

ANALISIS SENTIMEN HAK ANGKET PEMILU 2024 DI MEDIA SOSIAL X DENGAN NAIVE BAYES MULTINOMIAL

Achmad Rizki Nur Fauzie¹, Dewi Kusumaningsih^{2*}

^{1,2}Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: ¹2011500713@student.budiluhur.ac.id, ^{2*}dewi.kusumaningsih@budiluhur.ac.id
(* : corresponding author)

Abstrak-Analisis sentimen adalah area studi yang bertujuan untuk menilai sikap publik terhadap berbagai isu seperti kejadian, masalah, dan layanan. Hak angket adalah wewenang DPR untuk menyelidiki penerapan undang-undang serta kebijakan pemerintah mengenai isu-isu krusial yang mempengaruhi kesejahteraan masyarakat dan negara. Pemilu 2024 menimbulkan ketidakpuasan di kalangan masyarakat yang merasa hasilnya tidak adil dan menuduh adanya kecurangan. Sebagai akibatnya, banyak masyarakat menolak keputusan hak angket pemilu 2024, dan beragam pandangan mengenai pemilu tersebut muncul. Hal ini memerlukan sistem untuk menganalisis sentimen dalam setiap pernyataan. Penelitian ini bertujuan untuk mengkategorikan dan menilai opini masyarakat melalui jejaring sosial X, serta menentukan tingkat akurasi tertinggi dari algoritma yang digunakan dalam sistem online. Analisis opini dilakukan dengan pelabelan manual yang dibagi menjadi dua kategori: label negatif dan label positif. *Dataset* yang dianalisis dikumpulkan dari media sosial X. Pengumpulan data menggunakan kata kunci “hak angket pemilu” pada periode 24 April 2024 - 21 Mei 2024. Hasil uji coba menggunakan algoritma *Multinomial Naive Bayes* menunjukkan *accuracy* 73,75%, *recall* 53,84%, *precision* 68,29%, dan *f1 score* 60,21% dari 702 data. Dengan 20% data digunakan untuk pengujian dan 80% untuk pelatihan, diperoleh 141 data uji. Kesimpulannya, analisis sentimen menggunakan metode *Multinomial Naive Bayes* mampu memberikan *accuracy* dan *precision* yang memadai

Kata Kunci: *Multinomial Naive Bayes*, Hak Angket, Pemilu, Media sosial X, Analisis Sentimen

SENTIMENT ANALYSIS OF 2024 ELECTION VOICE RIGHTS ON SOCIAL MEDIA X WITH NAIVE BAYES

Abstract-*Sentiment analysis is a field of study aimed at assessing public attitudes toward various issues such as events, problems, and services. The right of inquiry is the authority of the DPR (Indonesian Parliament) to investigate the implementation of laws and government policies regarding crucial issues that impact the welfare of the people and the nation. The 2024 elections sparked dissatisfaction among the public, who felt that the results were unfair and accused of fraud. As a result, many people disagreed with the decision on the 2024 election inquiry, leading to diverse opinions about the election. This necessitates a system to analyze sentiment in each statement. The purpose of this study is to categorize and evaluate public opinions through social network X and to determine the highest accuracy level of the algorithm used in the online system. The sentiment analysis was conducted through manual labeling, dividing the data into two categories: negative and positive labels. The dataset analyzed was sourced from social network X, collected using the keyword “election inquiry” during the period from April 24, 2024, to May 21, 2024. Testing results using the Multinomial Naive Bayes algorithm showed an accuracy of 73.75%, recall of 53.84%, precision of 68.29%, and an F1 score of 60.21% from 702 data points. With 20% of the data used for testing and 80% for training, 141 test data were obtained. Therefore, it can be concluded that sentiment analysis using the Multinomial Naive Bayes method achieves a satisfactory level of accuracy and precision.*

Keywords: *Multinomial Naive Bayes, Right of Inquiry, Election, Social Media X, Sentiment Analysis*

1. PENDAHULUAN

Dalam sistem demokrasi, pengawasan terhadap pemerintah oleh lembaga legislatif sangat penting untuk memastikan akuntabilitas dan transparansi. Salah satu instrumen pengawasan yang dimiliki oleh legislatif Indonesia adalah hak angket, yang memungkinkan DPR untuk mengawasi implementasi undang-undang atau aturan pemerintah yang dianggap memiliki dampak signifikan, strategis, dan besar, serta yang diduga melanggar peraturan [1]. Keputusan mengenai hak angket terkait Pemilu 2024 telah menjadi topik perdebatan, mengindikasikan adanya tantangan dari pihak eksekutif dan konflik di antara partai-partai politik di DPR [1]. Situasi ini menimbulkan keresahan di masyarakat, mengingat pemilu merupakan proses yang melibatkan partisipasi publik. Masyarakat diharapkan aktif memberikan kritik objektif, saran, dan rekomendasi kepada pemerintah, serta memantau pelaksanaan pemerintahan [2].

Pada penelitian ini, *dataset* diperoleh dari media sosial x, yang sebelumnya bernama twitter, dengan topik hak angket pemilu. Data tersebut disimpan dalam format CSV agar lebih mudah diproses dan mempercepat pengembangan. Selanjutnya, teknik text mining diperlukan untuk mengolah data yang cenderung tidak terstruktur, yaitu teks dalam bahasa alami atau *natural language*, dari data teks [3]. Oleh karena itu, penelitian ini melakukan berbagai tahapan, seperti pengumpulan data, *casefolding*, *cleansing*, *slangword*, *tokenization*, *stopword*, dan *stemming*. Penelitian sebelumnya yang menggunakan algoritma *Multinomial Naive Bayes* untuk klasifikasi menghasilkan performa yang cukup memuaskan, dengan menggunakan 2971 data, di mana 2238 digunakan sebagai data latih dan 733 sebagai data uji, mencapai akurasi sebesar 83% [4].

Penelitian ini bertujuan untuk mendapatkan pemahaman umum tentang pandangan masyarakat terkait keputusan hak angket pemilu 2024, menilai efektivitas dan akurasi metode *Multinomial Naive Bayes* dalam menganalisis kasus hak angket dan sentimen di media sosial X, serta mengimplementasikan model yang telah dikembangkan ke dalam aplikasi berbasis *web*. Selain itu, penelitian ini juga akan menilai bagaimana pendekatan ini dapat digunakan untuk memprediksi tren sentimen masyarakat di masa depan. Dengan demikian, hasil dari penelitian ini diharapkan dapat menawarkan pemahaman baru tentang pandangan publik terhadap isu-isu politik yang penting. Akhirnya, implementasi model dalam aplikasi berbasis *web* diharapkan memudahkan aksesibilitas dan penggunaan analisis sentimen oleh berbagai pemangku kepentingan.

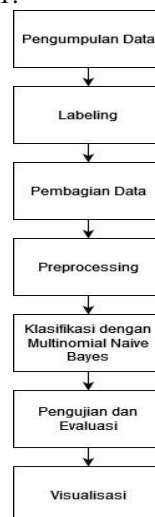
2. METODE PENELITIAN

2.1 Data Penelitian

Data dari penelitian ini didapatkan dari media sosial x, dan digunakan dalam penelitian ini. Data dikumpulkan selama periode 24 April 2024 hingga 21 Mei 2024. Total data yang berhasil diperoleh berjumlah 702. Pengumpulan data dilakukan menggunakan Python dengan teknik *web scraping*. Data yang terkumpul didasarkan pada parameter dengan kata kunci "hak angket pemilu".

2.2 Penerapan Metode

Untuk membangun aplikasi menggunakan analisis sentimen menggunakan algoritma *Multinomial Naive Bayes*, berikut sistem dapat dilihat dari Gambar 1.



Gambar 1. Penerapan Metode

Aplikasi ini dibuat dengan algoritma *Multinomial Naive Bayes* dan dilakukan beberapa tahap dalam sistem, yaitu: Untuk tahapan pengumpulan data, menggunakan teknik *web scraping*. Berikut penjelasan secara detailnya:

- a. Langkah awal dilakukan untuk mengumpulkan informasi dari Media Sosial X dengan memanfaatkan teknik *web scraping*, *web scraping merupakan metode untuk mengambil dokumen sebagian terstruktur dari internet dan dianalisis untuk mengekstrak informasi spesifik yang dapat digunakan dalam berbagai aplikasi* [5]. Setelah data berhasil dikumpulkan, data tersebut disimpan dalam format CSV.
- b. Setelah itu, dilakukan labelisasi secara manual yang dilabeli dengan ahli / pakar pada bidangnya dengan 2 jenis label, yaitu positif dan negatif.

- c. Setelah pembagian data, dilakukan alokasi sebesar 80% digunakan sebagai data pelatihan, dan 20% sebagai data pengujian.. Data yang dialokasikan ke dalam data uji tidak lagi dimasukkan ke dalam data latih untuk memastikan ketidaksamaan dalam penggunaan data uji dan latih.
- d. Setelah itu, data latih serta data uji dilakukan *preprocessing* untuk membersihkan data yang menghasilkan banyak noise atau tidak relevan
- e. Setelah *preprocessing*, klasifikasi dilakukan dengan algoritma *multinomial naive bayes* pada data uji, dengan memanfaatkan *model* yang dihasilkan dari data latih. Pendekatan ini memastikan bahwa evaluasi dan penerapan model dilakukan secara terpisah antara data uji dan data latih, sehingga integritas hasil klasifikasi tetap terjaga.
- f. Setelah proses klasifikasi dengan metode *multinomial naive bayes* selesai, hasilnya dievaluasi menggunakan confusion matrix untuk menilai kinerja model. Selain itu, hasil klasifikasi juga bisa divisualisasikan dalam bentuk grafik untuk mendapatkan pemahaman yang lebih mendalam tentang distribusi klasifikasi yang telah dilakukan. Pendekatan ini memungkinkan analisis yang menyeluruh terhadap kinerja model dalam mengklasifikasikan data.
- g. Setelah dilakukan pengujian dan evaluasi, proses selanjutnya yaitu menampilkan hasil dari pengujian dan evaluasi tersebut dalam bentuk grafik dan chart.

2.3 Pengumpulan Data

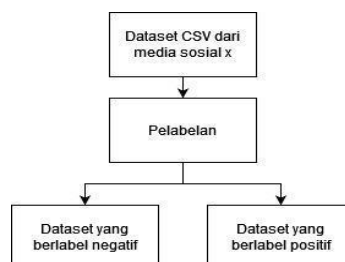
Pada langkah pengumpulan data, teknik *web scraping* digunakan. Proses ini menggunakan python untuk mengambil data dari media sosial X dengan kata kunci "hak angket pemilu". Data diambil melalui proses *web scraping* dari media sosial x dengan menggunakan *API* yang diperlukan, yaitu *api session*. Data yang terkumpul kemudian disimpan dalam format CSV. Gambar 2 menunjukkan proses pengumpulan data.



Gambar 2. Pengumpulan Data

2.4 Pelabelan

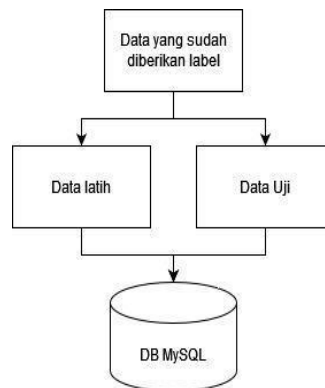
Pada tahap labeling, data dimasukkan ke dalam dua kategori: positif dan negatif. Jika data mengandung opini yang setuju atau mendukung keputusan hak angket Pemilu 2024, ahli di bidang tersebut melabelkannya sebagai data kelas positif. Jika tidak, data tersebut dikategorikan sebagai data kelas negatif. Seorang pakar politik, Pak Dosen Bambang Pujiyono, MM, M.Si, memverifikasi label secara manual. Saat ini, sebanyak 702 data telah dikumpulkan. Proses pelabelan digambarkan dalam Gambar 3.



Gambar 3. Pelabelan

2.5 Pembagian Data

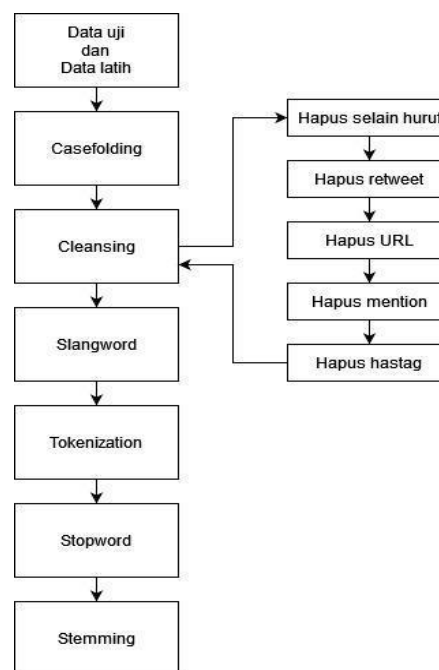
Dalam proses ini, *dataset* yang memiliki label akan dibagi menjadi dua bagian.: data uji dan data latih. Ini dicapai dengan membagi dua puluh persen *dataset* sebagai data uji dan delapan puluh persen sisanya sebagai data latihan. Proses pembagian data ini digambarkan pada Gambar 4.



Gambar 4. Pembagian Data

2.6 Preprocessing

Pada tahap *preprocessing*, data disaring dan dilakukan pembuangan kata melalui berbagai tahapan. Langkah ini sangat penting untuk mendapatkan data yang lebih akurat dan terstruktur. *Preprocessing* berfungsi untuk mempersiapkan data dengan menghilangkan *noise* yang ada, sehingga proses klasifikasi menjadi lebih mudah dan efektif [6]. Berikut Gambar 5 yang sesuai dengan tahapan *preprocessing*.



Gambar 5. Pembagian Data

Berikut penjelasan mengenai proses *preprocessing* yang dibuat oleh penulis:

- Case Folding*; Merubah semua huruf menjadi huruf kecil adalah tahap *Case Folding* [7].
- Cleansing*; Tahap *Cleansing* merupakan proses pembersihan, teks dibersihkan dari kata-kata yang tidak relevan, tanda baca seperti delimiter, dan simbol-simbol lain yang tidak diperlukan [7].
- Slangword*; Proses mengubah kata slang, jargon, dan bahasa informal menjadi bahasa asli formal disebut sebagai tahap *slangword*.
- Tokenization*; Tahap *tokenization* merupakan proses membagi kalimat menjadi sejumlah kata yang terpisah [7].
- Stopword*; Tahap *stopword* merupakan proses menyaring kata-kata atau istilah yang tidak penting dari isi dokumen teks [8].
- Stemming*; Tahap *stemming* merupakan proses menghapus kata imbuhan pada awal atau akhir kata sehingga kata tersebut menjadi kata dasar [7].

2.7 Klasifikasi dengan Multinomial Naïve Bayes

Dalam tahap ini, menggunakan metode *multinomial naïve bayes*, metode ini digunakan untuk memprediksi apakah data uji termasuk dalam kategori tertentu. Berikut rumus (1) algoritma *multinomial naïve bayes*:

$$P(c_j) = \frac{\text{count}(c_j)}{N} \quad (1)$$

Keterangan:

count (c_j) : Banyaknya kelas c_j pada dokumen latih

N : Jumlah keseluruhan kelas pada dokumen latih

Dalam perhitungan berikutnya, kata i termasuk kategori atau kelas tertentu, Sehingga harus menentukan kemungkinan bahwa. Ini dapat dilakukan dengan menggunakan metode (2) berikut::

$$P(w_o, c_k) = \frac{\text{count}(w_o, c_k) + 1}{\sum \text{count}(w_o, c_k) + |V|} \quad (2)$$

Keterangan:

count(w_o, c_k) : Jumlah kata w_o yang muncul di kelas c_k .

$\sum \text{count}(w_o, c_k)$: Total jumlah kata yang muncul dalam kelas c_k .

|V| : Total jumlah kata unik di seluruh kelas

Perhitungan selanjutnya untuk mencari probabilitas, menggunakan rumus (3) seperti berikut:

$$P(c_k | w_o) = P(c_k) \times P(w_o | c_k) \times \dots \times P(w_n | c_n) \quad (3)$$

Keterangan:

$P(c_k | w_o)$: Peluang kategori k jika kata o muncul.

$P(c_k)$: Probabilitas prior untuk kategori k .

$P(w_o | c_k)$: Tingkat kemungkinan kemunculan kata o dalam kategori j .

2.8 Pengujian dan Evaluasi

Pada tahap ini, Untuk menghitung metrik pengujian, evaluasi dilakukan dengan menggunakan *matrix confusion*, seperti *accuracy*, *precision*, dan *f1 score*. Setelah itu, hasil prediksi ditampilkan dalam bentuk *donut chart*. Metode yang digunakan untuk mengevaluasi kinerja klasifikasi adalah dengan *confussion matrix*, yang diterapkan untuk membandingkan hasil dari sistem dan dari hasil yang seharusnya [9]. Proses *confussion matrix* dapat ditunjukkan pada Tabel 1.

Tabel 1. Pembagian Data

	Actual Positive	Actual Negative
Predicted Positive	TP (<i>True Positive</i>)	FP (<i>False Positive</i>)
Predicted Negative	FN (<i>False Negative</i>)	TN (<i>True Negative</i>)

Keterangan:

TP (*True Positive*) = Total semua dokumen yang diprediksi positif dan juga yang aktual positif.

TN (*True Negative*) = Total semua dokumen yang diprediksi positif dan juga yang aktual positif.

FP (*False Positive*) = Total semua dokumen yang diprediksi negative, dan juga yang aktual positif.

FN (*False Negative*) = Total semua dokumen yang diprediksi positif, dan juga yang aktual negative.

Untuk mengukur *f1 score*, *precision*, *recall*, serta *accuracy* sebagai berikut:

a. *Accuracy*

Accuracy mengukur proporsi prediksi yang benar dibandingkan dengan jumlah total prediksi. Untuk mencari nilai *accuracy*, dapat digunakan sebagai rumus (4) berikut:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

b. *Precision*

Precision adalah ukuran yang menggambarkan proporsi dokumen relevan dari seluruh dokumen yang diidentifikasi oleh sistem, yaitu proporsi jumlah label positif yang diklasifikasikan dengan tepat dari jumlah keseluruhan prediksi sampel positif [10]. Untuk menghitung nilai presisi, rumus (5) yang digunakan.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

c. *Recall*

Recall adalah Ukuran yang menghitung persentase dokumen relevan yang berhasil diidentifikasi oleh sistem, yaitu Total seluruh label positif yang diklasifikasikan dengan tepat dibandingkan dengan total keseluruhan seluruh label positif yang tersedia [10]. Untuk menghitung nilai recall, rumus (6) digunakan.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

d. *F1 Score*

F1 Score adalah hasil rata - rata kombinasi dari *recall* dan *precision* [11]. Untuk mencari hasil nilai dari *F1 Score* menggunakan rumus (7) sebagai berikut:

$$F1\ Score = \frac{2 \times recall \times precision}{recall + precision} \quad (7)$$

3. HASIL DAN PEMBAHASAN

Berdasarkan proses segmentasi terhadap *dataset*, terdapat 702 data yang telah diproses sebelumnya dan diberi label secara manual. Data tersebut dibagi dengan proporsi 20% untuk pengujian dan 80% untuk pelatihan. Hasil pembagiannya adalah 561 data untuk pelatihan dan 141 data untuk pengujian.

3.1 Pengujian Metode

Data yang di latih dan di uji dalam penelitian ini didapat dari media sosial x. Proses pengumpulan data dikembangkan menggunakan Python dengan metode *scraping*. Data dikumpulkan berdasarkan parameter dan kata kunci “hak angket pemilu”. Data tersebut di filter dari tanggal 24 April 2024 – 21 Mei 2024. Data yang didapat bisa dilihat pada Tabel 2.

Tabel 2. Data Pengujian Metode

No	Teks	Label Actual	Label Prediction
1	@puanmaharani_ri Hak angket gimana nih buuu??? Sanksi dari pemilih itu berat dan kejam loh dan bisa berimbas ke partai anda jangan bikin saya DAN mereka kecewa!	Positif	Positif
2	Hak angket udah basi nggak ada yg dukung selain gerombolan sontoloyo	Negatif	Negatif
...
141	@PngAdilnR4kyt Seberapa kerasnya suara Tim GAMA untuk hak angket jika mbak Mega diam dan Puan pura-pura tidak tahu percuma saja	Negatif	Negatif

Berdasarkan hasil pengujian *confussion matrix* diatas dapat dilihat pada tabel 3:

Tabel 3. Hasil pengujian Confussion Matrix

	Actual Positive	Actual Negative
Predicted Positive	28	13
Predicted Negative	24	76

Berdasarkan tabel *confussion matrix* dapat diperoleh nilai *recall*, *accuracy*, *presisi* dan *f1 score* dapat dilihat pada Tabel 4:

Tabel 4. Hasil pengujian *Multinomial Naive Bayes*

Hasil Pengujian Multinomial Naive Bayes		
<i>Accuracy</i>	$\frac{28+76}{28+13+24+76}$	73%
<i>Preccision</i>	$\frac{28}{28+13}$	68%
<i>Recall</i>	$\frac{28+24}{28}$	53%
<i>F1 Score</i>	$\frac{2 \times recall \times precision}{recall+precision}$	60%

3.2 Dashboard

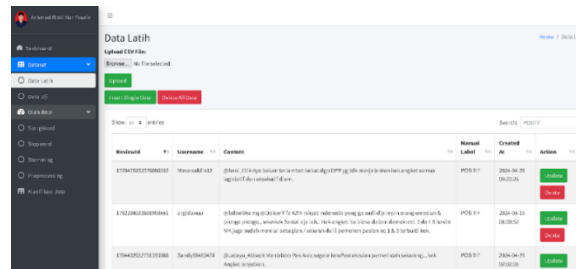
Berikut merupakan tampilan layar *dashboard*. Pada menu ini, terdapat judul skripsi penulis, hingga jumlah data latih dan data uji. Terdapat cara penggunaan aplikasi tersebut. Berikut tampilan layar *dashboard* pada Gambar 6.



Gambar 6. Dashboard

3.3 Dataset Latih

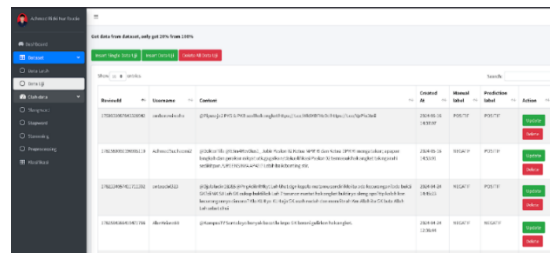
Berikut merupakan tampilan layar data latih. Pada menu ini, pengguna bisa melihat semua data yang telah diinput, serta bisa melakukan *input*, *update*, dan *delete data*. Pengguna juga bisa *upload* file .csv untuk melakukan *batch insert* sesuai dari data yang telah di *upload*. Berikut tampilan layar *dataset* latih pada Gambar 7.



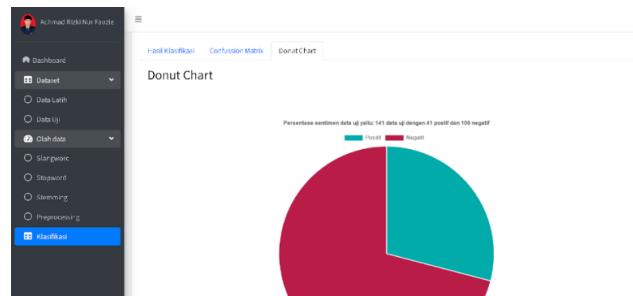
Gambar 7. Dataset Latih

3.4 Dataset Uji

Berikut merupakan tampilan layar data uji. Pada menu ini, pengguna bisa melakukan pembagian data, serta bisa melihat data uji yang sudah diinput. Pengguna juga bisa melakukan *insert*, *update*, dan *delete* data uji pada halaman data uji ini. Berikut tampilan layar *dataset* latih pada Gambar 8.



Gambar 8. Dataset Uji



Gambar 12. Klasifikasi Data

4. KESIMPULAN

Berdasarkan hasil aplikasi yang telah dikembangkan, dengan menggunakan algoritma dan *dataset* yang tersedia, dapat disimpulkan bahwa Program ini menggunakan 702 kumpulan data, 80% data pelatihan dan 20% data pengujian. Data eksperimen yang digunakan berjumlah 141 data. Hasil klasifikasi polinomial Naive Bayes sebesar 73,75%, recall 53,84%, presisi 68,29% dan skor f1 60,21%. Alhasil, mayoritas masyarakat menolak keputusan pertanyakan kekuasaan hingga pemilu 2024. Oleh karena itu, analisis sentimen dengan metode polinomial sederhana Bayes dapat dikatakan dapat menghasilkan akurasi dan prediktabilitas yang tinggi. Dalam penelitian ini, terdapat beberapa tahap utama, seperti pengambilan data dari web (*web scraping*), proses pra-pemrosesan (*preprocessing*), pemberian label (*labelling*), dan pengklasifikasian dengan algoritma Multinomial Naive Bayes serta dilakukan pengujian. Proses *preprocessing* yang efektif sangat penting karena berpengaruh langsung terhadap hasil optimal pada tahap berikutnya

DAFTAR PUSTAKA

- [1] S. Rabiha, H. N. Nathanael, and N. P. Fauzyyah, "Peran Hak Angket DPR dalam Upaya Penyelesaian Sengketa Hasil Pemilu", *Jurnal BATAVIA*, vol. 1, no. 02, pp. 89-95, 2024. Available: <https://journal.zhatainstitut.org/index.php/batavia/article/view/33>.
- [2] M. Shofyan, "Keterlibatan Publik dalam Pengawasan Pemilu 2024," *'Asabiyah: Jurnal Pengabdian Hukum*, vol. 2, no. 1, pp. 43-47, 2024. Available: <https://ojs.um-palembang.ac.id/index.php/asabiyah/article/view/151>
- [3] D. Reddy, D. Arifianto, and D. Lusiana, "Analisis Sentimen Pada Pelayanan Jaringan Internet Indihome Dengan Metode Multinomial Naïve Bayes Masa Pandemi Covid-19," *Jurnal Smart Teknologi**, vol. 3, no. 6, pp. 612-623, 2022. Available: <http://jurnal.unmuhjember.ac.id/index.php/JST/article/view/8485>
- [4] D. R. S. Pratama, T. A. Munandar, and K. F. Ramdhania, "Multinomial Naive Bayes Algorithm for Indonesian Language Sentiment Classification Related to Jakarta International Stadium (JIS)," *International Journal of Information Technology and Computer Science Applications*, vol. 2, no. 1, pp. 12-22, Jan. 2024. doi: 10.58776/ijitcsa.v2i1.118.
- [5] U. Mufidah, "Perancangan Aplikasi Perbandingan Harga Produk (Historical Data) Menggunakan Teknik *Scraping Web*", *Jurnal Teknologi dan Sistem Informasi*, vol. 1, no. 9, pp. 1-10, Aug. 2021.
- [6] O. H. Rahman, G. Abdillah, and A. Komarudin, "Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine," *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, vol. 5, no. 1, pp. 17-23, 2021. Available: <https://doi.org/10.29207/resti.v5i1.2700>.
- [7] H. Hidayat, F. Santoso, and L. F. Lidimillah, "Analisis Sentimen Pengguna YouTube Tentang Rohingya Menggunakan Algoritma SVM (Support Vector Machine)," *G-Tech: Jurnal Teknologi Terapan*, vol. 8, no. 3, pp. 1729-1738, 2024. doi: 10.33379/gtech.v8i3.4497.
- [8] E. Yuniar, D. S. Utsalinah, and D. Wahyuningsih, "Implementasi *Scraping Data* Untuk Sentiment Analysis Pengguna Dompet Digital dengan Menggunakan Algoritma Machine Learning," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 2, no. 1, pp. 35-42, Apr. 2022, doi: 10.25008/janitra.v2i1.14.
- [9] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier dan Confusion Matrix pada Analisis Sentimen Berbasis Teks pada Twitter," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 5, no. 2, pp. 697-711, 2021.
- [10] F. Farazilla, Y. D. Siregar, and M. Faturrahman, "Efektivitas Sistem Temu Kembali Informasi Perpustakaan Universitas Islam Negeri Sumatera Utara dalam Tinjauan Recall dan Precision," *Triwikrama: Jurnal Ilmu Sosial*, vol. 3, no. 6, pp. 122-132, 2024. doi: 10.6578/triwikrama.v3i6.2932.
- [11] M. I. Fikri, E. Budianita, I. Iskandar, and E. P. Cynthia, "Klasifikasi Tingkat Kecanduan Internet terhadap Remaja Pekanbaru melalui Pendekatan Algoritma Naïve Bayes," *ZONAsi: Jurnal Sistem Informasi*, vol. 6, no. 2, pp. 424-436, 2024. doi: 10.31849/zn.v6i2.20191.