

Klasifikasi Sentimen Aplikasi X Terhadap Gugatan Pemilu 2024 Menggunakan *Naïve Bayes* Dan *Textblob*

Nur'Ainun Suhharyani Azisa¹, Christian Dwi Suhendra²

nrainnsazs@gmail.com¹

^{1,2} Universitas Papua

Informasi Artikel

Diterima : 1 Jun 2024

Direvisi : 30 Jun 2024

Disetujui : 23 Jul 2024

Kata Kunci

Klasifikasi sentimen,
Naive Bayes, TextBlob,
Aplikasi X, gugatan
pemilu

Abstrak

Penelitian ini menganalisis sentimen publik terhadap gugatan hasil Pemilu 2024 di Mahkamah Konstitusi melalui aplikasi X (Twitter) menggunakan metode *Naïve Bayes* dan *TextBlob*. Dataset dikumpulkan melalui crawling dan *preprocessing* untuk menghapus data duplikat, membersihkan, dan menormalisasi tweet. Pelabelan dilakukan menggunakan TextBlob, kemudian klasifikasi sentimen diterapkan dengan algoritma Naïve Bayes. Hasil penelitian menunjukkan bahwa dari 898 tweet yang dianalisis, pelabelan TextBlob mengidentifikasi 340 tweet positif, 427 tweet netral, dan 131 tweet negatif. Sementara itu, klasifikasi *Naïve Bayes* menghasilkan 515 tweet positif, 281 tweet netral, dan 102 tweet negatif, dengan akurasi yang tinggi yaitu 95,29%. Visualisasi data melalui *word cloud* dan *bar chart* membantu memetakan distribusi sentimen secara jelas. Temuan ini memberikan wawasan berharga mengenai opini publik terhadap gugatan hasil pemilu dengan mayoritas sentimen bersifat positif dan netral.

Keywords

Sentiment classification,
Naive Bayes, TextBlob, X
application, election dispute

Abstract

This study analyzes public sentiment towards the 2024 Election Results Dispute at the Constitutional Court through the X (Twitter) application using the Naïve Bayes and TextBlob methods. The dataset was collected through crawling and preprocessing to remove duplicate data, clean, and normalize the tweets. Labeling was done using TextBlob followed by sentiment classification using the Naïve Bayes algorithm. The results show that out of 898 tweets analyzed, the TextBlob labeling identified 340 positive tweets, 427 neutral tweets, and 131 negative tweets. Meanwhile, the Naïve Bayes classification resulted in 515 positive tweets, 281 neutral tweets, and 102 negative tweets, demonstrating high accuracy with 95.29%. Data visualization through word clouds and bar charts helped map the sentiment distribution clearly. These findings provide valuable insights into public opinion on the election results dispute, with the majority of sentiments being positive and neutral.

A. Pendahuluan

Dunia telah memasuki abad ke-20 yang ditunjukkan dengan perkembangan teknologi yang semakin pesat. Teknologi seperti ponsel, komputer, dan internet telah muncul membentuk Cyberspace adalah ruang virtual luas dalam dunia nyata yang dimanfaatkan untuk berkomunikasi dan berinteraksi dengan orang lain di seluruh dunia[1]. Perkembangan teknologi internet juga mencakup teknologi informasi yang tercermin dalam fenomena layanan *microblogging* yang menghasilkan jutaan pesan setiap hari di situs-situs web populer seperti X/Twitter, Instagram, dan Facebook [2].

Aplikasi Twitter atau yang saat ini dikenal dengan nama X merupakan media sosial yang populer di Indonesia. Berdasarkan laporan We Are Social, pada Oktober 2023 terdapat sekitar 27,5 juta pengguna Aplikasi X/Twitter di Indonesia. Angka ini menempatkan Indonesia pada posisi keempat di dunia[3]. Melalui Aplikasi X, masyarakat dengan mudah dapat menyebarkan dan mengakses informasi secara bebas di internet. Selain itu, kebebasan beropini juga dapat dilakukan tanpa batas [4]. Di tahun ini salah satu topik yang sering diperbincangkan oleh warga X, yakni hasil pemilu 2024 yang digugat di Mahkamah Konstitusi.

Indonesia merupakan negara demokrasi yang mengadaptasi sistem pemilu. Pemilu nasional pertama di Indonesia diadakan pada tahun 1955 untuk memilih anggota DPR pada 29 September 1955 dan anggota Konstituante pada 25 Desember 1955.[5]. Kemudian, sampai saat ini Indonesia masih menggunakan pemilu sebagai sarana untuk memilih pemimpin-pemimpin di Indonesia, khususnya presiden dan wakil presiden. Seperti yang telah tertulis menurut Undang-Undang (UU) Nomor 7 Tahun 2017 tentang Pemilihan Umum, pemilu merupakan sarana kedaulatan rakyat untuk memilih presiden dan wakil presiden, anggota DPR, anggota DPD, serta anggota DPRD. Pemilu ini dilaksanakan secara langsung, umum, bebas, rahasia, jujur, dan adil dalam Negara Kesatuan Republik Indonesia berdasarkan Pancasila dan UUD 1945.[6]. Hal ini mengindikasikan bahwa demokrasi di Indonesia melibatkan partisipasi aktif masyarakat secara langsung dalam proses politik, khususnya dalam pemilihan pemimpin-pemimpin negara. Pelaksanaan pemilu yang berlangsung setiap 5 tahun membuat masyarakat ingin memastikan bahwa suara dan aspirasi dapat diwakili dengan baik oleh para pemimpin yang mereka pilih. Fakta buruk mengenai pelaksanaan pemilu di Indonesia adalah isu kecurangan dalam pemilu Indonesia semakin mencuat menuju hari pemungutan suara yang dilakukan pada 14 Februari 2024. Sejak peluncurannya pada 7 Januari 2024, situs kecuranganpemilu.com telah menerima 49 aduan publik. Sampai dengan 10 Februari 2024, telah ditemukan dan diverifikasi secara mandiri sedikitnya 53 masalah dan dugaan kecurangan pemilu. Hal ini menunjukkan kekhawatiran yang meningkat terhadap integritas dan kepercayaan dalam proses demokrasi. Potensi kecurangan menjadi perhatian serius seperti praktik politik uang, manipulasi oleh penyelenggara pemilu, kampanye terselubung, dan intimidasi pemilih melalui distribusi bantuan sosial. Bentuk-bentuk kecurangan tersebut menyebabkan adanya dugaan tindak pidana pemilu sehingga menjadi topik perbincangan hangat di X. Bahkan, keputusan hasil pemilu yang diumumkan oleh KPU dipertanyakan dan digugat di Mahkamah Konstitusi untuk pemeriksaan lebih lanjut [6].

Berdasarkan permasalahan tersebut pentingnya analisis sentimen dilakukan untuk memahami reaksi dan pandangan masyarakat. Penelitian ini menganalisis pendapat, sentimen, dan evaluasi, sikap, dan emosi seseorang yang diekspresikan dalam bahasa tertulis [7]. Penelitian ini bertujuan untuk mendapatkan wawasan dari data teks yang melimpah dari komentar dan cuitan di media sosial, khususnya di Aplikasi X. Dalam konteks ini, analisis sentimen dilakukan dengan menggunakan studi kasus untuk menilai opini masyarakat tentang berita gugatan pemilu 2024 di Mahkamah Konstitusi yang sedang ramai diperbincangkan. Analisis tersebut melibatkan pengelompokan komentar ke dalam tiga kategori utama: positif, netral, dan negatif [8]. Dengan klasifikasi tersebut para peneliti dapat memperoleh gambaran yang lebih jelas tentang bagaimana masyarakat bereaksi terhadap isu tersebut, serta mengidentifikasi tren dan pola sentimen yang dominan. Terdapat beberapa metode Algoritma yang dapat digunakan untuk melakukan analisis sentimen, seperti Support Vector Machine (SVM), K-NN, Decision Tree, dan lainnya. Namun, dalam penelitian ini metode yang digunakan adalah algoritma *Naïve Bayes* dengan metode pelabelan yaitu *TextBlob*[9].

Terdapat dua penelitian yang telah dilakukan untuk mendukung pemahaman yang lebih mendalam tentang dinamika opini masyarakat terkait isu-isu media sosial, khususnya pada aplikasi X/Twitter. *Pertama*, penelitian yang dilakukan oleh Dian Isnaeni Nurul Afra pada tulisan karya akhir Fakultas Ilmu Komputer, Program Studi Magister Teknologi Informasi, Universitas Indonesia yang berjudul “Analisis Sentimen dan Pemodelan Topik Dengan Data Media Sosial Twitter: Studi kasus Komisi Pemberantasan Korupsi”[10]. Penelitian tersebut menggunakan data dari media sosial X/Twitter yang mencakup pengguna Indonesia, terutama kelompok usia 25-34 tahun. Tujuannya adalah untuk mengembangkan model klasifikasi sentimen dan pemodelan topik untuk mengidentifikasi cuitan tentang LHKPN pejabat. Hasil pemodelan dengan *Latent Dirichlet Allocation* menghasilkan 10 topik pada kategori sentimen 'positif' dan 'negatif', dengan cuitan 'negatif' didominasi oleh kekecewaan dan kemarahan masyarakat terhadap peningkatan harta kekayaan pejabat selama pandemi Covid-19, sedangkan cuitan 'positif' mencakup berbagai pembahasan tentang aturan pembuktian terbalik hingga permintaan untuk transparansi laporan kekayaan [10].

Kedua, penelitian oleh Sisferi Hikmawan, Amsal Pardamean, dan Siti Nur Khasanah dalam jurnal berjudul “Sentimen Analisis Publik terhadap Joko Widodo terhadap Wabah Covid-19 Menggunakan Metode *Machine Learning*” mengkaji evaluasi pendapat publik melalui sentimen positif, negatif, dan netral di media X/Twitter terkait tindakan pencegahan Covid-19 oleh pemerintah Indonesia. Hasil penelitian ini mencakup upaya yang dapat menjadi acuan bagi

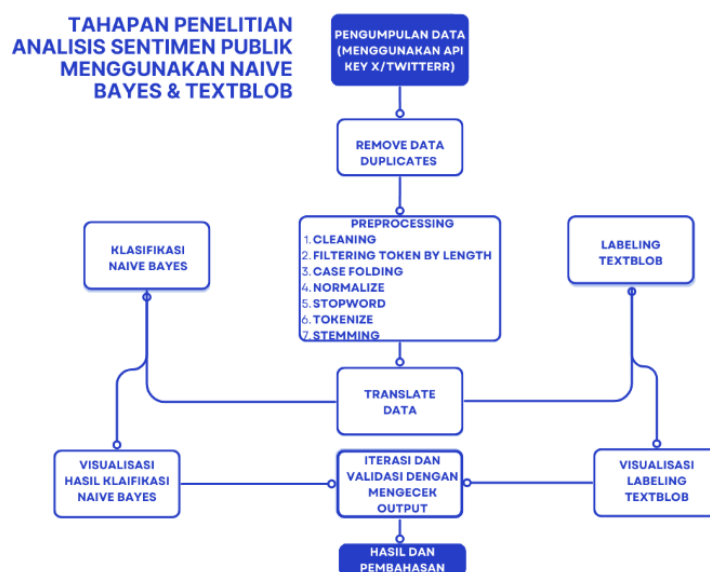
indikator keberhasilan pemerintah. Representasi dilakukan melalui text mining dengan metode *Support Vector Machine (SVM)*, *Naive Bayes*, dan *k-NN* untuk mengklasifikasikan label sentimen dari dataset [11].

Meskipun dua penelitian tersebut menggunakan metode yang berbeda, tetapi keduanya memberikan wawasan yang penting terkait penggunaan *machine learning* terhadap klasifikasi sentimen publik. Di sisi lain dalam penelitian saya yang berjudul “Analisis Sentimen Pengguna Aplikasi X Indonesia Terhadap Gugatan Pemilu 2024 Di Mahkamah Konstitusi Dengan Metode *Naïve Bayes*”, keputusan untuk menggunakan hanya satu metode (Algoritma Naive Bayes)

didasarkan pada beberapa pertimbangan, termasuk fokus yang lebih terarah pada evaluasi spesifik dari metode tersebut dan memperkuat konsistensi dalam analisis. Hal ini juga mempermudah interpretasi hasil penelitian dan penyajian, memungkinkan penjelasan yang lebih terperinci tentang relevansi hasil dengan tujuan penelitian.

B. Metode Penelitian

Metode penelitian ini berfokus pada tahap implementasi pengambilan dataset dari aplikasi X melalui data *crawling*. Setelah data terkumpul, dilakukan proses penghapusan data ganda dan *preprocessing* untuk membersihkan dan menormalisasi data *tweet* yang dilanjutkan dengan pelabelan menggunakan *TextBlob*. Data tersebut kemudian diklasifikasikan menggunakan algoritma *Naive Bayes* untuk menentukan sentimen setiap bagian data apakah positif, netral, atau negatif [12]. Adapun tahapan penelitian seperti di gambar 1.



Gambar 1. Bagan Tahapan Penelitian

1. *Crawling* (Pengumpulan) Data

Crawling data di Aplikasi X/Twitter adalah proses penting dalam pengumpulan informasi karena data diambil atau diunduh dari server X melalui bantuan Application Programming Integration (API) X, baik dalam bentuk data pengguna maupun data *tweet*[13]. Data yang diunduh terdiri dari kumpulan komentar pengguna X (*tweet*) dan kemudian disimpan dalam file berekstensi .csv untuk diproses pada tahap normalisasi (*preprocessing*) [12]. Dalam penelitian ini, proses pengunduhan data dilakukan melalui Google Colab menggunakan bahasa pemrograman Python dengan Library (X/Twitter) untuk melakukan *crawling* terhadap kumpulan opini yang terkait dengan kata kunci "Gugatan Pemilu 2024 di MK" pada aplikasi X [12]. Penggunaan Google Colab dan Python mempermudah pelaksanaan proses ini serta memungkinkan analisis yang lebih efektif dan efisien.

2. Remove data Duplicates

Setelah mendapatkan hasil crawling, tahap penting selanjutnya adalah menangani duplikasi dalam dataset. Proses ini melibatkan penyaringan *tweet* yang mungkin muncul berulang kali dalam kumpulan data yang diunduh sehingga dapat dipastikan bahwa data tersebut disimpan dalam dataset untuk menghindari duplikasi yang tidak perlu dan memastikan integritas data yang lebih baik untuk analisis lebih lanjut [14]. Berikut merupakan gambar 2 yang menunjukkan model data duplikat yang perlu dihilangkan.

full_text	username
Gugatan ditolak Mahkamah Konstitusi! Keputusan DKPP atas pelanggaran kode etik Ketua KPU tidak bisa membatalkan pencalonan Paslon Prabowo-Gibran https://t.co/VNha5srR40	dedefreefire
Gugatan ditolak Mahkamah Konstitusi! Keputusan DKPP atas pelanggaran kode etik Ketua KPU tidak bisa membatalkan pencalonan Paslon Prabowo-Gibran https://t.co/cOL3ue0yH7	dedefreefire
Gugatan ditolak Mahkamah Konstitusi! Keputusan DKPP atas pelanggaran kode etik Ketua KPU tidak bisa membatalkan pencalonan Paslon Prabowo-Gibran https://t.co/tV9ek5GZud	dedefreefire
Gugatan ditolak Mahkamah Konstitusi! Keputusan DKPP atas pelanggaran kode etik Ketua KPU tidak bisa membatalkan pencalonan Paslon Prabowo-Gibran https://t.co/flgTh50ORb	dedefreefire
FINAL! Semua Gugatan Kubu 01 di MK Ditolak karena Tidak Berlandaskan Hukum yang Jelas	

Gambar 2. Contoh Data yang Memiliki Duplikat Data

Penanganan duplikat juga dapat membantu dalam meminimalkan bias yang mungkin timbul dari duplikat informasi dan meningkatkan kualitas analisis data sentimen. Selain itu, mengidentifikasi dan menghapus duplikat dapat memberikan gambaran yang lebih akurat serta mewakili opini dan tren dalam kumpulan data sentimen.

Langkah ini juga memungkinkan penggunaan sumber daya yang lebih efisien dan mengurangi beban komputasi yang tidak diperlukan saat melakukan analisis data tingkat lanjut. Oleh karena itu, penanganan duplikat merupakan langkah penting dalam mempersiapkan kumpulan data untuk analisis sentimen yang lebih dalam dan akurat.

3. Preprocessing

Text preprocessing merupakan tahap mengubah bentuk data yang belum terstruktur berupa *term index* yang mewakili dataset dokumen[15]. *Preprocessing* akan mengeliminasi sejumlah permasalahan seperti *missing value*, *data redundant*, *outliers*, dan format data yang tidak sesuai dengan sistem yang berpotensi dapat mengurangi performa proses *data mining*. Setelah dilakukan *preprocessing* diharapkan data telah direpresentasikan dalam kondisi yang ideal untuk diproses [12]. *Preprocessing* memiliki beberapa tahapan antara lain[12][14][7]:

- a) *Cleaning* merupakan kegiatan menganalisa kualitas data dengan cara memodifikasi, mengubah, atau menghapus data-data yang dianggap tidak perlu, tidak lengkap, tidak akurat, atau format data yang salah dalam basis data guna menghasilkan data berkualitas tinggi[16], seperti html, koneksi, *hashtag*, *username*, dan konten. Selain itu, proses ini juga menghilangkan tanda-tanda non-alfanumerik dan tanda aksentuasi lainnya.
- b) *Filter Token by length* merupakan proses menyaring dan membatasi karakter dalam kata-kata yang digunakan dalam penelitian sehingga meminimalisir kata-kata yang tidak relevan[17].

- c) *Case folding* merupakan proses untuk mengubah kalimat atau menjadi huruf kecil (*lowercase*)[16]. Selanjutnya, terdapat proses ketika nilai '*false*' dalam kolom tersebut diganti dengan *pd.NA* untuk menangani nilai yang tidak relevan. Kemudian, dilanjutkan proses untuk mengkosongkan baris dengan nilai kosong dari dataset menggunakan *method dropna* sehingga memastikan data yang digunakan untuk analisis bersih dan lengkap.
- d) *Normalize* merupakan tahap yang digunakan untuk menormalisasi kata-kata khusus melibatkan penggantian singkatan, bentuk tidak baku, dan kata-kata khusus dalam teks dengan bentuk yang lebih umum atau formal. Hal ini membuat teks lebih konsisten dan lebih mudah dipahami untuk algoritma pemrosesan bahasa alami. Misalnya perubahan singkatan seperti "u" menjadi "anda" dan "gmna" menjadi "bagaimana". Normalisasi ini mempersiapkan data teks agar lebih mudah diolah pada tahap selanjutnya dalam analisis.
- e) *Tokenizing* adalah tahap pemotongan *string* berdasarkan kata yang membentuk sebuah dokumen. Proses ini mengurai teks menjadi kata-kata dengan memperhatikan tanda baca dan spasi.[15]
- f) *Filtering (stopword removal)* adalah tahap mengambil kata-kata penting dari hasil *tokenizing* dan penghilangan kata-kata yang dianggap tidak memiliki makna[15]
- g) *Stemming* merupakan proses pemetaan dan penguraian bentuk dari suatu kata menjadi bentuk dasarnya. Pada proses *stemming* digunakan library yang terdapat pada python yaitu *sastrawi*[18].

4. Translate dataset

Analisis sentimen menggunakan klasifikasi Naive Bayes dan TextBlob sering kali memerlukan terjemahan ke dalam bahasa Inggris karena beberapa alasan penting. *Pertama*, banyak model dan alat analisis sentimen termasuk TextBlob yang terutama dilatih pada teks berbahasa Inggris sehingga lebih akurat dan andal jika digunakan dengan data berbahasa Inggris. *Kedua*, sumber daya analisis sentimen mungkin kurang berkembang dalam bahasa lain sehingga membuat model bahasa Inggris lebih kuat. *Ketiga*, setiap bahasa mempunyai karakteristik unik dalam hal sintaksis dan morfologi, yang dapat membuat analisis menjadi lebih kompleks, dan penerjemahan ke dalam bahasa Inggris dapat mengurangi kompleksitas tersebut. Terakhir, penerjemahan data dapat membantu dalam standarisasi proses analisis, khususnya untuk kumpulan data multibahasa dan memastikan bahwa semua teks dianalisis dengan pendekatan yang konsisten. Oleh karena itu, hal pertama yang perlu dilakukan adalah proses penerjemahan. Pada penelitian ini terdapat data *tweet* dalam bahasa Indonesia dan kata bantu bahasa Inggris sehingga digunakan modul TextBlob untuk pengolahannya. Meskipun terjemahan otomatis dapat mengubah nuansa dan konteks sentimen asli,

penggunaan terjemahan masih merupakan solusi praktis yang memanfaatkan alat dan model yang ada[19].

5. Labeling Menggunakan TextBlob

Textblob adalah salah satu library yang disediakan oleh Python untuk pemrosesan di bidang *Natural Language Processing* yang dapat memberikan tag kata, ekstraksi kata, penerjemahan kata dan *sentiment analysis*. Dalam penelitian ini digunakan *TextBlob* untuk proses *labeling* awal dalam melakukan analisis sentimen berdasarkan opini publik di X/Twitter. Data tweet yang diambil diterjemahkan dahulu ke dalam bahasa Inggris menggunakan library *googletrans==4.0.0-rc1* dan diproses menggunakan *TextBlob* untuk menghitung nilai polaritas. Nilai polaritas digunakan untuk menentukan label sentimen positif, netral, atau negatif[20]. Analisis sentimen dengan *TextBlob* berfungsi dan bekerja untuk memberikan skor sentimen pada setiap kata dalam teks, kemudian menghitung skor sentimen keseluruhan dari teks tersebut. *TextBlob* menggunakan model pembelajaran mesin untuk menilai sentimen dari teks yang telah diproses sebelumnya[21][22].

6. Klasifikasi Menggunakan Naïve Bayes

Metode algoritma *Naïve Bayes* merupakan metode sederhana yang memiliki performa yang cepat dalam melatih data, mudah dalam implementasinya, serta memiliki efektifitas yang tinggi [15]. Algoritma *Naïve Bayes* merupakan salah satu teknik klasifikasi yang menggunakan metode probabilitas dan statistik yang dikembangkan oleh ilmuwan Inggris, Thomas Bayes. Metode ini memproyeksikan kemungkinan di masa depan berdasarkan pengalaman di masa lalu, yang dikenal sebagai Teorema Bayes. Teorema ini kemudian digunakan dengan konsep *Naïve*, di mana diasumsikan bahwa setiap atribut adalah independen satu sama lain. Dalam klasifikasi *Naïve Bayes*, diasumsikan bahwa keberadaan atau ketiadaan fitur tertentu dalam sebuah kelas tidak bergantung pada fitur-fitur dari kelas lain.[23]. Algoritma *Naive Bayes* dipilih karena metode klasifikasi dengan formula sederhana dan mudah diterapkan dan cukup tinggi dibandingkan dengan metode lain[24]. Dalam penggunaan *Naive Bayes* untuk menganalisis sentimen, data tweet digunakan sebagai bahan penelitian. Data ini diproses menggunakan metode Klasifikasi *Naive Bayes* untuk memisahkan sentimen positif, netral dan negatif.[25].

7. Visualisasi

Pada penelitian ini terdapat dua macam visualisasi, yaitu visualisasi menggunakan *bar chart* dan juga visualisasi menggunakan *word cloud*. *Word Cloud* digunakan untuk mengidentifikasi kata-kata yang paling sering digunakan dalam tweet dan menghasilkan bentuk kata yang dapat bergabung dengan kata-kata lain untuk menghasilkan informasi yang krusial[26], sedangkan *bar chart* digunakan untuk mempresentasikan hasil analisis sentimen. *Bar chart* dapat digunakan untuk menunjukkan distribusi frekuensi kata-kata yang digunakan dalam tweet[27].

8. Iterasi dan Validasi dengan Mengecek Output

Iterasi dan validasi dalam penelitian analisis sentimen dengan metode klasifikasi *Naive Bayes* dan *TextBlob* bertujuan utama menghindari *overfitting* sehingga memastikan kinerja model yang akurat dan mencegah *underfitting*. Dengan menguji model pada berbagai dataset dapat memastikan kegunaan model pada data baru dan meningkatkan generalisasinya sehingga dapat dibuktikan bahwa model tidak hanya sesuai dengan data pelatihan, tetapi juga dapat diterapkan secara luas.

9. Hasil dan Pembahasan

Pembahasan dilakukan setelah proses pengumpulan, prapemrosesan, *labeling*, dan klasifikasi data selesai, tahap evaluasi dilakukan untuk mengevaluasi kinerja model analisis sentimen. Evaluasi ini bertujuan untuk mengukur seberapa baik model *Naive Bayes* serta *TextBlob* dalam mengklasifikasikan sentimen berdasarkan data *tweet* yang diproses. Dengan melakukan evaluasi ini peneliti dapat menarik kesimpulan yang lebih tepat tentang kualitas analisis sentimen yang dilakukan dengan menggunakan model *Naive Bayes* dan *TextBlob*[9].

C. Hasil dan Pembahasan

Setelah melakukan seluruh proses analisis yang didasarkan pada metode penelitian dengan menggunakan *Naive Bayes*, *Textblob*, dan data *tweet* yang berisikan sentimen publik mengenai hasil pemilu 2024 yang digugat di Mahkamah Konstitusi dapat diuraikan sebagai berikut.

1. Hasil Crawling data

Berdasarkan hasil Crawling data, diperoleh 1023 data yang berisi *tweet* mengenai sentimen publik terhadap hasil pemilu 2024 yang digugat pada Mahkamah Agung. Data tersebut belum melewati proses pembersihan data sehingga masih banyak ditemukan data duplikat dan data tersebut belum melewati proses normalisasi. Berikut merupakan hasil crawling data yang dapat dilihat pada gambar 3.

1 to 25 of 937 entries Filter 🔍

index	full_text	username
0	@Miduk17 @ivolitus_hubert Hanya trik mereka saja tidak elok kalau semua gugatan ditolak. Jeruk makan jeruk.	sandyalfia96694
1	@simjaevv @waz_hereee @tubirfess + semua gugatan ditolak dg alasan2 yg bikin tambah jelas kalo hukum indo rusak	flavescens
2	@Summersibl36928 Bawa aja ke MK.. Biar semua gugatan ditolak	Jelli_cent
3	Gugatan 01 dan 03 Ditolak MK Nicho Silalahi angkat bicara untuk Segera Tumbangkan Jokowi Secepatnya! https://t.co/yDF3fU9S1F lewat @Keuangan News	keuangannews_id
4	ANIES CUMA PUJI PRABOWO USAI GUGATAN DITOLAK MK dan MASIH KRITIK ETIK GIBRAN? https://t.co/sTjLnyaNpf	venomencret
5	@BangRQ @prabowo Gugatan ditolak oleh MK tapi salut sama sikap pak anies yg tetap legowo #PrabowoPresiden	ishwarbike
6	@rixzuza @prabowo @gibran_tweet @Jokowi Yes sudah pasti #PrabowoPresiden karena gugatan Anies-Ganjar ditolak MK	arifhidayat55_
7	Alhamdulillah puas bngt sama hasil MK kmrn Gugatan pasangan lawan ditolak krn kurangnya bukti yg valid dimata hukum Finally pak @Prabowo Mas Gibran Pemenang Pilpres 2024 @tvOneNews #PrabowoPresiden https://t.co/lbzTI0sYa0N	missriantiii
8	Kontestasi Pilpres 2024 memasuki babak akhir setelah upaya gugatan PHPU Pilpres 2024 yang diajukan capres-cawapres nomor urut 01 dan 03 ditolak Mahkamah Konstitusi. #PrabowoPresiden	noernailaa
9	Gugatan 01 dan 03 ditolak oleh MK karena itu Paslon Prabowo gibran resmi jadi presiden hanya menunggu pelantikan. @Prabowo @Tribunnews @tvOne #PrabowoPresiden https://t.co/09TVOYYPNly	kadia_anin
10	Alhamdulillah puas bngt sama hasil MK kmrn Gugatan pasangan lawan ditolak krn kurangnya bukti yg valid dimata hukum Finally pak @Prabowo Mas Gibran Pemenang Pilpres 2024 @tvOneNews #PrabowoPresiden https://t.co/w2vWLeUX3B	NuelSubekti

Gambar 3. Hasil Data yang telah di Crawling di Aplikasi X

2. Hasil data yang telah Melewati Proses *Remove Data Duplicates* dan *Preprocessing*

Setelah data melalui proses penghapusan duplikat dan preprocessing, hasil yang diperoleh menunjukkan data yang lebih bersih dan juga terbebas dari data duplikat. Setelah penghapusan duplikat, data yang dapat digunakan berjumlah 937 data yang siap untuk dianalisis, seperti gambar 4 di bawah ini.

1 to 10 of 898 entries Filter 

	full_text
0	trik elok kalau semua gugat tolak jeruk makan jeruk
1	semua gugat tolak alas bikin tambah jelas kalau hukum indonesia rusak
2	bawa mahkamah konstitusi biar semua gugat tolak
3	gugat 01 03 tolak mahkamah konstitusi nicho silalahi angkat bicara segera tumbang jokowi cepat lewat news
4	anies cuma puji prabowo usai gugat tolak mahkamah konstitusi kritik etik gibran
5	gugat tolak mahkamah konstitusi salut sama sikap pak anies tetap legowo
6	yes gugat aniesganjar tolak mahkamah konstitusi
7	alhamdulillah puas bngt sama hasil mahkamah konstitusi kmrn gugat pasang lawan tolak kurang bukti valid mata hukum finally pak mas gibran menang pilpres 2024
8	kontestasi pilpres 2024 pasuk babak akhir upaya gugat phpu pilpres 2024 aju caprescapres nomor urut 01 03 tolak mahkamah konstitusi
9	gugat 01 03 tolak mahkamah konstitusi paslon prabowo gibran resmi jadi presiden tunggu lantik

Show 10 per page 1 2 10 80 90

Gambar 4. Potongan data yang telah dibersihkan dari data duplikat dan telah melewati proses

3. Hasil *Translate Data*

Data yang sudah melalui proses pembersihan dan normalisasi akan diterjemahkan ke bahasa Inggris. Hal ini dilakukan sebab akan menggunakan *TextBlob*. Proses translate akan menggunakan 'googletrans==4.0.0-rc1' yang berfungsi untuk menerjemahkan teks dari satu bahasa ke bahasa lain menggunakan API Google Translate.

Unnamed: 0		full_text	translated_text
0	0	trik elok kalau semua gugat tolak jeruk makan ...	Elok Tricks If all sue reject oranges eating o...
1	1	semua gugat tolak alas bikin tambah jelas kala...	All claims to refuse Alas make it more clear i...
2	2	bawa mahkamah konstitusi biar semua gugat tolak	Bring the Constitutional Court so that all sue...
3	3	gugat 01 03 tolak mahkamah konstitusi nicho si...	claim 01 03 Reject the Constitutional Court Ni...
4	4	anies cuma puji prabowo usai gugat tolak mahka...	Anies only praise Prabowo after the claim of t...
...	...	...	...
893	932	mahkamah konstitusi tolak gugat pilpres kubu amin	The Constitutional Court Refuses the Kubu Amin...
894	933	mahkamah konstitusi tolak seluruh gugat pilpre...	The Constitutional Court Rejected the entire c...
895	934	kecewa berat imbas mahkamah konstitusi tolak g...	Disappointed, the impact of the Constitutional...
896	935	mahkamah konstitusi tolak gugat pilpres amin d...	The Constitutional Court Refuses Amin's Presid...
897	936	mahkamah konstitusi tolak gugat anies muhaimin...	The Constitutional Court Refuses Anies Muhaimi...

898 rows x 3 columns

Gambar 5 Hasil data yang telah di Translate ke dalam Bahasa Inggris

4. Hasil Labeling Menggunakan *TextBlob*

Analisis sentimen ini dilakukan dengan menggunakan fungsi *polarity* dari *TextBlob* untuk menghitung polaritas sentimen dari setiap tweet dalam dataset yang telah diterjemahkan ke dalam bahasa Inggris. Setiap sentimen kemudian dikategorikan menjadi positif, netral, atau negatif berdasarkan nilai polaritasnya. Setelah menerjemahkan data ke bahasa Inggris, *TextBlob* dapat

diterapkan untuk melakukan *labeling* sentimen pada dataset tweet sehingga mendapat hasil berupa label sentimen untuk setiap tweet, yaitu positif, negatif, atau netral. Pada analisis ini didapatkan jumlah label dari setiap sentimen seperti gambar 6.

```
Hasil Analisis Data:
Positif= 340
Netral = 427
Negatif = 131

Total Data: 898
```

Gambar 6 Jumlah Hasil Analisis data TextBlob

Kemudian hasil detail dari labeling menggunakan TextBlob seperti gambar 7 di bawah ini:

Unnamed: 0		full_text	translated_text	klasifikasi
0	0	trik elok kalau semua gugat tolak jeruk makan ...	Elok Tricks If all sue reject oranges eating o...	Netral
1	1	semua gugat tolak alas bikin tambah jelas kala...	All claims to refuse Alas make it more clear l...	Positif
2	2	bawa mahkamah konstitusi biar semua gugat tolak	Bring the Constitutional Court so that all sue...	Netral
3	3	gugat 01 03 tolak mahkamah konstitusi nicho si...	claim 01 03 Reject the Constitutional Court NI...	Positif
4	4	anies cuma puji prabowo usai gugat tolak mahka...	Anies only praise Prabowo after the claim of t...	Netral
...	...	...	...	...
893	932	mahkamah konstitusi tolak gugat pilpres kubu amin	The Constitutional Court Refuses the Kubu Amin...	Netral
894	933	mahkamah konstitusi tolak seluruh gugat pilpre...	The Constitutional Court Rejected the entire c...	Netral
895	934	kecewa berat imbas mahkamah konstitusi tolak g...	Disappointed, the impact of the Constitutional...	Negatif
896	935	mahkamah konstitusi tolak gugat pilpres amin d...	The Constitutional Court Refuses Amin's Presid...	Positif
897	936	mahkamah konstitusi tolak gugat anies muhaimi...	The Constitutional Court Refuses Anies Muhaimi...	Netral

898 rows x 4 columns

Gambar 7 Hasil detail dari labeling menggunakan TextBlob

5. Hasil Klasifikasi Menggunakan Naïve Bayes

Hasil yang diperoleh menunjukkan distribusi sentimen dalam dataset tweet yang telah diproses. *Naïve Bayes* merupakan algoritma pembelajaran mesin berbasis probabilistik yang digunakan untuk mengklasifikasikan tweet ke dalam kategori sentimen positif, negatif, atau netral. Pada saat melakukan klasifikasi digunakan *'tuple'* untuk mengubah *DataFrame* menjadi format yang dapat digunakan untuk melatih model *Naïve Bayes* dan *'cl.accuracy'* (*cl=NaiveBayesClassifier*) dari *TextBlob* digunakan untuk melakukan klasifikasi sentimen dan mengevaluasi kinerjanya. Dalam penelitian ini model *Naïve Bayes* menunjukkan akurasi yang sangat tinggi pada dataset yang menunjukkan potensi model untuk mengklasifikasikan teks dengan baik. Adapun jumlah hasil dari klasifikasi setiap sentimen menggunakan *Naïve bayes* pada gambar 8..

```
Akurasi Test: 0.9529411764705882

Hasil Analisis Data:
Positif= 341
Netral = 434
Negatif = 123

Total Data: 898
```

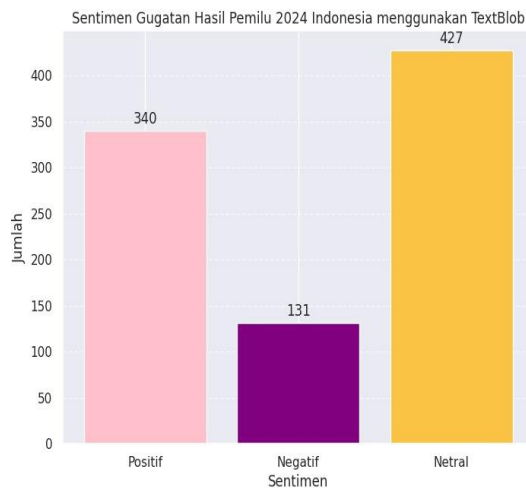
Gambar 8 Hasil tes akurasi dan jumlah keseluruhan sentimen model Naïve Bayes

Gambar 9 Hasil klasifikasi Naïve Bayes

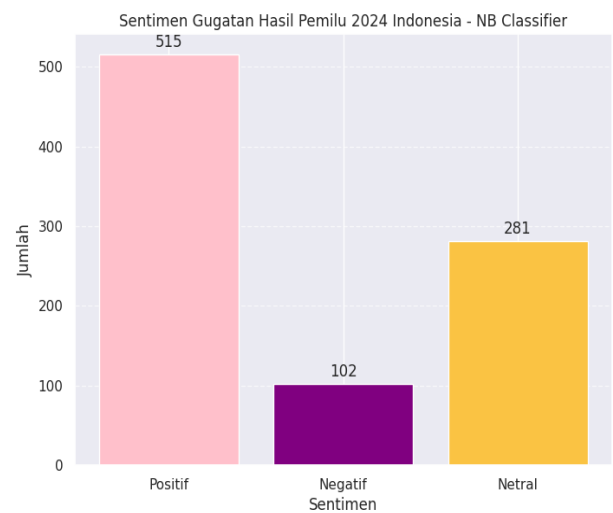


2. Visualisasi Menggunakan Bar Chart

Bar chart ini secara jelas dan informatif memvisualisasikan distribusi sentimen terhadap gugatan hasil pemilu 2024 di Indonesia. Dalam bar chart ini jumlah tweet yang terkait dengan masing-masing kategori sentimen, yakni positif, negatif, dan netral yang direpresentasikan oleh batang-batang chart berwarna yang mudah dipahami. Visualisasi ini memberikan wawasan yang mendalam tentang bagaimana opini publik terbagi terhadap topik tersebut sehingga memungkinkan pemahaman yang lebih baik tentang dinamika dan pandangan yang ada di masyarakat terkait dengan gugatan hasil pemilu. Terdapat dua visualisasi bar chart seperti pada gambar 11 dan gambar 12.



Gambar 11. Bar Chart hasil labeling menggunakan textblob



Gambar 12. Bar chart hasil Klasifikasi menggunakan Naïve Bayes

7. Hasil Iterasi dan Validasi dengan Mengecek Output

Iterasi digunakan untuk memeriksa setiap tuple dalam `data\_eval` dan melakukan validasi pada klasifikasi reguler dan klasifikasi *Naive Bayes*. Proses iterasi dilakukan melalui *loop for*, mengambil setiap tuple satu per satu dari dataset. Setiap tuple diproses dengan memeriksa panjangnya terlebih dahulu sebelum mencoba mengakses indeks untuk klasifikasi reguler dan klasifikasi *Naive Bayes*. Jika panjang tuple kurang dari 4, pesan kesalahan "Tuple tidak memiliki cukup elemen" akan dicetak sehingga menandakan adanya kesalahan dalam data. Namun, jika panjang tuple mencukupi, kode akan memeriksa apakah klasifikasi reguler dan klasifikasi *Naive Bayes* sama. Jika keduanya berbeda, informasi tentang teks, klasifikasi reguler, dan klasifikasi *Naive Bayes* akan dicetak. Dengan demikian, output dari ini adalah informasi tentang teks, klasifikasi reguler, dan klasifikasi *Naive Bayes* untuk setiap tuple yang memenuhi kriteria validasi yang ditetapkan sesuai dengan gambar di bawah ini.

```

Text : mahkamah keluarga tunggu sumpah nya azab allah kalau salah putus kan
classifier: Negatif
classifier Bayes Nb: Positif

Text : syukur tolak semua emang gugat receh
classifier: Netral
classifier Bayes Nb: Positif

Text : banyak banget yak begini baik sama doa in moga tetap lindung moga jadi boomerang wkwkwk kalah terima gugat tolak makin panas xixixi dukung siapa tuch
classifier: Positif
classifier Bayes Nb: Negatif

Text : kalau terima gugat bukan mahkamah konstitusi nama nama mahkamah konstitusi apa alas bukti aju gugat salah bukti mahkamah konstitusi sejarah
mahkamah konstitusi kabul gugat sidang orang tahu hasil 1000 tolak
classifier: Negatif
classifier Bayes Nb: Netral

Text : gugat tolak seluruh hakim mahkamah konstitusi jimly asshiddiqie kemarin jabat ketua mahkamah konstitusimahkamah konstitusi periksa hakim anwar usman
classifier: Netral
classifier Bayes Nb: Positif

```

Gambar 13 hasil Iterasi dan validasi

D. Simpulan

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap gugatan hasil Pemilu 2024 di Mahkamah Konstitusi menggunakan metode Naïve Bayes. Data tweet dikumpulkan dari aplikasi X (sebelumnya Twitter) dan diproses melalui beberapa tahap: crawling data, preprocessing, penerjemahan, pelabelan dengan TextBlob, klasifikasi dengan Naïve Bayes, dan visualisasi hasil. Dari 1023 tweet yang dikumpulkan, 937 tweet digunakan setelah penghapusan duplikat dan preprocessing. Tweet diterjemahkan ke dalam bahasa Inggris menggunakan googletrans dan dilabeli dengan TextBlob, menghasilkan 340 tweet positif, 427 netral, dan 131 negatif.

Klasifikasi menggunakan Naïve Bayes menunjukkan 515 tweet positif, 281 netral, dan 102 negatif, dengan akurasi tinggi sebesar 95.29%. Mayoritas tweet menunjukkan sentimen positif terhadap gugatan hasil Pemilu 2024 di Mahkamah Konstitusi, diikuti oleh sentimen netral dan negatif. Tingginya akurasi metode Naïve Bayes menunjukkan keefektifannya dalam analisis sentimen berbasis teks di media sosial, meskipun masih ada kesalahan klasifikasi. Analisis ini memberikan wawasan berharga bagi pihak terkait untuk memahami reaksi masyarakat dan merancang strategi komunikasi yang lebih efektif dan responsif.

Secara keseluruhan, penelitian ini mengindikasikan bahwa metode Naïve Bayes dan TextBlob masih perlu dioptimalkan untuk analisis sentimen dari media sosial. Namun, hasilnya tetap memberikan gambaran jelas tentang opini publik terkait gugatan hasil Pemilu 2024 dan pentingnya analisis sentimen dalam memahami dinamika opini publik di era digital. Penelitian lebih lanjut diperlukan untuk meningkatkan akurasi dan efektivitas metode yang digunakan.

E. Referensi

- [1] U. Z. Aqila, "Universitas indonesia analisis sentimen pengguna twitter terhadap perilaku," 2023.
- [2] Yislam, "Analisis Sentimen Masyarakat Terhadap Pemerintahan Jokowi Menggunakan Data Twitter."
- [3] A. M. Cindy, "27 Juta Pengguna Twitter di Indonesia, Terbanyak ke-4 GlobalNo Title," Indonesia, 2024. [Daring]. Tersedia pada: <https://databoks.katadata.co.id/datapublish/2023/11/28/ada-27-juta-pengguna-twitter-di-indonesia-terbanyak-ke-4-global>.
- [4] Y. Nurdiansyah, F. Rahman, P. Pandunata, dan A. Infantono, "Undang-Undang

- Cipta Kerja pada Twitter,” vol. 3, no. November, hal. 24–25, 2021, doi: 10.54706/senastindo.v3.2021.158.
- [5] KPU, “Sejarah Pemilihan Umum Republik Indonesia.”
 - [6] M. S. Maarif dan Tajul Arifin, “Kecurangan Pemilu Tahun 2024 Dalam Perspektif Uu No. 7 Tahun 2017 Dan Hadits Riwayat Muslim Muhamad,” vol. 3, no. 3, 2024.
 - [7] R. Vindua dan A. U. Zailani, “Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python,” vol. 10, no. 2, hal. 479–487, 2024, doi: 10.30865/jurikom.v10i2.5945.
 - [8] D. Normawati dan S. A. Prayogi, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” vol. 5, no. November 2019, hal. 697–711, 2021.
 - [9] F. Fazrin, O. N. Pratiwi, dan R. Andreswari, “Perbandingan Algoritma K-Nearest Neighbor dan Logistic Regression pada Analisis Sentimen terhadap Vaksinasi Covid-19 pada Media Sosial Twitter dengan Pelabelan Vader dan Textblob,” vol. 10, no. 2, hal. 1596–1604, 2023.
 - [10] D. I. N. Afra., “Analisis Sentimen Dan Pemodelan Topik Dengan Data Media Sosial Twitter: Studi Kasus Komisi Pemberantasan Korupsi Fakultas,” 2022.
 - [11] S. Hikmawan, A. Pardamean, dan S. N. Khasanah, “Sentimen Analisis Publik Terhadap Joko Widodo Terhadap Wabah Covid-19 Menggunakan Metode Machine Learning,” vol. 20, no. 2, hal. 167–176, 2020.
 - [12] B. M. Akbar, A. T. Akbar, dan R. Husaini, “Jurnal Teknologi Terpadu,” vol. 7, no. 2, hal. 83–92, 2021.
 - [13] J. E. Sembodo, E. B. Setiawan, dan A. Baizal, “Data Crawling Otomatis pada Twitter,” no. October 2018, hal. 10–16, 2016, doi: 10.21108/INDOSC.2016.111.
 - [14] P. Arsi dan R. Waluyo, “Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (Svm) Sentiment Analysis On The Discussion Of Relocating I Ndonesia ’ S Capital City Using The Support Vector Machine (SVM),” vol. 8, no. 1, hal. 147–156, 2021, doi: 10.25126/jtiik.202183944.
 - [15] S. K. Wardani dan Y. A. Sari, “Analisis Sentimen menggunakan Metode Naïve Bayes Classifier terhadap Review Produk Perawatan Kulit Wajah menggunakan Seleksi Fitur N-gram dan Document Frequency Thresholding,” vol. 5, no. 12, hal. 5582–5590, 2021.
 - [16] D. Darwis, N. Siskawati, dan Z. Abidin, “Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional,” vol. 15, no. 1, hal. 131–145.
 - [17] A. P. M. Kom dan T. Baidawi, “SURAT TUGAS Analisis Sentimen Penghapusan Ujian Nasional pada Twitter Menggunakan Support Vector Machine dan Naïve Bayes berbasis Particle Swarm Optimization,” no. September 2020, 2021.
 - [18] N. Luh, P. Chandra, R. A. Rahman, R. Venyutzky, dan N. A. Rakhmawati, “Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning,” vol. 7, no. April, hal. 47–58, 2021.
 - [19] B. A. B. Iii, “Universitas Multimedia Nusantara,” hal. 42–59.

- [20] I. G. P. Megayasa, P. Pasek, O. Mahawardana, dan P. R. Nurbawa, "Analisis Sentimen Berdasarkan Opini Dari Media Sosial Twitter Terhadap ' Figure Pemimpin ' Menggunakan Python," vol. 13, no. 1, 2023.
- [21] A. Ilmiah, H. Riset, A. F. Febrianti, A. H. Cabral, dan G. Anuraga, "K-Means Clustering Dengan Metode Elbow Untuk Pengelompokan Kabupaten Dan Kota Di Jawa Timur," hal. 863–870, 2018.
- [22] I. Artikel, "Kinerja Auto Labelling Pada Analisis Sentimen Terhadap Pasangan Calon Presiden 2024 Di Media Sosial X," no. Ciastech 2023, hal. 618–633, 2024.
- [23] Y. Adani, P. Studi, S. Informasi, dan A. N. Baye, "Penerapan Algoritma Naïve Bayes Untuk Memprediksi," hal. 13–24.
- [24] A. Safira, F. N. Hasan, K. J. Timur, dan N. B. Classifier, "PAYLATER MENGGUNAKAN METODE NAIVE BAYES," vol. 5, no. 1, hal. 59–70, 1978.
- [25] A. M. Hidayat, M. Syafrullah, dan I. Pendahuluan, "Algoritma Naïve Bayes Dalam Analisis Sentimen Untuk Klasifikasi Pada Layanan Internet PT . XYZ," vol. 9, no. 2, hal. 91–95, 2017.
- [26] A. M. Jon, "Analisis Sentimen Pada Media Sosial Instagram Klub Persija Jakarta Menggunakan Metode Naive Bayes," vol. 958, 2022.
- [27] R. Pinka, B. Dakhi, A. Maezar, dan B. Aji, "Sentiment Analysis Of Public Opinions On The Effectiveness Of Online Learning," vol. 6, no. 1, hal. 273–279, 2022, doi: 10.52362/jisicom.v6i1.822.