positif dan 79,95% untuk sentimen negatif. Nilai *recall* juga menunjukkan performa yang baik dengan 79,71% untuk sentimen positif dan 80,93% untuk sentimen negatif. Meskipun hasil ini menunjukkan performa yang memuaskan, namun masih terdapat ruang untuk peningkatan, terutama dalam penggunaan metode klasifikasi yang lebih *advanced* seperti *Support Vector Machine* (SVM).

Berdasarkan hasil dan keterbatasan dari penelitian tersebut, maka saya mengusulkan sebuah penelitian baru yang akan menggunakan metode Support Vector Machine (SVM) untuk menganalisis sentimen masyarakat terhadap presiden dan wakil presiden terpilih Prabowo-Gibran. Pemilihan metode SVM didasarkan pada kemampuannya dalam menangani data yang lebih kompleks dan potensinya untuk memberikan hasil klasifikasi yang lebih akurat, terutama dalam konteks analisis sentimen politik yang sering kali melibatkan opini-opini subjektif, penggunaan bahasa yang beragam, serta ketidakseimbangan jumlah data antar kelas sentimen. Selain itu, SVM juga dikenal efektif dalam memproses data dalam dimensi tinggi, seperti representasi teks dari media sosial, sehingga diharapkan mampu menggambarkan kecenderungan opini publik secara lebih representatif dan reliabel.

2.2 Landasan Teori

2.2.1 Prabowo-Gibran Presiden



Gambar 2. 1 Hasil Rekapitulasi KPU Pilpres 2024

Prabowo Subianto dan Gibran Rakabuming Raka merupakan presiden dan wakil presiden terpilih periode 2024–2029 yang berhasil memenangkan Pemilu 2024 dengan perolehan suara mencapai 58,6% atau sekitar 96,2 juta suara. Pasangan ini menjadi fokus perhatian publik setelah pengumuman resmi oleh KPU pada 20 Maret 2024, mengundang berbagai respons dan tanggapan masyarakat di media sosial, khususnya Twitter/X. Dinamika opini publik terhadap kepemimpinan baru ini menjadi objek penting dalam analisis sentimen untuk memahami persepsi masyarakat secara lebih mendalam, baik dari segi penerimaan terhadap kebijakan awal, ekspektasi terhadap program kerja, hingga kritik atau kekhawatiran yang muncul di tengah diskursus digital yang berlangsung secara dinamis dan real-time.[11].

2.2.2 Analisis Sentimen



Gambar 2. 2 Contoh Aplikasi X

Analisis sentimen, juga dikenal sebagai opinion mining, adalah proses menganalisis teks digital untuk menentukan sentimen atau emosi yang terkandung di dalamnya, biasanya diklasifikasikan sebagai positif, negatif, atau netral. Dalam konteks media sosial, analisis sentimen membantu memahami opini publik terhadap suatu topik atau tokoh tertentu. Teknik ini memanfaatkan kombinasi antara pemrosesan bahasa alami (*Natural Language Processing/NLP*) dan *machine learning* untuk mengidentifikasi dan mengekstrak informasi subjektif dari sumber data[22].

Proses analisis sentimen melibatkan beberapa tahapan kunci yang saling terintegrasi. Tahap pertama adalah *preprocessing* data yang terdiri dari pembersihan teks (menghilangkan URL, username, tanda baca),

normalisasi (mengubah kata informal menjadi formal), dan *tokenisasi* (memecah teks menjadi unit-unit kata). Selanjutnya, dilakukan ekstraksi fitur menggunakan berbagai metode seperti Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), atau word embeddings untuk mengubah teks menjadi representasi numerik yang dapat diproses oleh algoritma *machine learning*[18].

2.2.3 Twitter (X)



Gambar 2. 3 Logo Aplikasi X

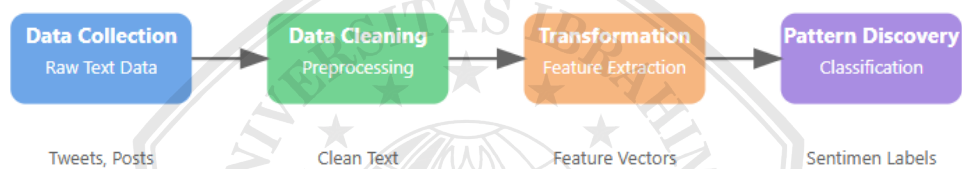
Twitter, yang kini dikenal juga sebagai X, adalah platform media sosial mikroblog yang memungkinkan penggunanya berbagi pesan singkat atau "*tweet*" dengan batasan 280 karakter. Platform ini menjadi sumber data yang sangat berharga untuk analisis sentimen karena beberapa karakteristik uniknya. Pertama, *Twitter/X* menyediakan data *real-time* yang mencerminkan opini publik terkini[23]. Kedua, pengguna *Twitter/X* cenderung mengekspresikan pendapat mereka secara spontan dan natural, memberikan gambaran yang autentik tentang sentimen masyarakat.

Twitter/X juga memiliki fitur-fitur seperti *hashtag*, *mention*, dan *retweet* yang membantu dalam mengorganisir dan menyebarkan informasi[24].

2.2.4 Data Mining

Data mining merupakan proses ekstraksi dan analisis pola yang bermakna dari kumpulan data besar. Dalam konteks analisis sentimen, data mining memungkinkan kita untuk menemukan pola dan tren dari data teks

Proses Data Mining dalam Analisis Sentimen



Gambar 2. 4 Gambaran Proses Data Mining

yang tidak terstruktur, seperti *tweet* atau postingan media sosial.

Seperti yang ditunjukkan dalam diagram, proses data mining meliputi beberapa tahap penting:

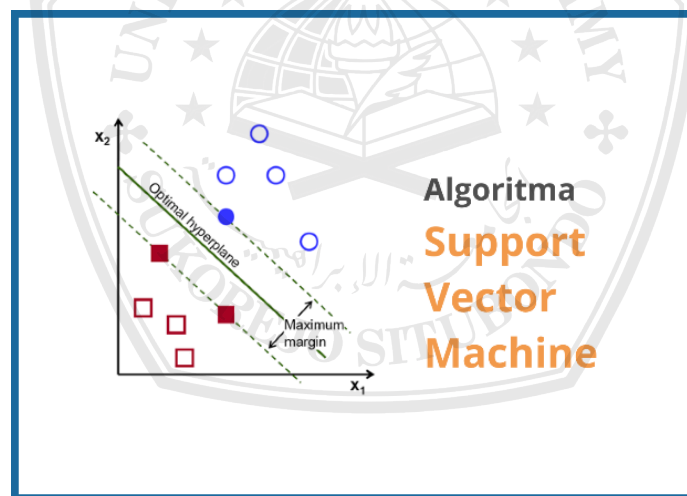
- 1) Pengumpulan Data (Data Collection): Mengumpulkan data mentah berupa *tweet* atau postingan media sosial
- 2) Pembersihan Data (Data Cleaning): Membersihkan dan mempreprocessing data untuk menghilangkan noise dan data yang tidak relevan
- 3) Transformasi Data: Mengubah data teks menjadi format yang dapat dianalisis menggunakan teknik ekstraksi fitur

- 4) Pattern Discovery: Menerapkan algoritma klasifikasi untuk menemukan pola dan mengkategorikan sentimen

Dalam analisis sentimen, data mining membantu mengidentifikasi dan mengklasifikasikan opini publik menjadi kategori positif, negatif, atau netral berdasarkan karakteristik teks yang dianalisis[25]. Setiap tahap dalam proses ini saling terkait dan berkontribusi pada akurasi hasil klasifikasi akhir.

2.2.5 Support Vector Machine

Support Vector Machine (SVM) adalah salah satu algoritma pembelajaran mesin yang powerful untuk klasifikasi dan regresi. Algoritma



Gambar 2. 5 Ilustrsi Algoritma SVM

ini bekerja dengan mencari *hyperplane* (garis pemisah) optimal yang memisahkan data ke dalam kelas-kelas yang berbeda dengan *margin* maksimal[4].

Secara matematis, *hyperplane* optimal dalam SVM dapat didefinisikan sebagai:

$$f(x) = w \cdot x + b = 0 \quad (1)$$

di mana w adalah vektor bobot normal terhadap *hyperplane*, x adalah vektor input, dan b adalah bias. Untuk klasifikasi biner, prediksi kelas dapat ditentukan dengan:

$$y = \text{sign}(w \cdot x + b) \quad (2)$$

Untuk memaksimalkan *margin* antara dua kelas, SVM menyelesaikan masalah optimasi:

$$\text{minimize: } (1/2)\|w\|^2 \quad (3)$$

$$\text{subject to: } y_i(w \cdot x_i + b) \geq 1, \text{ untuk semua } i$$

di mana y_i adalah label kelas (+1 atau -1) untuk data *training* x_i .

Dalam kasus data yang tidak dapat dipisahkan secara linear, SVM menggunakan fungsi kernel $K(x_i, x_j)$ untuk memetakan data ke dimensi yang lebih tinggi. Fungsi kernel yang umum digunakan adalah RBF (Radial Basis Function):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

Dalam konteks analisis sentimen, SVM dapat mengklasifikasikan teks berdasarkan fitur-fitur linguistik seperti kata kunci, frase, atau pola bahasa tertentu. Keunggulan SVM terletak pada kemampuannya menangani data dimensi tinggi dan memberikan hasil klasifikasi yang akurat, terutama untuk kasus-kasus yang memiliki pola kompleks. SVM juga efektif dalam mengatasi masalah *overfitting* karena fokusnya pada memaksimalkan *margin* antara kelas-kelas yang berbeda[26].

2.2.6 Python Notebook



Gambar 2. 6 Jupyter Notebook

Python Notebook atau *Jupyter Notebook* merupakan aplikasi web open-source yang memungkinkan pengguna untuk membuat dan berbagi dokumen yang berisi kode langsung, persamaan, visualisasi, dan teks naratif[5]. Dalam konteks analisis sentimen, Python Notebook menjadi alat yang sangat berguna karena memungkinkan peneliti untuk menulis dan menjalankan kode Python secara interaktif, melakukan *preprocessing* data, membangun model *machine learning*, dan memvisualisasikan hasil analisis dalam satu lingkungan yang terintegrasi[27].

Keunggulan utama *Jupyter Notebook* adalah kemampuannya menggabungkan kode dan dokumentasi dalam satu file, sehingga memudahkan peneliti dalam mencatat setiap tahap analisis. Tool ini mendukung berbagai library Python yang sering digunakan dalam analisis sentimen, seperti *pandas* untuk mengolah data, *scikit-learn* untuk membuat model SVM, dan *matplotlib* untuk membuat grafik. *Jupyter Notebook* juga memungkinkan peneliti untuk menjalankan kode per bagian (*cell*), sehingga

lebih mudah untuk mengecek hasil setiap tahap pengolahan data dan memperbaiki kesalahan jika terjadi[28].

a. Confusion Matrix

Confusion Matrix merupakan alat evaluasi yang digunakan untuk mengukur performa model klasifikasi. Matrix ini menunjukkan perbandingan antara hasil prediksi model dengan nilai aktual dalam bentuk matrix 2x2 yang terdiri dari empat komponen utama:

- 1) *True Positive* (TP): Data positif yang diprediksi benar sebagai positif
- 2) *True Negative* (TN): Data negatif yang diprediksi benar sebagai negatif
- 3) *False Positive* (FP): Data negatif yang salah diprediksi sebagai positif
- 4) *False Negative* (FN): Data positif yang salah diprediksi sebagai negatif

Dari komponen-komponen ini, dihitung empat metrik evaluasi utama:

- 1) Akurasi: Mengukur ketepatan model secara keseluruhan

$$\text{Akurasi} = (TP + TN) / (TP + TN + FP + FN)$$

- 2) Presisi: Mengukur ketepatan model dalam memprediksi kelas positif

$$\text{Presisi} = TP / (TP + FP)$$

- 3) *Recall*: Mengukur kemampuan model dalam mengenali seluruh data positif

$$\text{Recall} = TP / (TP + FN)$$

- 4) F1-Score: Mengukur keseimbangan antara presisi dan *recall*

$$\text{F1-Score} = 2 \times (\text{Presisi} \times \text{Recall}) / (\text{Presisi} + \text{Recall})$$

Confusion Matrix dan Metrik Evaluasi

	TP	FP
Aktual	FN	TN
	Prediksi	

$$\begin{aligned} \text{Akurasi} &= (TP + TN) / (TP + TN + FP + FN) \\ \text{Presisi} &= TP / (TP + FP) \\ \text{Recall} &= TP / (TP + FN) \\ \text{F1-Score} &= 2 \times (\text{Presisi} \times \text{Recall}) / (\text{Presisi} + \text{Recall}) \end{aligned}$$

Gambar 2. 7 Perhitungan Confusion Matrix

Dalam analisis sentimen, metrik-metrik ini membantu menilai seberapa baik model dalam mengklasifikasikan tweet ke dalam kategori sentimen yang benar. Akurasi memberikan gambaran umum tentang performa model secara keseluruhan dalam mengklasifikasikan data dengan tepat. Presisi menunjukkan sejauh mana prediksi positif yang dibuat oleh model benar-benar relevan atau akurat. Recall mengukur kemampuan model dalam mengenali seluruh kasus positif yang ada di dalam data. Sementara itu, F1-score merupakan rata-rata harmonis dari presisi dan recall, yang berguna untuk menyeimbangkan trade-off antara keduanya, terutama ketika terdapat ketidakseimbangan kelas dalam data.[29].

BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1 Analisis Masalah

Dalam pengembangan sistem analisis sentimen terhadap presiden dan wakil presiden terpilih Prabowo-Gibran, terdapat beberapa permasalahan utama yang perlu diidentifikasi dan diselesaikan. Pertama, kompleksitas dalam menganalisis sentimen masyarakat di *Twitter/X* terkait Prabowo-Gibran muncul karena beragamnya opini yang disampaikan dalam berbagai bentuk bahasa, mulai dari bahasa formal hingga informal [30]. Masyarakat sering menggunakan singkatan, emoji, dan bahasa campuran (code-mixing) antara Bahasa Indonesia dengan bahasa asing dalam *tweet* mereka. Hal ini menyebabkan tantangan tersendiri dalam mengidentifikasi makna sebenarnya dari setiap *tweet* dan mengklasifikasikannya ke dalam sentimen yang tepat [31].

Tantangan kedua muncul dalam proses pengolahan data *Twitter/X* berbahasa Indonesia. Karakteristik unik Bahasa Indonesia di media sosial, seperti penggunaan slang, dialek daerah, dan variasi penulisan kata yang sama, membuat proses normalisasi teks menjadi lebih kompleks. Sebagai contoh, kata "presiden" bisa ditulis dalam berbagai bentuk seperti "presiden", "presiden.", "preSiden", atau bahkan disingkat menjadi "pres"[15]. Selain itu, banyaknya penggunaan kata serapan dan istilah baru yang muncul di media sosial menambah kompleksitas dalam proses pengolahan bahasa.

3.2 Alur Proses

Penelitian analisis sentimen terhadap presiden dan wakil presiden terpilih Prabowo-Gibran dilaksanakan melalui beberapa tahapan yang sistematis dan terstruktur. Setiap tahapan dirancang untuk memastikan kualitas dan akurasi hasil analisis[32]. Berikut adalah penjelasan detail dari setiap tahapan penelitian:

3.2.1 Pengumpulan Data

Tahap pengumpulan data merupakan langkah awal yang krusial dalam penelitian ini. Penelitian ini tidak menggunakan API Twitter/X resmi karena keterbatasan akses dan kuota pada versi gratis, melainkan dilakukan dengan pendekatan web scraping menggunakan library Selenium dan ChromeDriver. Teknik ini memungkinkan peneliti untuk mengakses halaman pencarian Twitter/X secara langsung dan otomatis mengambil data publik dari postingan pengguna.



Gambar 3. 1 Tahapan Pengumpulan Data

Berikut adalah tahapan teknis pengumpulan data yang digunakan:

1. Pembuatan dan Persiapan Lingkungan Scraping

Peneliti menyiapkan lingkungan scraping menggunakan Python dengan library seperti:

- selenium: untuk mengontrol browser secara otomatis
- pandas: untuk menyimpan hasil scraping ke dalam dataset

- chromedriver: untuk menjembatani antara Selenium dan browser Google Chrome
- Google Chrome profile login: untuk memastikan sudah login akun X (Twitter) secara manual agar dapat mengakses konten tanpa hambatan

2. Penentuan Kata Kunci dan Rentang Waktu

Scraping dilakukan berdasarkan kata kunci tertentu yang relevan dengan objek penelitian, yaitu:

- "100 hari presiden"
- "danantara"
- "kabinet indonesia"
- "makan siang gratis"
- "program kerja presiden"
- "swasembada".

Rentang waktu yang digunakan adalah 20 Oktober 2024 hingga 28 Januari 2025, yang mewakili 100 hari pertama masa jabatan Presiden dan Wakil Presiden terpilih (Prabowo-Gibran). Rentang ini dipilih karena merupakan masa awal kepemimpinan nasional yang sering menjadi sorotan publik di media sosial.

3. Proses Scraping dengan Selenium

Tahapan scraping dilakukan sebagai berikut:

- Browser dibuka dan diarahkan ke URL pencarian Twitter/X berdasarkan kata kunci dan rentang waktu.

- Scroll otomatis dilakukan secara berulang untuk memuat tweet sebanyak mungkin.
 - Setiap tweet yang dimuat kemudian diekstrak elemen pentingnya: “Tanggal posting (timestamp)”, “Nama pengguna (username)”, “Isi tweet (text)”
 - Tweet disimpan ke dalam struktur data dictionary dan dikonversi menjadi *DataFrame*.
4. Setelah seluruh data terkumpul, dataset disimpan dalam file Excel (.xlsx) atau CSV (.csv) menggunakan library Pandas. Format yang digunakan adalah “Tanggal”, “Username”, “Tweet”. Dataset ini menjadi dasar untuk proses preprocessing, labeling, dan analisis sentimen di tahap berikutnya.

3.2.2 Preprocessing

1. Case Folding

Seluruh teks diubah menjadi huruf kecil untuk memastikan konsistensi dalam analisis.

Contoh: "TweET Bagus" → "tweet bagus"

2. Cleaning

Menghapus elemen yang tidak relevan dengan analisis sentimen:

- a) URL: Dihapus menggunakan regex agar tidak mengganggu teks utama.
- b) Username (@mention): Menghapus mention seperti @user agar analisis fokus pada isi tweet.
- c) Hashtag (#): Menghapus simbol #, tetapi bisa mempertahankan kata jika relevan.

- d) Emoji & Karakter Khusus: Menghapus emoji dan simbol yang tidak berkontribusi pada analisis[24].

Contoh:

Sebelum: "@user Wah keren banget! 😍 Cek di <https://example.com>
#keren"

Sesudah: "wah keren banget cek keren"

3. Tokenisasi

Memecah teks menjadi kata-kata individual agar lebih mudah dianalisis[37].

Contoh: "Wah keren banget" → ["wah", "keren", "banget"]

4. Stopwords Removal

Menghapus kata-kata umum yang tidak memiliki nilai signifikan dalam analisis sentimen. Menggunakan daftar stopwords bahasa Indonesia yang telah disesuaikan.

Contoh:

Sebelum: ["wah", "keren", "banget", "dan", "sangat"]

Sesudah: ["keren", "banget"]

5. Stemming

Mengubah kata ke bentuk dasar menggunakan library Sastrawi untuk mengurangi variasi kata[38].

Contoh: "memilih" → "pilih", "berjalan" → "jalan"



Gambar 3. 2 Tahapan Preprocessing Data

3.2.3 Transformation

1. Proses TF-IDF

Proses transformasi teks ke dalam bentuk numerik dilakukan menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF), yang merupakan salah satu teknik representasi teks paling populer dalam pemrosesan bahasa alami. Dalam tahap ini, setiap tweet direpresentasikan sebagai vektor numerik yang mencerminkan seberapa sering kata-kata tertentu muncul dalam tweet tersebut (term frequency), serta seberapa penting kata-kata tersebut dalam keseluruhan kumpulan data (inverse document frequency). Dengan pendekatan ini, kata-kata umum yang sering muncul di banyak tweet akan memiliki bobot yang lebih rendah, sementara kata-kata yang lebih spesifik dan jarang muncul, namun relevan, akan diberi bobot yang lebih tinggi. Hasil dari proses ini adalah representasi numerik dari setiap tweet yang dapat digunakan sebagai input bagi algoritma pembelajaran mesin untuk melakukan analisis sentimen secara lebih efektif dan akurat.[26]

2. *Splitting Data*

Dataset yang telah melalui tahapan preprocessing dan proses vektorisasi menggunakan metode TF-IDF selanjutnya dibagi menjadi dua bagian, yaitu data training dan data testing, dengan rasio pembagian sebesar 80:20. Pembagian ini dilakukan menggunakan teknik stratified split, yang bertujuan untuk memastikan bahwa proporsi distribusi label sentimen (misalnya positif, negatif, dan netral) tetap seimbang di kedua subset data. Dengan demikian, data training yang berjumlah 80% dapat merepresentasikan karakteristik keseluruhan dataset secara proporsional, sementara data testing sebesar 20% dapat memberikan evaluasi kinerja model yang lebih akurat dan adil berdasarkan distribusi kelas yang sama dengan data asli.[39].

3.2.4 Data Mining

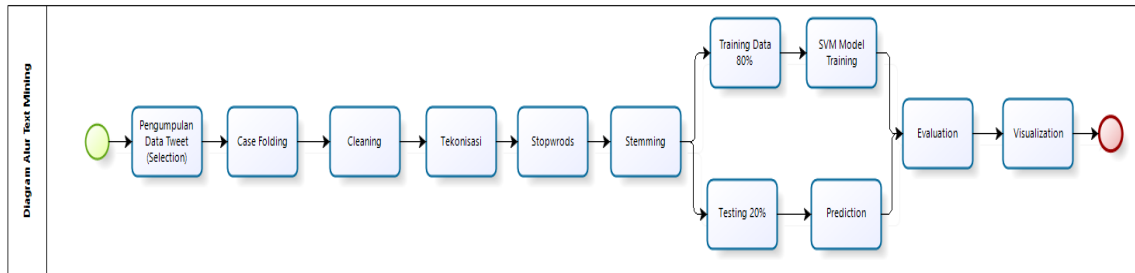
Model Support Vector Machine (SVM) dilatih menggunakan data training yang telah dipersiapkan melalui proses data splitting. Data dibagi menjadi training set dan testing set untuk mengevaluasi performa model [5]. Model Support Vector Machine (SVM) dilatih menggunakan data training yang telah dipersiapkan melalui proses data splitting. Data dibagi menjadi training set (80%) dan testing set (20%) untuk mengevaluasi performa model [5]. Dalam tahap pelatihan, model dilatih menggunakan 80% data training, sementara 20% data testing digunakan sebagai validation set untuk memastikan model tidak mengalami overfitting. Pendekatan ini memberikan estimasi performa model yang lebih akurat dan stabil

dibandingkan dengan pembagian data secara acak satu kali. Selain itu, dengan pembagian data 80:20, setiap data memiliki kesempatan yang sama untuk digunakan dalam pelatihan dan evaluasi, sehingga hasil evaluasi lebih representatif terhadap distribusi data keseluruhan [40]. Metode ini juga merupakan kompromi yang baik antara waktu komputasi dan akurasi evaluasi model, menjadikannya pilihan yang efisien dalam banyak skenario pelatihan model machine learning. Selain itu, Grid Search digunakan untuk menemukan kombinasi parameter terbaik guna meningkatkan performa model. Kernel RBF (Radial Basis Function) dipilih karena kemampuannya dalam menangani data non-linear dengan baik [18].

3.2.5 Evaluasi Model

Evaluasi performa model dilakukan menggunakan data testing yang telah disiapkan sebelumnya, yaitu data yang tidak dilibatkan dalam proses pelatihan model untuk menjaga objektivitas penilaian. Salah satu metode evaluasi yang digunakan adalah confusion matrix, yang disusun untuk menganalisis secara rinci distribusi hasil prediksi model terhadap label aktual. Confusion matrix ini mencakup empat komponen utama, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) untuk masing-masing kelas sentimen, seperti sentimen positif, negatif, dan netral. Melalui analisis confusion matrix, dapat diketahui sejauh mana model mampu membedakan antar kelas dengan tepat, serta mengidentifikasi kesalahan prediksi yang mungkin terjadi,

sehingga menjadi dasar dalam menilai akurasi, presisi, recall, dan F1-score dari model yang dibangun.



Gambar 3. 3 Diagram Alur Analisis Sentimen

Metrik evaluasi yang dihitung meliputi:

- Accuracy: ketepatan klasifikasi secara keseluruhan
- Precision: ketepatan model dalam memprediksi kelas positif
- Recall: kemampuan model mengenali seluruh kelas positif
- F1-score: harmonic mean dari *precision* dan *recall*

Hasil evaluasi divisualisasikan dalam bentuk heatmap confusion matrix, kurva ROC, dan grafik distribusi sentimen.

3.3 Perangkat yang Digunakan

Penelitian ini menggunakan beberapa perangkat keras dan lunak sebagai berikut:

3.3.1 Perangkat Keras

Laptop dengan spesifikasi:

- 1) Processor RYZEN 3 atau setara

- 2) RAM 4 GB DDR4
- 3) Storage 512GB SSD
- 4) Koneksi internet stabil

3.3.2 Perangkat Lunak

- 1) Sistem Operasi Windows 10 64-bit
- 2) Python 3.8 atau lebih tinggi
- 3) Jupyter Notebook untuk development
- 4) Chromedriver
- 5) Library Python:
 - *Selenium*
 - NLTK untuk processing bahasa
 - Pandas untuk manipulasi data
 - Scikit-learn untuk implementasi SVM
 - Matplotlib dan Seaborn untuk visualisasi
 - Sastrawi untuk stemming bahasa Indonesia

Dengan rancangan metodologi ini, diharapkan penelitian dapat menghasilkan model analisis sentimen yang akurat dan dapat diandalkan untuk menganalisis opini masyarakat terhadap Prabowo-Gibran di media sosial *Twitter/X*.

BAB IV

PEMBAHASAN

4.1 Program Kerja Presiden

Periode 100 hari kerja pertama pemerintahan Presiden Prabowo Subianto dan Wakil Presiden Gibran Rakabuming Raka merupakan fase strategis yang menjadi perhatian luas publik dan media. Masa ini terhitung sejak tanggal pelantikan pada 20 Oktober 2024 hingga 28 Januari 2025, dan sering dijadikan sebagai early performance benchmark bagi pemerintahan baru dalam menjalankan visi dan janji politiknya. Dalam konteks penelitian ini, periode tersebut dipilih sebagai dasar dalam mengumpulkan opini publik karena diyakini mencerminkan reaksi awal masyarakat terhadap implementasi program-program pemerintahan yang baru berjalan.



Gambar 4. 1 Presiden Prabowo dan Gibran

Salah satu program yang paling menonjol dan menjadi sorotan publik dalam 100 hari pertama adalah Program Makan Bergizi Gratis (MBG). Program ini merupakan janji kampanye utama pasangan Prabowo-Gibran, yang bertujuan untuk

memberikan makanan bergizi secara gratis kepada pelajar sekolah dasar dan menengah. Tujuan dari program ini tidak hanya untuk meningkatkan gizi anak-anak Indonesia, tetapi juga mendorong partisipasi dan konsentrasi belajar di sekolah. Namun, sejak diumumkan, MBG menuai respon yang beragam dari masyarakat. Di satu sisi, program ini diapresiasi sebagai bentuk kepedulian negara terhadap



Gambar 4. 2 Siswa menikmati makan siang gratis

generasi muda, namun di sisi lain juga menimbulkan kontroversi dan kritik, khususnya menyangkut isu besarnya anggaran yang diperlukan, kesiapan infrastruktur, serta potensi pengawasan dan korupsi dalam pelaksanaannya. Media sosial seperti Twitter/X menjadi tempat utama di mana opini-opini masyarakat tentang MBG berkembang sangat cepat dan dinamis.

Selain MBG, sejumlah topik lain yang ramai dibicarakan masyarakat dalam 100 hari pertama pemerintahan Prabowo-Gibran juga menjadi sorotan publik di Twitter/X. Misalnya, topik “100 hari presiden” sering digunakan sebagai penanda untuk mengevaluasi capaian dan janji-janji yang telah diumumkan selama masa kampanye. Di sisi lain, kata kunci seperti “danantara” yang merujuk pada media

atau opini tertentu sering muncul dalam diskusi yang berkaitan dengan narasi politik, opini elite, dan framing media. Pembahasan mengenai “kabinet Indonesia” juga menjadi ramai diperbincangkan, terutama terkait komposisi menteri, figur-figur yang dianggap kontroversial, serta isu kompetensi dan loyalitas. Kata kunci “program kerja presiden” dan “swasembada” juga sering muncul, menandakan adanya perhatian masyarakat terhadap janji-janji strategis seperti ketahanan pangan, energi, dan sektor-sektor ekonomi domestik. Diskusi mengenai topik-topik tersebut menggambarkan keragaman perhatian publik, mulai dari program konkret seperti Makan Bergizi Gratis (MBG) hingga pada level konseptual dan strategis yang lebih luas, seperti visi besar swasembada nasional atau reformasi kabinet. Semua dinamika ini menjadi bagian penting dalam memahami opini publik secara menyeluruh, khususnya dalam periode awal pemerintahan baru yang sangat menentukan arah kebijakan ke depan.

Dalam penelitian ini, periode 100 hari dipilih tidak hanya karena alasan simbolik sebagai masa evaluasi awal pemerintahan, tetapi juga karena dalam rentang waktu tersebut terjadi dinamika sosial-politik yang sangat intens. Mulai dari pengumuman program-program kerja unggulan, perdebatan di DPR, hingga berbagai komentar dari akademisi, jurnalis, dan masyarakat sipil. Twitter/X menjadi kanal utama penyebaran informasi dan reaksi publik terhadap isu-isu tersebut. Dengan demikian, analisis sentimen terhadap tweet-tweet yang membahas program kerja 100 hari pertama pemerintahan ini sangat penting untuk memahami sejauh mana kebijakan awal diterima atau ditolak oleh masyarakat, serta wacana-wacana publik apa yang muncul di balik angka statistik dan pemberitaan media arus utama.

Data yang dikumpulkan melalui proses scraping dan dianalisis menggunakan metode Support Vector Machine (SVM) diharapkan mampu memberikan wawasan objektif dan terukur terkait persepsi publik terhadap pemerintahan baru di masa transisi awal kekuasaan.

4.2 Pengumpulan Data *Twitter/X*

Dengan demikian, analisis sentimen terhadap tweet-tweet yang membahas program kerja 100 hari pertama pemerintahan menjadi sangat penting sebagai salah satu instrumen untuk mengukur respons publik secara real-time terhadap kebijakan yang telah diluncurkan. Melalui pendekatan ini, dapat diperoleh pemahaman yang lebih mendalam mengenai sejauh mana kebijakan-kebijakan awal pemerintahan Presiden dan Wakil Presiden terpilih, Prabowo Subianto dan Gibran Rakabuming Raka, diterima, didukung, atau bahkan ditolak oleh masyarakat. Selain itu, analisis ini juga membantu mengidentifikasi isu-isu yang paling banyak dibicarakan serta persepsi publik terhadap wacana yang telah atau sedang direalisasikan, sehingga dapat menjadi masukan strategis bagi pemerintah dalam merumuskan dan menyesuaikan arah kebijakan selanjutnya.

4.2.1 Persiapan Scraping

Sebelum melakukan proses pengambilan data (scraping), peneliti melakukan berbagai tahapan persiapan untuk memastikan bahwa lingkungan scraping dapat berjalan stabil dan mendukung pengambilan data dalam jumlah besar dari platform *Twitter/X*. Adapun beberapa komponen utama yang dipersiapkan meliputi instalasi perangkat lunak, konfigurasi teknis, dan pengaturan debugging browser.

a. Bahasa Pemrograman dan Versi

Penelitian ini dibangun menggunakan bahasa pemrograman Python versi 3.10.11. Pemilihan versi ini didasarkan pada kestabilan serta kompatibilitas yang baik dengan pustaka-pustaka scraping dan analisis data seperti Selenium, Pandas, dan openpyxl.



Gambar 4.3 Python

b. Penggunaan Selenium dan ChromeDriver

Untuk melakukan web scraping langsung dari halaman Twitter/X, peneliti memanfaatkan Selenium WebDriver, pustaka Python yang memungkinkan otomatisasi interaksi dengan antarmuka web. Dengan Selenium, peneliti dapat meniru aktivitas pengguna seperti menggulir halaman, mengklik elemen, dan mengekstrak konten HTML yang dimuat secara dinamis. agar dapat berjalan di browser Google Chrome, digunakan ChromeDriver sebagai penghubung antara Selenium dan browser. Kombinasi ini memungkinkan pengambilan data

tweet secara efektif, mencakup teks, username, waktu posting, dan elemen lainnya yang relevan, terutama pada konten yang dimuat dengan JavaScript.



Gambar 4. 4 Silenium

Selenium dipadukan dengan ChromeDriver versi 113, yang disesuaikan dengan Google Chrome versi 113 yang telah terpasang di sistem. Pemilihan versi yang cocok sangat penting karena ketidakcocokan versi Chrome dengan ChromeDriver dapat menyebabkan error saat inisialisasi browser atau crash selama proses scraping.

c. Eksekusi Chrome dalam Mode Debugging

Untuk menghindari keharusan login berulang ke akun Twitter/X, peneliti menjalankan Chrome dalam mode remote debugging menggunakan perintah PowerShell berikut:

```
"C:\Program Files\Google\Chrome\Application\chrome.exe" --remote-debugging-port=9222 --user-data-dir="C:\SeleniumDebug"
```

Segmen 1 Chrome mode Debugging

Dengan konfigurasi ini, Chrome berjalan dengan profil pengguna khusus (user data) yang telah login ke akun X (Twitter). Hal ini memungkinkan

Selenium untuk langsung membuka halaman hasil pencarian tweet dan mengeksekusi scroll otomatis tanpa hambatan otentikasi.

```
import logging
import time
import random
import datetime
import pytz
import pandas as pd
from selenium import webdriver
from selenium.webdriver.chrome.service import Service
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.by import By

# Logging
logging.basicConfig(
    filename="twitter_scraping.log",
    level=logging.DEBUG,
    format="%(asctime)s - %(levelname)s - %(message)s",
)
console_handler = logging.StreamHandler()
console_handler.setLevel(logging.INFO)
formatter = logging.Formatter("%(asctime)s - %(levelname)s - %(message)s")
console_handler.setFormatter(formatter)
logging.getLogger().addHandler(console_handler)

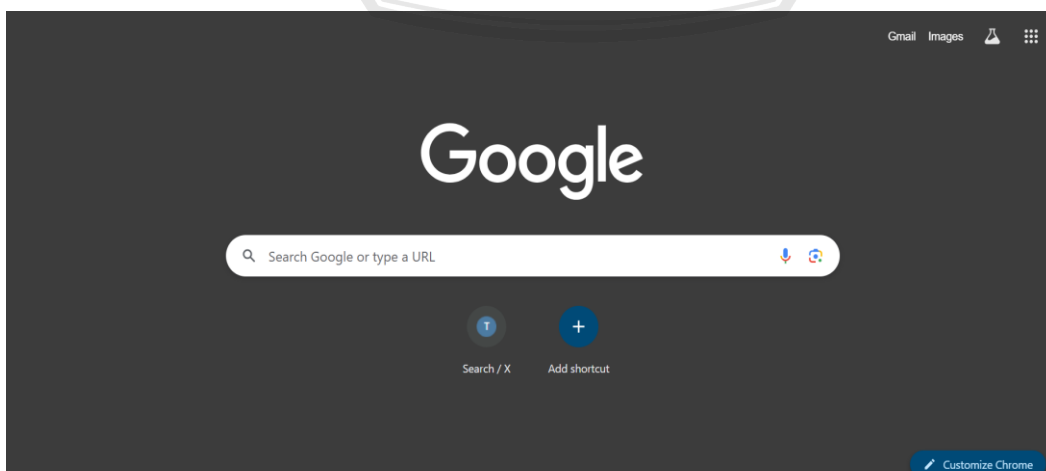
# ChromeDriver setup
driver_path = r"C:\Users\Mizoss\Music\chromedriver-win64\chromedriver.exe"
user_data_dir = r"C:\Users\Mizoss\AppData\Local\Google\Chrome\User Data"

options = Options()
options.add_argument(f"user-data-dir={user_data_dir}")
options.add_argument("--start-maximized")
options.add_argument("--disable-blink-features=AutomationControlled")
options.add_argument("--disable-extensions")
options.add_argument("--no-sandbox")
options.add_argument("--disable-gpu")
options.add_argument("--disable-dev-shm-usage")
options.add_argument("--remote-debugging-port=9222")
options.add_experimental_option("devToolsSwitches", {"enableLogging": 1})
```

Gambar 4. 5 Code Scrapping Data Twitter

4.2.2 Kata Kunci dan Rentang Waktu

Menjelaskan dasar pemilihan kata kunci ("makan siang gratis", "kabinet Indonesia", dll) dan periode waktu (20 Oktober 2024 – 28 Januari 2025) sesuai dengan 100 hari kerja presiden.



Gambar 4. 6 Chrome Mode Debugging

Dengan demikian, analisis sentimen terhadap tweet-tweet yang membahas program kerja 100 hari pertama pemerintahan ini sangat penting untuk memahami sejauh mana kebijakan awal diterima atau ditolak oleh masyarakat, serta wacana-wacana.

```

PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS

wo-gibran
📅Masukkan tanggal mulai (YYYY-MM-DD): 2024-10-20
📅Masukkan tanggal akhir (YYYY-MM-DD): 2025-01-28
2025-05-15 11:15:49,226 - INFO - ✅ Membuka halaman: https://twitter.com/search?q=kinerja 1
00 hari presiden prabowo-gibran%20lang%3Aid%20since%3A2024-10-20%20until%3A2025-01-28&src=t
yped_query&f=live
2025-05-15 11:15:59,544 - INFO - 📄 Round 1: Total tweet unik: 4
2025-05-15 11:16:04,297 - INFO - 📄 Round 2: Total tweet unik: 9
2025-05-15 11:16:08,934 - INFO - 📄 Round 3: Total tweet unik: 16
2025-05-15 11:16:12,767 - INFO - 📄 Round 4: Total tweet unik: 18
2025-05-15 11:16:16,778 - INFO - 📄 Round 5: Total tweet unik: 18
2025-05-15 11:16:20,837 - INFO - 📄 Round 6: Total tweet unik: 18
2025-05-15 11:16:24,717 - INFO - 📄 Round 7: Total tweet unik: 18
2025-05-15 11:16:28,910 - INFO - 📄 Round 8: Total tweet unik: 18
  
```

Gambar 4. 7 Masukkan Kata Kunci Dan Waktu

4.2.3 Pengumpulan Data *Twitter/X*

Setelah proses scraping selesai dilakukan dengan menggunakan Selenium, seluruh data tweet yang berhasil dikumpulkan disimpan dalam format CSV (*Comma Separated Values*). Penyimpanan data dilakukan secara otomatis melalui

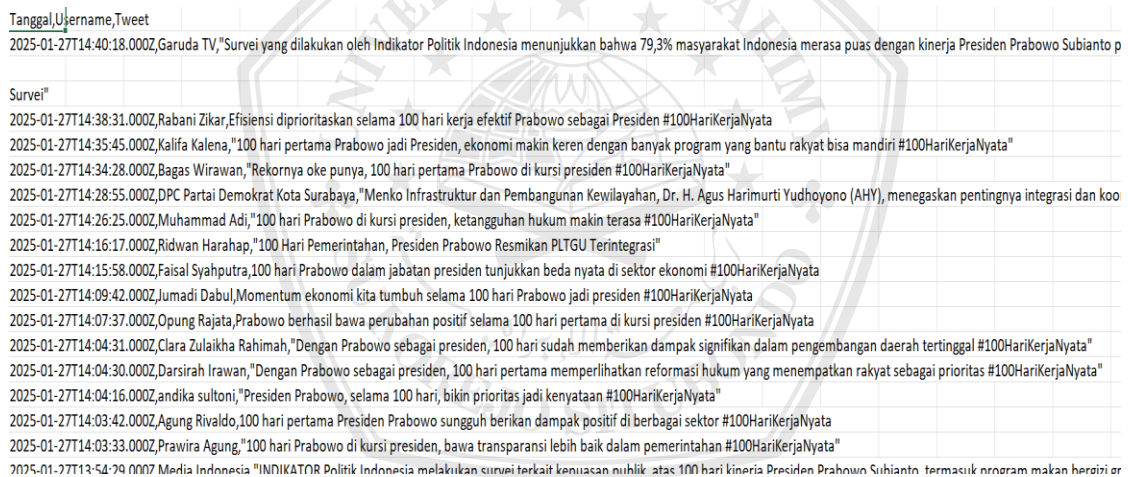
```

📄 scraped_tweets_100_hari_presiden_2024-10-20_to_2025-01-28.csv
📄 scraped_tweets_100_hari_presiden_2024-10-20_to_2024-12-20.csv
📄 scraped_tweets_danantara_2024-10-20_to_2025-01-28.csv
📄 scraped_tweets_kabinet_indonesia_2024-10-20_to_2025-01-28.csv
📄 scraped_tweets_makan_siang_gratis_2024-10-20_to_2024-11-20.csv
📄 scraped_tweets_makan_siang_gratis_2024-11-21_to_2024-12-21.csv
📄 scraped_tweets_makan_siang_gratis_2024-12-21_to_2025-01-28.csv
📄 scraped_tweets_makan_siang_gratis_2024-12-22_to_2025-01-15.csv
📄 scraped_tweets_makan_siang_gratis_2024-12-22_to_2025-01-20.csv
📄 scraped_tweets_makan_siang_gratis_2025-01-16_to_2025-01-19.csv
📄 scraped_tweets_prabowo_gibran_2024-10-20_to_2025-01-28.csv
📄 scraped_tweets_program_kerja_presiden_2024-10-20_to_2024-11-20.csv
📄 scraped_tweets_swasembada_2024-10-20_to_2025-01-28.csv
  
```

Gambar 4. 8 Nama file yang tersimpan

kode program, dengan menggunakan library pandas untuk menulis data ke dalam file. Setiap kata kunci seperti "makan siang gratis", "kabinet indonesia", "program kerja presiden", dan lainnya dilakukan secara terpisah (per kata kunci dan rentang waktu) kemudian hasilnya disimpan ke file CSV dengan format nama "scraped_tweets_makan_siang_gratis_2024-10-20_to_2025-01-28.csv"

Setelah semua proses scraping dilakukan untuk masing-masing kata kunci, seluruh data dari file CSV yang terpisah tersebut digabungkan (merged) menjadi satu file utama untuk dianalisis lebih lanjut.



Tanggal	Username	Tweet
2025-01-27T14:40:18.000Z	Garuda TV	"Survei yang dilakukan oleh Indikator Politik Indonesia menunjukkan bahwa 79,3% masyarakat Indonesia merasa puas dengan kinerja Presiden Prabowo Subianto p Survei"
2025-01-27T14:38:31.000Z	Rabani Zikar	"Efisiensi diprioritaskan selama 100 hari kerja efektif Prabowo sebagai Presiden #100HariKerjaNyata
2025-01-27T14:35:45.000Z	Kalifa Kalena	"100 hari pertama Prabowo jadi Presiden, ekonomi makin keren dengan banyak program yang bantu rakyat bisa mandiri #100HariKerjaNyata"
2025-01-27T14:34:28.000Z	Bagas Wirawan	"Rekornya oke punya, 100 hari pertama Prabowo di kursi presiden #100HariKerjaNyata"
2025-01-27T14:28:55.000Z	DPC Partai Demokrat Kota Surabaya	"Menko Infrastruktur dan Pembangunan Kewilayahan, Dr. H. Agus Harimurti Yudhoyono (AHY), menegaskan pentingnya integrasi dan koo
2025-01-27T14:26:25.000Z	Muhammad Adi	"100 hari Prabowo di kursi presiden, ketangguhan hukum makin terasa #100HariKerjaNyata"
2025-01-27T14:16:17.000Z	Ridwan Harahap	"100 Hari Pemerintahan, Presiden Prabowo Resmikan PLTGU Terintegrasi"
2025-01-27T14:15:58.000Z	Faisal Syahputra	"100 hari Prabowo dalam jabatan presiden tunjukkan beda nyata di sektor ekonomi #100HariKerjaNyata
2025-01-27T14:09:42.000Z	Jumadi Dabul	"Momentum ekonomi kita tumbuh selama 100 hari Prabowo jadi presiden #100HariKerjaNyata
2025-01-27T14:07:37.000Z	Opung Rajata	"Prabowo berhasil bawa perubahan positif selama 100 hari pertama di kursi presiden #100HariKerjaNyata
2025-01-27T14:04:31.000Z	Clara Zulaikha Rahimah	"Dengan Prabowo sebagai presiden, 100 hari sudah memberikan dampak signifikan dalam pengembangan daerah tertinggal #100HariKerjaNyata"
2025-01-27T14:04:30.000Z	Darsirah Irawan	"Dengan Prabowo sebagai presiden, 100 hari pertama memperlihatkan reformasi hukum yang menempatkan rakyat sebagai prioritas #100HariKerjaNyata"
2025-01-27T14:04:16.000Z	andika sultoni	"Presiden Prabowo, selama 100 hari, bikin prioritas jadi kenyataan #100HariKerjaNyata"
2025-01-27T14:03:42.000Z	Agung Rivaldo	"100 hari pertama Presiden Prabowo sungguh berikan dampak positif di berbagai sektor #100HariKerjaNyata
2025-01-27T14:03:33.000Z	Prawira Agung	"100 hari Prabowo di kursi presiden, bawa transparansi lebih baik dalam pemerintahan #100HariKerjaNyata"
2025-01-27T13:54:29.000Z	Media Indonesia	"INDIKATOR Politik Indonesia melakukan survei terkait kepuasan publik atas 100 hari kinerja Presiden Prabowo Subianto, termasuk program makan bergizi ar

Gambar 4. 9 File CSV

Setelah proses scraping dilakukan secara terpisah berdasarkan enam kata kunci utama yang berkaitan dengan program kerja 100 hari Presiden dan Wakil Presiden terpilih, seluruh file hasil scraping dikonsolidasikan ke dalam satu dataset gabungan. Proses ini dilakukan menggunakan Python dan library pandas, yang

secara otomatis membaca seluruh file .csv dalam satu folder dan menggabungkannya menjadi satu file .xlsx bernama combined_tweets.xlsx.

```
import pandas as pd
import os

# 1. Path ke folder tempat file CSV berada
folder_path = "tes" # Ubah ini ke path folder kamu

# 2. Ambil semua file CSV dari folder
csv_files = [f for f in os.listdir(folder_path) if f.endswith(".csv")]

# 3. Gabungkan semua CSV jadi satu DataFrame
dfs = []
for file in csv_files:
    file_path = os.path.join(folder_path, file)
    df = pd.read_csv(file_path)
    dfs.append(df)

combined_df = pd.concat(dfs, ignore_index=True)

# 4. Simpan hasil gabungan ke Excel
output_file = "combined_tweets.xlsx"
combined_df.to_excel(output_file, index=False, engine="openpyxl")

print(f"✅ {len(csv_files)} file CSV berhasil digabung dan disimpan sebagai {output_file}")
```

Gambar 4. 10 Code

Penggabungan ini bertujuan untuk menyatukan semua data mentah agar mempermudah proses preprocessing dan analisis pada tahap selanjutnya. Hasil akhir dari proses ini berjumlah 5.136 data tweet, yang masing-masing berisi informasi tanggal postingan, nama pengguna, dan isi tweet.

Tanggal	Username	Tweet
2024-12-1	POLRESTA	Polresta Barelang Ungkap Kasus Judi Online, Dukung Program 100 Hari Asta Cita Presiden RI
2024-12-1	Sleepyy	Satu lagi, jangan suka buat klaim sendiri "sekarabg pada nyeseff", belum 100 hari, belum setahun yang baru terpilih menjabat jd presiden, wapresnya mau diapain terserah
2024-12-1	Santorini	UJIAN PERTAMA bagi pemerintahan Presiden @prabowo di 100 hari pertamanya adalah ini. Meski kita tau kalau ini adalah amanat UU dari pemerintahan sebelumnya & sdh disetujui
2024-12-1	lawren sis	Dimana wibawa penegak hukum dan pemerintahan presiden Prabowo ini ??? Keriminalisasi terjadi di pemerintahannya 100 hari ? Berarti kekuasaan Habib lebih tinggi dari hukum In
2024-12-1	jamil	sehu sumpah sedih dan marah banget sama kebijakan presiden yang belum 100 hari ini
2024-12-1	Sinta	Voters 02....Tolong kondisikan. 100 hari kerja janji kampanye dah dilanggar. Presiden blunder dan kena sindiran pers luar... Surplus demografi bakal jadi beban kalo gini mulu anjir
2024-12-1	aldho	dengerin tuh Pak Prabowo Yth...bpk sampai dikatain ma rakyatnya serius gak sih jadi presiden? wajar rakyat ngomong begitu karena hampir 100 hari, bpk tdk ada langkah yang bikin
2024-12-1	Miss Twee	Tidak suka Jokowi hari ini , berarti 2014 2019 harusnya Prabowo jadi Presiden, hari ini Prabowo Presiden ((Belum 100 Hari) Menyalahkan pemilih atas kebijakan Penguasa yang dian
2024-12-1	Ayah Dien	http://kaberehnews.com/2024/12/giat-sosial-satpolairud-polres-aceh.html...Giat Sosial Satpolairud Polres Aceh Timur Bagikan Sembako, Dukung Program 100 Hari Pertama Kinerja P
2024-12-1	humasace	DUKUNG PROGRAM 100 HARI PERTAMA KINERJA PRESIDEN, SATPOLAIRUD POLRES ACEH TIMUR BAGIKAN SEMBAKO Kegiatan ini sebagai bentuk upaya Polri dalam Mendukung Progra
2024-12-1	Repsus.co	Dukung Program 100 Hari Pertama Kinerja Presiden, Satpolairud Polres Aceh Timur Bagikan Sembako
2024-12-1	KRITIKPED	Secara kasat mata presiden @prabowo masih bersama jkw tapi kita liat setelah 100 hari, Macan tuh kan ga bisa dikendalikan, walau banyak oligarch ngelilingi jkw, tetep presiden ga ak
2024-12-1	Salam Ind	Kalau aku jadi Opa wowo sdh terlaksana ini 100 hari kerja. Jadi presiden.
2024-12-1	aku kamu	Presiden" lain pas naik langsung bikin gebrakan nguatin rupiah dgn nurunin harga dollar, ngasih insentif, ngasih bansos. Krn program kerja 100 hari biasanya masa "honeymoon" per
2024-12-1	cinta kead	Gebrakan 100 hari pertama jadi Presiden ya cuma PPN 12% ini. Kirain dpt meringankan beban rakyat ga tau ditambah lg bebannya buat gaji n fasilitas kabinet gemuk n program mak
2024-12-1	KJA SS Off	Gesss....Menjelang 100 hari kepemimpinan Presiden @prabowo, gie kasih RAPOR MERAH, Why ? Salah satunya adalah kenaikan Gaji guru tidak sesuai dgn harapan, POVUSD tembus
2024-12-1	AbdiNata	Presiden @prabowo sebaiknya jk tdk tll urgent, tetap tinggal di dagri sebab di 100 hari pemerintahannya hrsnya segera meletakkan kembali dasar2 bernegara kt sesuai dg konstitusi.
2024-12-1	nolkek air	Kankek Air Ratu AKP S TAMRIINAN SH dan Anazonta Mengunjungi Petani Cahai dan Labung di Desa Sei Alim Ilmu Ker Air Ratu Kah Acahan Dlm Ranaka Mendukung Program 100 Har

Gambar 4. 11 Merged Data

4.3 Preprocessing Data

Preprocessing merupakan tahap krusial dalam pengolahan data teks sebelum dilanjutkan ke proses analisis lebih lanjut, terutama ketika data diperoleh dari platform media sosial seperti Twitter/X. Data mentah dari Twitter/X umumnya mengandung banyak noise atau gangguan, seperti tautan URL, mention (@username), hashtag (#), simbol-simbol khusus, angka, serta penggunaan kata-kata tidak baku atau singkatan yang umum dalam komunikasi informal. Jika tidak dibersihkan, elemen-elemen ini dapat mengganggu kinerja algoritma machine learning dan menurunkan akurasi model dalam melakukan klasifikasi.

Oleh karena itu, dilakukan serangkaian proses preprocessing untuk mengubah data mentah menjadi bentuk yang lebih terstruktur dan siap dianalisis. Proses ini meliputi beberapa langkah utama, yaitu: pembersihan teks (cleaning) untuk menghapus elemen-elemen tidak relevan, tokenisasi untuk memisahkan teks menjadi unit-unit kata, stemming untuk mengembalikan kata ke bentuk dasarnya, serta penghapusan duplikat dan entri kosong yang tidak memberikan kontribusi informasi. Tahapan ini tidak hanya meningkatkan kualitas data, tetapi juga menjadi fondasi penting agar proses vektorisasi dan pembelajaran mesin dapat berjalan secara optimal dan efisien.

4.3.1 Cleaning Text

Langkah awal dari preprocessing adalah membersihkan teks dari elemen-elemen yang tidak relevan. Teks tweet dibersihkan menggunakan regular expression untuk menghapus:

- URL (tautan) seperti <https://...>

- Mention (penyebutan pengguna) seperti @username
- Hashtag seperti #pemilu2024
- Simbol dan angka, hanya menyisakan huruf dan spasi
- Konversi teks ke huruf kecil untuk menjaga konsistensi data

```
0 polresta barelang ungkap kasus judi online duk...
1 satu lagi jangan suka buat klaim sendiri sekar...
2 uji pertama bagi perintah presiden di hari per...
3 mana wibawa tegak hukum dan perintah presiden ...
4 sumpah sedih dan marah banget sama bijak presi...
5 voters tolong kondisi hari kerja janji kampany...
6 dengerin tuh pak prabowo yth bpk sampai dikata...
7 tidak suka jokowi hari ini arti harus prabowo ...
8 giat sosial satpolairud polres aceh timur bagi...
9 dukung program hari pertama kerja presiden sat...
Name: Stemmed_Tweet, dtype: object
```

Gambar 4.12 Head Data Cleanning Text

Langkah ini diimplementasikan dalam fungsi `preprocess_text()` yang diterapkan ke seluruh kolom Tweet dan hasilnya disimpan pada kolom baru bernama `Cleaned_Tweet`. Dengan membersihkan teks dari elemen-elemen tersebut, data menjadi lebih siap untuk dianalisis dan tidak bias oleh unsur-unsur teknis dari platform Twitter.

4.3.2 Tokenisasi dan Stemming

Setelah teks tweet dibersihkan dari elemen-elemen seperti URL, mention, hashtag, simbol, dan angka, langkah selanjutnya adalah melakukan proses tokenisasi dan stemming. Tahapan ini penting dalam preprocessing teks karena bertujuan untuk menyederhanakan kata menjadi bentuk dasarnya serta

mempermudah proses ekstraksi fitur dan klasifikasi pada tahap selanjutnya. Tokenisasi adalah proses memecah satu kalimat atau teks menjadi bagian-bagian terkecil yang disebut token, biasanya berupa kata per kata. Misalnya, kalimat "makan siang gratis sudah dimulai" akan diubah menjadi daftar token: ["makan", "siang", "gratis", "sudah", "dimulai"]. Stemming, di sisi lain, merupakan proses

```
Hasil Tokenisasi dan Stemming:

Cleaned_Tweet \
0 polresta barelang ungkap kasus judi online duk...
1 satu lagi jangan suka buat klaim sendiri sekar...
2 ujian pertama bagi pemerintahan presiden \n\n ...
3 dimana wibawa penegak hukum dan pemerintahan p...
4 sumpah sedih dan marah banget sama kebijakan p...

Original_Tokens \
0 [polresta, barelang, ungkap, kasus, judi, onli...
1 [satu, lagi, jangan, suka, buat, klaim, sendir...
2 [ujian, pertama, bagi, pemerintahan, presiden,...
3 [dimana, wibawa, penegak, hukum, dan, pemerint...
4 [sumpah, sedih, dan, marah, banget, sama, kebi...

Stemmed_Tokens \
0 [polresta, barelang, ungkap, kasus, judi, onli...
1 [satu, lagi, jangan, suka, buat, klaim, sendir...
2 [uji, pertama, bagi, perintah, presiden, di, h...
3 [mana, wibawa, tegak, hukum, dan, perintah, pr...
4 [sumpah, sedih, dan, marah, banget, sama, bija...

Stemmed_Tweet
0 polresta barelang ungkap kasus judi online duk...
1 satu lagi jangan suka buat klaim sendiri sekar...
2 uji pertama bagi perintah presiden di hari per...
3 mana wibawa tegak hukum dan perintah presiden ...
4 sumpah sedih dan marah banget sama bijak presi...
```

Gambar 4. 13 Hasil Tokenisasi & Stemming

mengembalikan setiap token ke bentuk dasarnya (root word).

Output tersebut menunjukkan:

- Cleaned_Tweet: teks setelah dibersihkan
- Original_Tokens: daftar kata asli setelah tokenisasi
- Stemmed_Tokens: kata-kata dasar hasil stemming
- Stemmed_Tweet: hasil penggabungan kembali kata dasar menjadi kalimat yang siap untuk proses selanjutnya

Dengan menggunakan pendekatan ini, model klasifikasi nantinya akan bekerja dengan teks yang telah dinormalisasi dan lebih terstruktur, sehingga performa dalam membaca dan memahami pola sentimen di dalam tweet akan lebih optimal.

4.3.3 Penghapusan Duplikat Data

Setelah proses tokenisasi dan stemming, langkah selanjutnya adalah menghapus tweet yang bersifat duplikat atau spam. Penghapusan ini penting untuk menjaga kualitas data latih dan uji yang digunakan dalam proses klasifikasi. Banyaknya tweet yang memiliki struktur teks serupa setelah diproses, seperti hasil retweet tanpa komentar atau konten viral yang disalin ulang oleh banyak akun, dapat menyebabkan data menjadi tidak representatif jika dibiarkan.

Dalam penelitian ini, penghapusan dilakukan berdasarkan kolom `Stemmed_Tweet`, yaitu bentuk akhir teks setelah melewati proses pembersihan, tokenisasi, dan stemming. Jika terdapat lebih dari satu baris dengan nilai `Stemmed_Tweet` yang sama, maka hanya satu yang dipertahankan. Selain itu, sistem juga memfilter kembali data yang kosong atau tidak memiliki konten setelah proses pembersihan dilakukan.

```
Jumlah data setelah penghapusan duplikat: 4631
Contoh data setelah pembersihan duplikat dan baris kosong:
                                Stemmed_Tweet
0 polresta barelang ungkap kasus judi online duk...
1 satu lagi jangan suka buat klaim sendiri sekar...
2 uji pertama bagi perintah presiden di hari per...
3 mana wibawa tegak hukum dan perintah presiden ...
4 sumpah sedih dan marah banget sama bijak presi...
```

Gambar 4. 14 Hasil Penghapusan duplikat data

Langkah ini menghasilkan dataset yang lebih bersih, ringkas, dan siap digunakan untuk proses pelabelan serta pelatihan model klasifikasi. Dengan menghindari data yang berulang dan tidak bermakna, akurasi dan generalisasi model yang dibangun dapat meningkat secara signifikan.

4.3.4 Labelling Sentimen

Proses pelabelan sentimen dilakukan untuk mengklasifikasikan setiap tweet ke dalam kategori sentimen tertentu, yaitu positif, negatif, atau netral. Dalam penelitian ini, metode lexicon-based digunakan untuk memberikan label secara otomatis berdasarkan kata-kata yang terkandung dalam setiap tweet.

Lexicon-based sentiment analysis adalah pendekatan berbasis kamus yang menggunakan daftar kata (lexicon) yang telah diberi nilai polaritas (positif atau negatif) untuk menghitung skor sentimen suatu teks. Dalam kasus ini, dua file kamus digunakan: `positive.tsv` untuk kata-kata bernilai positif dan `negative.tsv` untuk kata-kata bernilai negatif. Masing-masing kata dalam kamus tersebut diberi bobot tertentu yang mencerminkan tingkat kekuatan sentimennya.

```
Contoh hasil labelling sentimen (dengan label teks):  
                                Stemmed_Tweet  Sentiment  Sentimen_Label  
0  polresta barelang ungkap kasus judi online duk...      2      Positif  
1  satu lagi jangan suka buat klaim sendiri sekar...      0      Negatif  
2  uji pertama bagi perintah presiden di hari per...      0      Negatif  
3  mana wibawa tegak hukum dan perintah presiden ...      0      Negatif  
4  sumpah sedih dan marah banget sama bijak presi...      0      Negatif
```

Gambar 4. 15 Labelling Sentimen

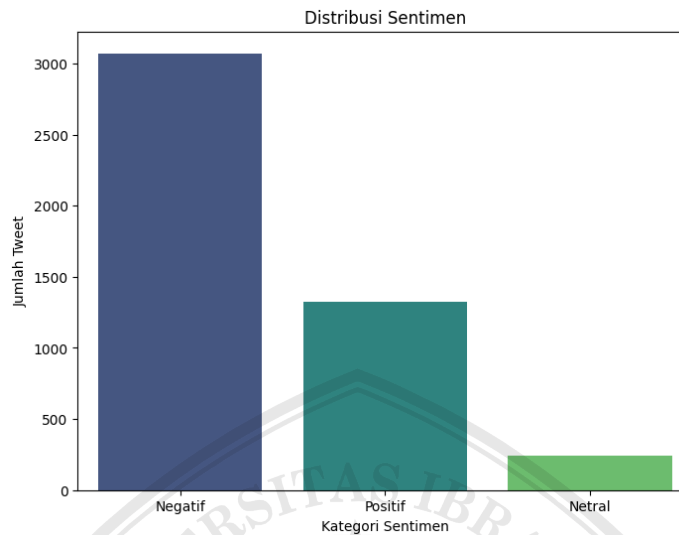
Setiap tweet, setelah melalui proses tokenisasi dan stemming, dipecah menjadi daftar kata dasar. Sistem kemudian memeriksa setiap kata dalam tweet tersebut terhadap kamus sentimen yang telah disusun sebelumnya. Jika sebuah kata ditemukan dalam kamus, maka skor sentimen tweet akan disesuaikan—bertambah jika kata tersebut bernuansa positif, atau berkurang jika kata tersebut bernuansa negatif—sesuai dengan bobot nilai yang telah ditetapkan untuk masing-masing kata. Hasil akhirnya adalah skor total sentimen dari masing-masing tweet yang mencerminkan kecenderungan opini, apakah cenderung positif, negatif, atau netral. Hasil akhirnya adalah:

- Skor positif → Label 2 (Positif)
- Skor negatif → Label 0 (Negatif)
- Skor netral (0) → Label 1 (Netral)

Dengan metode ini, proses pelabelan dapat dilakukan secara efisien tanpa intervensi manual, namun tetap menghasilkan dataset yang siap digunakan untuk pelatihan model klasifikasi seperti SVM. Pendekatan ini juga cocok untuk data sosial media seperti Twitter yang memiliki bahasa informal dan tidak baku.

Setelah proses pelabelan sentimen dilakukan menggunakan pendekatan berbasis kamus (lexicon-based), data kemudian dianalisis untuk mengetahui bagaimana persebaran sentimen publik terkait program 100 hari kerja Presiden dan Wakil Presiden Indonesia terpilih. Terdapat tiga kategori sentimen yang digunakan

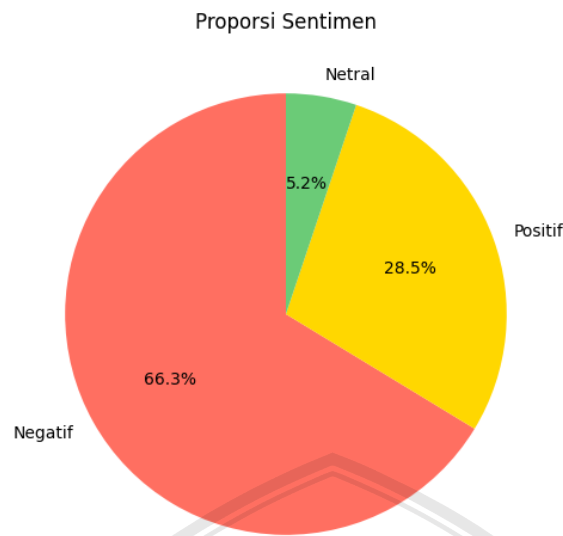
dalam penelitian ini, yaitu positif, netral, dan negatif, yang masing-masing direpresentasikan dengan label 2, 1, dan 0.



Gambar 4. 16 Diagram Batang Distribusi Sentimen

Visualisasi distribusi ini dilakukan dengan menggunakan diagram batang dan pie chart untuk menggambarkan jumlah serta proporsi masing-masing kategori sentimen—positif, negatif, dan netral—terhadap total 5.136 data tweet yang berhasil dikumpulkan. Hasil visualisasi menunjukkan bahwa masing-masing kategori sentimen memiliki representasi yang berbeda-beda, sehingga memberikan gambaran awal mengenai kecenderungan opini publik terhadap program kerja 100 hari pertama pemerintahan Prabowo-Gibran. Hasil visualisasi menunjukkan bahwa:

- Sentimen negatif mendominasi dengan jumlah tweet mencapai 66.3% dari total data.
- Sentimen positif berada pada posisi kedua dengan 28.5%.
- Sentimen netral hanya menyumbang sekitar 5.2%.



Gambar 4. 17 Diagram Distribusi Sentimen

Temuan ini mengindikasikan bahwa sebagian besar masyarakat menyuarakan opini yang cenderung negatif terhadap pelaksanaan program-program awal pemerintah, termasuk "makan siang gratis", "swasembada", serta kinerja "kabinet Indonesia" selama periode 100 hari kerja. Namun demikian, tidak sedikit pula pengguna Twitter/X yang memberikan apresiasi dan opini positif, yang tercermin dalam proporsi sentimen positif yang cukup signifikan.

4.3.5 TF-IDF Vectorization

Setelah proses preprocessing dan pelabelan sentimen selesai dilakukan, langkah selanjutnya adalah mengubah data teks menjadi representasi numerik agar dapat diproses oleh algoritma klasifikasi mesin, khususnya Support Vector Machine (SVM). Dalam penelitian ini, metode TF-IDF (Term Frequency – Inverse Document Frequency) digunakan sebagai teknik ekstraksi fitur utama. TF-IDF bekerja dengan menghitung bobot pentingnya suatu kata dalam sebuah dokumen relatif terhadap seluruh korpus dokumen. Semakin sering sebuah kata muncul

dalam satu tweet tetapi jarang muncul di tweet lain, maka bobotnya akan semakin tinggi. Dengan kata lain, TF-IDF dapat membantu menekankan kata-kata yang memiliki karakteristik khas dari sebuah opini atau sentimen tertentu.

Proses TF-IDF dilakukan menggunakan pustaka *TfidfVectorizer* dari Scikit-

```

akankah ambisius anaknya bagaikan bapaknya baru bermata \
0 0.132453 0.132453 0.132453 0.132453 0.132453 0.132453 0.132453

    besar    bumn  danantara  ...  saksikan  selengkapnya  senilai  \
0 0.264906 0.132453 0.132453  ...  0.132453    0.132453 0.132453

    seperti  siang  sungguh  ugalugalan  wkwk  yang  yg
0 0.132453 0.132453 0.132453 0.132453 0.132453 0.132453 0.39736

[1 rows x 46 columns]
```

Gambar 4. 18 TF-IDF vectorization

learn dengan jumlah maksimum fitur dibatasi hingga 5000 kata unik (vocabulary) untuk menjaga efisiensi dan menghindari dimensi data yang terlalu tinggi.

4.4 Implementasi SVM

4.4.1 Splitting Data

Setelah melalui tahapan preprocessing dan pelabelan sentimen, langkah selanjutnya dalam proses analisis sentimen adalah melakukan pembagian data (data splitting). Tujuan utama dari tahap ini adalah untuk memisahkan dataset menjadi dua bagian utama, yaitu data latih (training set) dan data uji (testing set). Data latih digunakan untuk melatih model, di mana algoritma klasifikasi akan mempelajari pola-pola yang terkandung dalam data, seperti hubungan antara fitur-fitur teks (misalnya frekuensi kata, tokenisasi, atau vektor representasi) dengan label sentimen yang sesuai.

Sementara itu, data uji berfungsi untuk mengevaluasi kinerja model terhadap data yang belum pernah dilihat sebelumnya selama proses pelatihan. Hal ini penting dilakukan untuk mengetahui seberapa baik model dapat melakukan generalisasi terhadap data baru yang belum dikenal, serta menghindari permasalahan overfitting, yaitu ketika model terlalu menyesuaikan diri dengan data latih namun gagal memberikan prediksi yang akurat pada data uji. Oleh karena itu, pembagian data merupakan langkah krusial dalam membangun sistem klasifikasi sentimen yang andal dan dapat diandalkan dalam berbagai skenario aplikasi dunia nyata, seperti analisis opini publik, pemantauan media sosial, dan pengambilan keputusan berbasis data.

```
Jumlah Data Total: 4631
Jumlah Data Latih: 3704
Jumlah Data Uji: 927

Contoh data latih (label):
713      2
1917     0
4275     1
4566     1
4237     2
Name: Sentiment, dtype: int64

Bentuk matriks TF-IDF:
Train: (3704, 5000) | Test: (927, 5000)
```

Gambar 4. 19 Hasil Splitting Data

Pada penelitian ini, pembagian dilakukan dengan perbandingan 80% data latih dan 20% data uji, menggunakan fungsi `train_test_split` dari library `scikit-learn`. Proporsi ini dipilih karena merupakan standar umum yang cukup untuk menghasilkan model yang stabil dan dapat diandalkan. Selain itu, parameter

stratify=y digunakan untuk memastikan bahwa distribusi kelas sentimen (positif, negatif, dan netral) tetap seimbang di kedua subset data. Dengan kata lain, proses ini menjaga agar proporsi label dalam data latih dan data uji mencerminkan distribusi keseluruhan dataset. Secara keseluruhan, dataset yang digunakan dalam penelitian ini terdiri dari 4.631 tweet, sehingga menghasilkan:

- 3.704 tweet sebagai data latih
- 927 tweet sebagai data uji

4.4.2 Klasifikasi Support Vector Machine

Dalam penelitian ini, metode klasifikasi yang digunakan adalah Support Vector Machine (SVM), salah satu algoritma pembelajaran mesin yang terkenal efektif dalam menangani permasalahan klasifikasi teks, termasuk analisis sentimen. SVM bekerja dengan mencari hyperplane terbaik yang mampu memisahkan data ke dalam kelas-kelas berbeda berdasarkan fitur yang telah diekstraksi. Dataset yang telah diproses terdiri dari 4.631 tweet yang kemudian dibagi menjadi dua bagian: 80% untuk data latih dan 20% untuk data uji. Proses pembagian ini dilakukan secara acak namun konsisten menggunakan parameter `random_state`, guna memastikan reproduktibilitas hasil. Model SVM dilatih menggunakan kernel linear, yang umumnya cocok untuk data teks karena dapat memisahkan kelas dalam ruang dimensi tinggi secara efisien. Setelah pelatihan, model dievaluasi menggunakan

data uji, dan hasil prediksi dibandingkan dengan label asli untuk mengukur performa model.

```
Akurasi Model: 0.85
Laporan Klasifikasi:
      precision    recall  f1-score   support

Negatif    0.86    0.95    0.90     616
Netral     0.00    0.00    0.00      42
Positif    0.83    0.75    0.79     269

accuracy          0.85     927
macro avg    0.56    0.57    0.57     927
weighted avg 0.81    0.85    0.83     927
```

Gambar 4. 20 Implementasi Algoritma SVM

Berdasarkan hasil evaluasi, model SVM menunjukkan kinerja yang baik dengan tingkat akurasi sebesar 85%. Kategori sentimen negatif memiliki performa terbaik dengan precision 0.86, recall 0.95, dan F1-score 0.90. Kategori positif juga cukup baik dengan F1-score sebesar 0.79. Namun, kategori netral tidak terdeteksi secara akurat oleh model karena jumlah datanya yang sangat kecil, sehingga menghasilkan nilai 0 pada seluruh metrik evaluasi. Secara keseluruhan, model berhasil memetakan sentimen publik dengan cukup baik. Hasil ini menunjukkan bahwa SVM merupakan algoritma yang tepat dalam menangani data teks berbahasa Indonesia, terutama untuk klasifikasi opini publik yang bersumber dari media sosial seperti Twitter/X.

Tanggal	Username	Tweet	anned_Tw	med_Tw	med_Tok	inal_Tok	Sentiment	ted_Sentim
2025-01-1	Indonesia	Lobi besar	lobi besar	lobi besar	['lobi', 'be:	['lobi', 'be:	0	0
2024-12-2	Al-Biruni.l	wkwk yg n	wkwk yg n	wkwk yg a	['wkwk', 'y	['wkwk', 'y	0	0
2025-01-2	Abdurrauf	Kasus pen	kasus pen	kasus gela	['kasus', 'g	['kasus', 'p	0	0
2025-01-1	A.S	Apaan kor	apaan kon	apa konto	['apa', 'ko	['apaan', 'l	0	0
2024-12-1	Anwar Re	belum ger	belum ger	belum ger	['belum', 'l	['belum', 'l	0	0
2025-01-1	LKMN	Penjelas	penjelas	jelas panj	['jelas', 'p	['penjelas	0	0
2025-01-2	Adhi_Bek:	Waduh lag	waduh lag	waduh lag	['waduh', 'l	['waduh', 'l	0	0
2025-01-1	•	"Makan Si	makan sia	makan sia	['makan', 'l	['makan', 'l	2	2
2025-01-2	Friren An	Kalo buat	kalo buat	kalo buat	['kalo', 'bu	['kalo', 'bu	2	2
2024-11-1	Polres Tul:	SOSIALISA	sosialisasi	sosialisasi	['sosialisa:	['sosialisa:	0	0
2025-01-2	SIMOh	Sebenarny	sebenarny	benar beli	['benar', 't	['sebenarr	0	0
2025-01-2	Seputar B	100 hari k	hari kerja	hari kerja	['hari', 'ke	['hari', 'ke	0	2
2025-01-1	dian	Program n	program r	program r	['program	['program	2	2
2025-01-1	Jiheon	JOROK, mi	rorok mini	rorok mini	['rorok', 'r	['rorok', 'r	0	0
2025-01-2	Reynaldi P	Mantap jiv	mantap jiv	mantap jiv	['mantap', 'l	['mantap', 'l	2	2
2025-01-2	20DETIK	Video: Ses	video sesu	video sesu	['video', 's	['video', 's	0	0
2025-01-1	?????	Saya mau	saya mau	saya mau	['saya', 'm	['saya', 'm	1	0
2024-11-0	wawa	Impact da	impact da	impact da	['impact', 'l	['impact', 'l	2	2

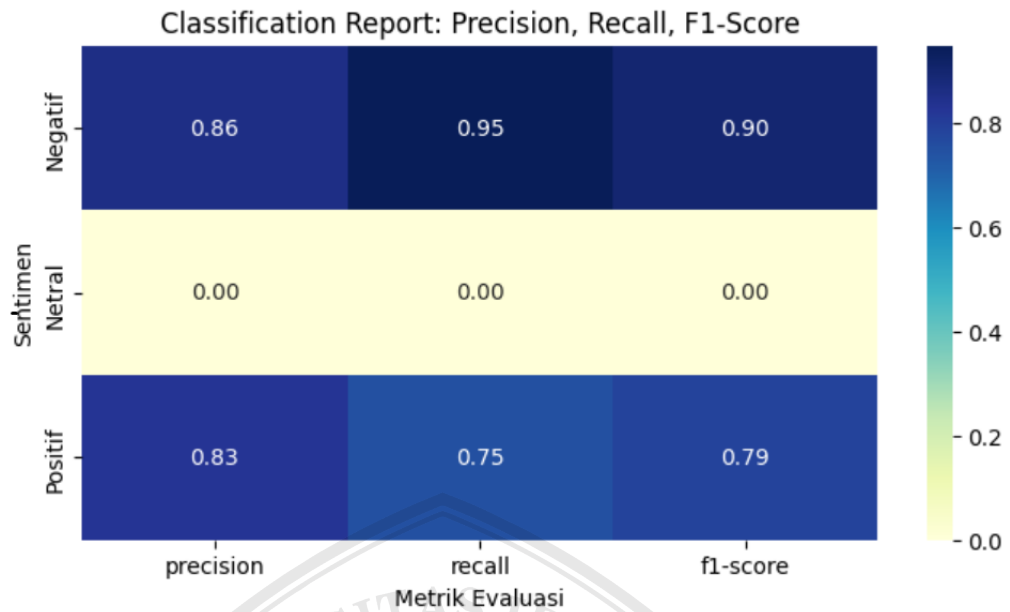
Gambar 4. 21 Output Implementasi SVM

4.4.3 Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja algoritma klasifikasi Support Vector Machine (SVM) dalam mengelompokkan data tweet ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Evaluasi ini mencakup metrik akurasi, precision, recall, dan fl-score, serta visualisasi dalam bentuk confusion matrix dan heatmap classification report.

a. Akurasi Model

Berdasarkan hasil prediksi pada data uji sebanyak 927 tweet (yang merupakan 20% dari total 4.631 tweet), model menghasilkan akurasi sebesar 0,85 atau 85%. Nilai akurasi ini menunjukkan bahwa model berhasil memprediksi label sentimen dengan benar pada sebagian besar data uji, yang menandakan bahwa performa model cukup andal dan layak digunakan untuk analisis sentimen lebih lanjut terhadap data serupa.



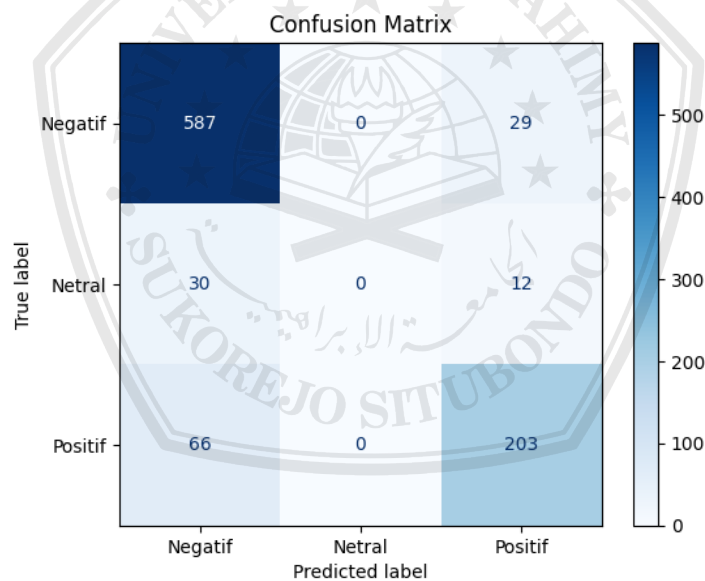
Gambar 4. 22 Akurasi Model

Dari gambar di atas dapat dijelaskan bahwa:

- Precision untuk kelas negatif adalah 0.86, artinya 86% dari prediksi negatif benar-benar negatif.
- Recall untuk kelas negatif mencapai 0.95, menunjukkan hampir semua tweet negatif berhasil diklasifikasikan dengan benar.
- Untuk sentimen netral, model gagal dengan precision, recall, dan f1-score sebesar 0.00, kemungkinan karena ketidakseimbangan data atau fitur yang kurang representatif.
- Kelas positif memiliki precision 0.83, recall 0.75, dan f1-score 0.79, yang menunjukkan performa cukup baik meski masih ada beberapa tweet positif yang salah klasifikasi.

b. Confusion Matrix

Confusion matrix menunjukkan bahwa dari 616 tweet negatif, sebanyak 587 berhasil diprediksi dengan benar, namun 29 salah diklasifikasikan sebagai positif. Untuk tweet netral, hanya 12 yang berhasil dikenali sebagian, sedangkan sebagian besar lainnya (30 tweet) salah diprediksi. Sementara dari 269 tweet positif, sebanyak 203 berhasil diklasifikasikan dengan tepat, sedangkan 66 salah diklasifikasikan sebagai negatif.



Gambar 4. 23 Confusion Matrix

4.5 Pembahasan Hasil

4.5.1 Interpretasi Sentimen

Berdasarkan hasil klasifikasi terhadap 5.136 tweet yang diperoleh dari proses scraping Twitter/X selama periode 20 Oktober 2024 hingga 28 Januari 2025, persebaran sentimen masyarakat terhadap program-program 100 hari kerja Presiden

Prabowo Subianto dan Wakil Presiden Gibran Rakabuming Raka terbagi ke dalam tiga kategori utama, yaitu negatif, positif, dan netral.

a. Sentimen Negatif

Tweet dengan sentimen negatif mendominasi data yang dikumpulkan, dengan total sekitar 3.000 lebih tweet, atau setara dengan lebih dari 60% dari keseluruhan data. Dominasi sentimen negatif ini menunjukkan bahwa mayoritas masyarakat pengguna media sosial memberikan respons yang kritis terhadap jalannya pemerintahan di awal masa kepemimpinan. Kritik ini dapat mencakup isu-isu seperti:

- Pelaksanaan Program Makan Siang Gratis (MBG) yang dianggap belum merata atau masih sebatas retorika,
- Kebijakan swasembada pangan yang dianggap ambisius tanpa roadmap yang jelas

Serta berbagai komentar terhadap komposisi kabinet dan gaya komunikasi pemerintah yang dinilai belum memenuhi ekspektasi publik. Ciri-ciri tweet negatif umumnya mengandung kata-kata sindiran, ekspresi pesimis, hingga pernyataan frontal terhadap kebijakan maupun sosok Presiden dan Wakil Presiden.

b. Sentimen Positif

Sentimen positif mencakup sekitar 25%–30% dari keseluruhan tweet yang dianalisis. Tweet dalam kategori ini umumnya menunjukkan dukungan dan harapan terhadap realisasi janji kampanye Prabowo-Gibran, khususnya terhadap program-program unggulan seperti:

- Program Makan Siang Gratis (MBG) yang dinilai sebagai langkah nyata untuk mengatasi stunting dan meningkatkan kesejahteraan siswa,
- Rencana swasembada pangan, yang dipuji sebagai visi jangka panjang untuk ketahanan nasional
- beberapa pujian terhadap gaya kepemimpinan tegas atau figur muda dalam kabinet.

Tweet positif biasanya mengandung kosakata bernada optimis, seperti hebat, kerja nyata, terbaik, atau lanjutkan, dan cenderung disebarakan ulang melalui retweet atau like secara massif oleh simpatisan.

c. Sentimen Netral

Sentimen netral menempati porsi paling kecil dalam distribusi data, yakni kurang dari 10% dari total keseluruhan dataset. Tweet-tweet yang termasuk dalam kategori ini umumnya bersifat informatif, objektif, atau deskriptif, tanpa menyiratkan emosi, penilaian, maupun opini pribadi yang kuat dari penulisnya. Konten dalam tweet netral biasanya hanya menyampaikan fakta atau peristiwa secara langsung, seperti jadwal peluncuran program kerja pemerintah, kutipan dari pidato Presiden atau Wakil Presiden, serta pemberitaan resmi dari media tanpa tambahan ekspresi dukungan maupun kritik. Kehadiran sentimen netral ini penting sebagai penyeimbang dalam analisis, karena mencerminkan adanya segmen masyarakat yang memilih untuk menyampaikan informasi apa adanya tanpa menunjukkan afiliasi emosional tertentu.

4.5.2 Evaluasi Kinerja Model

Setelah dilakukan pelatihan terhadap model Support Vector Machine (SVM) menggunakan fitur teks hasil transformasi TF-IDF dan label sentimen yang diperoleh dari pendekatan lexicon-based, model berhasil mencapai akurasi sebesar 85% pada data uji. Angka ini menunjukkan bahwa model memiliki kinerja yang cukup baik dalam mengklasifikasikan sentimen publik terhadap 100 hari kerja Presiden dan Wakil Presiden terpilih, terutama mengingat karakteristik data yang tidak terstruktur, bersifat informal, dan sangat kontekstual seperti yang umumnya ditemukan di media sosial Twitter/X.

a. Akurasi Model

Akurasi model yang mencapai angka 0,85 menunjukkan bahwa sebanyak 85% dari seluruh tweet dalam data uji berhasil diklasifikasikan ke dalam kategori sentimen yang sesuai, baik itu positif maupun negatif. Nilai akurasi yang tinggi ini mengindikasikan bahwa model memiliki kemampuan klasifikasi yang cukup andal secara keseluruhan. Meskipun terdapat ketidakseimbangan distribusi kelas dalam dataset—dengan dominasi sentimen negatif—model tetap mampu memberikan prediksi yang akurat. Hal ini mencerminkan bahwa model tidak terlalu bias terhadap satu kelas tertentu dan tetap mampu menangani tantangan dari data yang tidak seimbang.

b. Precision

Precision merupakan metrik yang menggambarkan sejauh mana model tepat dalam memberikan label pada suatu kelas, tanpa terlalu banyak menghasilkan

kesalahan berupa prediksi positif palsu (false positive). Pada model ini, nilai precision tertinggi dicapai oleh kelas sentimen Negatif, yaitu sebesar 0,86. Ini berarti bahwa sebagian besar tweet yang diprediksi oleh model sebagai negatif, memang benar-benar termasuk dalam kategori negatif. Dengan kata lain, model memiliki tingkat ketepatan yang sangat baik dalam mengidentifikasi tweet bernada negatif. Sementara itu, precision pada kelas Positif juga tergolong tinggi, yaitu sebesar 0,83, yang mengindikasikan bahwa model juga cukup akurat dalam mengenali dan mengklasifikasikan tweet dengan sentimen positif. Kedua nilai ini menunjukkan bahwa model memiliki keandalan yang baik dalam menghindari klasifikasi yang keliru.

c. Recall

Recall adalah ukuran seberapa baik model mampu mengenali semua data yang termasuk dalam suatu kelas tertentu. Dalam hal ini, recall tertinggi kembali ditemukan pada kelas Negatif, dengan nilai mencapai 0,95. Artinya, dari seluruh tweet negatif yang benar-benar ada dalam data uji, sebanyak 95% berhasil dikenali dan diklasifikasikan secara tepat oleh model. Ini menunjukkan bahwa model sangat efektif dalam mendeteksi tweet yang memiliki muatan sentimen negatif, yang sangat penting khususnya dalam konteks pemantauan opini publik, kritik, atau isu-isu sensitif. Untuk kelas Positif, nilai recall tercatat sebesar 0,75, yang berarti 75% dari seluruh tweet positif dalam data uji berhasil diklasifikasikan secara benar. Meskipun sedikit lebih rendah dibandingkan kelas negatif, nilai ini tetap menunjukkan bahwa performa model dalam mendeteksi tweet positif masih berada dalam kategori baik.

d. F1-Score

F1-score merupakan metrik evaluasi yang menggabungkan dua komponen penting, yaitu precision (ketepatan) dan recall (kelengkapan), sehingga memberikan gambaran yang lebih seimbang terhadap performa model klasifikasi, terutama ketika menghadapi dataset dengan distribusi kelas yang tidak merata. Metrik ini sangat berguna dalam analisis sentimen, karena mampu menangkap keseimbangan antara kemampuan model dalam mengenali sentimen yang relevan dan menghindari kesalahan klasifikasi. Pada hasil evaluasi model, nilai F1-score tertinggi tercatat pada kelas Negatif, yakni sebesar 0,90, yang menunjukkan bahwa model memiliki tingkat akurasi dan konsistensi yang sangat tinggi dalam mengenali dan memprediksi tweet dengan nuansa negatif. Sementara itu, kelas Positif memperoleh nilai F1-score sebesar 0,79, yang meskipun lebih rendah, tetap mencerminkan performa yang baik dan stabil dalam mendeteksi sentimen positif. Secara keseluruhan, nilai F1-score yang cukup tinggi pada kedua kelas utama ini mengindikasikan bahwa model mampu melakukan klasifikasi sentimen secara efektif dan konsisten, sehingga layak digunakan dalam analisis lanjutan untuk menggambarkan persepsi publik secara lebih akurat.

4.5.3 Pembahasan Hasil

	PREDIKSI NEGATIF	PREDIKSI NETRAL	PREDIKSI POSITIF
ASLI NEGATIF	587	7	29
ASLI NETRAL	17	12	12
ASLI POSITIF	66	4	193

Tabel 1 Hasil Prediksi Model

Analisis confusion matrix merupakan metode yang efektif untuk memahami performa klasifikasi model tidak hanya secara agregat (melalui metrik seperti akurasi), tetapi juga secara rinci berdasarkan jumlah prediksi benar dan salah antar kategori. Dalam konteks penelitian ini, confusion matrix digunakan untuk mengukur kemampuan model Support Vector Machine (SVM) dalam mengklasifikasikan sentimen tweet ke dalam tiga kelas: Negatif (0), Netral (1), dan Positif (2). Berdasarkan hasil klasifikasi terhadap 927 tweet pada data uji, confusion matrix menunjukkan performa sebagai berikut:

a. Kelas Negatif (0)

Model menunjukkan kinerja yang sangat kuat dalam mengenali tweet negatif:

- Sebanyak 587 tweet negatif berhasil diklasifikasikan dengan benar (True Negative).
- Hanya terdapat 7 kesalahan klasifikasi ke netral dan 29 ke positif.

- Tingginya akurasi ini sejalan dengan tingginya nilai recall dan precision yang diperoleh pada evaluasi sebelumnya.

b. Kelas Netral (1)

Kategori netral adalah yang paling menantang bagi model:

- Dari seluruh tweet netral pada data uji, hanya 12 tweet yang berhasil dikenali dengan benar.
- 17 tweet netral diklasifikasikan sebagai negatif, sedangkan 12 lainnya diklasifikasikan sebagai positif.
- Rendahnya akurasi ini menunjukkan bahwa tweet netral sulit dibedakan secara tekstual.

c. Kelas Positif (2)

Model memiliki performa yang cukup baik dalam mendeteksi tweet positif:

- Sebanyak 193 tweet positif berhasil diklasifikasikan dengan benar, namun.
- Terdapat 66 tweet positif yang justru diklasifikasikan sebagai negatif, yang menjadi sumber utama kesalahan klasifikasi.
- Selain itu, 4 tweet positif diklasifikasikan sebagai netral, yang bisa jadi akibat gaya bahasa ambigu atau netralitas ekspresi.

BAB V

PENUTUP

5.1 Kesimpulan

Penelitian ini dilakukan untuk menganalisis sentimen masyarakat terhadap 100 hari kerja pertama Presiden Prabowo Subianto dan Wakil Presiden Gibran Rakabuming Raka dengan memanfaatkan data dari media sosial Twitter/X. Data dikumpulkan menggunakan metode scraping berbasis Selenium dan ChromeDriver selama periode 20 Oktober 2024 hingga 28 Januari 2025, yang mencerminkan rentang waktu 100 hari pertama masa kepemimpinan mereka. Dari proses scraping yang dilakukan terhadap beberapa kata kunci yang mewakili topik-topik penting seperti “makan siang gratis”, “kabinet Indonesia”, “program kerja presiden”, hingga “swasembada”, berhasil dihimpun sebanyak 5.136 tweet dalam bahasa Indonesia.

Data yang diperoleh kemudian melalui tahapan preprocessing yang mencakup pembersihan teks dari unsur-unsur non-relevan seperti URL, simbol, angka, mention, dan hashtag, serta konversi ke huruf kecil. Setelah itu dilakukan tokenisasi dan stemming menggunakan pustaka Sastrawi, serta penghapusan duplikat dan baris kosong untuk menjamin kualitas data. Proses pelabelan sentimen dilakukan secara otomatis dengan pendekatan lexicon-based menggunakan kamus sentimen positif dan negatif, menghasilkan tiga kategori sentimen yaitu negatif, netral, dan positif. Hasil klasifikasi awal menunjukkan bahwa sentimen negatif mendominasi opini publik terhadap kinerja awal pemerintahan, disusul oleh sentimen positif, sementara jumlah sentimen netral relatif kecil.

Model klasifikasi yang digunakan dalam penelitian ini adalah Support Vector Machine (SVM) dengan kernel linear. Model ini menunjukkan performa yang cukup baik, dengan akurasi mencapai 85%, serta F1-score tertinggi pada kelas negatif yaitu 0.90. Precision dan recall juga tinggi pada kelas negatif, mengindikasikan bahwa model mampu mengidentifikasi kritik masyarakat dengan baik. Namun, performa pada kelas netral masih tergolong rendah, yang menunjukkan bahwa model kesulitan dalam mengenali teks yang bersifat informatif atau ambigu. Confusion matrix yang dihasilkan mendukung temuan ini, di mana sebagian besar kesalahan klasifikasi berasal dari salah interpretasi terhadap tweet netral dan positif.

5.2 Saran

Penelitian ini masih memiliki beberapa keterbatasan yang dapat menjadi acuan untuk pengembangan ke depan. Salah satunya adalah pendekatan pelabelan sentimen yang masih berbasis lexicon, yang kurang mampu menangkap konteks dan ironi dalam bahasa media sosial. Oleh karena itu, disarankan untuk penelitian selanjutnya menggunakan metode pelabelan manual berbasis crowdsourcing atau memanfaatkan model deep learning seperti LSTM atau BERT untuk klasifikasi yang lebih akurat. Selain itu, penggunaan platform selain Twitter/X juga dapat dipertimbangkan guna mendapatkan perspektif yang lebih luas, seperti YouTube atau TikTok yang kini juga banyak digunakan untuk mengekspresikan opini politik. Dari sisi teknis, proses scraping dapat dioptimalkan dengan menerapkan teknik anti-detection dan manajemen IP agar dapat menjangkau lebih banyak data dalam waktu yang lebih efisien. Diharapkan hasil penelitian ini dapat menjadi landasan

untuk membangun sistem pemantauan opini publik secara real-time yang dapat membantu pengambilan keputusan kebijakan publik secara responsif.



DAFTAR PUSTAKA

- [1] F. Solikhah, M. Febianah, A. L. Kamil, W. A. Arifin, and Shelly Janu Setyaning Tyas, “Analisis Perbandingan Algoritma Naive Bayes Dan C.45 Dalam Klasifikasi Data Mining Untuk Memprediksi Kelulusan,” *Tematik*, vol. 8, no. 1, pp. 96–103, 2021, doi: 10.38204/tematik.v8i1.576.
- [2] Nurhayati, *Pemodelan K- Means Algoritma Dan Big Data Analysis (Pemetaan Data Mustahiq)*. 2022.
- [3] Jimmy, E. H. Hermaliani, and L. Kurniawati, “Analisis Klasifikasi Sentimen Pengguna Media Sosial Twitter Terhadap Penundaan Pemilu Presiden Tahun 2024,” *J. Indones. Manaj. Inform. dan Komun.*, vol. 4, no. 2, pp. 570–579, 2023, doi: 10.35870/jimik.v4i2.243.
- [4] H. Alrasyid, A. Homaidi, M. Kom, Z. Fatah, and M. Kom, “Comparison Support Vector Machine and Random Forest Algorithms in Detect Diabetes,” vol. 1, no. 1, pp. 447–453, 2024.
- [5] A. Jalil, A. Homaidi, and Z. Fatah, “Implementasi Algoritma Support Vector Machine Untuk Klasifikasi Status Stunting Pada Balita,” *G-Tech J. Teknol. Terap.*, vol. 8, no. 3, pp. 2070–2079, 2024, doi: 10.33379/gtech.v8i3.4811.
- [6] T. Krisdiyanto, “Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers,” *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 7, no. 1, p. 32, 2021, doi: 10.24014/coreit.v7i1.12945.
- [7] D. F. Setiawan, A. Erlansari, and J. P. Sari, “Penerapan Data Mining pada Review TIX ID Menggunakan Naïve Bayes Berbasis SMOTE & PSO,” *J. Eksplora Inform.*, vol. 12, no. 1, pp. 37–45, 2023, doi: 10.30864/eksplora.v12i1.1034.
- [8] A. D. E. Rian, “ANALISIS SENTIMEN PADA MEDIA SOSIAL DENGAN MENGGUNAKAN ALGORITMA RADIAL BASIS FUNCTION NEURAL NETWORKS (RBFNN) SKRIPSI OLEH :

FAKULTAS TEKNIK UNIVERSITAS MEDAN AREA MEDAN ANALISIS SENTIMEN PADA MEDIA SOSIAL DENGAN MENGGUNAKAN ALGORITMA RADIAL BASIS FUNC,” 2024.

- [9] E. Suryati and A. A. Aldino, “Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma Support Vector Machine (SVM),” vol. 4, no. 1, pp. 96–106, 2023.
- [10] A. Mukti, A. D. Hadiyanti, A. Nurlaela, and J. Panjaitan, “Sistem Analisa Sentiment Bakal Calon Presiden 2024 Menggunakan Metode NLP Berbasis Web,” *Jurikom*, vol. 6, no. 1, p. p-ISSN, 2023.
- [11] Nabila and F. S. H. Nandeak, “Analysis of The Prabowo-Gibran Campaign Model in the 2024 Presidential and Vice Presidential Elections,” *Sociae Polites*, vol. 25, no. 2, pp. 34–49, 2024.
- [12] A. Wanto *et al.*, *Data Mining : Algoritma dan Implementasi*. 2023.
- [13] I. Werdiningsih, B. Nuqoba, and Muhammadun, *Data Mining Menggunakan Android, WEKA, dan SPSS*. 2020.
- [14] Ika Anikah, Agus Surip, Nela Puji Rahayu, Muhammad Harun Al- Musa, and Edi Tohidi, “Pengelompokan Data Barang Dengan Menggunakan Metode K-Means Untuk Menentukan Stok Persediaan Barang,” *KOPERTIP J. Ilm. Manaj. Inform. dan Komput.*, vol. 4, no. 2, pp. 58–64, 2022, doi: 10.32485/kopertip.v4i2.120.
- [15] S. S. Salim and J. Mayary, “Analisis Sentimen Pengguna Twitter Terhadap Dompot Elektronik Dengan Metode Lexicon Based Dan K – Nearest Neighbor,” *J. Ilm. Inform. Komput.*, vol. 25, no. 1, pp. 1–17, 2020, doi: 10.35760/ik.2020.v25i1.2411.
- [16] D. A. Alzahra, U. Enri, and Y. U. Maidah, “Analisis Sentimen Ulasan Pengguna Klik Indomaret Pada Google Play Menggunakan Support Vector Machine,” *Innov. J. Soc. Sci. Res.*, vol. 3, no. 4, pp. 2173–2185, 2023,

[Online]. Available: <https://jinnovative.org/index.php/Innovative/article/view/3715>

- [17] J. M. Polgan, S. Japit, Y. Risyani, T. Selamat, C. Bombongan, and S. V. Machine, “Deteksi Anomali Transaksi E-Commerce Menggunakan Support Vector Machine Berbasis Data Mining,” vol. 13, pp. 1976–1980, 2024.
- [18] T. Tukino and F. Fifi, “Penerapan Support Vector Machine Untuk Analisis Sentimen Pada Layanan Ojek Online,” *J. Desain Dan Anal. Teknol.*, vol. 3, no. 2, pp. 104–113, 2024, doi: 10.58520/jddat.v3i2.59.
- [19] A. Rahman and S. Y. Irianto, “Analisis Sentimen Pemilihan Calon Presiden Indonesia di Media Sosial Menggunakan Algoritma Support Vector,” vol. 1, no. 1, pp. 39–45, 2024.
- [20] L. Nursinggah, R. Ruuhwan, and T. Mufizar, “Analisis Sentimen Pengguna Aplikasi X Terhadap Program Makan Siang Gratis Dengan Metode Naïve Bayes Classifier,” *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4336.
- [21] T. A. Siddiq and M. Ikhsan, “Analisis Sentimen X Terhadap Pemilihan Presiden Indonesia 2024 dengan Metode K-Nearest Neighbor,” vol. 5, no. 4, pp. 1064–1078, 2024, doi: 10.47065/josyc.v5i4.5802.
- [22] I. A. Gellysa Urva, Desyanti, *PENERAPAN DATA MINING DI BERBAGAI BIDANG : Konsep, Metode, dan Studi Kasus*. 2023.
- [23] I. T. Julianto, D. Kurniadi, M. R. Nashrulloh, and A. Mulyani, “TWITTER SOCIAL MEDIA SENTIMENT ANALYSIS AGAINST BITCOIN ANALISIS SENTIMEN MEDIA SOSIAL TWITTER TERHADAP TRENDATA,” vol. 3, no. 5, pp. 1183–1187, 2022.
- [24] Fuad Amirullah, Syariful Alam, and M. Imam Sulistyono, “Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini Twitter Menggunakan Naïve Bayes,” *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 3, pp. 69–76, 2023, doi: 10.55123/storage.v2i3.2293.

- [25] Ismai, *Data Mining: Pengolahan Data Menjadi Informasi dengan RapidMiner*. 2017. [Online]. Available: https://www.google.co.id/books/edition/DATA_MINING/rTlmDwAAQB-AJ?hl=id&gbpv=0
- [26] A. R. Isnain *et al.*, “JAKARTA MENGGUNAKAN ALGORITMA SVM Media sosial menjadikan masyarakat mengalami pergeseran perilaku baik budaya , etika dan norma yang ada , sehingga mereka dapat mengeluarkan opini - opini yang mereka miliki . Opini merupakan suatu pendapat dari pemikiran,” vol. 2, no. 1, pp. 31–37, 2021.
- [27] A. Pajankar, P. Python, and D. Visualization, “Exploring Jupyter Notebook,” pp. 17–29.
- [28] H. S. D. Suparwito, R. (Universitas S. D. Gunawan, I. (Universitas S. D. Binanto, R. (Universitas S. D. Arum Kumalasanti, and W. (Universitas S. D. Widyastuti, *Pengantar Pembelajaran Mesin Menggunakan Bahasa Pemrograman Python*. 2023.
- [29] A. Wasik *et al.*, “Implementasi data mining untuk memprediksi penjualan accessoris handphone dan handphone terlaris menggunakan metode k-nearest neighbor (k-nn) 1,” vol. 1, no. 2, pp. 469–479, 2024.
- [30] Y. P. Dewi, Y. Abidin, and K. Kurniawan, “Penggunaan Twitter terhadap Pembelajaran Bahasa Indonesia,” *Semin. Int. Riksa Bhs.*, p. 307, 2021, [Online]. Available: <http://proceedings.upi.edu/index.php/riksabahasa/article/view/1692>
- [31] A. N. Sabrina, “INTERNET SLANG CONTAINING CODE-MIXING OF ENGLISH AND INDONESIAN USED BY MILLENNIALS ON TWITTER (Slang Internet Mengandung Campur-Kode Bahasa Inggris dan Indonesia yang Digunakan oleh Milenial di Twitter),” *Kandai*, vol. 17, no. 2, p. 153, 2021, doi: 10.26499/jk.v17i2.3422.
- [32] A. Rosadi *et al.*, “Analisis Sentimen Berdasarkan Opini Pengguna pada

Media Twitter Terhadap BPJS Menggunakan Metode Lexicon Based dan Naïve Bayes Classifier Twitter Text Mining,” vol. 20, pp. 39–52, 2021.

- [33] N. L. Amalia, *Analisis Sentimen Pada Perpindahan Ibukota Indonesia Dengan Algoritma Support Vector Machine: Evaluasi Lesikon, Metode Ekstraksi Fitur, Dan Kernel Trick*. 2023. [Online]. Available: <https://repository.uinjkt.ac.id/dspace/handle/123456789/68835>
- [34] A. Agustian, T. Tukiro, and F. Nurapriani, “Penerapan Analisis Sentimen Dan Naive Bayes Terhadap Opini Penggunaan Kendaraan Listrik Di Twitter,” *J. TIKA*, vol. 7, no. 3, pp. 243–249, 2022, doi: 10.51179/tika.v7i3.1550.
- [35] S. Pokhrel, “No TitleEΛENH,” *Ayan*, vol. 15, no. 1, pp. 37–48, 2024.
- [36] R. Parluka, S. Ilham Pradika, A. Muhammad Hakim, and K. Rachman N.M, “Analisis Sentimen Twitter Terhadap Bitcoin dan Cryptocurrency Berbasis Python TextBlob,” *J. Ilm. Teknol. Inf. dan Robot.*, vol. 2, no. 2, pp. 33–37, 2020, doi: 10.33005/jifti.v2i2.22.
- [37] S. Analisis, A. Satusehat, D. Wardhani, R. Astuti, and D. D. Saputra, “Optimasi Feature Selection Text Mining: Stemming dan Stopword,” *Innov. J. Soc. Sci. Res.*, vol. 4, pp. 7537–7548, 2024.
- [38] A. Prasetya, Y. N. Kunang, E. Surya, and W. Chandra, “Sentiment Analisis Terhadap Cryptocurrency Berdasarkan Comment Dan Reply Pada Platform Twitter,” vol. 3, no. 2, pp. 268–277, 2021.
- [39] V. R. Joseph, A. Vakayil, V. R. Joseph, and A. Vakayil, “SPLIT : An Optimal Method for Data Splitting SPLIT : An Optimal Method for Data Splitting ABSTRACT,” *Technometrics*, vol. 0, no. 0, pp. 1–23, 2022, doi: 10.1080/00401706.2021.1921037.
- [40] M. R. Saputra and H. Irsyad, “Klasifikasi Tingkat Kemanisan Alpukat Berdasarkan Fitur Hue Saturation Value (HSV) dengan Menggunakan Support Vector Machine (SVM),” *J. Algoritm.*, vol. 2, no. 2, pp. 113–119,

2022, doi: 10.35957/algorithm.v2i2.2361.



LAMPIRAN-LAMPIRAN

A. Surat Penelitian

**PANITIA TUGAS AKHIR DAN SKRIPSI
UNIVERSITAS IBRAHIMY
FAKULTAS SAINS DAN TEKNOLOGI**

JL. KHR. Syamsul Arifin No. 1-2 PO. Box 2 Phone (0852-3333-7581)
Fax. (0338) 453068 Situbondo 68374 website: www.ibrahimy.ac.id e-mail : umb2018@ibrahimy.ac.id
SUMBEREJO BANYUPUTIH SITUBONDO JAWA TIMUR

Nomor : 0828/380.265/071.095/M.3/V/2025
Prihal : Permohonan Izin Penelitian

6 Mei 2025

Kepada

Yth. Kepala Desa Wonobooyo 1

Di Tempat

Assalamu'alaikum warahmatullahi wabarakatuh.

Disampaikan dengan hormat, dalam rangka pelaksanaan penelitian guna penyusunan Skripsi. Judul "**Analisis Sentimen Presiden terpilih Prabowo - Gibran Menggunakan Metode SVM di Media X/Twitter**" dengan ini kami bermaksud mengajukan permohonan izin penelitian bagi mahasiswa kami :

Nama : SYARIFUL ABRORI
NIM : 2021503054
Program Studi : S1 Teknologi Informasi
Waktu : 7 - 15 Mei 2025

Untuk melakukan penelitian di lembaga yang Bapak/Ibu pimpin. Sebagai bahan pertimbangan bersama ini kami lampirkan :

1. Proposal Skripsi
2. Fotocopy Kartu Tanda Mahasiswa

Demikian permohonan kami, atas berkenannya kami sampaikan *Jazakumullahu khairan.*

Wassalamu'alaikum warahmatullahi wabarakatuh.

Ketua Panitia,


**PANITIA
TA & SKRIPSI
FAKULTAS SAINS & TEKNOLOGI
UNIVERSITAS IBRAHIMY**

Dr. Ach. Khumaidi, M.P.

Lampiran 1 Surat Penelitian

C. Dataset Twitter

Tanggal	Username	Tweet
2024-12-19T06:55:07.000Z	POLRESTA BARELANG	Polresta Barelang Ungkap Kasus Judi Online, Dukung Program 100 Hari Asta Cita Presiden RI
2024-12-19T06:47:48.000Z	Sleepyy	Satu lagi, jangan suka buat klaim sendiri “sekarabg pada nyesel”, belum 100 hari, belum setahun yang baru terpilih menjabat jd presiden, wapresnya mau diapain terserah
2024-12-19T02:57:16.000Z	Santorini	UJIAN PERTAMA bagi pemerintahan Presiden @prabowo di 100 hari pertamanya adalah ini. Meski kita tau kalau ini adalah amanat UU dari pemerintahan sebelumnya & sdh disetujui oleh anggota Dewan sebelumnya... Tapi rakyat MENUNTUT kepekaan dari Pemerintah dgn melihat kondisi saat ini
2024-12-19T02:39:16.000Z	lawren siswoyo	Dimana wibawa penegak hukum dan pemerintahan presiden Prabowo ini ??? Keriminalsasi terjadi di pemerintahannya 100 hari ? Berarti kekuasaan Habib lebih tinggi dari hukum Indonesia, penyelesaiannya paling2 satu MATERAI 10K, sedang harta dan modal masyarakat yg dibuly amblas
2024-12-19T02:28:13.000Z	jamil sehung jaehyun	sumpah sedih dan marah banget sama kebijakan presiden yang belum 100 hari ini
2024-12-18T16:30:47.000Z	Sinta	Voters 02..... Tolong kondisikan. 100 hari kerja janji kampanye dah dilanggar. Presiden blunder dan kena sindiran pers luar... Surplus demografi bakal jadi beban kalo gini mulu anjir
2024-12-18T14:57:38.000Z	aldho	dengerin tuh Pak Prabowo Yth... bpk sampai dikatain ma rakyatnya serius gak sih jadi presiden? wajar rakyat ngomong begitu karena hampir 100 hari, bpk tdk ada langkah yang bikin rakyat seneng. yg ada malah tambah keblangsak susahny. harga bahan pangan sdh pada naik pak... https://x.com/vanc1Bozz/stat/vanc1Bozz/status/1869036449351557589...

2024-12-18T07:30:12.000Z	Miss Tweet	Tidak suka Jokowi hari ini , berarti 2014 2019 harusnya Prabowo jadi Presiden, hari ini Prabowo Presiden !(Belum 100 Hari) Menyalahkan pemilih atas kebijakan Penguasa yang dianggap tidak pro rakyat adalah keGBlkan Ber-Demokrasi Hantam Pilihannya Bukan Pemilihnya
2024-12-18T06:11:28.000Z	Ayah Dien	http://kaberehnews.com/2024/12/giat-sosial-satpolairud-polres-aceh.html... Giat Sosial Satpolairud Polres Aceh Timur Bagikan Sembako, Dukung Program 100 Hari Pertama Kinerja Presiden

LINK DATASET :	https://bit.ly/4eCs4Es
QR QODE DATASET :	

Lampiran 3 Dataset Twitter/X

D. Lampiran Program Scapping

```
import logging
import time
import random
import datetime
import pytz
import pandas as pd
from selenium import webdriver
from selenium.webdriver.chrome.service import Service
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.by import By

# ◆ Logging
logging.basicConfig(
    filename="twitter_scraping.log",
    level=logging.DEBUG,
    format="%(asctime)s - %(levelname)s - %(message)s",
)
console_handler = logging.StreamHandler()
console_handler.setLevel(logging.INFO)
formatter = logging.Formatter("%(asctime)s - %(levelname)s - %(message)s")
console_handler.setFormatter(formatter)
logging.getLogger().addHandler(console_handler)

# ◆ ChromeDriver setup
driver_path = r"C:\Users\Lenovo\Music\chromedriver-win64\chromedriver.exe"
user_data_dir = r"C:\Users\Lenovo\AppData\Local\Google\Chrome\User Data"
```

```
options = Options()
options.add_argument(f"user-data-dir={user_data_dir}")
options.add_argument("--start-maximized")
options.add_argument("--disable-blink-features=AutomationControlled")
options.add_argument("--disable-extensions")
options.add_argument("--no-sandbox")
options.add_argument("--disable-gpu")
options.add_argument("--disable-dev-shm-usage")
options.add_argument("--remote-debugging-port=9222")
options.add_experimental_option('excludeSwitches', ['enable-logging'])

service = Service(driver_path)
driver = webdriver.Chrome(service=service, options=options)

# ◆ Fungsi scroll cepat & parsing
def collect_tweets_until_stable(max_rounds=70, min_wait=2.5, max_wait=4):
    tweet_data = {}
    stable_rounds = 0

    for round in range(max_rounds):
        driver.execute_script("window.scrollTo(0, document.body.scrollHeight);")
        time.sleep(random.uniform(min_wait, max_wait))

        tweets = driver.find_elements(By.XPATH, '//article[@data-testid="tweet"]')
        new_tweet_found = False
```

```
for tweet in tweets:
    try:
        timestamp = tweet.find_element(By.XPATH,
            './time').get_attribute("datetime")
        content = tweet.find_element(By.XPATH, './div[@data-
            testid="tweetText"]').text
        username = get_username(tweet)
        tweet_key = hash((timestamp, username, content))
        if tweet_key not in tweet_data:
            tweet_data[tweet_key] = (timestamp, username, content)
            new_tweet_found = True
    except Exception:
        continue

logging.info(f"🔄 Round {round+1}: Total tweet unik: {len(tweet_data)}")

if not new_tweet_found:
    stable_rounds += 1
    if stable_rounds >= 4:
        logging.info("🛑 Tidak ada tweet baru selama 4 scroll. Selesai.")
        break
    else:
        stable_rounds = 0

return list(tweet_data.values())
```

Lampiran 4 Program Scapping

E. Lampiran Program Algoritma

```
# 8. Visualisasi Heatmap Classification Report
from sklearn.metrics import classification_report
import seaborn as sns

# Buat classification report dalam bentuk dictionary
report_dict = classification_report(y_test, y_pred, target_names=["Negatif",
"Netral", "Positif"], output_dict=True)

# Konversi ke DataFrame
report_df = pd.DataFrame(report_dict).transpose().iloc[:3, :3] # Hanya ambil per kelas
report_df = report_df.round(2)

# Visualisasi heatmap
plt.figure(figsize=(8, 4))
sns.heatmap(report_df, annot=True, cmap="YlGnBu", fmt=".2f")
plt.title("Classification Report: Precision, Recall, F1-Score")
plt.ylabel("Sentimen")
plt.xlabel("Metrik Evaluasi")
plt.show()
```

```
import pandas as pd
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.metrics import (
    accuracy_score,
    precision_score,
    recall_score,
    f1_score,
    confusion_matrix,
    classification_report,
    ConfusionMatrixDisplay
)
import matplotlib.pyplot as plt

# 1. Muat Dataset dari Excel
file_path = "tweets_with_sentiment_labels.xlsx" # Ganti dengan path file Excel Anda
df = pd.read_excel(file_path)

# Pastikan kolom 'Stemmed_Tweet' dan 'Sentiment' ada
if "Stemmed_Tweet" not in df.columns or "Sentiment" not in df.columns:
    raise ValueError("Kolom 'Stemmed_Tweet' atau 'Sentiment' tidak ditemukan dalam dataset.")

# Hapus baris dengan nilai NaN di kolom 'Stemmed_Tweet' atau 'Sentiment'
df = df.dropna(subset=["Stemmed_Tweet", "Sentiment"])

# 2. Konversi Teks ke Fitur Numerik menggunakan TF-IDF
vectorizer = TfidfVectorizer(max_features=5000) # Batasi jumlah fitur untuk efisiensi
X = vectorizer.fit_transform(df["Stemmed_Tweet"]) # Fitur numerik dari teks
```

Lampiran 5 Program Algoritma

F. Letter of Acceptance Journal



ISSN : 2540-8984

Letter of Acceptance

Tulungagung, 16 Juni 2025

No : 107/JIPI.PTI.UBHI/XI.II/V/2025
Lamp : -
Hal : Penerimaan artikel JIPI Vol. 11 No.2 2026

Kepada
Syariful Abrori, Firman Santoso, Farihin Lazim
Di Tempat

Assalamu'alaikum Wr. Wb.

Bersama surat ini, redaksi Jurnal Ilmiah Penelitian dan Pembelajaran Informatika (JIPI) Program Studi Pendidikan Teknologi Informasi Universitas Bhinneka PGRI menginformasikan kepada Bapak/Ibu bahwa naskah dengan judul : "**ANALISIS SENTIMEN PRESIDEN TERPILIH PRABOWO-GIBRAN MENGGUNAKAN METODE SVM DI MEDIA TWITTER/X**" telah diterima untuk diterbitkan pada Jurnal Ilmiah Penelitian dan Pembelajaran Informatika (JIPI) Vol.11 No.2 2026.

Kami mengucapkan terima kasih dan selamat atas diterimanya artikel tersebut. Kami juga mengharapkan artikel – artikel berikutnya untuk diterbitkan pada JIPI

Demikian surat kami, atas perhatian dan kerjasamanya kami sampaikan ucapan terimakasih.

Wassalamu'alaikum Wr. Wb.



Fahrur Rozi, M.Kom.

Program Studi Pendidikan Teknologi Informasi
Universitas Bhinneka PGRI
Jl. Mayor Sujadi Tim. No. 24 Plosokandang
Tulungagung, Jawa Timur 66229
E-mail : jipi@ubhi.ac.id
Website : jurnal.stkipppgritulungagung.ac.id/index.php/jipi



Lampiran 6 Letter of Acceptance Journal

G. Lembar Keaslian Tulisan



PONDOK PESANTREN SALAFIYAH SYAFI'YAH SUKOREJO
UNIVERSITAS IBRAHIMY
PERPUSTAKAAN IBRAHIMY

NPP. 3512142F2006567

Jl. KHR. Syamsul Arifin No. 1-2 PO. Box. 2 Kode Pos. 68374 Phone (0338) 452666 Fax. (0338) 453068
SUMBEREJO BANYUPUTIH SITUBONDO JAWA TIMUR



SURAT KETERANGAN HASIL PEMERIKSAAN PLAGIASI

Yang bertanda tangan di bawah ini

Nama : Muhammad Ali Ridla, M.Kom.

Jabatan : Kepala Perpustakaan

Menyatakan dengan sebenarnya bahwa:

NPM : 2021503054

Nama : SYARIFUL ABRORI

Fakultas : Sains dan Teknologi

Prodi : Teknologi Informasi

Kecamatan : KLABANG

Kabupaten : KAB. BONDOWOSO

Provinsi : JAWA TIMUR

Judul Skripsi : ANALISIS SENTIMEN PRESIDEN TERPILIH
PRABOWO GIBRAN MENGGUNAKAN
METODE SVM DI MEDIA TWITTER/X

Dengan dosen Pembimbing :

1. Firman Santoso, M.Kom.

2. Farihin Lazim, M.Tr.T

Telah dilakukan cek plagiasi di Perpustakaan Universitas Ibrahimy dengan persentase plagiasi terakhir sebesar **20%**.

Demikian Surat Keterangan ini dibuat untuk dipergunakan sebagaimana mestinya.

Sukorejo, 28 Juni 2025

Kepala Perpustakaan,



Dokumen ini telah ditandatangani secara elektronik.

Muhammad Ali Ridla, M.Kom.



UU ITE No.11 Tahun 2008 Pasal 5 Ayat 1
"Informasi Elektronik dan/atau Dokumen Elektronik dan/atau hasil cetaknya merupakan alat bukti yang sah."

© www.lib.ibrahimy.ac.id

© library@ibrahimy.ac.id

f @ Perpustakaan Ibrahimy

t @ibrahimy_lib

Lampiran 7 Lembar Keaslian Tulisan

CURRICULUM VITAE



Nama : Syariful Abrori
Tempat, Tanggal Lahir : Situbondo, 23 Juli 2003
Alamat : Dusun Krajan, Desa Wonobojo, Kecamatan
Klabang, Kabupaten Bondowoso, Provinsi Jawa
Timur, Negara Indonesia
Jenis Kelamin : Laki – Laki
Agama : Islam
Hobi : Sesuai Musimnya Ajah
Email : Syarifulabrori2307@gmail.com

A. Riwayat Pendidikan

TK : TK PGRI WONOBOYO 1
SD/MI : SDN 1 WONOBOYO 1
SLTP : SMPN 1 KAPONGAN
SLTA/SMA/SMK : SMAN 1 PANJI
Perguruan Tinggi : UNIVERSITAS IBRAHIMY

B. Keahlian

- ✓ Mengoperasikan Microsoft Office (Word, Excel, PowerPoint)
- ✓ Dasar-dasar Pemrograman
- ✓ Manajemen Database Dasar
- ✓ Instalasi perangkat lunak dasar dan driver
- ✓ Pemahaman dasar jaringan komputer dan internet