

Analisis Sentimen Terhadap Kinerja Kpu Menjelang Pilkada Serentak 2024 Berdasarkan Opini X Menggunakan Naive Bayes

Fanny Susanto Putra Tama¹, Iedam Fardian Anshori²

^{1,2}Program Teknik Informatika, Universitas Adhirajasa Reswara Sanjaya
e-mail: ¹fannycindy922@gmail.com, ²iedam@ars.ac.id

Abstrak

Abstrak Penelitian ini menganalisis sentimen publik terhadap kinerja Komisi Pemilihan Umum (KPU) menjelang Pilkada Serentak 2024, menggunakan data tweet dari media sosial X. Dalam menghadapi pesta demokrasi terbesar di Indonesia, KPU diharapkan menjalankan pemilu secara langsung, umum, bebas, rahasia, jujur, dan adil. Data tweet dianalisis menggunakan metode Naïve Bayes Classifier, yang efektif dalam klasifikasi teks, serta metode TF-IDF (*Term Frequency-Inverse Document Frequency*) untuk membobotkan kata-kata penting dalam data. Hasil penelitian menunjukkan model Naïve Bayes Classifier mampu mengklasifikasikan opini publik dengan akurasi yang baik, memisahkan sentimen positif, negatif, dan netral terkait kinerja KPU dengan akurasi 92.54%. Temuan ini memberikan masukan berharga bagi KPU untuk memahami pandangan masyarakat dan meningkatkan kualitas pelaksanaan pemilu di masa mendatang.

Kata kunci—Sentimen Publik, KPU, Pilkada Serentak 2024, Naïve Bayes, TF-IDF, Media Sosial X.

Abstract

This study analyzes public sentiment toward the performance of the Indonesian General Elections Commission (KPU) ahead of the 2024 Regional Head Elections (Pilkada), utilizing tweet data from the social media platform X. As Indonesia's largest democratic event approaches, the KPU is expected to conduct elections that are direct, general, free, confidential, honest, and fair. Tweets were analyzed using the Naïve Bayes Classifier, an effective method for text classification, alongside Term Frequency-Inverse Document Frequency (TF-IDF) to weigh important words within the dataset. The findings indicate that the Naïve Bayes Classifier effectively distinguishes positive, negative, and neutral public sentiment regarding the KPU's performance, achieving high accuracy 92.54%. This insight is valuable for the KPU to understand public opinion and enhance the quality of election implementation in the future.

Keywords—Public Sentiment, KPU, 2024 Regional Head Elections, Naïve Bayes, TF-IDF, Social Media X.

Corresponding Author:

Iedam Fardian Anshori,

Email: iedam@ars.ac.id

1. PENDAHULUAN

Tahun 2024 merupakan pesta demokrasi terbesar bagi rakyat republik Indonesia karena pada tahun ini Komisi Pemilihan Umum (KPU) sebagai lembaga penyelenggara pemilihan umum menyelenggarakan dua kali pesta demokrasi yang pertama [1]. Pemilihan Umum

Presiden Indonesia 2024 pada 14 Februari 2024 dan pada 27 November 2024 KPU akan menyelenggarakan Pemilihan Kepala Daerah (Pilkada). Sebagai lembaga yang bertanggung jawab atas pemilihan umum di Indonesia KPU harus terus memperbaiki kinerjanya agar Pemilu berjalan sesuai dengan asas langsung, umum, bebas, rahasia, jujur, dan adil [2].

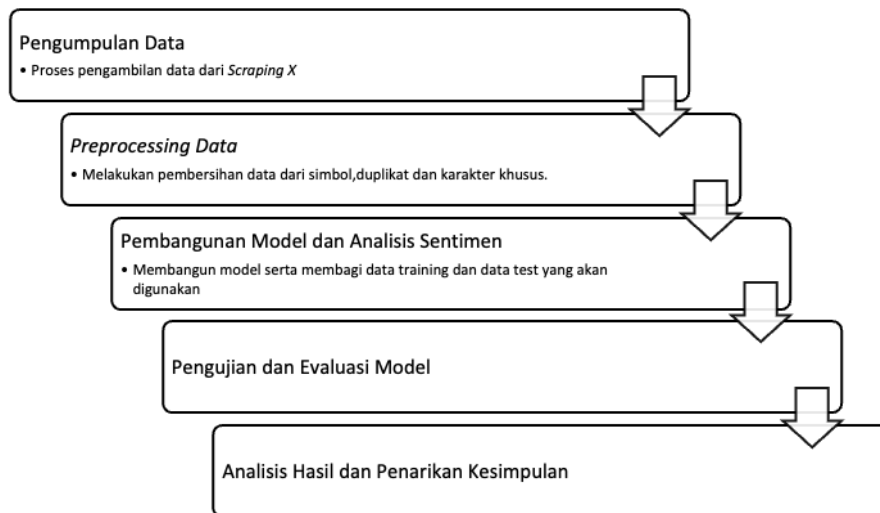
Pada era digital, data merupakan salah satu sumber daya baru yang dapat dimanfaatkan untuk tujuan tertentu. Salah satu contoh dari data tersebut adalah data Tweet yang ada pada Media Sosial X (sebelumnya Twitter). Tweet tersebut antara lain opini pengguna media sosial X terhadap suatu persoalan yang disampaikan melalui media sosial X yang digunakan sebagai data analisis sentimen terhadap persoalan publik dengan contoh Pilkada Serentak 2024 [3]. Analisis sentimen adalah teknik yang digunakan untuk memahami opini dan emosi masyarakat menggunakan data teks guna mengetahui opini seseorang terhadap suatu topik atau objek positif, negatif atau netral [4].

Naïve Bayes Classifier merupakan metode klasifikasi yang berakar pada teorema Bayes yang dapat digunakan untuk melakukan analisa sentimen dengan metode pengklasifikasian menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, dengan memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari Naïve Bayes Classifier ini adalah asumsi yang sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian. Penelitian yang dilakukan oleh Muhammad Khaerul Khatami [5], analisis terhadap KPU pada Pemilu Presiden 2024 menunjukkan penggunaan algoritma Naïve Bayes Classifier memberikan hasil akurasi tertinggi dengan menerapkan skenario data uji sebesar 10% dan data latih sebesar 90% [6]. Pada penelitian ini peneliti menggunakan metode Naïve Bayes Classifier dan data tweet dari media sosial X karena platform media sosial X dinilai mempresentasikan opini publik secara luas serta metode Naïve Bayes Classifier memiliki tingkat *accuracy* yang baik dalam melakukan analisis sentimen.

2. METODE PENELITIAN

Pendekatan yang digunakan dalam penelitian ini adalah pendekatan empiris dengan analisis statistik dan studi literatur dengan melibatkan proses pengumpulan data opini dari media sosial X dan pengolahan data untuk mendapatkan hasil sentimen masyarakat terhadap kinerja KPU. Metode Naïve Bayes dipilih karena keakuratannya dalam klasifikasi teks dan analisis sentimen, yang bertujuan untuk mengkategorikan sentimen menjadi positif, negatif, atau netral berdasarkan data opini yang tersedia [10].

Dalam Penelitian ini dilakukan dengan beberapa tahapan yang saling terkait untuk memastikan hasil yang akurat dalam analisis sentimen terhadap kinerja KPU menjelang Pilkada Serentak 2024. Adapun tahapan penelitian yang dilakukan disajikan pada gambar II.1 sebagai berikut :



Gambar 1. Tahapan Penelitian

2. HASIL DAN PEMBAHASAN

3.1 Dataset Penelitian

Data diperoleh dari media sosial X informasi yang dapat dimanfaatkan dari Tweet tersebut antara lain opini pengguna media sosial X terhadap suatu persoalan yang disampaikan melalui media sosial X oleh masyarakat [3] dengan menggunakan metode scraping bantuan library Node JS bernama Tweet-Harvest. Proses mengolah atau mengambil data dari web atau media sosial disebut *Scraping* [7]. Data dikumpulkan dengan kata kunci "kpu" dan disaring berdasarkan rentang waktu dari 26 Januari 2024 hingga 20 Oktober 2024 dengan jumlah data sebanyak 3010 data.

3.2 Preprocessing Data

Tahap preprocessing data bertujuan untuk membersihkan dan mengoptimalkan dataset agar siap digunakan dalam analisis sentimen dengan langkah preprocessing meliputi Pembersihan teks (*text cleansing*), Normalisasi text, *Tokenization* dengan proses pemisahan atau pembagian paragraf menjadi kalimat dan kalimat menjadi kata-kata [8]. *Tokenization* berfungsi untuk memotong dokumen menjadi potongan-potongan yang disebut *token* [9], Penghapusan *stop words* dan *Stemming*.

3.3 Labeling

Labeling adalah proses pemberian label atau kategori pada data untuk menentukan jenis sentimen yang terkandung dalam teks. Pada penelitian ini, labeling digunakan untuk mengklasifikasikan sentimen dari data tweet yang membahas kinerja Komisi Pemilihan Umum (KPU) menjelang Pilkada Serentak 2024. Penelitian ini memanfaatkan metode labeling otomatis dengan bantuan pustaka python TextBlob untuk sebagian besar data. Proses ini kemudian dilanjutkan dengan verifikasi manual sebanyak 10% data untuk memastikan akurasi labeling pada sampel data yang menghasilkan 76 data positif, 21 data negatif dan 1204 data netral.

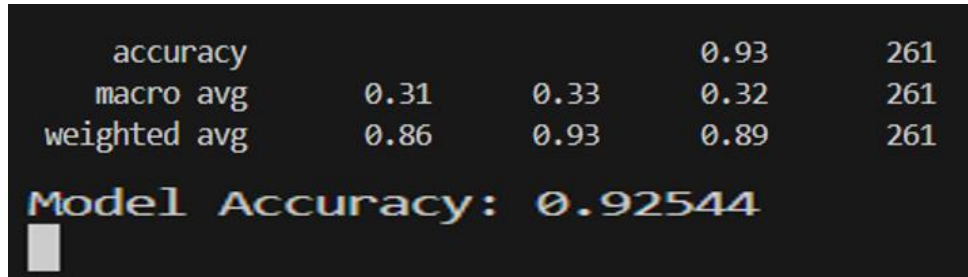
3.4 Ekstraksi Fitur/Pembobotan Kata dengan TF-IDF

Pada analisis sentimen, data teks mentah perlu diubah ke dalam bentuk numerik agar dapat diolah oleh model pembelajaran mesin. Dengan metode TF-IDF (*Term Frequency-Inverse Document Frequency*).

3.5 Pengembangan dan Pengujian Model Klasifikasi Sentimen

Model klasifikasi Naïve Bayes diterapkan untuk menganalisis sentimen terhadap kinerja KPU menjelang Pilkada Serentak 2024 berdasarkan opini yang diambil dari platform X.

Algoritma Naïve Bayes dipilih karena efisiensinya dalam memproses data teks dan keakuratannya yang cukup baik pada masalah klasifikasi berbasis teks, termasuk analisis sentimen. Dalam penelitian ini, pembagian data dilakukan dengan menggunakan metode train-test split dari library Scikit-learn, dengan komposisi 80% data latih dan 20% data uji.

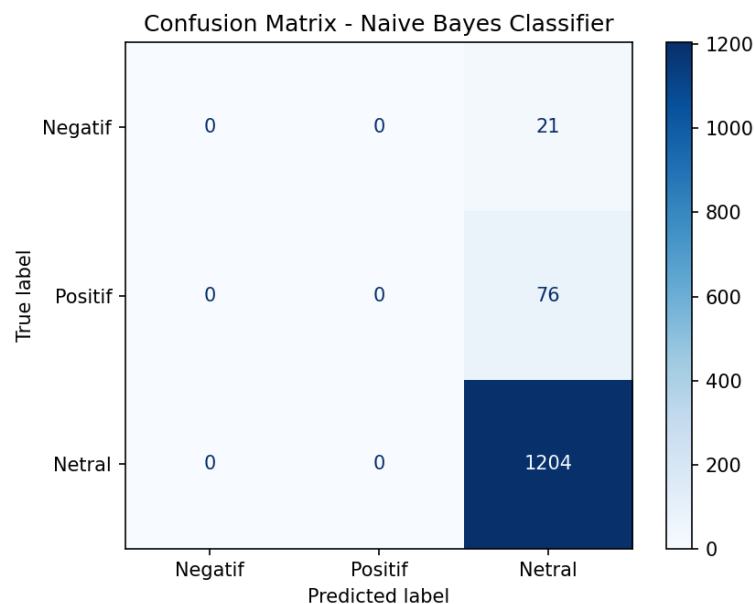


Gambar 2. Hasil Penerapan Model Klasifikasi Naïve Bayes

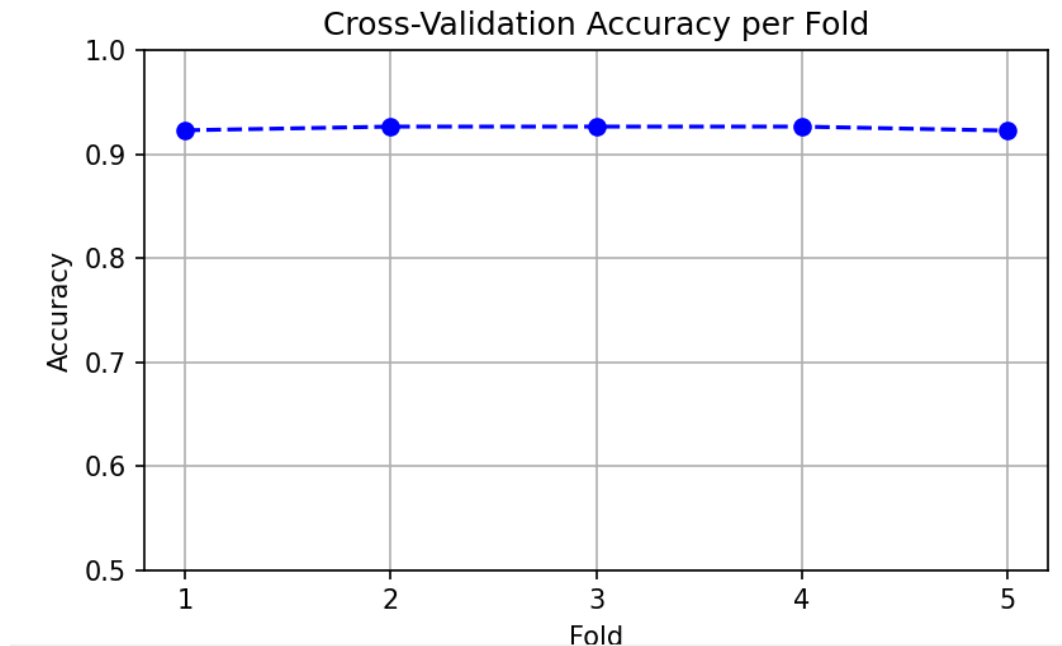
3.6 Evaluation Confusion Matrix

Evaluasi bertujuan untuk mengecek kebenaran hasil klasifikasi yang dilakukan oleh model Naïve Bayes terhadap data sentimen masyarakat mengenai kinerja KPU menjelang Pilkada Serentak 2024. Pada tahap ini, evaluasi dilakukan menggunakan metode cross-validation, yaitu dengan cara membagi data secara acak menjadi beberapa subset, di mana setiap subset secara bergantian digunakan sebagai data uji dan sisanya sebagai data latih.

Hasil penelitian menunjukkan bahwa model Naïve Bayes memiliki akurasi sebesar 92,54% dalam mengklasifikasikan sentimen terhadap kinerja KPU menjelang Pilkada Serentak 2024. Untuk mengevaluasi hasil klasifikasi, digunakan metrik Confusion Matrix. Confusion Matrix adalah tabel yang menunjukkan jumlah prediksi benar dan salah yang dibuat oleh model untuk setiap kelas sentimen (positif, negatif, dan netral). Matriks ini memberikan gambaran yang jelas tentang kesalahan prediksi model, seperti berapa banyak data positif yang diklasifikasikan salah sebagai negatif atau netral, dan sebaliknya.

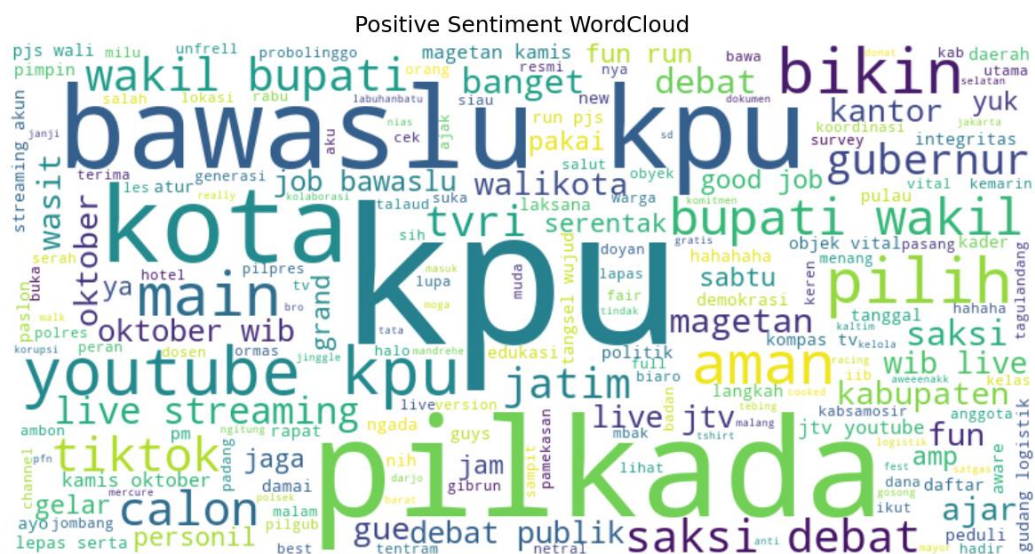


Gambar 3. Hasil Hasil Confusion Matrix



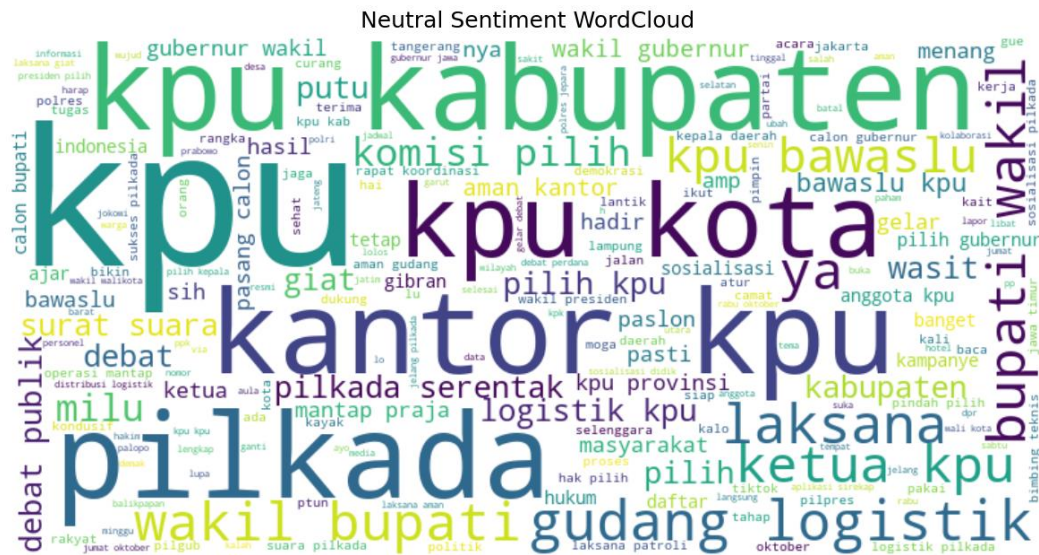
Gambar 4. Hasil *Cross validation Accuracy*

Grafik *Cross validation Accuracy* tersebut memberikan gambaran hasil akurasi model pada proses validasi silang dengan menggunakan 5 lipatan (fold). Dari grafik menunjukkan Akurasi Tinggi: Model menunjukkan performa yang sangat baik dengan akurasi yang konsisten di setiap lipatan, yaitu sekitar 0,9 (90%), Konsistensi Performa: Nilai akurasi pada kelima lipatan hampir tidak mengalami fluktuasi, menunjukkan bahwa model memiliki stabilitas yang baik ketika diuji pada berbagai subset data dan Indikasi Generalisasi yang Baik: Stabilitas akurasi ini menunjukkan bahwa model mampu menggeneralisasi dengan baik pada data baru, tanpa adanya indikasi overfitting (akurasi terlalu tinggi pada data tertentu) atau underfitting (akurasi rendah secara keseluruhan)



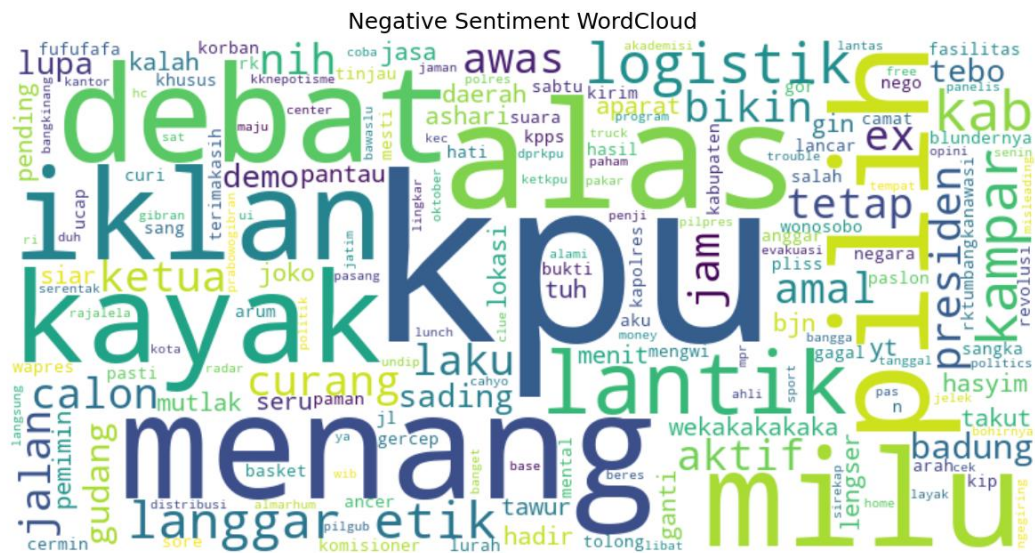
Gambar 5. *Word Cloud* Setimen Positif

Word Cloud Sentimen Positif ini mencerminkan sentimen positif yang terkait dengan topik pemilu atau pilkada, dengan fokus pada peran KPU, platform digital, dan aspek-aspek penting seperti keamanan dan kelancaran proses pemilu. Dominasi kata-kata seperti "KPU" dan "pilkada" menunjukkan bahwa mereka adalah tema sentral dalam data yang dianalisis.



Gambar 6. *Word Cloud* Setimen Netral

Word Cloud Sentimen Netral menggambarkan kata netral yang terkait dengan topik pemilu atau pilkada kata-kata seperti "KPU" dan "Pilkada" menunjukkan bahwa kata yang tergabung dalam kelompok sentimen netral.



Gambar 7 *Word Cloud* Setimen Negatif

Word Cloud Sentimen Negatif menggambarkan kata negatif yang terkait dengan topik pemilu atau pilkada kata-kata seperti "Curang" dan "Langgar Etik" menunjukkan bahwa kata yang tergabung dalam kelompok sentimen negatif.

4. KESIMPULAN

Penelitian ini menganalisis sentimen terhadap kinerja Komisi Pemilihan Umum (KPU) menjelang Pilkada Serentak 2024 menggunakan model Naïve Bayes. Beberapa kesimpulan utama yang dapat diambil dari hasil penelitian ini adalah: 1). Persepsi Masyarakat di Media Sosial Analisis menunjukkan bahwa sentimen netral mendominasi dataset, yang mencerminkan opini publik cenderung netral terhadap kinerja KPU menjelang Pilkada 2024. 2) Klasifikasi Sentimen Model Naïve Bayes berhasil mengklasifikasikan sentimen dari 1301 data tweet menjadi 76 data positif, 21 data negatif, dan 1204 data netral. 3) Tingkat Akurasi Model menunjukkan kinerja yang andal dengan rata-rata akurasi sebesar 92,54% berdasarkan evaluasi Cross-Validation. Hal ini menunjukkan efektivitas model dalam menganalisis sentimen pada dataset yang diberikan. Tren Sentimen terhadap KPU Pola sentimen masyarakat teridentifikasi cenderung netral menjelang pelaksanaan Pilkada Serentak 2024. Dominasi sentimen netral ini menunjukkan bahwa opini publik relatif stabil terhadap KPU. Kesimpulan ini mengindikasikan bahwa model Naïve Bayes adalah alat yang efektif untuk mengkaji opini publik berdasarkan data media sosial.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada kedua orang tua dan istri atas dukungan yang diberikan dalam penelitian ini. Kontribusi tersebut memungkinkan peneliti untuk mengumpulkan data dan menganalisis temuan yang telah disajikan dalam penelitian ini. Penelitian ini tidak akan berhasil tanpa dukungan dari kedua orang tua dan istri. Penulis juga ingin menyampaikan terima kasih kepada Iedam Fardian Anshori, S.T, M.Kom, M.M. atas saran dan panduan akademik yang berharga dalam penyusunan penelitian ini.

DAFTAR PUSTAKA

- [1] S. Zaenab, "Strategi komunikasi KPU Bangkalan dalam meningkatkan partisipasi masyarakat pada pelaksanaan pilkada serentak." 2019.
- [2] Matera, "Peran dan Tanggung Jawab KPU dalam Penyelenggaraan Pemilihan Umum di Indonesia." 2022.
- [3] N. Legiawati and Y. Ramadhan, "Analisis sentimen opini pengguna Twitter terhadap perusahaan jasa ekspedisi menggunakan algoritma Naïve Bayes berbasis PSO," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, p. 930, 2022.
- [4] A. D. Dayani, Yuhandri, and G. W. Nurcahyo, "Analisis sentimen terhadap opini publik pada sosial media Twitter menggunakan metode Support Vector Machine," *J. KomtekInfo*, vol. 11, no. 1, pp. 1–10, 2024.
- [5] M. K. Khatami, "Analisis sentimen Twitter menggunakan Naive Bayes dan Support Vector Machine terhadap KPU pada pemilihan umum presiden 2024." 2024.
- [6] A. M. Taufiqi and A. Nugroho, "Sentimen pengguna Twitter mengenai isu kebocoran data dengan algoritma Naïve Bayes," *J. Nas. Ilmu Komput.*, vol. 4, no. 1, pp. 1–11, 2023.
- [7] E. Yunia, "Web Scraping: Teknik dan Implementasinya dalam Analisis Data Media Sosial." 2022.
- [8] A. F. Anees, A. Shaikh, A. Shaikh, and S. Shaikh, "Survey paper on sentiment analysis: Techniques and challenges," no. 2389. 2020.
- [9] C. D. Manning, P. Raghavan, and H. Schütze, *An introduction to information retrieval*. Cambridge University Press, 2009.
- [10] A. Safira and F. N. Hasan, "Analisis sentimen masyarakat terhadap PayLater menggunakan metode Naive Bayes Classifier," *ZONasi: Jurnal Sistem Informasi*, vol. 5, no. 1, pp. 59, 2023.