

ANALISIS SENTIMEN MASYARAKAT TERHADAP TREN KPOPFICATION DALAM MASA KAMPANYE DI MEDIA SOSIAL X

Hutami Jane Antonia .M^{*1)}, Evangs Mailoa²⁾

1. Universitas Kristen Satya Wacana, Salatiga
2. Universitas Kristen Satya Wacana, Salatiga

Article Info

Kata Kunci: Analisis Sentimen; Kpopfication; Lexicon Based; Naïve Bayes; Pemilu

Keywords: Election; Kpopfication; Lexicon Based; Naïve Bayes; Sentiment Analysis;

Article history:

Received 29 September 2024

Revised 13 Oktober 2024

Accepted 4 November 2024

Available online 4 December 2024

DOI :

<https://doi.org/10.29100/jupi.v9i4.5638>

* Corresponding author.

Hutami Jane Antonia .M

E-mail address:

hutamijane29@gmail.com

ABSTRAK

Perkembangan teknologi informasi dalam bidang politik menjadikan media sosial sebagai *platform* untuk berkampanye dan berinteraksi antara politisi dan pemilih. Penelitian ini bertujuan untuk menerapkan Algoritma Naïve Bayes dalam menganalisis sentimen pada *tweet* terkait dengan tren "Kpopfication" dalam pemilihan umum. "Kpopfication" adalah istilah yang digunakan untuk menggambarkan tindakan para penggemar K-Pop yang mengintegrasikan elemen-elemen dari dunia K-Pop dalam berbagai aspek kehidupan di luar dunia K-Pop itu sendiri. Langkah pertama adalah pengumpulan data dari *tweet* pengguna X melalui metode *crawling* dengan kata kunci 'Kpopfication Pemilu lang:id'. Langkah kedua melakukan Pre-Processing data untuk membersihkan dan mempersiapkan data agar dapat diklasifikasikan. Langkah ketiga adalah melakukan pelabelan dengan menggunakan Lexicon Based dan yang terakhir adalah penerapan Algoritma Naïve Bayes. Dari analisis sentimen di dapatkan hasil akurasi 86%, presisi 84%, recall 96%, dan *f1-score* 90%, serta evaluasi kinerja model yang baik dalam mengenali sentimen positif dan negatif. Hasil penelitian ini akan menjadi landasan untuk penelitian lebih lanjut dalam konteks pemilu dan inovasi dalam berkampanye.

ABSTRACT

The development of information technology in politics has made social media a platform for campaigning and interaction between politicians and voters. The research aims to apply the Naïve Bayes algorithm in analyzing sentiment on tweets related to the "Kpopfication" trend in the general election. "Kpopfication" is a term used to describe the actions of K-Pop fans that integrate elements of the K-pop world into various aspects of life outside the K -Pop world itself. The first step is to collect data from user X's tweets through a crawling method with the keyword 'Kpopfication Selection lang:id'. The second step is pre-processing the data to clear and prepare the data for classification. The third step is to do labelling using Lexicon Based and the last is to apply the Naïve Bayes algorithm. From the sentiment analysis obtained accuracy results of 86%, precision 84%, recall 96%, and *f1-score* 90%, as well as evaluation of the model's good performance in recognizing positive and negative sentiments. The results of this research will be the basis for further research in the context of elections and innovation in campaigning.

I. PENDAHULUAN

PERKEMBANGAN teknologi informasi telah membawa banyak perubahan dalam berbagai aspek kehidupan bersosialisasi. Salah satunya adalah munculnya media sosial sebagai alat komunikasi yang memungkinkan individu untuk saling berkomunikasi, bertukar informasi, dan berbagi momen meskipun dari jarak jauh. Hal ini juga menjadikan media sosial memiliki dampak yang besar dalam membentuk opini publik [1]. Di bidang politik media sosial seringkali dimanfaatkan sebagai alat kampanye menjelang pemilihan umum untuk memperkenalkan visi, misi serta program kerja kedepannya dari para calon kepada masyarakat[2]. Penelitian di Amerika Serikat mengungkapkan bahwa media sosial adalah sarana kampanye yang efektif dan efisien, bahkan sebelum perkembangan media sosial menjadi seperti sekarang, para politisi di negara tersebut sudah menggunakan media sebagai alat kampanye[3]. Para kandidat politik menggunakan media sosial untuk berinteraksi secara langsung dengan pemilih, mendengarkan umpan balik dari pemilih, dan menyesuaikan strategi kampanye sesuai dengan masukan yang diterima [4]. Salah satu media sosial yang populer dikalangan masyarakat adalah X, yaitu platform media sosial berbasis *microblogging* yang memungkinkan penggunaanya untuk berbagi pesan singkat atau "*Tweet*" dengan panjang maksimal 280 karakter[5]. X menjadi platform yang sangat dinamis untuk berbagi informasi,

berpartisipasi dalam sebuah diskusi dan membangun jaringan.

Pemilu 2024 yang dilaksanakan serentak pada tanggal 14 Februari menarik banyak perhatian, para pendukung dari calon kandidat yang berpartisipasi turut meramaikan tren yang sedang ramai[6] Salah satu tren yang muncul adalah "Kpopfication", istilah yang digunakan untuk menggambarkan tindakan penggemar K-pop yang mengintegrasikan elemen-elemen dari dunia K-pop ke dalam berbagai aspek kehidupan di luar K-pop itu sendiri, seperti membuat *merch photocard*, membuat akun *fanbase*, mengedit calon kandidat dengan idol *Boygrouop* atau *Girlgroup* dan berbagai aktivitas lain yang berhubungan dengan K-pop. Tren ini Menimbulkan perdebatan, terutama di kalangan penggemar K-pop yang tidak ingin artis idola mereka dikaitkan dengan politik di Indonesia. Selain itu, menggunakan budaya asing sebagai strategi kampanye dianggap tidak etis karena mengurangi esensi dari kampanye itu sendiri[7]. Secara umum, tren ini memiliki dampak positif dan negatif. Dampak positifnya adalah meningkatkan kreativitas dan partisipasi politik dikalangan generasi muda, serta memperluas jangkauan kampanye. Di sisi lain, dampak negatifnya adalah berpotensi terjadinya pergeseran fokus dari isu-isu penting dalam kampanye ke hal-hal yang bersifat superfisial.

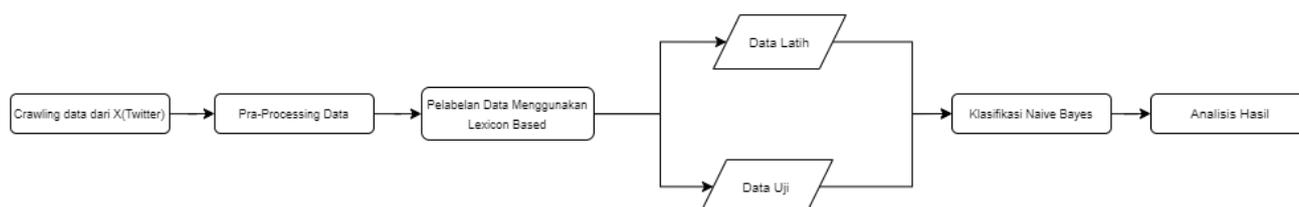
Berdasarkan latar belakang tersebut, dilakukan penelitian dengan menganalisis sentimen masyarakat di media sosial X untuk mengetahui segmentasi dari pendapat masyarakat terkait tren "Kpopfication" pada masa kampanye. Langkah-langkah penelitian ini meliputi pengumpulan data *tweet* dari X dengan proses *crawling*, *pra-processing* data untuk membersihkan data dari *noise*, pelabelan menggunakan Kamus InSet (*Indonesian Sentiment*) *Lexicon Based* dan yang terakhir adalah klasifikasi sentimen menggunakan Algoritma Naïve Bayes. Keunggulan yang dimiliki oleh Algoritma Naïve Bayes adalah kemampuannya untuk memberikan hasil yang baik dalam klasifikasi sentimen dengan hanya memerlukan sedikit data latih untuk menentukan pola dan modelnya.

Penelitian pertama berjudul Analisis Sentimen Masyarakat pada Twitter Terhadap Pemilihan Umum 2024 Menggunakan Algoritma Naïve Bayes. Data diperoleh melalui metode *crawling*, dengan penerapan pembobotan kata TF-IDF untuk meningkatkan akurasi model, serta pelabelan menggunakan kamus *lexicon*. Penelitian ini menguji berbagai skenario pembagian data latih dan data uji. Dalam skenario dengan 90% data latih dan 10% data uji, akurasi terbaik yang diperoleh adalah sebesar 66%. Data yang dianalisis mencakup 331 *tweet* negatif, 261 *tweet* positif, dan 825 *tweet* netral. Kekurangan dari penelitian ini terletak pada akurasi yang relatif rendah dan tingginya jumlah data netral.[8]. Penelitian kedua berjudul Analisis Sentimen Pemilu 2024 dengan Naïve Bayes Berbasis *Particle Swarm Optimization* (PSO). Data dikumpulkan menggunakan metode *crawling*, dengan pembagian 70% untuk data latih dan 30% untuk data uji. Penelitian ini membandingkan dua hasil, yaitu akurasi sebesar 73,67% ketika hanya menggunakan Naïve Bayes dan akurasi sebesar 78,33% ketika menggabungkan Naïve Bayes dengan PSO.[9].

Perbedaan penelitian pertama dan kedua dengan penelitian ini adalah dalam hal pelabelan sentimen, yang menggunakan *InSet Lexicon Based*. Metode ini lebih relevan dan akurat untuk konteks lokal. Selain itu, objek penelitian "Kpopfication" dalam masa kampanye belum pernah diteliti dan dibahas dalam jurnal lain, sehingga diharapkan penelitian ini dapat memberikan kontribusi baru dan orisinal dalam bidang ini. Tujuan penelitian ini adalah untuk mengidentifikasi persepsi dan sikap masyarakat di aplikasi X terhadap elemen K-pop dalam konteks pemilu melalui analisis sentimen. hasilnya diharapkan dapat membantu politisi memahami tren opini publik secara lebih baik, sehingga dapat merancang strategi kampanye politik yang sesuai dengan preferensi pemilih.

II. METODE PENELITIAN

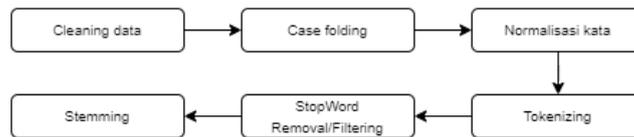
Data yang digunakan dalam penelitian ini diperoleh dari *tweet* di *platform X* melalui proses *crawling*, yaitu metode otomatis untuk mengambil data dari situs *web* menggunakan mesin pencari[10]. Proses pengumpulan data dilakukan dalam rentang waktu 20 Maret sampai 15 April 2024 dengan menggunakan bantuan Google Colaboratory dan *Tweet-Harvest*. Data yang berhasil dikumpulkan sebanyak 3352 data. Data tersebut nantinya akan dianalisis dengan mempertimbangkan nilai sentimen yang terkandung di dalamnya. Proses ini disebut analisis sentimen yang perhitungannya memanfaatkan Algoritma Klasifikasi Naive Bayes[11]. Alur penelitian dapat dilihat pada gambar 1.



Gambar 1 Tahapan Penelitian

Tahap pertama adalah pengumpulan data yang didapatkan dari *tweet* pengguna X yang dikumpulkan

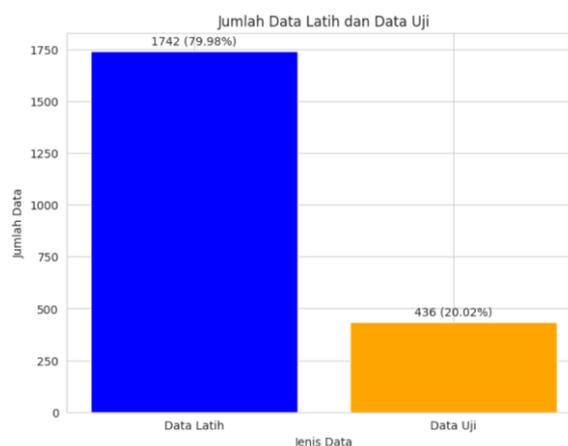
menggunakan Metode *crawling* dengan kata kunci ‘Kpopfication Pemilu lang:id’ dengan jumlah data sebanyak 3352 data. Tahap kedua, dilakukan *pre-processing* untuk membersihkan dan menstrukturkan data agar data tersebut diklasifikasikan. Terdapat enam tahapan dalam melakukan *pre-processing* yaitu, *cleaning* data bertujuan untuk menghilangkan *emoticon*, *url*, dan *username*. *Case folding* bertujuan untuk mengubah huruf kapital menjadi huruf kecil. Normalisasi kata bertujuan mengubah dari kata tidak baku menjadi kata baku. *Tokenizing* bertujuan untuk memisahkan setiap kata berdasarkan spasi. *StopWord Removal* atau *filtering* bertujuan untuk menghilangkan kata yang tidak penting dalam proses klasifikasi. Dan yang terakhir adalah proses *Stemming* yang bertujuan untuk mengubah kata yang memiliki imbuhan menjadi kata dasar. Setelah dilakukan *pre-processing*, data yang tersisa berjumlah 2966 data. Gambaran tahapan *pre-processing* dapat dilihat pada gambar 2.



Gambar 2 Tahapan *Pre-Processing*

Tahap kedua adalah pelabelan data menggunakan Metode Lexicon Based, yaitu melabeli setiap kata dalam kalimat berdasarkan klasifikasi positif dan negatif yang mengacu pada kamus InSet Lexico Based[12]. Kamus InSet Lexicon berisi daftar kata-kata dengan nilai sentimen yang sudah ditentukan, di mana kata-kata tersebut telah dikategorikan sebagai positif atau negatif. Dalam proses ini, setiap kata dalam kalimat diidentifikasi dan diberi label sesuai dengan dalam kamus tersebut[13]. Penilaian positif dan negatif dilakukan dengan memberikan skor berdasarkan jumlah kata positif dan negatif dalam kalimat. Skor sentimen dihitung dengan menjumlahkan nilai dari semua kata dalam kalimat. Hasil skor ini digunakan untuk menentukan sentimen keseluruhan dari kalimat tersebut apakah positif atau negatif. Jumlah data setelah dilakukan pelabelan adalah 2178.

Pembagian dataset dilakukan menjadi dua bagian, yaitu data latih dan data uji. Data latih digunakan untuk melatih sistem agar dapat mengidentifikasi pola yang diperlukan [14] sedangkan data uji digunakan untuk menguji hasil pelatihan yang telah dilakukan[15]. Dataset dibagi menjadi 80% data latih dan 20% data uji. Proses pembagian dataset ini dilakukan secara acak untuk memastikan bahwa setiap subset adalah representatif dari keseluruhan data dan tidak ada bias dalam pembagian data. Teknik stratifikasi digunakan untuk menjaga distribusi kelas yang sama antara data latih dan data uji, sehingga proporsi sentimen positif dan negatif tetap konsisten di kedua subset. Pembagian dataset dapat dilihat pada gambar 3.



Gambar 3 Pembagian Data latih dan Data Uji

Tahap terakhir pengklasifikasian menggunakan Algoritma Naive Bayes *Classifier* berdasarkan data latih dan data uji yang telah dibagi. Naive bayes *Classifier* menghitung probabilitas awal dari setiap label berdasarkan frekuensi kemunculan label pada data latih dan pengaruh masing-masing fitur. Berdasarkan sifat probabilitasnya, *Naive Bayes* bisa belajar secara efisien dalam pengaturan yang diawasi[16]. Dalam praktiknya, parameter untuk Model *Naive Bayes* sering dihitung dengan Metode *Maximum Likelihood*[17].

Performa model dievaluasi menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Metode validasi yang digunakan adalah *split validation*, dengan pembagian dataset menjadi 80% data latih dan 20% data uji. Alasan pemilihan metrik-metrik tersebut adalah untuk mendapatkan gambaran yang komprehensif tentang kinerja model, di mana akurasi memberikan informasi umum tentang prediksi yang benar, presisi mengukur ketepatan prediksi positif, *recall* menilai kemampuan model dalam menangkap semua kasus positif, dan *F1-score* memberikan keseimbangan antara presisi dan *recall*.

$$P(H|X) = \frac{P(X|H)P(H)}{P(H)} \tag{1}$$

Variabel X dalam persamaan (1) merujuk pada kelas data yang tidak dikenali, sedangkan H adalah hipotesis data yang mewakili kelas spesifik. $P(H|X)$ menggambarkan probabilitas hipotesis berdasarkan kondisi, atau yang dikenal sebagai probabilitas posteriori. $P(H)$ adalah probabilitas hipotesis itu sendiri, $P(X|H)$ adalah probabilitas kondisional pada hipotesis, dan yang terakhir, $P(X)$, mengacu pada probabilitas X.

III. HASIL DAN PEMBAHASAN

Pengumpulan data menggunakan metode *crawling* dengan bantuan *library* Tweet-Harvest[18]. Dengan menggunakan kata kunci ‘Kpopfication Pemilu lang:id’. Data yang berhasil diperoleh sebanyak 3325 data dan dilakukan *drop duplicates* untuk menghapus data yang duplikat dan *null* sehingga tersisa 2966 data. Proses *pre-processing* dapat dilihat pada tabel 1.

TABEL I
 PREPROCESSING DATA

Preprocessing	Text
Cleaning	@starfess Gara2 kebanyakan Kpopfication politik, gw sampe mikir sbs gayo nih maksudnya disamain sm acara politik yg mana 🤔🤔 Gara kebanyakan Kpopfication politik gw sampe mikir sbs gayo nih maksudnya disamain sm acara politik yg mana
Case Folding	gara kebanyakan Kpopfication politik gw sampe mikir sbs gayo nih maksudnya disamain sm acara politik yg mana
Normalisasi	gara kebanyakan Kpopfication politik gue sampai mikir sbs gayo nih maksudnya disamain sama acara politik yang mana
Tokenize	['gara', 'kebanyakan', 'Kpopfication', 'politik', 'gue', 'sampai', 'mikir', 'sbs', 'gayo', 'nih', 'maksudnya', 'disamain', 'sama', 'acara', 'politik', 'yang', 'mana']
Stopword Removal	['gara', 'kebanyakan', 'Kpopfication', 'politik', 'gue', 'mikir', 'sbs', 'gayo', 'nih', 'maksudnya', 'disamain', 'acara', 'politik']
Stemming	gara banyak Kpopfication politik gue mikir sbs gayo nih maksud disamain acara politik

Gambar 4 menampilkan visualisasi data dalam bentuk *WordCloud* yang bertujuan untuk mengidentifikasi kata apa saja yang sering muncul dalam topik “Kpopfication”. Kata yang berukuran besar menunjukkan frekuensi


```

1 #baca kamus lexicon positif dan negatif
2 positive_lexicon = set(pd.read_csv("positive.tsv", sep="\t", header=None)[0])
3 negative_lexicon = set(pd.read_csv("negative.tsv", sep="\t", header=None)[0])
4
5 #fungsi untuk menentukan sentimen
6 def determine_sentiment(text):
7     positive_count = sum(1 for word in text.split() if word in positive_lexicon)
8     negative_count = sum(1 for word in text.split() if word in negative_lexicon)
9     if positive_count > negative_count:
10        return "Positive"
11    elif positive_count < negative_count:
12        return "Negative"
13 #tentukan sentimen setiap ulasan
14 data['Sentiment'] = data['stemming_word'].apply(determine_sentiment)
15 data.head()
    
```

Gambar 7 Codingan Proses Pelabelan Menggunakan InSet *Lexicon Based*

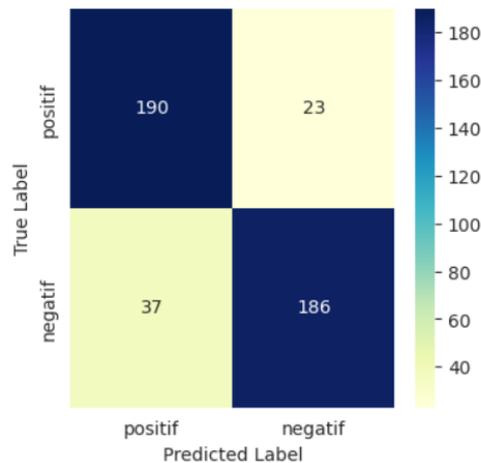
Berdasarkan hasil dari pelabelan, dilakukan klasifikasi menggunakan Algoritma Naïve Bayes, data dibagi menjadi 20% (436 data) data uji dan 80% (1742 data) data latih. Pelatihan pada dataset akan membentuk sebuah pola atau model yang akan diterapkan pada data testing. Setelah model telah dipelajari maka proses berikutnya adalah penerapan model Algoritma Naïve Bayes untuk mengklasifikasikan data uji guna menentukan probabilitas sentimen. Hasilnya dapat dilihat pada gambar 8.

Accuracy: 0.8623853211009175				
	precision	recall	f1-score	support
Negative	0.84	0.89	0.86	213
Positive	0.89	0.83	0.86	223
accuracy			0.86	436
macro avg	0.86	0.86	0.86	436
weighted avg	0.86	0.86	0.86	436

Gambar 8 Hasil klasifikasi

Dari gambar di atas, telah didapatkan hasil dari klasifikasi menggunakan Algoritma Naïve Bayes dengan akurasi 86% yang menunjukkan bahwa model dapat mengklasifikasikan *tweet* dengan benar 86% dari seluruh data uji. Presisi 89% yang berarti dari semua *tweet* yang diprediksi sebagai positif, 89% di antaranya memang benar-benar positif. *Recall* 83% menunjukkan bahwa dari semua *tweet* yang sebenarnya positif, model dapat mengidentifikasi 83% di antaranya dengan benar. Dan yang terakhir *f1-score* 86% menunjukkan keseimbangan antara kemampuan model dalam mengidentifikasi *tweet* positif dengan tepat dan menangkap semua *tweet* yang benar-benar positif. Hasil yang telah diperoleh menunjukkan bahwa Model Naïve Bayes bekerja dengan baik dalam mengklasifikasikan sentimen *tweet*. Keseimbangan antara presisi dan *recall* yang baik dan juga akurasi dan *F1-score* yang tinggi menunjukkan bahwa model ini dapat digunakan secara efektif. Dalam penelitian ini, *confusion matrix* juga digunakan sebagai alat evaluasi kinerja model.

Confusion matrix adalah alat pengujian yang digunakan untuk mengukur tingkat kinerja atau keakuratan suatu proses klasifikasi. Sehingga dapat dianalisa seberapa baik pengklasifikasi dapat mengidentifikasi data dari berbagai kelas[20]. Hasil yang diperoleh adalah, True Positive (TP): 190, menunjukkan jumlah *tweet* positif yang benar-benar diklasifikasikan sebagai positif. False Positive (FP): 37, menunjukkan jumlah *tweet* negatif yang salah diklasifikasikan sebagai positif. True Negative (TN): 186, menunjukkan jumlah *tweet* negatif yang benar-benar diklasifikasikan sebagai negatif. dan yang terakhir False Negative (FN): 23, menunjukkan jumlah *tweet* positif yang salah diklasifikasikan sebagai negatif. hasil ini dapat digunakan untuk mengevaluasi kinerja model secara keseluruhan. True Positives dan True Negatives menunjukkan keakuratan model dalam klasifikasi yang benar, sementara False Positives dan False Negatives menunjukkan kesalahan klasifikasi yang perlu diminimalkan. Elemen-elemen ini dapat digunakan untuk menghitung metrik seperti akurasi, presisi, *recall*, dan *F1-score*, dan memberikan gambaran lengkap tentang seberapa baik model dalam memprediksi sentimen *tweet*. Hasil *confusion matrix* seperti terlihat pada gambar 9.



Gambar 9 Confusion Matrix

IV. KESIMPULAN

Meskipun mengalami keterbatasan kosakata dalam kamus InSet Lexicon dan 788 data gagal diidentifikasi kelasnya, penelitian ini tetap berhasil mencapai tingkat akurasi yang baik, yakni sebesar 86%. Evaluasi kinerja model menunjukkan hasil memuaskan dengan berhasil mengidentifikasi 213 sentimen negatif dan 223 sentimen positif dari total 436 data uji. Keseimbangan evaluasi model menunjukkan kemampuan yang baik dalam mengenali kedua kelas sentimen dengan presisi yang tinggi. Hasil penelitian ini juga menunjukkan potensi pengembangan lebih lanjut, termasuk eksplorasi fitur lebih mendalam untuk meningkatkan kinerja model, mempertimbangkan penggunaan algoritma klasifikasi lainnya untuk perbandingan kinerja dengan Naïve Bayes, serta memperluas cakupan sentimen yang diprediksi atau menggunakan dataset yang lebih besar untuk eksplorasi lebih lanjut. Untuk penelitian mendatang, disarankan untuk menggunakan kamus *lexicon* yang lebih komprehensif guna mengurangi jumlah data yang tidak terklasifikasi. Selain itu, teknik *pre-processing* yang lebih mutakhir dapat meningkatkan kualitas data. Penerapan metode klasifikasi alternatif seperti Random Forest atau teknik Deep Learning untuk menyajikan perbandingan kinerja yang lebih baik. Dan yang terakhir memperluas cakupan sentimen yang diprediksi atau menggunakan dataset yang lebih besar sehingga dapat meningkatkan akurasi model dalam analisis sentimen.

DAFTAR PUSTAKA

- [1] C. N. Zempi, A. Kuswanti, and S. Maryam, "Analisis Peran Media Sosial Dalam Pembentukan Pengetahuan Politik Masyarakat," *Eksprei Dan Persepsi J. Ilmu Komun.*, vol. 6, no. 1, pp. 116–123, 2023, doi: 10.33822/jep.v6i1.5286.
- [2] F. M. Tamim, "Strategi Komunikasi Politik Partai Gerindra dalam Meningkatkan Partisipasi Politik Pemilih Muda melalui Media Sosial Twitter pada Pemilu Tahun 2024," *JIIP - J. Ilm. Ilmu Pendidik.*, vol. 6, no. 10, pp. 8040–8046, 2023, doi: 10.54371/jiip.v6i10.3028.
- [3] A. S. Alim and D. E. Rahmawati, "Komunikasi Politik Anies Baswedan Melalui Sosial Media Twitter," *J. Acad. Praja*, vol. 4, no. 2, pp. 441–453, 2021, doi: 10.36859/jap.v4i2.334.
- [4] N. Faradis, N. Ainya, A. Fauzah, M. Ichsan, and A. Anshori, "Media Sosial dan Persepsi Publik: Analisis Strategi Kampanye Digital Calon Presiden Indonesia 2024," *Pros. Semin. Nas.*, pp. 643–652, 2023.
- [5] B. Mas Pintoko and K. Muslim, "Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naïve Bayes Classifier," *e-Proceeding Eng.*, vol. 5, no. 3, pp. 8121–8130, 2018.
- [6] W. Kustiawan, N. W. Kartika, C. I. Kesuma, F. A. Silalahi, and A. M. Nasution, "Kampanye dalam Pemasaran Politik dan Brand Politik," *J. Ilmu Komputer, Ekon. dan Manaj.*, vol. 2, no. 1, pp. 959–964, 2022.
- [7] M. I. Dwitama, F. A. Hakiki, E. Sulastris, U. Usni, and D. Gunanto, "Media Sosial Dan Pengaruhnya Terhadap Partisipasi Politik Masyarakat Di Pilkada 2020 Tangerang Selatan," *Indep. J. Polit. Indones. dan Glob.*, vol. 3, no. 1, p. 53, 2022, doi: 10.24853/independen.3.1.53-66.
- [8] S. Puad, G. Garno, and A. Susilo Yuda Irawan, "Analisis Sentimen Masyarakat Pada Twitter Terhadap Pemilihan Umum 2024 Menggunakan Algoritma Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1560–1566, 2023, doi: 10.36040/jati.v7i3.6920.
- [9] T. D. Putra, E. Utami, and M. P. Kurniawan, "Analisis Sentimen Pemilu 2024 dengan Naive Bayes Berbasis Particle Swarm Optimization (PSO)," *Explore*, vol. 13, no. 1, pp. 1–5, 2023, doi: 10.35200/ex.v11i2.13.
- [10] A. P. Putra, A. N. Alivia, M. Meilani, N. J. Azzahra, and N. A. Rakhmawati, "Analisis Sentimen Warga Twitter Terhadap Game Shopee Cocoki Dengan Metode Naive Bayes Classifier," *J. Ilm. Inform. Komput.*, vol. 28, no. 2, pp. 137–148, 2023, doi: 10.35760/ik.2023.v28i2.9494.
- [11] I. Oktavia and A. R. Isnain, "Analisis Sentimen Opini Terhadap Tools Artificial Intelligence (AI) Berdasarkan Twitter Menggunakan Algoritma Naïve Bayes," vol. 8, no. April, pp. 777–787, 2024, doi: 10.30865/mib.v8i2.7524.
- [12] Liza Wikarsa, "Paper_3_15-24," vol. 18, no. 1, pp. 15–24, 2022.
- [13] A. R. Ismail and Raden Bagus Fajriya Hakim, "Implementasi Lexicon Based Untuk Analisis Sentimen Dalam Menentukan Rekomendasi Pantai Di DI Yogyakarta Berdasarkan Data Twitter," *Emerg. Stat. Data Sci. J.*, vol. 1, no. 1, pp. 37–46, 2023, doi: 10.20885/esds.vol1.iss.1.art5.
- [14] V. Karenina, M. F. Erinsyah, and D. S. Wibowo, "Klasifikasi Rentang Usia Dan Gender Dengan Deteksi Suara Menggunakan Metode Deep Learning Algoritma Cnn (Convolutional Neural Network)," *Komputika J. Sist. Komput.*, vol. 12, no. 2, pp. 75–82, 2023, doi: 10.34010/komputika.v12i2.10516.
- [15] F. Nur Fajri, A. Tholib, and W. Yuliana, "Application of Machine Learning Algorithm for Determining Elective Courses in Informatics Study Program," *J. Tek. Inform. dan Sist. Inf.*, vol. 8, no. 3, pp. 485–496, 2022, doi: 10.28932/jutisi.v8i3.3990.
- [16] D. A. Punkastyo, F. Septian, and A. Syaripudin, "Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Prediksi Kelulusan Siswa," *J. Syst. Comput. Eng.*, vol. 5, no. 1, pp. 24–35, 2024, doi: 10.61628/jsce.v5i1.1073.

- [17] L. Indriyani and W. Susanto, “Analisis Penerapan Naïve Bayes Untuk Memprediksi Resiko Kredit Anggota Koperasi Keluarga Guru,” *J. Inform.*, vol. 6, no. 2, pp. 262–270, 2019, doi: 10.31311/ji.v6i2.5724.
- [18] D. F. Rahman, “Analisis Chatgpt Tweet Menggunakan Eda Dan Sentiment Analysis: Data Pengguna Twitter Di Indonesia,” no. September, 2023, [Online]. Available: <https://www.researchgate.net/publication/376356578>
- [19] Y. Nooryuda Prasetya, D. Winarso, and Syahril, “Penerapan Lexicon Based Untuk Analisis Sentimen Pada Twiter Terhadap Isu Covid-19,” *J. Fasilkom*, vol. 11, no. 2, pp. 97–103, 2021.
- [20] S. Anggina, N. Y. Setiawan, and F. A. Bachtiar, “Analisis Ulasan Pelanggan Menggunakan Multinomial Naïve Bayes Classifier dengan Lexicon-Based dan TF-IDF Pada Formaggio Coffee and Resto,” *is Best Account. Inf. Syst. Inf. Technol. Bus. Enterp. this is link OJS us*, vol. 7, no. 1, pp. 76–90, 2022, doi: 10.34010/aisthebest.v7i1.7072.