



KLASIFIKASI TEKS *QUICK COUNT* PEMILIHAN PRESIDEN 2024 PADA TWITTER MENGGUNAKAN METODE *TF-IDF* DAN *NAIVE BAYES*

Aditya Pranata¹, Rudiman², Naufal Azmi Verdikha³

^{1,2,3} Teknik Informatika, Universitas Muhammadiyah Kalimantan Timur
Samarinda, Kalimantan Timur, Indonesia 75124
2011102441052@umkt.ac.id, rud959@umkt.ac.id, nav651@umkt.ac.id

Abstract

The 2024 Indonesian Presidential Election generated various responses on X Twitter platform related to the Quick Count. The large number of diverse opinions makes identifying and categorizing sentiments difficult. This study aims to evaluate the accuracy of the Naive Bayes method with TF-IDF weighting in text classification regarding the Quick Count of the 2024 Presidential Election on X Twitter. Data was obtained through crawling, resulting in 2113 tweets, which experts in data labelling then labelled. The preprocessing stage includes case folding, cleansing, stopword removal, and stemming. Words are weighted using TF-IDF, and then the data is divided into 80% for training and 20% for testing. Text classification using the Naive Bayes algorithm achieved an accuracy of 74.46%, indicating a pretty good accuracy in classifying text related to the 2024 Presidential Election Quick Count on X Twitter.

Keywords: Naive Bayes, Quick Count, Text Classification, TF-IDF Weighting, X Twitter

Abstrak

Pemilihan Presiden Indonesia 2024 menghasilkan berbagai tanggapan di platform X Twitter terkait *Quick Count*. Banyaknya pendapat yang beragam menyulitkan proses identifikasi dan kategorisasi sentimen secara tepat. Penelitian ini bertujuan mengevaluasi keakuratan metode *Naive Bayes* dengan pembobotan *TF-IDF* dalam klasifikasi teks mengenai *Quick Count* Pemilihan Presiden 2024 pada X Twitter. Data diperoleh melalui *crawling*, menghasilkan 2113 *tweet* yang kemudian diberi label oleh ahli dalam pelabelan data. Tahap *preprocessing* mencakup *case folding*, *cleansing*, penghapusan *stopword*, dan *stemming*. Kata-kata diberi bobot menggunakan *TF-IDF*, lalu data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Klasifikasi teks menggunakan algoritma *Naive Bayes* mencapai akurasi 74,46%, menunjukkan akurasi yang cukup baik dalam mengklasifikasikan teks terkait *Quick Count* Pemilihan Presiden 2024 pada X Twitter.

Kata kunci: Klasifikasi Teks, *Naive Bayes*, Pembobotan *TF-IDF*, *Quick Count*, X Twitter

1. PENDAHULUAN

Pemilihan umum (Pemilu) di Indonesia dikelola oleh Komisi Pemilihan Umum (KPU), lembaga independen yang bertugas mengatur dan mengawasi proses pemilu. KPU bersama Badan Pengawas Pemilihan Umum (Bawaslu) berperan penting dalam menjamin Pemilu yang langsung, umum, bebas, rahasia, jujur, dan adil (LUBER), serta bebas dari intervensi eksternal demi menjaga demokrasi [1]. *Quick Count* (Hitung Cepat) digunakan untuk memberikan estimasi awal hasil pemilihan presiden berdasarkan sampel suara dari sejumlah TPS. Metode ini tidak menghitung seluruh suara, melainkan sebagian kecil sampel yang diperkirakan mencerminkan hasil akhir nasional dan sering menjadi topik pembicaraan masyarakat dan media [2]. Meskipun bukan hasil resmi, *Quick Count* sering dianggap cukup akurat dan penting bagi kepercayaan publik. Namun,

pengumuman hasil yang tidak sesuai harapan berpotensi memicu ketegangan politik, protes, atau kerusuhan [3].

Media sosial telah memberikan dampak besar dalam kehidupan masyarakat, memudahkan komunikasi jarak jauh dan sosialisasi virtual. Namun, konsekuensinya termasuk penyebaran bahasa sarkasme atau ungkapan kasar oleh *netizen* yang menggunakannya sebagai sarana ekspresi diri, terkadang melanggar norma kesantunan berbahasa [4]. Dalam konteks pemilihan presiden, media sosial berperan penting dalam kampanye politik, memungkinkan kandidat menyampaikan visi dan misi mereka [5]. Perkembangan platform digital telah menciptakan wadah bagi opini masyarakat, baik pro maupun kontra, yang tidak hanya muncul di media konvensional tetapi juga di media sosial seperti X Twitter [6]. Platform X Twitter telah menjadi

saluran penting bagi masyarakat untuk mengekspresikan pandangan dan opini terkait isu politik, termasuk pemilihan presiden [7].

Penelitian ini menekankan pentingnya klasifikasi teks terkait *Quick Count* Pemilihan Presiden Indonesia 2024 di X Twitter untuk memahami respons masyarakat terhadap perkembangan politik. Menurut data *We Are Social* yang dikutip oleh Tirto.id, pengguna media sosial di Indonesia mencapai 167 juta orang pada 2023, dengan mayoritas pengguna adalah kaum muda yang mewakili lebih dari 50% pemilih. Minat terhadap informasi politik cukup tinggi, dengan 46,07% pengguna menunjukkan ketertarikan signifikan [8].

Text mining adalah proses ekstraksi informasi dari teks media sosial, melibatkan analisis tren, pola, dan pembobotan kata untuk mendapatkan wawasan berharga [9]. Klasifikasi, sebagai teknik *text mining*, bertujuan mengelompokkan objek berdasarkan karakteristik serupa untuk menganalisis pendapat, sikap, dan perasaan [10]. *Text mining* didefinisikan sebagai proses menambang data berupa teks, bertujuan mencari kata-kata yang mewakili isi dokumen untuk analisis hubungan antar dokumen menggunakan algoritma *machine learning* [11]. Metode *Naive Bayes* dan *TF-IDF* telah menunjukkan efektivitas tinggi dalam klasifikasi teks, khususnya dalam studi tentang kenaikan harga bahan pokok di X Twitter, mencapai akurasi tinggi setelah *preprocessing* menyeluruh [12].

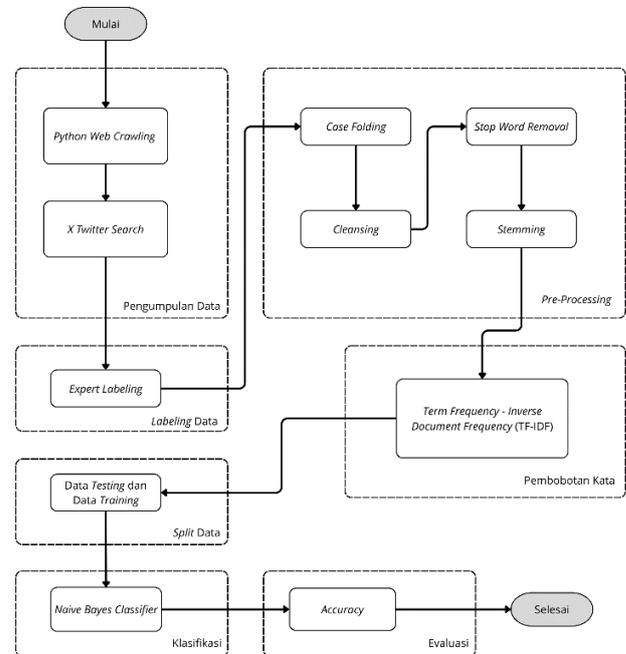
Penelitian sebelumnya [13] pada ulasan aplikasi Ruang Guru, metode ini menghasilkan prediksi sentimen dominan positif dengan *presisi* 71%, *recall* 69%, *F1-score* 69%, dan akurasi 69%. Penelitian lain [14] pada sentimen Pelayanan Dinas Kependudukan dan Pencatatan Sipil menunjukkan hasil sentimen positif 70,20%, negatif 29,80%, dengan akurasi 65,39%. Meskipun pendekatan ini telah digunakan dalam berbagai bidang, belum ada penelitian yang secara khusus menerapkannya untuk mengklasifikasi teks terkait *Quick Count* Pemilihan Presiden Indonesia 2024 pada X Twitter.

Penelitian ini bertujuan untuk mengukur akurasi metode *Naive Bayes* dengan pembobotan *TF-IDF* dalam mengklasifikasikan teks *Quick Count* Pemilihan Presiden Indonesia 2024 di media sosial X Twitter. Dengan menggabungkan pembobotan *TF-IDF* dan klasifikasi *Naive Bayes*, penelitian ini diharapkan dapat memberikan wawasan berharga tentang sentimen dan opini publik seputar *Quick Count*, yang dapat membantu pemangku kepentingan dalam memahami dan merespons dinamika politik yang terjadi selama periode pemilihan presiden.

2. METODE PENELITIAN

Penelitian ini melibatkan klasifikasi teks data X Twitter melalui serangkaian tahap. Dimulai dengan pengumpulan data menggunakan pencarian X Twitter dan *web crawling* berbasis *Python*, diikuti *labeling* oleh ahli. Data kemudian

melalui *preprocessing* yang mencakup *case folding*, *cleansing*, *stopwords removal*, dan *stemming*. Ekstraksi fitur dilakukan dengan metode *TF-IDF* untuk pembobotan kata. Selanjutnya, data dibagi menjadi set pelatihan dan pengujian. Klasifikasi teks menggunakan *Naive Bayes classifier*, dan kinerja model dievaluasi berdasarkan akurasi. Gambar 1 menyajikan alur penelitian secara komprehensif.



Gambar 1. Alur Penelitian

2.1 Pengumpulan Data

Twitter menyediakan kunci antarmuka pemrograman aplikasi yang memungkinkan pengguna untuk mengakses data Twitter secara terprogram, namun dengan keterbatasan jumlah data yang dapat diambil. Pengumpulan data dilakukan menggunakan *Tweet Harvest*, sebuah alat yang memungkinkan pengguna untuk mengumpulkan lebih banyak data Twitter dapat diakses dan dijalankan melalui *Command Line Interface (CLI)* hanya dengan menggunakan *auth_token*. Proses *crawling* dilakukan dengan menentukan kata kunci pencarian, rentang tanggal pencarian, dan batas jumlah data yang akan diambil. Setelah proses *crawling* selesai, data yang berhasil diambil akan disimpan dalam format CSV untuk digunakan dalam klasifikasi selanjutnya.

2.2 Labeling Data

Labeling data dilakukan oleh ahli dalam pelabelan data yang memiliki kualifikasi dan pengalaman yang relevan. Proses ini mengklasifikasikan *tweets* ke dalam kategori sentimen positif atau negatif. Acuan pelabelan yang digunakan adalah sebagai berikut:

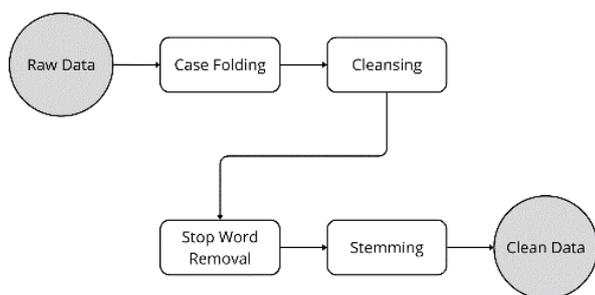
- Positif: *Quick Count* terbukti akurat dan terpercaya dalam berbagai pemilihan, memberikan gambaran awal valid tentang hasil pemungutan suara. Metode ini mendukung transparansi demokrasi, disambut antusias

oleh masyarakat dan media. *Hashtag #QuickCount* digunakan untuk berbagi informasi terkini, sementara media menyajikan berita objektif untuk membantu pemahaman publik.

- b. Negatif: Beberapa pihak meragukan metodologi dan akurasi *Quick Count*, mengkritik potensi pengiringan opini atau manipulasi data yang menimbulkan perdebatan di media sosial. Sebagian masyarakat kecewa dengan hasil yang tidak sesuai ekspektasi, mengekspresikan dengan bahasa kasar. Keraguan juga muncul mengenai objektivitas lembaga pelaksana dan adanya kepentingan tertentu di balik penyelenggaraan *Quick Count*.

2.3 Pre-Processing

Preprocessing adalah tahap yang krusial sebelum memulai sebuah penelitian. *Text preprocessing* dilakukan untuk memastikan bahwa data awal melewati serangkaian tahapan sehingga menjadi siap digunakan sepenuhnya dalam proses klasifikasi [15]. Tahapan *preprocessing* ditunjukkan pada Gambar 2 dan meliputi:



Gambar 2. Pre-Processing

- Folding*: Mengubah semua huruf dalam teks menjadi huruf kecil menggunakan *Python string lower method*, memastikan konsistensi format.
- Cleansing*: Menghapus karakter dan elemen tidak relevan seperti *username*, *URL*, *emoji*, angka, dan simbol-simbol yang tidak berkontribusi pada klasifikasi teks.
- Stopword Removal*: Menghapus kata-kata umum (*stopwords*) yang tidak memberikan makna signifikan dalam klasifikasi sentimen.
- Stemming*: Mengubah kata-kata ke bentuk dasarnya menggunakan algoritma *Sastrawi*, yang dirancang khusus untuk Bahasa Indonesia. Proses ini membantu dalam penyederhanaan kata-kata dengan akar yang sama.

2.4 Pembobotan Kata

Term Frequency-Inverse Document Frequency (TF-IDF) adalah metode penting dalam analisis teks untuk pembobotan kata. *TF* mengukur frekuensi kemunculan term dalam dokumen, sementara *DF* menghitung jumlah

dokumen yang memuat term tersebut [16]. Salah satu *transformer* yang mengubah data teks menjadi representasi numerik adalah *TfidfVectorizer* untuk melakukan ekstraksi fitur kata dan menghitung frekuensi kemunculannya. Alat ini tidak hanya mengonversi teks tetapi juga menghitung bobot *TF-IDF* untuk setiap kata dalam *dataset* [17]. Rumus perhitungan *TF-IDF* disajikan dalam Persamaan (2.1).

$$TF - IDF(t, d) = TF(t, d) * IDF(t) \quad (2.1)$$

Keterangan:

- a. $TF(t, d)$ adalah *term frequency* dari kata t dalam dokumen d , yang dihitung sebagai berikut:

$$TF(t, d) = \frac{\text{jumlah kemunculan kata } t \text{ dalam dokumen } d}{\text{total jumlah kata dalam dokumen } d} \quad (2.2)$$

- b. $IDF(t)$ adalah *inverse document frequency* dari kata t , yang dihitung sebagai berikut:

$$IDF(t) = \log \frac{\text{total jumlah dokumen dalam koleksi}}{\text{jumlah dokumen mengandung kata } t + 1} \quad (2.3)$$

2.5 Split Data

Split Data adalah proses pembagian *dataset* menjadi data pelatihan (*training data*) dan data pengujian (*testing data*). Data pelatihan digunakan untuk melatih model klasifikasi *Naive Bayes* agar dapat mempelajari pola dan karakteristik data, sedangkan data uji digunakan untuk menilai performa model pada data baru [18]. Dalam penelitian klasifikasi komentar *toxic* menggunakan *TF-IDF* dan *Naive Bayes*, rasio 80:20 menghasilkan akurasi terbaik. Rasio ini dipilih karena memberikan keseimbangan optimal antara jumlah data untuk melatih dan menguji model secara akurat. Hal ini menunjukkan bahwa rasio 80:20 efektif dalam menjaga keseimbangan antara data pelatihan dan pengujian, memungkinkan evaluasi model yang baik pada data yang belum pernah dilihat sebelumnya [19]. Dalam proses pembagian data, digunakan *random state* sebesar 42 untuk memastikan hasil yang dapat diulang dan konsistensi dalam pemilihan sampel acak, sehingga eksperimen dapat dijalankan kembali dengan hasil yang sama.

2.6 Klasifikasi

Klasifikasi *Naive Bayes* adalah metode yang berdasarkan *Teorema Bayes*, diusulkan oleh *Thomas Bayes*. Metode ini menggunakan pendekatan probabilitas dan statistik untuk memproyeksikan kemungkinan kejadian masa depan berdasarkan pengalaman masa lalu. *Naive Bayes Classifier* dikenal dengan proses yang cepat, mudah diterapkan, struktur sederhana, dan efektivitas tinggi [20]. Salah satu pengembangan dari algoritma klasifikasi *Naive Bayes* adalah *Multinomial Naive Bayes* yaitu model di mana kelas tidak hanya ditentukan dengan kata yang muncul, tapi juga dengan jumlah kemunculannya sehingga cocok untuk klasifikasi teks atau dokumen [21]. Rumus *Teorema Bayes* disajikan dalam Persamaan (2.4).

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (2.4)$$

Keterangan:

- X: Data yang kelasnya belum diketahui
- C: Hipotesis bahwa data X merupakan kelas tertentu
- $P(C|X)$: Probabilitas hipotesis C benar jika diberikan data X (*probabilitas posterior*)
- $P(C)$: Probabilitas awal hipotesis C sebelum melihat data X (*probabilitas prior*)
- $P(X|C)$: Probabilitas munculnya data X jika hipotesis C benar (*likelihood*)
- $P(X)$: Probabilitas kemunculan data X secara umum (*probabilitas evidence*)

2.7 Evaluasi

Confusion Matrix adalah alat evaluasi yang digunakan untuk mengukur performa model klasifikasi. *Confusion Matrix* memberikan gambaran jumlah prediksi yang akurat dan tidak akurat yang dibuat oleh model [22]. Struktur *Confusion Matrix* untuk klasifikasi biner ditunjukkan pada Tabel 1.

Tabel 1. *Confusion Matrix*

	Predicted Class	
	Positive	Negative
Actual Class	True Positive (TP)	False Negative (FN)
	False Positive (FP)	True Negative (TN)

Keterangan:

- True Positive* (TP): Jumlah data positif yang diprediksi dengan akurat sebagai positif.
- False Positive* (FP): Jumlah data negatif yang keliru diprediksi sebagai positif.
- False Negative* (FN): Jumlah data positif yang keliru diprediksi sebagai negatif.
- True Negative* (TN): Jumlah data negatif yang diprediksi dengan tepat sebagai negatif.

Akurasi adalah cara untuk mengukur seberapa tepat suatu model dalam membuat prediksi. Metode ini membandingkan seberapa dekat hasil prediksi dengan nilai yang sebenarnya untuk keseluruhan data [23]. Akurasi model dihitung dengan Persamaan (2.5) sebagai berikut:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (2.5)$$

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pengumpulan data penelitian menghasilkan *dataset* komprehensif tentang diskusi *Quick Count* Pemilihan Presiden Indonesia 2024 di X Twitter. Menggunakan *Tweet Harvest*, terkumpul 2113 *tweet* relevan pada 14 Februari 2024. Proses *crawling* dilakukan selama 29 menit pada 21 April 2024, dari 21:11 hingga 21:40 WITA. Sampel data di Tabel 2 menunjukkan keragaman sudut pandang dan topik diskusi, mencakup hasil sementara dan dampak *Quick Count* terhadap proses pemilu selanjutnya.

Tabel 2. Hasil Pengumpulan Data

No	created_at	full_text
1	Wed Feb 14 22:24:25 +0000 2024	@m4retha Bilas muka gosok gigi liat quick count lagi
2	Wed Feb 14 22:24:19 +0000 2024	@nvtamarsaoly @convomfs Bener bgt. Ga jauh2 hasil real count nanti sama quick count. Udah nder capee ngomong sama yang ga mau kalah sipaling dicurangi dan tersakiti.
...
2113	Wed Feb 14 16:29:25 +0000 2024	@m4retha Update 23.29: 02 unggul dengan suara diatas 50% pada 29 provinsi. Total quickcount di 38 provinsi + Luar Negeri: 01 Amin 32 50% 02 Pagi 51 21% 03 Gama 16 29%

Dari 2113 *tweet* yang terkumpul mengenai *Quick Count*, hanya konten *full_text* yang diambil dari setiap *tweet*. Selanjutnya, data ini akan melalui proses *labeling* oleh seorang ahli pelabelan data yang berkualifikasi dan berpengalaman.

3.2 Labeling Data

Proses *labeling* data untuk klasifikasi teks *tweet* dilakukan oleh ahli berkualifikasi dengan acuan yang jelas. *Dataset* terdiri dari 2113 *tweet* total, dengan 1485 *tweet* (70.3%) berlabel negatif dan 628 *tweet* (29.7%) berlabel positif. Distribusi ini menunjukkan ketidakseimbangan kelas yang signifikan, dengan dominasi sentimen negatif. Meskipun tidak seimbang, jumlah data ini masih cukup representatif untuk klasifikasi teks terkait *Quick Count* Pemilihan Presiden Indonesia 2024 di X Twitter. Gambar 3 menampilkan distribusi sentimen dalam bentuk *pie chart*.

Total Vocabulary: 4999

Term Frequency-Inverse Document Frequency (TF-IDF):

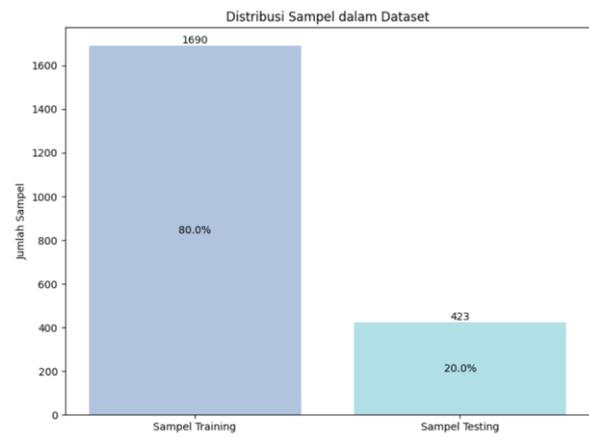
(0, 838)	0.06886996842891753
(0, 3669)	0.06886996842891753
(0, 2466)	0.23852635668010155
(0, 1458)	0.484358630847651
(0, 1494)	0.484358630847651
(0, 2868)	0.4516208951228151
(0, 536)	0.5103440335708471
(1, 3861)	0.37087090191925937
(1, 880)	0.22228298442708463
(1, 4169)	0.46581517195266436
(1, 2001)	0.2259455060552625
(1, 3067)	0.39160913667826514
(1, 716)	0.4908057348679603
(1, 3720)	0.1733645454710436
(1, 1583)	0.13154268294619828
(1, 449)	0.28542760295740477
(1, 838)	0.13246662346017476
(1, 3669)	0.06623331173008738
(2, 1841)	0.3231369472829371
(2, 146)	0.2701786628891291
(2, 1950)	0.29363684743878893
(2, 4029)	0.2851545569615561
(2, 295)	0.4865155590556142
(2, 3411)	0.19556787121704372
(2, 4820)	0.36762908696095403
(2, 1054)	0.37704700921422624
(2, 3720)	0.17184915083145502
(2, 838)	0.1313087211303153
(2, 3669)	0.06565436056515765
(2, 2466)	0.2273893219790304
(3, 1129)	0.3293850799960921
(3, 987)	0.3821001212629606
(3, 2414)	0.36264456799210665
(3, 4851)	0.30992952672523816
(3, 3461)	0.24906216927890243
(3, 350)	0.22395499746235517
(3, 1432)	0.3821001212629606
(3, 718)	0.2537856486666287
(3, 3463)	0.282321657274917
(3, 3411)	0.3071906166874691
...	...
(2112, 3670)	0.1995555639791592
(2112, 4317)	0.18476864059004822
(2112, 146)	0.2451273162975515

Gambar 7. Tampilan Hasil Pembobotan Kata

Matriks *TF-IDF* menggambarkan bobot setiap *term* dalam setiap dokumen, mencerminkan kepentingan *term* relatif terhadap seluruh kumpulan dokumen. Visualisasi matriks membantu identifikasi pola dan tema utama dalam kumpulan dokumen.

3.6 Split Data

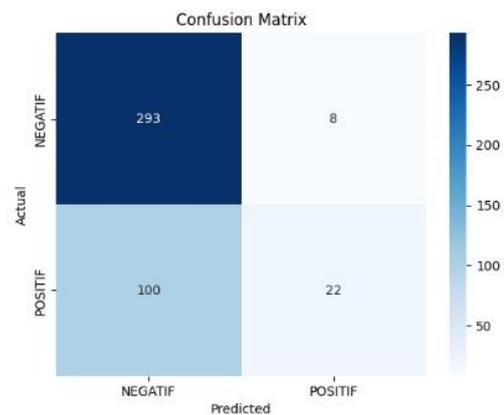
Proses *split* data membagi *dataset* menjadi 1690 sampel (80%) untuk data *training* dan 423 sampel (20%) untuk data *testing* dari total 2113 sampel. Rasio 80:20 dipilih untuk keseimbangan optimal antara data pelatihan dan pengujian model *Naive Bayes*. Visualisasi *bar chart* pada Gambar 8 menampilkan distribusi sampel *training* dan sampel *testing* beserta persentasenya.



Gambar 8. Tampilan Hasil Split Data

3.7 Klasifikasi & Evaluasi

Hasil analisis model klasifikasi *Naive Bayes* divisualisasikan melalui *confusion matrix* pada Gambar 9. Matriks ini menunjukkan performa model dalam memprediksi sentimen 293 *true negative*, 22 *true positive*, 100 *false negative*, dan 8 *false positive*. Visualisasi ini memberikan gambaran rinci tentang distribusi prediksi model secara lebih mendalam terhadap akurasinya.



Gambar 9. Tampilan Hasil Confusion Matrix

Hasil akurasi berdasarkan data dari *confusion matrix*, yaitu 74,46%. Meskipun hasil ini menunjukkan performa yang cukup baik, masih ada ruang untuk peningkatan, misalnya dengan mengoptimalkan fitur, menambah data atau mencoba algoritma klasifikasi lainnya.

4. KESIMPULAN

Penelitian ini berhasil mencapai tujuannya untuk mengukur akurasi metode *Naive Bayes* dengan pembobotan *TF-IDF* dalam klasifikasi teks *Quick Count* Pemilihan Presiden Indonesia 2024 di X Twitter. Hasil analisis menunjukkan model mencapai akurasi 74,46% dalam mengklasifikasikan teks. Meskipun akurasi keseluruhan cukup baik, namun memerlukan penyempurnaan lebih lanjut untuk meningkatkan keseimbangan performa antar kelas sentimen. Temuan ini mengindikasikan bahwa metode *Naive Bayes* dengan pembobotan *TF-IDF* memiliki potensi

yang baik dalam mengklasifikasi teks terkait *Quick Count*, namun memerlukan penyempurnaan lebih lanjut untuk meningkatkan keseimbangan performa antar kelas sentimen. Untuk meningkatkan kinerja model, beberapa saran dapat dipertimbangkan, termasuk mengatasi ketidakseimbangan kelas melalui teknik *oversampling* atau *undersampling*, menggunakan metode ekstraksi fitur alternatif seperti *word embeddings*, membandingkan performa dengan algoritma klasifikasi lain, meningkatkan proses *preprocessing* terutama untuk bahasa informal, memperbesar *dataset*, dan menerapkan *cross-validation*. Implementasi saran-saran ini diharapkan dapat meningkatkan akurasi dan keseimbangan model secara signifikan, sehingga meningkatkan efektivitas klasifikasi teks terkait *Quick Count* di masa mendatang.

DAFTAR PUSTAKA

- [1] Hananto Widodo and Dicky Eko Prasetyo, "Penataan Kewenangan KPU dan Bawaslu dalam Melakukan Pengawasan dan Menangani Sengketa Proses Pemilu," *Perspekt. Huk.*, vol. 21, no. 2, pp. 17–38, 2021, doi: 10.30649/ph.v21i2.93.
- [2] N. B. Nugraha, M. Suhaidi, and J. S., "Aplikasi Quick Count Pada Pemilu Legislatif Dengan Metode Sainte Lague Berbasis Mobile," *Inform. J. Inform. Manaj. DAN Komput.*, vol. 11, no. 1, p. 76, 2019, doi: 10.36723/juri.v11i1.158.
- [3] A. P. Rhima Indria Saraswati, "Kepercayaan Masyarakat Terhadap Hasil Quick Count Pada Pemilihan Presiden Dan Legislatif Tahun 2019," *J. Kewarganegaraan*, vol. 4, no. 1, p. 25, 2020.
- [4] Andi Saadillah, Andi Haryudi, Muhammad Reskiawan, and Alam Ikhsanul Amanah, "Penggunaan Bahasa Sarkasme Netizen di Media Sosial," *J. Onoma Pendidikan, Bahasa, dan Sastra*, vol. 9, no. 2, pp. 1437–1447, 2023, doi: 10.30605/onoma.v9i2.2367.
- [5] K. Ulfa, E. P. Purnomo, and A. N. Kasiwi, "The Campaign Strategy of 2019 Presidential and Vice-Presidential Elections on Social Media," *Society*, vol. 8, no. 2, pp. 284–297, 2020, doi: 10.33019/society.v8i2.137.
- [6] D. F. Zhafira, B. Rahayudi, and I. Indriati, "Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube," *J. Sist. Informasi, Teknol. Informasi, dan Edukasi Sist. Inf.*, vol. 2, no. 1, pp. 55–63, 2021, doi: 10.25126/justsi.v2i1.24.
- [7] L. Wulandari, "Framing and Sentiment Analysis of Gerindra Party Political Communication on Twitter (X) Towards the 2024 Elections Framing and Sentiment Analysis of Gerindra Party Political," *J. Wacana Polit.*, vol. 8, no. January, pp. 2–10, 2024, doi: 10.24198/jwp.v8i1.51967.
- [8] I. Amin, "Media Sosial antara Potensi & Bahaya dalam Kampanye Pemilu 2024," *tirto.id*. [Online]. Available: <https://tirto.id/media-sosial-antara-potensi-bahaya-dalam-kampanye-pemilu-2024-gNwH>
- [9] E. A. Sosiawan and R. Wibowo, "Kontestasi Berita Hoax Pemilu Presiden Tahun 2019 di Media Daring dan Media Sosial," *J. Ilmu Komun.*, vol. 17, no. 2, p. 133, 2020, doi: 10.31315/jik.v17i2.3695.
- [10] M. Rizki, M. Fikri Hidayattullah, and Dwi Intan Af'idah, "Klasifikasi Opini Publik di Twitter Terhadap Bakal Calon Presiden Indonesia Tahun 2024 Menggunakan LSTM Secara Realtime Berbasis Website," *Infotekmesin*, vol. 14, no. 2, pp. 285–295, 2023, doi: 10.35970/infotekmesin.v14i2.1908.
- [11] D. Nuralasari and H. Ribut Yuliantoro, "Implementasi Ekstraksi Fitur untuk Pengelompokan Dokumen Proposal Menggunakan Algoritma Naïve Bayes," *J. Komput. Terap.*, vol. 8, no. 1, pp. 194–203, 2022, doi: 10.35143/jkt.v8i1.5351.
- [12] M. Muslimin and V. Lusiana, "Analisis Sentimen Terhadap Kenaikan Harga Bahan Pokok Menggunakan Metode Naive Bayes Classifier," *J. Media Inform. Budidarma*, vol. 7, no. 3, pp. 1200–1209, 2023, doi: 10.30865/mib.v7i3.6418.
- [13] I. Novitasari, T. B. Kurniawan, D. A. Dewi, and Misinem, "Analisis sentimen masyarakat terhadap tweet ruang guru menggunakan algoritma naive bayes classifier (NBC) [Analysis of public sentiment towards ruang guru's tweets using the Naive Bayes Classifier (NBC) algorithm]," *J. Mantik*, vol. 6, no. 3, pp. 2685–4236, 2022.
- [14] E. Salim and A. Solichin, "ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP PELAYANAN DINAS KEPENDUDUKAN DAN PENCATATAN SIPIL MENGGUNAKAN ALGORITMA NAÏVE BAYES," 2022. [Online]. Available: <http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/indexEmilSalim%7Chttp://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index%7C>
- [15] M. U. Albab, Y. Karuniawati, and M. N. Fawaiq, "Optimization of the Stemming Technique on Text preprocessing President 3 Periods Topic," *J. Transform.*, vol. 20, no. 2, pp. 1–10, 2023, [Online]. Available: <https://journals.usm.ac.id/index.php/transformatika/page1>
- [16] M. Hamka, N. Alfajari, and D. Ratna Sari, "Analisis Sentimen Produk Kecantikan Jenis Serum Menggunakan Algoritma Naïve Bayes Classifier," *J. Sist. Komput. dan Inform.*, vol. 4, no. 1, p. 64, 2022, doi: 10.30865/json.v4i1.4740.
- [17] M. R. A. Surya, Martanto, and U. Hayati, "Analisis

- Sentimen Ulasan Pengguna Ovo Menggunakan Algoritma Naive Bayes Pada Google Play Store,” *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 3, pp. 2780–2786, 2024.
- [18] A. Putri *et al.*, “Komparasi Algoritma K-NN, Naive Bayes dan SVM untuk Prediksi Kelulusan Mahasiswa Tingkat Akhir,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 1, pp. 20–26, 2023, doi: 10.57152/malcom.v3i1.610.
- [19] R. P. Sidiq, B. A. Dermawan, and Y. Umaidah, “Sentimen Analisis Komentar *Toxic* pada Grup Facebook Game Online Menggunakan Klasifikasi Naïve Bayes,” *J. Inform. Univ. Pamulang*, vol. 5, no. 3, p. 356, 2020, doi: 10.32493/informatika.v5i3.6571.
- [20] L. S. Y. Sriani, Suhardi, “ANALISIS SENTIMEN PENGGUNA APLIKASI MOBILE JKN MENGGUNAKAN ALGORITMA NAÏVE BAYES,” *J. Sci. Soc. Res.*, vol. 4307, no. May, pp. 555–563, 2024.
- [21] I. N. Husada and H. Toba, “Pengaruh Metode Penyeimbangan Kelas Terhadap Tingkat Akurasi Analisis Sentimen pada *Tweets* Berbahasa Indonesia,” *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 400–413, 2020, doi: 10.28932/jutisi.v6i2.2743.
- [22] M. I. Fikri, T. S. Sabrila, Y. Azhar, and U. M. Malang, “Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter,” *SMATIKA J. STIKI Inform. J.*, vol. 10, no. 2, pp. 71–76, 2020.
- [23] L. Farokhah, “Implementasi K-Nearest Neighbor Untuk Klasifikasi Bunga Dengan Ekstraksi Fitur Warna Rgb Implementation of K-Nearest Neighbor for Flower Classification With Extraction of Rgb Color Features,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 6, pp. 1129–1136, 2020, doi: 10.25126/jtiik.202072608.