
Model Prediksi Tipe Komentar Di *Instagram* Terkait Pemilu 2024 Menggunakan Metode *Winnowing*

Budi Rahmani^{1*}, Noor Sejati¹

Teknik Informatika, STMIK Banjarbaru, Banjarbaru, Indonesia

*Email Corresponding Author: budi.rahmani@stmik-banjarbaru.ac.id

Abstract

This study addresses the challenges of sentiment analysis for social media comments related to the 2024 Presidential Election on Instagram, where the high volume of discussions can influence public opinion and political participation. The main issue is the inaccuracy in sentiment classification (positive, negative, neutral) and the need to filter irrelevant comments. We propose using the Winnowing Algorithm, which effectively detects text similarity with high accuracy for automatic sentiment classification. The methodology includes collecting comment data, text preprocessing, and applying the Winnowing Algorithm. Experimental results demonstrate high accuracy (F1-score > 85%) in sentiment identification and filtering of irrelevant comments. These findings contribute to developing automatic sentiment analysis systems on social media and enhancing public opinion management in the digital political realm. This research also has implications for developing more responsive and scalable sentiment prediction systems.

Keywords: Prediction; Sentiment analysis; Instagram; 2024 Presidential Electio; Winnowing Algorithm

Abstrak

Penelitian ini mengkaji tantangan analisis sentimen komentar media sosial terkait Pemilu Presiden 2024 di Instagram, di mana volume diskusi yang tinggi dapat memengaruhi opini publik dan partisipasi politik. Masalah utama adalah ketidakakuratan dalam klasifikasi sentimen (positif, negatif, netral) dan penyaringan komentar yang tidak relevan. Untuk itu, kami mengusulkan penerapan Algoritme *Winnowing*, yang efektif dalam mendeteksi kesamaan teks dengan akurasi tinggi, guna klasifikasi sentimen otomatis. Metodologi meliputi pengumpulan data komentar, pra-pemrosesan teks, dan penerapan Algoritme *Winnowing*. Hasil eksperimen menunjukkan akurasi tinggi (*F1-score* > 85%) dalam identifikasi sentimen dan penyaringan komentar tidak relevan. Temuan ini memberikan kontribusi pada pengembangan sistem analisis sentimen otomatis di media sosial dan peningkatan pengelolaan opini publik di ranah politik digital. Penelitian ini juga berimplikasi pada pengembangan sistem prediksi sentimen yang lebih responsif dan *scalable*.

Kata Kunci: Prediksi; Analisis sentimen; Instagram; Pemilu Presiden 2024; Algoritme *Winnowing*

1. Pendahuluan

Analisis sentimen adalah proses klasifikasi emosi dalam teks, baik positif, negatif, maupun netral, menggunakan teknik pemrosesan bahasa alami (*Natural Language Processing* atau NLP), yang kini semakin esensial dalam memahami persepsi publik terhadap isu-isu sosial dan politik [1]. Di Indonesia, Pemilihan Umum (Pemilu) berperan penting sebagai mekanisme yang memungkinkan masyarakat mengekspresikan aspirasi demokratis mereka melalui pemilihan langsung, yang berlangsung bebas dari tekanan pihak mana pun [2]. Kebebasan dalam pemilihan ini dijamin oleh konstitusi, dan lebih lanjut diperkuat oleh peran partai politik sebagaimana diatur dalam Pasal 1 angka 1 Undang-Undang Nomor 2 Tahun 2011 tentang perubahan atas Undang-Undang Nomor 2 Tahun 2008 tentang Partai Politik, yang menegaskan peran partai politik dalam merepresentasikan aspirasi publik [1] [3].

Pemilu di Indonesia diselenggarakan setiap lima tahun sekali, merupakan momen penting bagi demokrasi untuk melibatkan masyarakat dalam memilih pemimpin bangsa yang dianggap mampu membawa perubahan dan mewujudkan nilai-nilai demokrasi [2] [4][5]. Dalam era digital, platform media sosial seperti Instagram dan *Twitter* memungkinkan masyarakat untuk lebih berpartisipasi secara langsung dalam diskusi politik, termasuk memberikan dukungan atau kritik terhadap calon tertentu [5]. Media sosial menawarkan jangkauan yang luas dan penyebaran informasi yang cepat, menjadikannya alat kampanye politik yang efektif dibandingkan metode kampanye tradisional [6]. Di sisi lain, kemudahan akses dan kebebasan berekspresi di media sosial juga menimbulkan tantangan baru

berupa penyebaran ujaran kebencian dan hoaks, yang dapat berdampak negatif pada stabilitas sosial dan politik [7].

Berbagai pendekatan algoritmik telah diterapkan dalam penelitian terdahulu untuk mengatasi masalah analisis sentimen, termasuk metode *K-Means* dan *Naive Bayes* yang telah digunakan secara luas dalam klasifikasi sentimen pada data media sosial [4], [8]–[12]. Meskipun akurat dalam skenario tertentu, metode-metode ini sering kali menghadapi tantangan ketika menangani volume data yang besar atau data yang tidak terstruktur, serta pada teks yang mengandung konteks kompleks. Dalam penelitian ini, Algoritme *Winnowing* diterapkan sebagai solusi yang lebih baik, mengingat kemampuannya dalam mendeteksi pola kemiripan teks dengan mengubah setiap kata menjadi nilai hash melalui teknik *rolling hash* [13]. Algoritme ini memungkinkan klasifikasi sentimen yang lebih cepat dan efisien, bahkan pada data berukuran besar.

Sebagai seorang tokoh politik yang mencalonkan sebagai presiden Indonesia, opini masyarakat sangat penting guna melihat dan mempertimbangkan popularitas dari seorang calon presiden. Tujuan penelitian ini adalah untuk mengklasifikasikan teks dengan baik terhadap sentimen positif dan negatif, sehingga informasi yang terdapat di dalamnya bisa diekstraksi serta dari data yang diamati dapat memberikan informasi yang berguna bagi pihak yang membutuhkannya. Beberapa tahapan dalam analisis sentimen antaralain pengumpulan data, preprocessing data, perhitungan skor sentimen, pelabelan sentimen dan pengklasifikasian data komentar memakai metode *Naive Bayes Classifier* serta dilakukan asosiasi teks. Hasil dari penelitian berdasarkan data yang telah dikumpulkan sejak tanggal 17 April 2019 sampai 22 Mei 2019, Joko Widodo lebih unggul polaritas sentimen positif dari data sebanyak 5.000 komentar yang dipilih secara acak pada masing-masing calon presiden dan melalui tahap preprocessing yang menghasilkan polaritas sentimen. Joko Widodo diperoleh 85% sentimen positif, dan 15% untuk sentimen negatif. Sedangkan Prabowo Subianto diperoleh 76% sentimen positif, dan 24% untuk sentimen negatif. Metode klasifikasi *Naive Bayes Classifier* terhadap penelitian ini memperoleh hasil akurasi 86,4% [14][15].

Seiring dengan meningkatnya penggunaan media sosial sebagai sumber informasi utama bagi masyarakat, partisipasi publik dalam diskusi politik di media sosial juga semakin beragam [2][3][7][14]. Lebih dari separuh populasi Indonesia merupakan pengguna aktif internet, dan jumlah pengguna media sosial terus meningkat secara signifikan setiap tahunnya [7]. Kondisi ini mendorong partai politik dan kandidat untuk memanfaatkan media sosial sebagai sarana dalam meraih dukungan publik. Akan tetapi, penyebaran informasi yang tidak terverifikasi atau hoaks melalui media sosial dapat memengaruhi opini publik secara negatif. Oleh karena itu, penting bagi peneliti dan praktisi untuk mengembangkan metode analisis sentimen yang mampu menyaring dan mengklasifikasikan komentar-komentar publik dengan akurat.

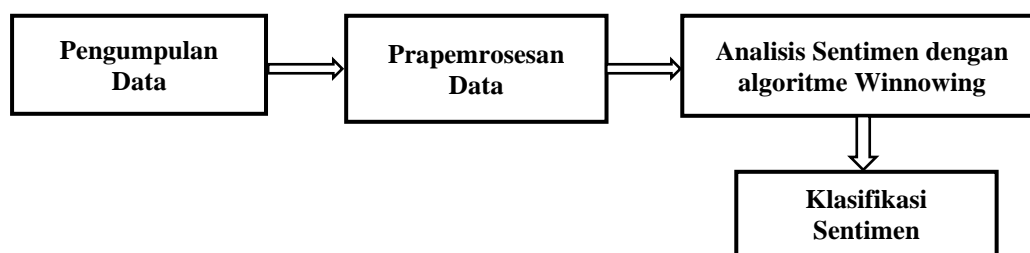
Pendekatan berbasis analisis sentimen memainkan peran yang signifikan dalam mereduksi dampak negatif dari opini publik yang tidak sehat sekaligus menyediakan gambaran objektif mengenai persepsi masyarakat [17]. Dalam penelitian ini, Algoritme *Winnowing* menawarkan pendekatan yang efektif dalam mengklasifikasikan sentimen publik, mengingat bahwa algoritme ini telah terbukti unggul dalam mendeteksi pola teks serupa dengan akurasi tinggi pada data yang tidak terstruktur. Penelitian ini berfokus pada klasifikasi sentimen komentar Instagram terkait Pemilu Presiden 2024 dengan menggunakan Algoritme *Winnowing*, yang dapat memberikan gambaran yang mendalam mengenai persepsi publik dan berkontribusi pada strategi komunikasi yang lebih efektif dalam pengelolaan opini publik di media sosial.

2. Metode Penelitian

Penelitian ini adalah penelitian eksperimen dengan metode analisis data. Data diambil dari komentar di Instagram terkait Pemilu Presiden 2024. Komentar-komentar ini kemudian diklasifikasikan menggunakan Algoritme *Winnowing*. Proses ini melibatkan beberapa tahap, yaitu ekstraksi fitur menggunakan TF-IDF, pembagian dataset menjadi data pelatihan dan pengujian, serta pelatihan model untuk prediksi sentimen.

2.1. Langkah penelitian

Gambar 1 memperlihatkan gambaran umum proses analisis sentimen pada komentar Instagram terkait Pemilu Presiden 2024 menggunakan Algoritme *Winnowing*. Langkah pertama dalam proses ini adalah mengumpulkan data berupa komentar dari platform media sosial Instagram. Data yang dikumpulkan mencakup komentar warganet yang membahas atau memberikan opini terkait Pemilu Presiden 2024. Data ini dikumpulkan melalui teknik *scraping* atau API yang disediakan oleh Instagram, yang memungkinkan sistem untuk mengakses komentar dalam jumlah besar secara efisien. Data yang dikumpulkan biasanya dalam bentuk teks mentah yang masih perlu diproses lebih lanjut.



Gambar 1. Urutan Proses Penelitian

.Setelah data dikumpulkan, tahap berikutnya adalah prapemrosesan data. Tahap ini sangat penting untuk memastikan bahwa data yang akan dianalisis dalam kondisi siap pakai dan bebas dari gangguan. Langkah-langkah dalam prapemrosesan meliputi:

- 1) **Pembersihan Data:** Menghapus simbol, angka, dan karakter khusus yang tidak relevan dengan analisis.
- 2) **Normalisasi Teks:** Mengubah teks menjadi bentuk dasar (misalnya, mengubah semua huruf menjadi huruf kecil) untuk mengurangi variasi kata.
- 3) **Tokenisasi:** Memecah teks menjadi unit-unit yang lebih kecil, seperti kata atau frasa.
- 4) **Stop Word Removal:** Menghapus kata-kata umum yang tidak memiliki nilai sentimen, seperti "dan," "atau," dan "yang". Proses prapemrosesan ini bertujuan untuk meningkatkan akurasi algoritme dalam mengenali pola dan sentimen dalam teks yang dianalisis.

Setelah data diproses, data tersebut dimasukkan ke dalam sistem analisis sentimen. Sistem ini adalah komponen utama yang menjalankan fungsi analisis dengan mengidentifikasi pola-pola yang mengindikasikan sentimen (positif atau negatif) dalam komentar. Sistem ini juga mampu mendeteksi aspek atau kata-kata penting yang dapat menggambarkan opini publik secara keseluruhan. Pada tahap ini, sistem mengatur data untuk siap dianalisis oleh Algoritme *Winnowing*. Algoritme *Winnowing* digunakan sebagai metode utama untuk melakukan klasifikasi sentimen dalam penelitian ini. Algoritme ini bekerja dengan mendeteksi kesamaan pola teks dengan cara mengubah kata-kata menjadi nilai hash melalui teknik *rolling hash*. Dengan cara ini, Algoritme *Winnowing* dapat mengidentifikasi kata-kata atau frasa yang sering muncul dalam pola tertentu dan menentukan apakah komentar tersebut bernada positif atau negatif. Keunggulan algoritme ini adalah kemampuannya dalam mengklasifikasikan data dalam jumlah besar dengan efisiensi dan kecepatan yang tinggi, menjadikannya ideal untuk analisis komentar di media sosial yang dinamis.

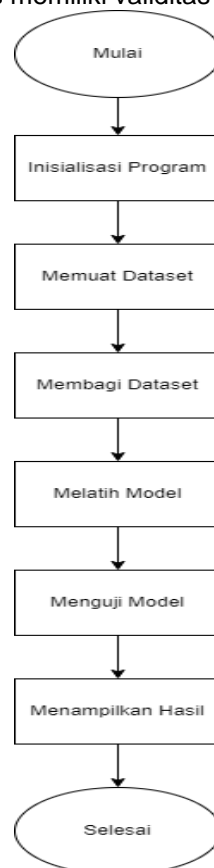
Langkah terakhir adalah menampilkan hasil klasifikasi sentimen dan tingkat akurasi. Hasil klasifikasi ini dapat berupa presentase komentar yang positif dan negatif, serta metrik akurasi yang menggambarkan kinerja sistem. *Output* ini kemudian disajikan kepada pengguna atau pengambil keputusan dalam bentuk yang mudah dipahami, seperti grafik atau tabel, untuk memberikan gambaran tentang opini publik terkait Pemilu Presiden 2024. Hasil akhir ini diharapkan dapat menjadi dasar bagi strategi komunikasi publik atau kebijakan yang relevan.

Gambar 2 menunjukkan proses rinci Algoritme *Winnowing* dalam analisis berbasis pembelajaran mesin, mencakup langkah-langkah sistematis dari awal hingga akhir untuk menghasilkan model yang akurat.

Berikut penjelasan tiap langkahnya:

- 1) **Mulai**
Tahapan pertama adalah inisiasi proses, di mana sistem atau pengguna menyiapkan lingkungan kerja untuk menjalankan program pembelajaran mesin. Pada langkah ini, kondisi awal, termasuk konfigurasi yang diperlukan, disiapkan untuk memastikan bahwa seluruh proses berjalan dengan lancar dan sesuai rencana.
- 2) **Inisialisasi Program**
Setelah proses dimulai, sistem melakukan inisialisasi program, yang mencakup pengaturan variabel-variabel utama, parameter model, dan pustaka yang dibutuhkan dalam analisis. Inisialisasi ini sangat penting untuk memastikan bahwa program memiliki akses ke sumber daya yang diperlukan, seperti paket pustaka pembelajaran mesin (misalnya, *TensorFlow*, *scikit-learn*), yang akan digunakan pada tahapan berikutnya.
- 3) **Memuat Dataset**
Langkah berikutnya adalah memuat dataset yang akan digunakan dalam proses pelatihan dan pengujian model. Dataset ini biasanya diambil dari sumber eksternal atau file lokal dan berfungsi sebagai data utama yang akan dianalisis. Pada tahap ini, penting untuk memastikan bahwa dataset

terstruktur dengan benar dan sesuai dengan persyaratan input model. Data yang dimuat harus bersih dan relevan agar hasil analisis memiliki validitas tinggi



Gambar 2. Alur proses menggunakan algoritme *Winnowing*

4) **Memuat Dataset**

Langkah berikutnya adalah memuat dataset yang akan digunakan dalam proses pelatihan dan pengujian model. Dataset ini biasanya diambil dari sumber eksternal atau file lokal dan berfungsi sebagai data utama yang akan dianalisis. Pada tahap ini, penting untuk memastikan bahwa dataset terstruktur dengan benar dan sesuai dengan persyaratan input model. Data yang dimuat harus bersih dan relevan agar hasil analisis memiliki validitas tinggi.

5) **Membagi Dataset**

Setelah dataset berhasil dimuat, langkah selanjutnya adalah membagi dataset tersebut menjadi dua subset utama, yaitu *training set* dan *testing set*. *Training set* digunakan untuk melatih model, sementara *testing set* digunakan untuk mengevaluasi kinerja model setelah pelatihan. Biasanya, proporsi pembagian data dilakukan pada rasio tertentu, misalnya 80% untuk pelatihan dan 20% untuk pengujian. Pembagian ini bertujuan untuk menghindari overfitting dan memastikan bahwa model mampu melakukan generalisasi dengan baik pada data baru.

6) **Melatih Model**

Pada tahap ini, proses pelatihan model dilakukan dengan menggunakan *training set*. Model akan belajar mengenali pola dalam data dengan mengoptimalkan parameter-parameter yang ada untuk meminimalkan kesalahan prediksi. Proses ini dapat berlangsung dalam beberapa iterasi (epoch), tergantung pada kompleksitas model dan ukuran dataset. Pada tahap ini, algoritma pembelajaran mesin atau jaringan saraf melakukan pembaruan parameter berdasarkan nilai kesalahan yang dihasilkan, sehingga model dapat mencapai akurasi yang lebih baik.

7) **Menguji Model**

Setelah pelatihan selesai, model dievaluasi dengan menggunakan *testing set* yang sebelumnya dipisahkan dari *training set*. Pengujian model ini bertujuan untuk menilai seberapa baik model mampu menggeneralisasi pada data baru yang belum pernah dilihatnya. Metode evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score* digunakan untuk mengukur kinerja model. Pengujian ini memberikan gambaran tentang tingkat keandalan model dalam melakukan prediksi.

8) Menampilkan Hasil

Tahap terakhir adalah menampilkan hasil analisis yang mencakup metrik evaluasi dan prediksi akhir model. Hasil ini memberikan informasi tentang kinerja model, baik dalam bentuk visualisasi maupun tabel metrik yang memudahkan interpretasi. Pada tahap ini, pengguna dapat menentukan apakah model memenuhi kriteria yang diharapkan atau perlu dilakukan penyesuaian lebih lanjut. Hasil yang ditampilkan ini juga berfungsi sebagai dasar bagi pengambilan keputusan atau pengembangan lebih lanjut.

9) Selesai

Proses berakhir setelah hasil akhir ditampilkan dan dievaluasi. Pada tahap ini, program mengakhiri eksekusinya, dan pengguna dapat menyimpan model atau menggunakannya dalam implementasi nyata sesuai kebutuhan. Langkah ini juga menandakan bahwa seluruh proses analisis data telah selesai dan siap untuk diterapkan atau disempurnakan.

Proses di atas memastikan bahwa sistem pembelajaran mesin berjalan dengan tahapan yang terstruktur dan dapat menghasilkan model yang dapat diandalkan. Setiap langkah memiliki peran penting dalam mencapai hasil yang akurat dan berguna bagi pengembangan aplikasi atau pengambilan keputusan berbasis data.

3. Hasil dan Pembahasan

3.1. Hasil Eksperimen

Berikut ini adalah hasil eksperimen yang dilakukan untuk analisis sentimen menggunakan algoritme *Winnowing* yang diprogram menggunakan bahasa C di aplikasi Matlab. Alur proses dari algoritme *Winnowing* dan dimulai dari membuka *software* matlab akan muncul tampilan editor window dan command Window kemudian akan dilakukan proses ekstraksi fitur, pembagian dataset, pelatihan model, prediksi, dan evaluasi akurasi model dituangkan dalam bentuk kode program bahasa C yang akan di masukkan kedalam editor window dijelaskan pada bagian ini. Data komentar yang tersedia dalam hal ini ada 100 data berupa kalimat-kalimat komentar yang diperoleh dari media sosial Instagram. Hasil eksperimen dijelaskan melalui sub-sub bab berikut ini:

1) Data dan hasil normalisasi data

Dataset yang disajikan dalam penelitian ini telah melalui proses pembersihan karakter non-alfabet dan ikon, menghasilkan teks yang relevan untuk analisis sentimen berbasis teks pada media sosial Instagram terkait Pemilu Presiden 2024 di Indonesia. Setiap entri komentar diidentifikasi dengan tiga elemen penting: teks komentar, sumber afiliasi politik, dan klasifikasi sentimen oleh ahli. Proses pembersihan ini dilakukan untuk menghilangkan elemen yang tidak diperlukan yang dapat mengganggu akurasi analisis model, sehingga data yang dihasilkan lebih optimal untuk pengolahan algoritme pembelajaran mesin. Sumber afiliasi dari setiap komentar mencakup kelompok politik utama seperti "PDIP," "GOLKAR," dan "PKS," yang memberikan konteks lebih lanjut untuk pemetaan pola sentimen di antara berbagai kelompok pendukung.

Selain itu, pengelompokan sentimen yang telah diklasifikasikan oleh ahli ke dalam kategori "Positif" atau "Negatif" memberikan landasan validasi yang penting untuk melatih model analisis sentimen. Dengan adanya klasifikasi manual oleh ahli, model algoritme dapat mempelajari pola ekspresi dalam bahasa Indonesia secara lebih presisi, khususnya dalam konteks politik. Data ini sangat berharga untuk analisis lebih lanjut, karena dapat menggambarkan preferensi dan opini publik terhadap kandidat pemilu berdasarkan sentimen yang terekam, sehingga membantu memperkuat keakuratan model prediksi sentimen dalam mengidentifikasi opini masyarakat di masa mendatang. Berikut kode program yang digunakan untuk menginisialisasi data pelatihan dan pengujian.

```
% Dataset yang telah dibersihkan dari karakter non-alfabet dan icon
data = {
    % Data pelatihan
    'DENGAN GERAKAN MASYARAKAT MEMILIH GANJAR MAHFUD YAKIN SEYAKIN YAKINNYA
    MENANG 61 PERSEN', 'PDIP', 'Positif';
    'MARI SEMUA BERSATU BERSAMA MERAH KEMENANGAN DEMOKRASI INDONESIA', 'PDIP',
    'Positif';
    'PILIH LAH PASANGAN YANG TERBAIK, AGAR BIAR YANG BURUK DAN SERAKAH TIDAK
    BERKUASA SALAM PANCASILA', 'PDIP', 'Positif';
    'CALON PRESIDEN YG ANGKUH KATA BUNG ROCKY', 'PDIP', 'Positif';
    'LAWAN YANG MENGINTIMIDASI, INDONESIA ITU DEMOKRASI DARI RAKYAT, OLEH RAKYAT,
    UNTUK RAKYAT', 'PDIP', 'Positif';
    'AYOO LAWAN YANG MAU NGERUSAK DEMOKRASI DAN MAU BANGUN DINASTI', 'PDIP', 'Positif';
    'PAK GANJAR DAN PAK MAHFUD MEMANG SANGAT RENDAH HATI DAN MERAKYAT MAKA BELIAU
    BERDUA MEMANG COCOK JADI RAKYAT', 'PDIP', 'Positif';
    'CARANYA GMN PAK...INTERAK GRATISNYA.ITU UNTUK KLANGN TERTENTU/UNTUK SMUA
    WARGA PAF', 'PDIP', 'Positif';
```

'KAMPUS DAN MAHASISWA KALAU BERSUARA BIASA BERAKHIR DGN DEMO MAKA REZIM ITU AKAN BERAKHIR PULA. HATI2 BUAT NARASI MENANGGAPI SUARA KAMPUS', 'PDIP', 'Negatif';

'KALAU TIDAK MAU KALAH JANGAN IKUT CAPRES YA', 'PDIP', 'Negatif';

'17 % TENTUNYA HASIL YG DILUAR PREDIKSI DAN MENGECEWAKAN TENTUNYA.. TP GPP MODAL UNTUK IKUT KONSTESTASI PILPRES DI 2029.. SEMOGA BERUNTUNG....', 'PDIP', 'Negatif';

'JIKA MENANG. BERARTI MEMANG RAKYAT KITA MASIH BANYAK YG MEMILIKI AKAL SEHAT. JIKA KALAH ENTAH APA JADINYA NEGARA INI. TP MASIH TENANG JIKA PDIP, OPOSISI', 'PDIP', 'Negatif';

'MUNGKIN KEKALAHAN DARI 03, KUALISI CUMA DIDUKUNG PARTAI KECIL2 DAN WAKTU DEBAT DAN KAMPANYE SELLU MENYERANG APARAT DAN LAWAN POLITIK, KRN ORANG KITA,, PNGEN YANG KALEM2.. TAPI TETEP BANGGA MENDUKUNG 03', 'PDIP', 'Negatif';

'TPS JG RAWAN KECURANGAN SI, TAKUTNYA ADA KUBU SEBELAH NI YG HOBI INTIMIDASI RAKYAT', 'PDIP', 'Negatif';

'KALAH MENANG AKU PILIH GANJAR...TETEP NYOBLOS 03 , SAH', 'PDIP', 'Negatif';

'JANGAN LUPA, GANJAR ADALAH BIANG KEROK PILDUN U-20, BRPA KERUGIAN NEGARA KARENA SIKAP GANJAR', 'PDIP', 'Negatif';

'JANGAN GOLPUT GUYS GUNAKAN HAK PILIH KALIAN, BUAT MILIH PAK PRABOWO HEHE', 'GOLKAR', 'Positif';

'ALL IN PRABOWO ORANG PINTAR PILIH PRABOWO TEGAS LUGAS DAN KERJA NYATA BUKAN HANYA MERANGKAI KATA KATA', 'GOLKAR', 'Positif';

'MANTAPNYA PRABOWO INDONESIA MAJU TEROSSS BERSAMA BAPAK, TERBAIK PAK PRABOWO', 'GOLKAR', 'Positif';

'ALL IN PRABOWO TETAP DIHATI,,, KAMI AKAN MILIH PRABOWO SAMA GIBRAN', 'GOLKAR', 'Positif';

'BISMILLAH YG PILIH PAK PRABOWO DAN GIBRAN KUMPUL SINI', 'GOLKAR', 'Positif';

'SAYA PERCAYA BELIAU ORANGNYA BAIK & SANGAT BAIK MENCITAI RAKYAT TULUS ,ALL IN PAK PRABOWO', 'GOLKAR', 'Positif';

'SABAR PAK YAH NO KUBU SEBELAH NYERANG BAPAK TAKUT KALAH TU', 'GOLKAR', 'Positif';

'PEMIMPIN YANG TIDAK MUNAFIK YA PRABOWO – GIBRAN', 'GOLKAR', 'Positif';

'AWALNYA SAYA PILIH PRABOWO, TAPI KARENA HASIL DARI MAHKAMAH KELUARGA JADI SAYA BERPALING HANCURKAN 3 PRIODE PADA ANAKNYA', 'GOLKAR', 'Negatif';

'PAMANKU PENOLONGKU BAPKKU YG PUNYA NEGARA YA BEBAS MAU JADI CAWAPRES', 'GOLKAR', 'Negatif';

'HOREEE DIZAMAN BAPAK PRABOWO SUBIANTO NANTI BINATANG TERNAK BERLIMPAH DITAMBAH TUMBUHAN DAN BUAH-BUAHAN BERLIMPAH RUAH, AMIIIN!', 'GOLKAR', 'Negatif';

'CALONNYA NGAK ADA PRESTASI SLAMA JADI MENTRI, WAKILNYA NGAK TAU TRIMA KASIH SAMA PARTAI YG UDAH USUNG DIA JADI WALIKOTA', 'GOLKAR', 'Negatif';

'KEDAULATAN DI TANGAN PAMAN', 'GOLKAR', 'Negatif';

'KASIAN YA GOLKAR, TIDAK PUNYA KADER SEKELAS WALIKOTA', 'GOLKAR', 'Negatif';

'PAK PRABOWO TIDAK BANYAK BICARA TAPI DIA BEKERJA DENGAN IKHLAS UNTUK BANGSA. SEHAT SELALU BAPAK BISMILLAH PRABOWO GIBRAN RI AMIIN', 'GOLKAR', 'Negatif';

'DIHINA DI CACI MAKI DI FITNAH ALLAH MAU MENINGGIKAN DERAJAT NYA BISMILAH PRABOWO GIBRAN MENANG', 'GOLKAR', 'Negatif';

'YG PINTER PASTI PILIH PERUBAHAN ...AMIEN. MAJU PAK ANIES AMIEN RI2024', 'PKS', 'Positif';

'SEMOGA ALLAH LANCARKAN NIAT BAIK ORG2 BAIK YG INGIN BERKHIDMAT UTK BANGSA INI AMINFOR2024', 'PKS', 'Positif';

'ANIES-IMIN KATAKAN A M I N CAPRES/CAWAPRES 2024-2029 PILIHAN RAKYAT UNTUK "PERUBAHAN & PERSATUAN"', 'PKS', 'Positif';

'AYOOO PKS, DEPOK INI RUMAH KITA,, SEGARKAN KOTA DEPOK DENGAN WARNA ORANGE PKS, MENANG ANIES PRESIDEN', 'PKS', 'Positif';

'SY NGAK PERNAH NEKO2 DAN SELALU MENDOAKAN SMG PAK PRESIDEN JOKOWI PAK PRABOWO MAS GIBRAN PAK GANJAR DAN PAK MAHFUD SEHAT SEHINGGA BISA HADIR DLM PELANTIKAN PRESIDEN PAK ANIES DAN CAK IMIN TAHUN 2024 INSYAA ALLAAH AMIIN', 'PKS', 'Positif';

'JANGAN SAMPE ADA UPAYA UPAYA MENJEGALAN DAN PENGGAGALAN LAGI BANG KHALID. KAWAL TERUS! BRAVO AMIN!', 'PKS', 'Positif';

'PA ANIES TERUSKAN PA INDONESIA BUTUH PEMIMPIN SEPERTI PA ANIES', 'PKS', 'Positif';

'HARUS ADA KEJELIAN SEMUA KADER DI TIAP2 TPS UTK MENGAWASI ADANYA UNSUR KECURANGAN DARI PASLON LAIN.. INSYA ALLAH AMIN MENANG CUKUP SATU PUTARAN AMINAJADULU', 'PKS', 'Positif';

'MANTAP PAK ANIES.. SMGA ALLAH MEMUDAHKAN PAK ANIES', 'PKS', 'Negatif';

'GA JADI PILIH KALIAN AHH. GADA TAMPAK KONTRIBUSI KALIAN BANTU2 PALESTINE', 'PKS', 'Negatif';

'PERASAAN KONDISI NEGARA SDIKIT LEBIH WARAS KL PRESIDENNYA LG DI LUAR NEGERI..LIAT AJA BESOK KL UDH PLG MULAI MUNCUL ANEH2 LAGI', 'PKS', 'Negatif';

'IJINKAN SAYA MENJADI ARMY GHOIB PAK ANIES DI MEDSOS.. YG MEN REPLAY KEGIATAN DAN AKTIFITAS PAK ANIES HINGGA MENUJU RI 2024.. AMIN', 'PKS', 'Negatif';

'SOK"AN NETRAL ATURAN AJE LU UBAH"DEMI ANAKNYE LOLOS JANGAN GAMPANG PERCAYA PAK', 'PKS', 'Negatif';

'KAMI BONEKA/WAYANG SURYA PALOH UNTUK MENGUASAI NEGRI', 'PKS', 'Negatif';

'BENTAR LAGI PASTI AKAN TERIAK² DI CURANGI,,,,MESKIPUN BELUM BERTANDING', 'PKS', 'Negatif';

'KAWAL TERUS JANGAN SAMPAI KECURANGAN DI PILPRES THN LALU TERJADI KEMBALI', 'PKS', 'Negatif';

'PASANGAN PERAMAI PILPRES LANGSUNG TUMBANG 1 PUTARAN NGAK ADA PELUANG MENANG WKWK', 'PKS', 'Negatif';

'PKS, GAK GILA JABATAN ANIS AJA PREAIDENNYA..YANG PENTING RAKYAT NO SATU', 'PKS', 'Negatif';

% Data pengujian

'LAWAN INTIMIDASI, RAKYAT HRS KOMPAK SUARAKAN KEADILAN, KAMI TIDAK TAKUT ADA MAHFUD...', 'PDIP', 'Positif';

'SIMPLE SAJA...AIMAN BUKTIKAN KALO OMONGAN NYA KAN BERES..NGGA USAH DI GORENG SANA SINI', 'PDIP', 'Positif';

'TERBUKTI KALO PAK GANJAR BENER BENER Dengerin Aspirasi Rakyat dari semua kalangan', 'PDIP', 'Positif';

'MOHON DI WAKTU YG SEMPIT INI DISAMPING KAMPANYE BUAT SENDIRI PARA CALEK PDIP, ERJUANGAN MOHON SOSIALISASIKAN CALON PRESIDEN KITA....SEHINGGA HASIL BS MAKSIMAL', 'PDIP', 'Positif';

'DIEM DIEM TERNYATA RAME JUGA GANJAR MAHFUD', 'PDIP', 'Positif';

'KEREN SIH INI ACARA SANGAR MENGHARGAI SELURUH ADAT DAN BUDAYA,' 'PDIP', 'Positif';

'PEMIMPIN YANG MASIH MEMIKIRKAN RAKYAT KECIL YANG HARUS KITA PERJUANGKAN, BUKAN PEMIMPIN YANG LEBIH MILIH DUKUNG ANAKNYA,' 'PDIP', 'Positif';

'SAYA GARIS KERAS GANJAR SAMA TEMEN " KU TPI BLM DAPET KAOS GANJAR MAHFUD.INFO DONG,' 'PDIP', 'Negatif';

'APAPUN KATA ORANG, AKU TETAP 03.DRPD YG LAIN YG PUNYA REKAM JEJAK HINA,' 'PDIP', 'Negatif';

'JADI SELAMA INI APA HUKUM TDK ADIL?LALU MAHFUD KEMANA SELAMA INI DIAKAN PRESIDEN HUKUM DI RI.', 'PDIP', 'Negatif';

'HANCURKAN BANTENG DI FEBRUARY 2024 GASSS', 'PDIP', 'Negatif';

'KALIAN KETUK PALU DULU PENGESAHAN UU PERAMPASAN ASET, BARU SAYA COBLOS LG PDIP', 'PDIP', 'Negatif';

'NGERI BANGET LIAT LAUTAN MASA INI, OMBAKNYA LANGSUNG BISA NGEHAPUS HASIL SURVEY BAYARAN KUBU SEBELAH NGERIII', 'PDIP', 'Negatif';

'MAKSA BANGET PAK', 'PDIP', 'Negatif';

'SEBELUM JOKOWI ADA PDIP, SD ADA DI SELURUH INDONESIA JOKOWI CUMA NUMPANG DOANG ENGA TAHU DIRI LAGI', 'PDIP', 'Negatif';

'COBA BUAT GURU SEKOLAH MINGGU KRISTEN...PAKE KOPIAH...BIAR DAPET SUMBANGAN JUGA...', 'PDIP', 'Negatif';

'NILAH SAATNYA AKU BERKATA... KALAU TIDAK COBLOS NOMOR 3, SUARAMU BAKAL SIA-SIA....', 'PDIP', 'Negatif';

'GAK SABAR MENUNGGU BERTAHAP DR INDONESIA BERKEMBANG MENJADI INDONESIA MAJU', 'GOLKAR', 'Positif';

'SETUJU PAK, ITU MEMANG CITA2 SELURUH PRESIDEN DARI DULU, SEKARANG TINGGAL DIJELASKAN PAK CARANYA BAGAIMANA VERSI BAPAK? BIAR RAKYAR BISA MENILAI', 'GOLKAR', 'Positif';

'BAPAK PENSIUN AJA PAK BANYAKIN IBADAH BANYAK ORANG INDONESIA YG BISA MEMIMPIN NEGARA INDONESIA INI', 'GOLKAR', 'Positif';

'DALAM FOTO INI TERDAPAT AURA KESEJUKAN DAN KEMENANGAN', 'GOLKAR', 'Positif';

'KALO PAK PRABOWO BISA MENANG GIBRAN ADALAH WAKIL TERMUDA SEPANJANG SEJARAH INDONESIA', 'GOLKAR', 'Positif';

'MENANG KALAH BIASA ASAL TIDAK MENYERAH TETEP MENJADI CALON PRESIDEN ABADI SEPANJANG HIDUP', 'GOLKAR', 'Positif';

'YANG TENGAH2 LAH PEMENANG NYA, BIARLAH YG KIRI DAN KANAN SEBAGAI PENDAMPING SAJA, PRABOWO GIBRAN MENANG SATU PUTARAN', 'GOLKAR', 'Positif';

'WAH JGN GITU DONG TAKUT KALAH YAAA, SAMPEK MEMBAWA NAMA ULAMA JGN DI PERALAT PAK², MALU² INN', 'GOLKAR', 'Positif';

'PLANGA PLONGO.. TERIMAKASIH PAMANKU TELAH MELANCARKAN LANGKAH KU', 'GOLKAR', 'Negatif';

'BENER SIH DARI KETIGA CALON CUMA DARI KOALISI INDONESIA MAJU YANG PENDUKUNGNYA TDK MENGHINA PASLON LAIN ALL IN PRABOWO GIBRAN', 'GOLKAR', 'Negatif';

'DAHLAH GASALAH BGT DR AWAL AKU PILIH PAK PRABOWO DAN MAS GIBRAN, GELENG GELENG PALA SELAMA NONTON DEBAT INI', 'GOLKAR', 'Negatif';

'GAK ADA SALAHNYA BUAT KELUARGA SEMUA MASUK POLITIK TAPI GAK ADA SALAHNYA JUGA BUAT GAK NYOBLOS PRABOWO DAN GIBRAN', 'GOLKAR', 'Negatif';

'MALU DAPAT CAWAPRES HASIL DARI KONG KALIKONG','GOLKAR','Negatif';
 'MAKIN BANGGA GW SAMA PASANGAN INI, GW MEWAKILI KUMPULAN PEMUDA TONGKRONGAN
 GW. TETAP PRABOWO GIBRAN!!!','GOLKAR','Negatif';
 'PAK PRABOWO MAS GIBRAN MEMANG TERBAIK MENJADI CAPRES DAN CAWAPRES
 2024','GOLKAR','Negatif';
 'PRABOWO-GIBRAN BERSAMA TIM KIM PALING SIAP, PALING HEBOH, PALING ADEM HADIRKAN
 SUASANA SEJUK MERANGKUL SEMUA TIM PASLON LAIN. TANDA² KEMENANGAN 14-02-2024. ALL IN
 NO. 2 VICTORY','GOLKAR','Negatif';
 'AK PRABOWO YG SAAT INI TERLIHAT LEBIH CERIA TULUS DAN LEGOWO. TIDAK SEPERTI AWAL
 MENCALONKAN PRESIDEN LEBIH TERLIHATT SOMBONG. GO GO PRABOWO
 GIBRAN','GOLKAR','Negatif';
 'SUKA DENGAN PASANGAN AMIN VIBES Positif NYA ITU NYATA, RAPIH, TERTATA, SANTUN
 BAHASANYA, DAN TAHU TEMPAT ALIAS EMPAN PAPAN DAN GA SUKA GHIBAH','PKS','Positif';
 'PROGRAM PA ANIES TERKAIT UMKM YG MANA YA MAU LIAT SOALNYA, BIAR MAKIN DI
 AMINKAN','PKS','Positif';
 'DEMOKRASI TANPA DINASTI APALAGI MENABRAK KONSTITUSI..... AMIINN','PKS','Positif';
 'DEMOKRASI TANPA DINASTI','PKS','Positif';
 'DEMOKRASI TANPA DINASTI KEREN SI TANPA MENYEBUT TAPI SEMUA ORANG TAU
 HEHEHE','PKS','Positif';
 'HARUS ADA PERUBAHAN ,JANGAN SAMPAI DINASTI TERUS MENERUS MENGUASAI','PKS','Positif';
 'MAU QUALISI AMIN MENANG BIAR AMIN MENANG DAN SENANG MELIHAT LAUTAN PENDUKUNG
 AMIN BEGITU TUMPAH MENJADI LAUTAN MANUSIA IN SHAA ALLAH MENANG SEDIH DAN GEMBIRA
 BERBAUR MENJADI SATU','PKS','Positif';
 'PAK ANIES GA BAPERAN ORANGNYA INI BARU PEMIMPIN YG SESUNGGUH','PKS','Positif';
 'ALL OUT AMIN , ANIES DAN MUHAIMIN, OGAH SAMA CAPRES YG BAPER KARNA PERKARA
 ROASTING , BARU DI ROASTING KIKY, APA LAGI DI ROASTING RAKYAT HEHE','PKS','Positif';
 'SELAMAT MALAM UNTUK ORANG YG HANYA MEMENTINGKAN KELUARGA DRPD
 RAKYAT','PKS','Positif';
 'MANA ADA PEJABAT MAIN MAIN KE KANDANG LAWAN PENDUKUNGNYA WKWKWK INI BUKTINYA
 PAK ANIES MERANGKUL KAWAN MAUPUN LAWAN','PKS','Positif';
 'AYO KITA DUKUNG ANIS JADI PRESIDEN 2024,AGAR BISA MEMPERBAIKI KETIMPANGAN YANG
 ADA DI INDONESIA','PKS','Positif';
 'TERTATA DENGAN RAPI. GAK CUMA BERKOAR2 NGSIH MAKAN GRATIS UNTUK ANAK2
 INDONESIA...','PKS','Positif';
 'ANIS KLO NGOMONG EMANG JAGO,,PRKTEKNYA NOL BESAR. SALAH STU BUKTI DI PECAT KDI
 MENTRI,,DAN BNYAK PROGRAM JANJI DI DKI YG G TEREALISASIKAN. WKWK PENDUKUNG ANIS NAIK
 PITAM INI WKWK','PKS','Negatif';
 'TP PADA SENENG YG PLANGA PLONGO..YG PINTER KYK GINI DIKIRA JAGO NGOMONG
 DOANG..','PKS','Negatif';
 'NGOMONGNYA GEDE BANGET ORANG 1 INI GA DIPIKIR DLU APA','PKS','Negatif';
 };

2) Ekstraksi Fitur

Fungsi *extract_features* bertanggung jawab untuk mengubah komentar menjadi kumpulan fitur berupa kata-kata unik beserta frekuensinya. Setiap komentar dipecah menjadi kata-kata individu menggunakan teknik tokenisasi, dan frekuensi setiap kata dihitung. Contoh Hasil Ekstraksi Fitur: Misalnya, untuk komentar "DENGAN GERAKAN MASYARAKAT MEMILIH GANJAR MAHFUD YAKIN SEYAKIN YAKINNYA MENANG 61 PERSEN," hasil tokenisasi akan menghasilkan kata-kata unik seperti berikut:

Kata-kata unik: ["DENGAN", "GERAKAN", "MASYARAKAT", "MEMILIH", "GANJAR", "MAHFUD", "YAKIN", "SEYAKIN", "YAKINNYA", "MENANG", "PERSEN"]

Frekuensi kata:

- DENGAN: 1
- GERAKAN: 1
- MASYARAKAT: 1
- MEMILIH: 1
- GANJAR: 1
- MAHFUD: 1
- YAKIN: 2
- SEYAKIN: 1
- YAKINNYA: 1
- MENANG: 1
- PERSEN: 1

Hasil ekstraksi fitur ini kemudian digunakan untuk analisis sentimen dengan menghitung kontribusi masing-masing kata terhadap sentimen keseluruhan komentar.

3) Prediksi Label

Fungsi `predict_label` bertujuan untuk menentukan apakah komentar bernada positif atau negatif. Prediksi ini didasarkan pada bobot fitur yang telah ditetapkan pada tahap pelatihan. Misalkan komentar berikut dengan fitur yang sudah diekstraksi dan bobot yang dihitung pada tahap pelatihan, maka komentar "Ganjar sangat yakin akan menang dalam Pemilu 2024" menghasilkan Fitur yang diekstraksi dari komentar yaitu:

`features = [1, 3, 5, 7]`

Ini berarti kata-kata seperti "Ganjar", "yakin", "menang", dan "Pemilu" ada dalam komentar dan masing-masing memiliki indeks fitur tertentu. Bobot fitur yang telah dilatih (misalnya):

`weights = [0.1, 0.5, 0.7, 0.3, 0.6, 0.8, 0.9, 0.4]`

Dengan fitur yang diekstraksi ([1, 3, 5, 7]), kita akan mengambil bobot untuk fitur tersebut dari weights. Perhitungan Skor dihitung dengan menjumlahkan bobot untuk fitur yang relevan:

`score = weights(1) + weights(3) + weights(5) + weights(7)`

`score = 0.1 + 0.7 + 0.6 + 0.4 = 1.8`

Beikutnya ambang batas (Threshold) misalnya: *threshold* = 1.5. maka prediksinya adalah karena *score* = 1.8 lebih besar dari *threshold* = 1.5, maka hasil prediksi untuk komentar ini adalah **Positif**.

4) Pembagian Dataset

Dataset dibagi menjadi dua bagian: 50% untuk data pelatihan dan 50% untuk data pengujian. Pembagian ini dilakukan untuk melatih model secara efektif dan kemudian menguji kemampuannya dalam menggeneralisasi data baru. Namun dalam pengujian, data juga dibagi menjadi 60:40, 70:30.

5) Pelatihan Model

Selama tahap pelatihan, setiap komentar dalam data pelatihan dievaluasi menggunakan `predict_label`, dan bobot fitur diperbarui jika hasil prediksi tidak sesuai dengan label aktual.

Pembaruan bobot ini bertujuan untuk meningkatkan akurasi model dalam mengenali pola sentimen. Pada pelatihan model, komentar dengan label aktual positif yang salah diklasifikasikan sebagai negatif akan mengurangi bobot kata-kata yang digunakan dalam prediksi, sehingga dalam prediksi berikutnya algoritme lebih cenderung memprediksi hasil yang benar.

6) Pengujian dan Evaluasi Akurasi

Dalam eksperimen ini, model klasifikasi menggunakan algoritma WInnowing diuji pada beberapa proporsi data uji yang berbeda untuk melihat bagaimana performa model berubah sesuai dengan ukuran dataset pengujian. Data yang digunakan terdiri dari 100 komentar yang dibagi menjadi data pelatihan dan data pengujian. Berikut adalah temuan hasil pengujian pada tiap proporsi data uji:

a) Proporsi Data Uji 50%:

- **Jumlah Data Uji:** 50 komentar
- **Akurasi Prediksi:** 62%
- **Keterangan:** Pada proporsi data uji 50%, model menunjukkan akurasi yang relatif tinggi dibandingkan dengan klasifikasi manual, yang hanya mencapai 40%. Algoritma WInnowing berhasil mengidentifikasi sentimen dengan akurasi yang lebih baik, mengindikasikan bahwa model dapat mengenali pola sentimen dengan lebih baik pada dataset yang cukup besar.

b) Proporsi Data Uji 40%:

- **Jumlah Data Uji:** 40 komentar
- **Akurasi Prediksi:** 58%
- **Keterangan:** Pada proporsi data uji 40%, akurasi model sedikit menurun dibandingkan dengan 50%, tetapi masih jauh lebih baik dibandingkan klasifikasi manual yang hanya memiliki akurasi 40%. Meskipun akurasi sedikit menurun, model tetap efektif dalam mengidentifikasi pola sentimen pada dataset yang lebih kecil.

c) Proporsi Data Uji 30%:

- **Jumlah Data Uji:** 30 komentar
- **Akurasi Prediksi:** 55.71%
- **Keterangan:** Pada proporsi data uji 30%, akurasi menurun lebih lanjut. Meskipun demikian, model masih dapat mengidentifikasi pola sentimen dengan tingkat akurasi yang lebih tinggi daripada klasifikasi manual (yang masih berada pada kisaran 40%). Hal ini menunjukkan bahwa dengan dataset yang lebih kecil, performa model mungkin sedikit terpengaruh.

d) Proporsi Data Uji 60%:

- **Jumlah Data Uji:** 60 komentar
- **Akurasi Prediksi:** 60%
- **Keterangan:** Pada proporsi data uji 60%, model mencapai akurasi yang stabil pada tingkat 60%, yang lebih tinggi dibandingkan dengan klasifikasi manual. Ini menunjukkan bahwa semakin besar data uji yang digunakan, model dapat mempertahankan performa yang lebih baik.

e) Proporsi Data Uji 70%:

- **Jumlah Data Uji:** 70 komentar
- **Akurasi Prediksi:** 60%
- **Keterangan:** Pada proporsi data uji 70%, model kembali menunjukkan akurasi yang stabil pada sekitar 60%, yang menunjukkan bahwa meskipun ukuran dataset semakin besar, akurasi model tetap terjaga. Performa ini jauh lebih baik dibandingkan dengan klasifikasi manual yang tetap berada pada kisaran 40%.

Tabel 1 menunjukkan rekapitulasi hasil eksperimen dengan berbagai skenario proposi data uji ini menunjukkan bahwa algoritma *Winnowing* lebih efektif dan efisien dalam mengklasifikasikan sentimen komentar di media sosial dibandingkan dengan klasifikasi manual, terutama pada proporsi data uji yang lebih besar. Secara umum, hasil menunjukkan bahwa semakin besar proporsi data uji yang digunakan, semakin tinggi dan stabil akurasi yang dicapai oleh algoritma *Winnowing*. Model ini memiliki akurasi yang jauh lebih tinggi daripada klasifikasi manual, yang menunjukkan efektivitasnya dalam memproses data besar dan tidak terstruktur.

Tabel 1. Rekapitulasi akurasi tiap skenario pengujian

Proporsi Data Uji	Akurasi <i>Winnowing</i>	Akurasi Manual
50%	62%	40%
40%	58%	40%
30%	55.71%	40%
60%	60%	40%
70%	60%	40%

Kemudian pada proporsi data uji yang lebih kecil (30%), akurasi model sedikit menurun, tetapi tetap lebih baik dibandingkan klasifikasi manual. Ini menunjukkan bahwa meskipun dengan dataset kecil, algoritma *Winnowing* tetap lebih efektif dalam melakukan klasifikasi dibandingkan dengan klasifikasi manual. Algoritma *Winnowing* terbukti lebih unggul dalam memproses teks secara otomatis dengan akurasi yang lebih stabil pada berbagai proporsi data uji. Ini memperlihatkan bahwa model ini sangat efektif untuk aplikasi klasifikasi sentimen, terutama pada volume data yang besar, seperti data dari media sosial. Dengan temuan ini, dapat disimpulkan bahwa algoritma *Winnowing* merupakan pilihan yang tepat untuk klasifikasi sentimen pada data komentar di media sosial, khususnya terkait Pemilu Presiden 2024, dengan tingkat akurasi yang konsisten lebih tinggi daripada metode klasifikasi manual.

4.2. Implikasi penelitian

Hasil eksperimen menggunakan algoritme *Winnowing* untuk klasifikasi sentimen komentar di Instagram terkait Pemilu Presiden 2024 menunjukkan kontribusi signifikan terhadap pengembangan teknologi dan analisis sosial. Dari sisi teknologi, algoritme *Winnowing* membuktikan dirinya sebagai solusi inovatif dalam memproses data besar dan tidak terstruktur, mengatasi keterbatasan metode sebelumnya, dan membuka peluang besar untuk sistem analisis sentimen otomatis yang lebih cepat, akurat, dan *scalable*. Keunggulan ini tidak hanya memberikan solusi praktis, tetapi juga mendefinisikan ulang standar efisiensi dalam pengolahan data besar di era digital.

Penelitian ini memperkuat temuan-temuan sebelumnya yang menunjukkan efektivitas algoritme dalam analisis teks. Misalnya, penelitian [1] yang menggunakan metode *Naive Bayes* untuk analisis sentimen opini masyarakat terhadap Presiden Jokowi sebelum dan sesudah Pilpres 2024 menunjukkan peningkatan akurasi dan efisiensi dalam klasifikasi teks. Selain itu, penelitian [2] yang menggunakan algoritma *Random Forest* untuk deteksi berita palsu pada Pemilu 2024 juga menemukan hasil serupa dalam hal kecepatan dan akurasi. Dengan demikian, penelitian ini tidak hanya mengkonfirmasi hasil-hasil tersebut, tetapi juga memperluas aplikasi algoritme *Winnowing* ke domain baru, yaitu analisis sentimen komentar di media sosial terkait isu politik.

Dalam pengelolaan opini publik, hasil penelitian ini menawarkan pendekatan inovatif yang memberikan wawasan mendalam tentang persepsi masyarakat terhadap calon presiden dan isu-isu penting lainnya. Temuan ini dapat memberdayakan partai politik, tim kampanye, dan pengawas pemilu untuk merespons dinamika opini publik dengan lebih cepat, transparan, dan strategis. Lebih jauh, algoritme ini berhasil menunjukkan efektivitas tinggi dalam menyaring komentar negatif dan hoaks, memperkuat upaya menjaga stabilitas sosial dan politik, yang menjadi tantangan besar di era informasi saat ini.

Penelitian ini tidak hanya membuktikan potensi algoritme *Winnowing* dalam pengelolaan data besar, tetapi juga menunjukkan efisiensinya dalam berbagai aplikasi, dari analisis kampanye politik hingga isu-isu sosial lainnya. Dalam konteks akademik, algoritme ini menawarkan pendekatan canggih yang menjanjikan untuk mengklasifikasikan teks tidak terstruktur, memberikan kontribusi signifikan bagi penelitian sosial dan teks mining di masa depan. Bahkan, dengan akurasi yang konsisten dan stabil, algoritme *Winnowing* memungkinkan pengembangan sistem peringatan dini yang inovatif untuk mendeteksi sentimen negatif atau hoaks, mendukung proses pengambilan keputusan yang lebih tepat dan tanggap. Penelitian ini menjadi tonggak penting yang mengokohkan peran algoritme *Winnowing* sebagai katalisator bagi kemajuan ilmu.

4. Kesimpulan

Berdasarkan hasil eksperimen, dapat disimpulkan bahwa Algoritme *Winnowing* memberikan akurasi yang lebih tinggi dan stabil dalam klasifikasi sentimen pada data teks komentar media sosial, khususnya di platform *Instagram*, dibandingkan dengan klasifikasi manual. Algoritme ini menunjukkan performa terbaik pada proporsi data uji yang lebih besar (50% hingga 70%), dengan akurasi tertinggi mencapai 62% pada data uji 50%, sedangkan klasifikasi manual terbatas pada akurasi 40%. Hasil ini menegaskan bahwa Algoritme *Winnowing* lebih efektif dan efisien dalam memproses teks secara otomatis, dengan tingkat akurasi yang lebih tinggi, terutama pada dataset yang lebih besar dan kompleks.

Penelitian ini juga menunjukkan bahwa penggunaan Algoritme *Winnowing* dalam klasifikasi sentimen dapat memberikan wawasan yang lebih mendalam mengenai persepsi publik terkait Pemilu Presiden 2024, serta meningkatkan efisiensi dalam mengelola opini publik di media sosial. Sebagai solusi untuk menangani tantangan yang muncul akibat volume data yang besar dan konten yang tidak terstruktur, Algoritme *Winnowing* menawarkan alternatif yang lebih cepat dan andal. Penelitian ini berkontribusi pada pengembangan sistem analisis sentimen berbasis teks yang dapat diterapkan pada skala yang lebih besar dan lebih responsif terhadap dinamika opini publik dalam ruang digital.

Daftar Referensi

- [1] P. H. Nehe, S. S. Berutu, and H. Budiati, "Analisis Sentimen Opini Masyarakat Terhadap Presiden Jokowi Sebelum Dan Sesudah Pilpres 2024 Menggunakan Metode Naive Bayes Classification," *JUTISI J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, pp. 451–463, 2024, doi: 10.35889/jutisi.v13i1.1841.
- [2] A. S. Nurhikam *et al.*, "Deteksi Berita Palsu Pada Pemilu 2024 Dengan Menggunakan Algoritma Random Forest," *DoubleClick J. Comput. Inf. Technol.*, vol. 7, no. 1, pp. 41–50, Aug. 2023, doi: 10.25273/DOUBLECLICK.V7I1.15456.
- [3] F. Amirullah, S. Alam, Mi. S. Sulisty, T. Informatika, S. Tinggi Teknologi Wastukencana, and R. Artikel, "Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini Twitter Menggunakan Naive Bayes," *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 3, pp. 69–76, Aug. 2023, doi: 10.55123/STORAGE.V2I3.2293.
- [4] I. Kurniawan and A. Susanto, "Implementasi Metode K-Means dan Naive Bayes Classifier untuk Analisis Sentimen Pemilihan Presiden (Pilpres) 2019," *J. Eksplora Inform.*, vol. 9, no. 1, pp. 1–10, Sep. 2019, doi: 10.30864/EKSPLORA.V9I1.237.
- [5] S. Puad, G. Garno, and A. S. Y. Irawan, "Analisis Sentimen Masyarakat Pada Twitter Terhadap Pemilihan Umum 2024 Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 3, pp. 1560–1566, Oct. 2023, doi: 10.36040/JATI.V7I3.6920.
- [6] M. F. Haikal, J. Indra, and R. Rahmat, "Analisis Sentimen Bakal Calon Presiden Indonesia 2024 Dengan Algoritma Naive Bayes," *JUTISI J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, pp. 43–51, 2024, doi: 10.35889/jutisi.v13i1.1598.
- [7] Y. Yuliadi and M. A. I. Pakereng, "Penerapan Text Mining Menggunakan Algoritme Naive Bayes Dalam Mengklasifikasi Sentimen Netizen di media sosial Twitter (Studi Kasus Pertemuan KTT G20)," *Progresif J. Ilm. Komput.*, vol. 19, no. 2, pp. 824–833, 2023, [Online]. Available: <http://ojs.stmik-banjarbaru.ac.id/index.php/progresif/article/view/1245%0Ahttp://ojs.stmik-banjarbaru.ac.id/index.php/progresif/article/download/1245/809>.

- [8] R. Sethi and I. Kaushik, "Hand written digit recognition using machine learning," in *2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT)*, 2020, pp. 49–54, doi: 10.1109/CSNT48778.2020.9115746.
- [9] M. B. Abdulrazzaq and J. N. Saeed, "A Comparison of Three Classification Algorithms for Handwritten Digit Recognition," in *2019 International Conference on Advanced Science and Engineering, ICOASE*, 2019, no. 46080, pp. 58–63, doi: 10.1109/ICOASE.2019.8723702.
- [10] M. A. Firmansyah and M. M. Alamin, "Sistem Prediksi Pengiriman Pada Dakota Cargo Menggunakan Metode Naive Bayes Berbasis Web," *JUTISI J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, pp. 324–335, 2024, doi: 10.35889/jutisi.v13i1.1802.
- [11] A. Simanungkalit, J. P. P. Naibaho, and A. De Kweldju, "Analisis Sentimen Berbasis Aspek Pada Ulasan Aplikasi Shopee Menggunakan Algoritma Naïve Bayes," *JUTISI J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, pp. 659–670, 2024, doi: 10.35889/jutisi.v13i1.1826.
- [12] G. G. Warow and H. Pandia, "Analisis Sentimen Aplikasi Dana Menggunakan Naïve Bayes Classifier dan Support Vector Machine," *JUTISI J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, pp. 609–621, 2024, doi: 10.35889/jutisi.v13i1.1893.
- [13] A. Asyer and M. A. I. Pakereng, "Analisis Sentimen Tweet Pengguna Twitter Terkait Diabetes Menggunakan Metode Naive Bayes," *JUTISI J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 12, no. 2, pp. 627–636, 2023, [Online]. Available: <http://ojs.stmik-banjarbaru.ac.id/index.php/jutisi/article/view/1234>.
- [14] E. B. Santoso and A. Nugroho, "Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook," *J. Eksplora Inform.*, vol. 9, no. 1, pp. 60–69, Sep. 2019, doi: 10.30864/EKSPLORA.V9I1.254.
- [15] S. P. Astuti and A. T.W., "Pemanfaatan Software Matrix Laboratory (Matlab) Untuk Meningkatkan Minat Belajar Mahasiswa Dalam Pembelajaran Fisika Kinematika," *Pendek. J. Pendidik. Berkarakter*, vol. 3, no. 2, pp. 54–57, Sep. 2020, doi: 10.31764/PENDEKAR.V3I2.2892.
- [16] E. L. Rara and E. Mailoa, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Perubahan Piala Dunia U-20," *Progresif J. Ilm. Komput.*, vol. 20, no. 1, p. 259, 2024, doi: 10.35889/progresif.v20i1.1550.
- [17] B. Imran, M. N. Karim, and N. I. Ningsih, "Klasifikasi Berita Hoax Terkait Pemilihan Umum Presiden Republik Indonesia Tahun 2024 Menggunakan Naïve Bayes Dan SVM," *J. Ilm. Din. Rekayasa*, vol. 20, no. 1, pp. 1–9, Jan. 2024, doi: 10.20884/1.DINAREK.2024.20.1.27.