

**CESS**

**(Journal of Computer Engineering, System and Science)**

Available online: <https://jurnal.unimed.ac.id/2012/index.php/cess>

ISSN: 2502-714x (Print) | ISSN: 2502-7131 (Online)



## **Perbandingan Analisis Sentimen Presiden 2024 Menggunakan Algoritma Support Vector Machine dan K-Nearest Neighbor**

### **Comparison of Sentiment Analysis for the 2024 Presidential Election Using Support Vector Machine and K-Nearest Neighbor Algorithms**

**Muhammad Hanafi<sup>1\*</sup>, Mhd.Furqan<sup>2</sup>**

<sup>1,2</sup> Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera  
Utara Medan

Email: <sup>1</sup>[hanafimuhammad009@gmail.com](mailto:hanafimuhammad009@gmail.com), <sup>2</sup>[mfurqan@uinsu.ac.id](mailto:mfurqan@uinsu.ac.id)

*\*Corresponding Author*

#### **ABSTRAK**

Di era digital, media sosial menjadi wadah utama bagi masyarakat untuk menyampaikan opini mereka terhadap berbagai isu, termasuk Pemilihan Presiden 2024 di Indonesia. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap pasangan calon "Prabowo-Gibran" berdasarkan 944 tweet yang dikumpulkan selama periode Maret hingga Mei 2024. Metode klasifikasi sentimen yang digunakan dalam penelitian ini adalah K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM), dengan tujuan untuk membandingkan tingkat akurasi kedua algoritma tersebut dalam mengklasifikasikan sentimen publik. Dataset yang digunakan dibagi menjadi dua bagian, yaitu 80% untuk pelatihan dan 20% untuk pengujian. Model KNN diterapkan dengan jumlah tetangga terdekat sebanyak lima ( $k=5$ ) menggunakan `KNeighborsClassifier(n_neighbors=5)`, sedangkan model SVM menggunakan kernel linear untuk memisahkan data sentimen. Proses analisis dilakukan menggunakan Python dan Google Colab, mencakup tahapan seperti pelabelan data, preprocessing teks, dan ekstraksi fitur. Evaluasi model dilakukan menggunakan Confusion Matrix, yang mengukur akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma SVM memiliki tingkat akurasi sebesar 52%, sedangkan KNN hanya mencapai akurasi 51% berdasarkan 189 sampel data uji. Temuan ini mengindikasikan bahwa SVM lebih efektif dibandingkan KNN dalam mengklasifikasikan sentimen publik terkait Pemilihan Presiden 2024 di Indonesia. Meskipun demikian, akurasi yang diperoleh masih tergolong rendah, sehingga penelitian lebih lanjut diperlukan untuk meningkatkan performa model, misalnya dengan optimasi parameter, peningkatan kualitas dataset, atau penerapan teknik machine learning yang lebih canggih.

**Kata Kunci:** Analisis Sentimen; Support Vector Machine; K-Nearest Neighbor; Presiden 2024



## ABSTRACT

In the digital era, social media has become the primary platform for people to express their opinions on various issues, including the 2024 Indonesian Presidential Election. This study aims to analyze public sentiment toward the candidate pair "Prabowo-Gibran" based on 944 tweets collected between March and May 2024. The sentiment classification methods used in this research are K-Nearest Neighbor (KNN) and Support Vector Machine (SVM), with the goal of comparing the accuracy of these two algorithms in classifying public sentiment. The dataset is divided into 80% for training and 20% for testing. The KNN model is implemented with five nearest neighbors ( $k=5$ ) using `KNeighborsClassifier(n_neighbors=5)`, while the SVM model applies a linear kernel to separate sentiment data. The analysis process is conducted using Python and Google Colab, covering stages such as data labeling, text preprocessing, and feature extraction. Model evaluation is performed using a Confusion Matrix, which measures accuracy, precision, recall, and F1-score. The results indicate that the SVM algorithm achieves an accuracy of 52%, while KNN only reaches 51% based on 189 test samples. These findings suggest that SVM is more effective than KNN in classifying public sentiment regarding the 2024 Indonesian Presidential Election. However, the obtained accuracy remains relatively low, indicating the need for further research to enhance model performance. Possible improvements include parameter optimization, dataset quality enhancement, or the implementation of more advanced machine learning techniques.

**Keywords:** *Sentiment Analysis; Support Vector Machine; K-Nearest Neighbor; 2024 Presidential*

---

## 1. PENDAHULUAN

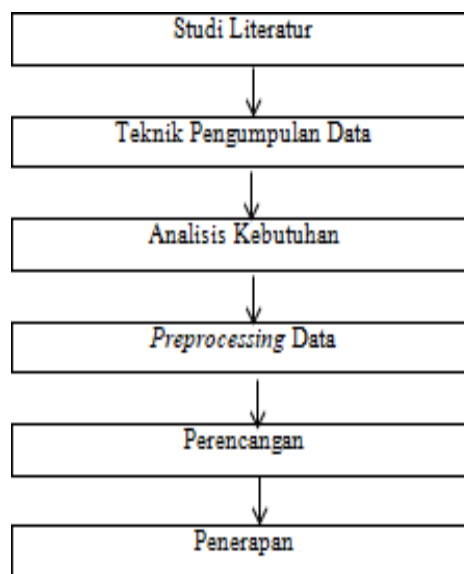
Setiap lima tahun sekali, sistem demokrasi Indonesia dihidupkan kembali melalui pemilihan umum atau "Pemilu," sebuah peristiwa penting yang membentuk tatanan politik negara. Latihan ekstensif ini melibatkan partai politik yang menampilkan kandidat, kampanye komprehensif, registrasi pemilih dan akhirnya, proses pemungutan suara itu sendiri. Sentimen positif dan dimungkinkan untuk memproses karakteristik negatif suatu opini secara manual, namun mengkategorikan polaritas suatu opini jelas akan memerlukan lebih banyak waktu dan upaya jika semakin banyak sumber opini. Pada penelitian sebelumnya yang dilakukan Susanto Perkembangan teknologi informasi mendorong inovasi dalam pelayanan publik. "Polri Super App" menjadi contoh sukses penerapan *e-government* di kepolisian, memanfaatkan teknologi digital untuk meningkatkan layanan. Tujuan dari penelitian ini adalah untuk mengevaluasi seberapa baik algoritma KNN dan SVM melakukan analisis sentimen pada "Polri Super App" yang tersedia di Google Play. Penelitian sebelumnya telah menunjukkan bahwa SVM dan Naive Bayes bekerja dengan baik pada tugas yang sebanding. Menurut hasil pengujian, SVM menggunakan kernel linear mencapai akurasi maksimum 89,67%. Selain itu, SVM dengan kernel RBF dan KNN menghasilkan hasil yang positif [1]. Gambaran yang lebih jelas tentang kinerja deteksi sentimen algoritma diberikan oleh penyelidikan positif dan negatif dari ulasan pengguna terhadap aplikasi tersebut [2]. Analisis sentimen adalah metode yang digunakan untuk mengevaluasi perasaan dan pendapat dalam teks yang diunggah oleh pengguna media sosial. Dalam kasus ini, analisis sentimen mengenai kandidat presiden 2024 dapat memberikan informasi berharga tentang bagaimana masyarakat merespons setiap calon. Untuk melakukan analisis sentimen, berbagai algoritma dapat digunakan, dengan dua

algoritma yang umum adalah *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (K-NN) [3][4].

Maka dari di usulkan Menggunakan teknik pembelajaran mesin untuk mengkategorikan polaritas opini dari berbagai sumber data sambil memanfaatkan keuntungan penambangan teks [7-8]. Di media sosial adalah platform utama bagi orang-orang untuk mengekspresikan pikiran dan ide mereka tentang berbagai topik, termasuk politik, di era digital saat ini. isu penting yang sering dibahas adalah pemilihan presiden, seperti Pemilihan Presiden 2024 di Indonesia. Opini publik tentang kandidat presiden dapat memengaruhi persepsi masyarakat dan hasil pemilihan, sehingga analisis sentimen publik menjadi sangat penting [6].

## 2. METODE PENELITIAN

Kerangka penelitian ini dapat dilihat terdiri dari Studi Literatur, Teknik Pengumpulan Data, Analisis Kebutuhan, Preprocessing Data, Perancangan dan Penerapan.



Gambar 2. Kerangka Penelitian

### 2.1 Studi Literatur

Dalam penelitian ini dilakukan pencarian dan pengumpulan dengan mencari jurnal, ebook, browsing internet, artikel dan lainnya yang relevan dengan topik penelitian serta observasi dilakukan untuk mengamati permasalahan yang ada.

### 2.2 Teknik Pengumpulan Data

Penelitian ini memerlukan data tentang Sentimen dengan kata kunci Prabowo-Gibran jumlah data yang diambil 944 tweet dengan cara melakukan *crawling data* menggunakan bahasa *Python* merupakan salah satu *library python* dan mudah untuk mengakses *API* dari X.

### 2.3 Analisis Kebutuhan

Dalam melakukan penelitian ini dibutuhkan beberapa analisis kebutuhan agar dapat terlaksana dengan baik. Dalam penelitian ini, data presiden 2024 menjadi focus utama. Data yang diperlukan meliputi data yang diambil bulan Maret hingga bulan Mei 2024, dengan kata kunci Prabowo-Gibran. Penerapan metode Algoritma K-Nearest Neighbor dan Support Vector

Machine dalam suatu sistem untuk melakukan analisis perbandingan mengklasifikasikan presiden 2024

### 2.4 Preprocessing Data

Setelah pengumpulan data dan analisis kebutuhan. Selanjutnya dilakukan tahap Preprocessing data melalui Google Colab dengan Python. Berikut alur yang menjelaskan tahapan sistem yang sedang melakukan proses preprocessing meliputi cleaning, stopword removal, filtering, dan tokenizing.

### 2.5 Perancangan

Dengan istilah analisis sentimen presiden 2024 menggunakan bahasa Python aplikasi Google Colab dengan Penerapan Algoritma K-Nearest Neighbor dan Support Vector Machine. klasifikasi dilakukan model perbandingan analisis sentimen melalui tahap pelabelan data, *text preprocessing*, pembagian data latih dan data uji.

### 2.6 Penerapan

*Support Vector Machine* adalah algoritma klasifikasi yang bertujuan menemukan hyperplane dengan margin terbesar untuk memisahkan dua kumpulan data secara optimal. Meskipun awalnya dirancang untuk klasifikasi data numerik, *Support Vector Machine* juga efektif dan cepat dalam menangani masalah data teks [1]. Data teks cocok untuk *Support Vector Machine* karena memiliki dimensi tinggi, dengan fitur-fitur yang sering kali tidak relevan tetapi saling berkorelasi, serta biasanya dapat dipisahkan secara linear [7] [8]. Istilah ini beroperasi pada premis sederhana untuk menentukan jalur terpendek antara sampel pelatihan dan sampel pengujian. Namun, biaya komputasi K-Nearest Neighbor dan kerentanan terhadap gangguan dan input yang tidak tepat merupakan kekurangannya [9] [10] [11].

## 3. HASIL DAN PEMBAHASAN

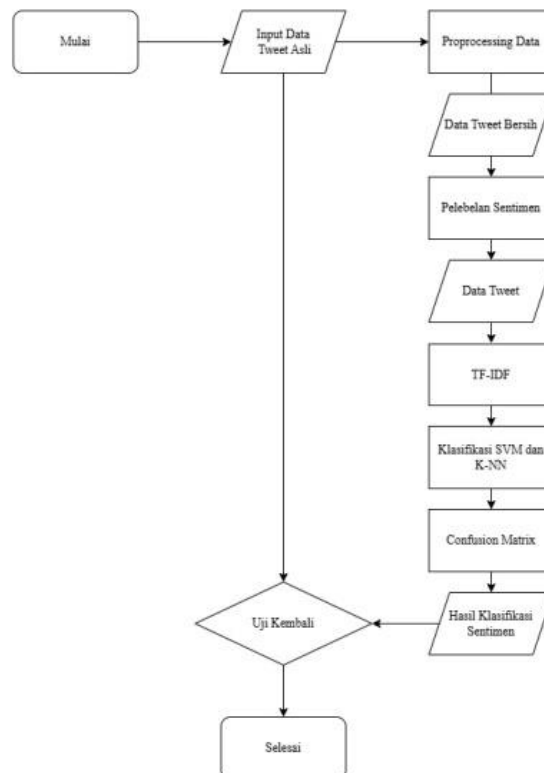
Dalam penelitian ini, analisis kebutuhan penting untuk kelancaran proses. Fokus utama adalah data presiden 2024 yang diambil dari Maret hingga Mei 2024 dengan kata kunci "Prabowo-Gibran." Metode ini untuk menganalisis serta membandingkan klasifikasi. Penelitian memerlukan 944 tweet tentang sentimen yang diperoleh melalui pengambilan data menggunakan Python. Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Model *K-Nearest Neighbors* tetangga terdekat (k) sebanyak 5, sedangkan model *Support Vector Machine* menggunakan *kernel linear* dan membandingkan akurasi kedua algoritma.

Tabel 1. Tampilan *Crawling* Data

Created_At	Full_Text	Username
Tue Apr 30	@BANGSAygSUJUD @jokowi @prabowo @gibran_tweet Mungkin loe lupa yang yang menjual Freeport ke asing dan Kita cuma dapet 9% itu mertuanya Prabowo	Toni7216
Tue Apr 30	@Jatayu_45back Aneh smua pasangan yg brkpetisi dan para kuasa hukumnya sdh mnerima kputusan MK bahkan anies & ganjar sdh ngucap selamat pd prabowo-gibran kok ini msh ada cecunguk dungu yg msh pnasaran ! Jokowi beres tugas di oktober tinggal brp bln l	coxhaez

Tue Apr 30	Prabowo Gibran pastikan langsung kerja seusai dilantik Presiden 20 Oktober <a href="https://t.co/q3iXeipime">https://t.co/q3iXeipime</a>	TmAmrul
Tue Apr 30	Hukum Mati Koruptor @jokowi @prabowo @gibran_tweet #DSM Saatnya Hukum Mati Buat Para Koruptor <a href="https://t.co/1LTPRmEehY">https://t.co/1LTPRmEehY</a>	EdySularno96489
Tue Apr 30	@BANGSAygSUJUD @jokowi @prabowo @gibran_tweet Hidup Jokowi.... Preet	ike_hw71
Tue Apr 30	<a href="https://t.co/q1ZryorAgR">https://t.co/q1ZryorAgR</a> Etika Hukum dan Masa Depan Demokrasi Politik: Evaluasi dan Refleksi Keputusan Mahkamah Konstitusi tentang Pilpres 2024 #pilpres2024 #MK #prabowo #prabowosubianto #prabowogibran	Jurnalpatroli_
Tue Apr 30	Lapor Pak Lurah ; Masyarakat Kita Pada Susah Suruh Pasang Foto Presiden & Wakil Presiden Baru. Pak Lurah ; Yg Tidak Mau Pasang Foto Prabowo-Gibran Coret Dari Daftar Penerima Bansos. <a href="https://t.co/46MvNFFVHM">https://t.co/46MvNFFVHM</a>	S4N_W1B1

Setelah pengumpulan data dan analisis kebutuhan. Selanjutnya dilakukan tahap *Preprocessing* meliputi *cleaning*, *Stop Removal*, *normalisasi*, *filtering* dan *tokenizing* data melalui *Google Colab* dengan *Python* [3]. Berikut alur yang menjelaskan tahapan sistem:



Gambar 2. Flowchart Sistem

Dataset dilabeli menjadi dua kategori: positif dan negatif, untuk mengidentifikasi sentimen tweet. Setelah itu, dataset dapat digunakan untuk model klasifikasi. Berikut contoh pelabelan:

Tabel 2. Pelabelan *Dataset*

<i>Dataset</i>	Kelas
BANGSA ygSUJUD jokowi prabowo gibran_ Mungkin loe lupa yang yang menjual Freeport ke asing dan Kita cuma dapet 9% itu mertuanya Prabowo	Netral
Prediksi Grab live pools Pastikan anda melihat dari situs acuan yang live xD35qQ1sYN #grablive #togeljitu #togelhariini #togeljitu #vincent #Putting #Tornado #PrabowoGibran	Positif
Prabowo kunjungi NasDem dan PKB artinya sudah pasti menjadi partai pendukung pemerintahan Saksikan Special Program #PanggungDemokrasi di Channel Youtube Metrotvnews KnnurAnvEz #WihadiWiyanto #Otakatik #Koalisi #Prabowo #Gibran #Pilpres2024 #Pemilu2024 eFSYYPrr0Z	Positif
Prabowo Gibran Gimana tanggapan kalean lur tulis di kolom komentar ya! #prabowo #pks #gibran #pemilu2024 suMms2cxny	Netral
ch_chotimah2 prabowo gibran_ ICPolri DivHumas_Polri Playing victim biar dikasihani lowbatt 16% semua ini adalah akibat dari keangkuhan & arogansi sebut saja angka gak usah takuuut!! gLm6Yh5Jza	Negatif
Dari fenomena PRD jadi PRIMA lalu kemudian jadi tim kampanye Prabowo Gibran aku ga kaget tiba tiba EXCOPARTAIBURUH ikut arus yang sama Tapi kecewa dikit boleh dong Besok besok awas aja u min kalo tiba tiba nge Hanya kaum buruh yang bisa mengubah nasibnya sendiri	Netral
atau oalah pantes anak Abah kenapa gak muji muji Prabowo Gibran nyebarluasin sisi positif mereka KdEJ7Rdl17	Netral
Bambangmulyonoo BANGSAygSUJUD jokowi prabowo gibran_ Pernah cek berapa hutang konglomerat di berbagai perusahaan top dunia belum Pernah cek berapa hutang negara jepang blmyg katanya negara maju Kalau belum beli kuota Sono	Positif
IDM mutlak dibutuhkan dlm proses pembangunan DESA #IDMKemendesBisa jokowi prabowo gibran_ halimiskandarnu Kemendespdtt malik_haramain imansyukri taufikmadjid71 yusradaridesa azzameldzikrie tppkemendes wzDWJ4M5Jc	Netral
cakimiNOW Selamat berjuang untuk presiden terpilih bpk Prabowo Gibran semngat menuju Indonesia emas timnas perjuangan sudah usai mungkin bisa berkontribusi atau ckup jadi penasehat/pengamat saja	Netral

### 3.1 Pelabelan Data

Untuk pelabelan data dengan pendekatan Pada tahap pelabelan data menggunakan *TextBlob*, pertama-tama, pustaka *pandas* dan *TextBlob* diimpor [12]. Data dimuat dari file CSV *hasil\_text\_processing.csv*, dan program memeriksa apakah DataFrame kosong. Kolom *full\_text* diambil, nilai *NaN* dihapus, dan disimpan dalam daftar data *tweet*. Setiap tweet dianalisis untuk mendapatkan nilai polaritas menggunakan *TextBlob* [13]. Status dikategorikan: 'Positif' jika polaritas > 0, 'Negatif' jika polaritas < 0, dan status netral diabaikan. Hasil akhir disimpan dalam DataFrame yang hanya mencakup tweet dengan status 'Positif' dan 'Negatif'. Dari total 944 data, terdapat 38 positif, 889 netral, dan 17 negatif.

### 3.2 Pemodelan Data

Tahap ini bertujuan untuk menghasilkan nilai akurasi dengan melatih model pada data yang memiliki label positif dan negatif.

	term	idf	tf_0	tf_1	tf_2	tf_3	tf_4	tf_5	tf_6	tf_7	...	tf_934	tf_935	tf_936	tf_937	tf_938	tf_939	tf_940	tf_941	tf_942	tf_943	
0	00	6.156979	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
1	01	4.365220	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	02	3.959754	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
3	03	4.452231	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
4	03_nakula	6.156979	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
5293	Sampah	6.156979	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
5294	Selamatkan	6.156979	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
5295	Uang	6.156979	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
5296	Yang	6.156979	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
5297	dari	6.156979	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	

5298 rows x 946 columns

Gambar 2. Hasil TF-IDF

Untuk pemodelan data menggunakan metode klasifikasi *Support Vector Machine* dan *K-Nearest Neighbors*. Pada tahap analisis teks menggunakan *TF-IDF*, pustaka yang diperlukan diimpor, termasuk *pandas*, *TfidfVectorizer* dari *sklearn*, dan *numpy* [14]. Data dimuat dari file CSV *hasil\_text\_processing.csv* ke dalam DataFrame, dan nilai hilang dalam kolom *full\_text* diganti dengan string kosong. Objek *TfidfVectorizer* diinisialisasi, dan matriks *TF-IDF* dihitung dengan *fit\_transform()*. Daftar istilah diambil dengan *get\_feature\_names\_out()*, dan nilai *Inverse Document Frequency (IDF)* dihitung dengan rumus  $IDF(t) = \log(N/df(t))$  [15]. DataFrame *tfidf\_df* dibuat untuk menyimpan istilah dan *IDF*, lalu dilakukan iterasi untuk menghitung *Term Frequency (TF)* untuk setiap dokumen dan menambahkan kolom baru ke DataFrame dengan nama *tf\_{i}*.

```

X = df.drop(columns = ['full_text'])
y = df['full_text']

print("X : ", X.shape)
print("y : ", y.shape)

X : (944, 10)
y : (944,)

[ ] x_train, x_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 42)

[ ]
print(f"x_train : {x_train.shape}")
print(f"y_train : {y_train.shape}")
print(f"x_test : {x_test.shape}")
print(f"y_test : {y_test.shape}")

x_train : (755, 10)
y_train : (755,)
x_test : (189, 10)
y_test : (189,)
    
```

Gambar 3. Data latihan dan uji dengan proporsi 80% dan 20% menggunakan *train\_test\_split*

Pada tahap persiapan data, perpustakaan yang diperlukan pertama kali diimpor, termasuk *pandas* untuk manipulasi data, *numpy* untuk operasi numerik, dan *train\_test\_split* dari *sklearn model\_selection* untuk membagi dataset. Data dimuat dari file CSV *hasil\_text\_processing.csv* ke dalam DataFrame *df* menggunakan *pd.read\_csv()*. Informasi mengenai DataFrame ditampilkan dengan mencetak bentuk, informasi kolom, statistik deskriptif, frekuensi nilai unik dalam kolom *full\_text*, dan beberapa baris pertama. Selanjutnya, objek *LabelEncoder* diinisialisasi untuk mengubah kolom *full\_text* menjadi

representasi numerik. Setelah *encoding*, kolom *full\_text* dihapus dari *DataFrame* dan disimpan dalam variabel X sebagai fitur, sementara kolom yang telah di-encode disimpan dalam variabel Y sebagai label. Bentuk dari X dan Y dicetak, lalu data dibagi menjadi data latih dan uji dengan proporsi 80% dan 20% menggunakan *train\_test\_split()*, dan hasilnya disimpan dalam variabel *x\_train*, *x\_test*, *y\_train*, dan *y\_test*. Bagian ini mencetak bentuk dataset: Data Latih (X\_train dan y\_train) 755 sampel Data Uji (X\_test dan y\_test) 189 sampel.

### 3.3 Klasifikasi

Confusion Matrix adalah dataset digunakan untuk menganalisis lainnya sebagai negative [10-13] [15-16] [18].

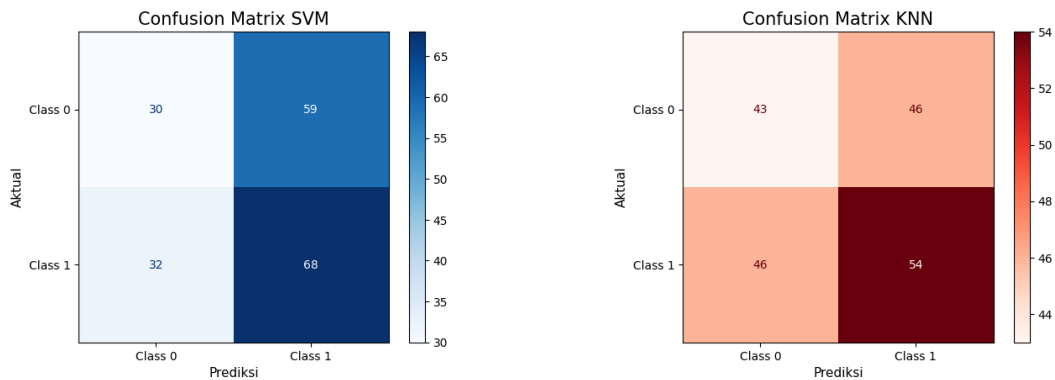
$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{4}$$

Berikut hasil *Confusion Matrix Support Vector Machine* dan *K-Nearest Neighbors* :



Gambar 4. *Confusion Matrix Support Vector Machine* dan *K-Nearest Neighbors*

*Support Vector Machine* dan *K-Nearest Neighbors* digunakan untuk memprediksi kelas sampel 189 dalam dataset Data Uji (X\_test dan y\_test). Hasil prediksi dari kedua model dicetak untuk membandingkan dan mengevaluasi kinerja mereka.



```

=== Laporan Klasifikasi SVM ===
      precision    recall  f1-score   support

   0         0.48         0.34         0.40         89
   1         0.54         0.68         0.60        100

 accuracy                   0.52         189
 macro avg                   0.51         189
 weighted avg                 0.51         189
    
```

Gambar 5. Klasifikasi *Support Vector Machine*

```

=== Laporan Klasifikasi KNN ===
      precision    recall  f1-score   support

   0         0.48         0.48         0.48         89
   1         0.54         0.54         0.54        100

 accuracy                   0.51         189
 macro avg                   0.51         189
 weighted avg                 0.51         189
    
```

Gambar 6. Klasifikasi *K-Nearest Neighbor*

Hasil klasifikasi menggunakan algoritma KNN dan SVM. Model KNN menunjukkan hasil precision, recall, dan f1-score yang seimbang, yaitu 0.51, dengan support 189. Sedangkan model SVM menunjukkan hasil f1-score yang lebih tinggi pada kelas 1 (0.60), namun recall pada kelas 0 lebih rendah (0.34). Secara keseluruhan, model SVM memiliki accuracy 0.52 dengan support 189.

### 3.4 Evaluasi Model

Hasil evaluasi perbandingan performa dua metode algoritma, khususnya K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM), yang dinilai berdasarkan metrik akurasi. Berdasarkan data yang ditampilkan, algoritma SVM memiliki akurasi sebesar 52%, sedikit lebih tinggi dibandingkan algoritma KNN yang memiliki akurasi sebesar 51%. Perbedaan nilai akurasi yang kecil ini menunjukkan bahwa kedua algoritma memiliki kemampuan klasifikasi yang hampir serupa pada dataset yang digunakan. Namun, secara keseluruhan, metode SVM menunjukkan performa yang lebih unggul dibandingkan KNN dalam konteks pengujian ini. Penilaian akurasi ini penting untuk memahami efektivitas masing-masing algoritma dalam menyelesaikan tugas klasifikasi.

Tabel 3. Perbandingan Metode *Support Vector Machine* dan *K-Nearest Neighbor*

Ukuran Evaluasi	<i>Support Vector Machine</i>	<i>K-Nearest Neighbor</i>
Akurasi	52%	51%

## 4. KESIMPULAN

Penelitian ini menganalisis sentimen publik terhadap Pemilihan Presiden 2024 di Indonesia menggunakan dua algoritma klasifikasi, yaitu Support Vector Machine (SVM) dan K-Nearest Neighbor (KNN). Berdasarkan hasil analisis, SVM menunjukkan akurasi yang lebih

tinggi, yaitu 52%, dibandingkan dengan KNN yang hanya mencapai 51% dari 189 sampel uji. Hal ini menunjukkan bahwa SVM lebih efektif dalam mengklasifikasikan sentimen publik dibandingkan KNN, terutama dalam menangani data berdimensi tinggi dan kompleks. Proses klasifikasi dilakukan melalui beberapa tahapan utama, yaitu pengumpulan data, pelabelan sentimen, preprocessing teks, pembagian data latih dan uji, serta evaluasi model menggunakan Confusion Matrix. Dataset yang digunakan terdiri dari 944 tweet yang dikumpulkan antara Maret hingga Mei 2024, dengan proporsi 80% untuk pelatihan dan 20% untuk pengujian. Model KNN diterapkan dengan jumlah lima tetangga terdekat ( $k=5$ ), sementara SVM menggunakan kernel linear untuk memisahkan data sentimen secara optimal. Meskipun SVM memiliki akurasi yang lebih baik dibandingkan KNN, hasil yang diperoleh masih tergolong rendah. Hal ini menunjukkan bahwa kedua algoritma masih memiliki keterbatasan dalam mendeteksi sentimen publik terkait Pemilihan Presiden 2024. Oleh karena itu, diperlukan penelitian lebih lanjut untuk meningkatkan performa model, seperti menggunakan teknik optimasi hyperparameter, data balancing, serta menerapkan algoritma machine learning yang lebih kompleks atau pendekatan hybrid. Penelitian ini memberikan wawasan yang berharga dalam penerapan machine learning untuk analisis sentimen politik, yang dapat digunakan sebagai dasar dalam memahami persepsi publik terhadap suatu isu. Ke depan, pendekatan yang lebih canggih dan dataset yang lebih luas diharapkan dapat meningkatkan akurasi prediksi serta memberikan hasil analisis yang lebih akurat dan mendalam.

#### DAFTAR PUSTAKA

- [1] M. I. Petiwi, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Gofood Berdasarkan Twitter Menggunakan Metode Naïve Bayes dan Support Vector Machine," *Jurnal Media Informatika Budidarma*, vol. 6, no. 1, p. 542, 2022, doi: 10.30865/mib.v6i1.3530.
- [2] L. A. Susanto, "Komparasi Model Support Vector Machine Dan K-Nearest Neighbor Pada Analisis Sentimen Aplikasi Polri Super App," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4152.
- [3] W. P. Ali, "Analisis Sentimen Masyarakat Terhadap Kinerja Presiden Indonesia Dalam Aspek Ekonomi, Kesehatan, dan Pembangunan Berdasarkan Opini dari Twitter," *e-Proceeding of Engineering*, vol. 6, no. 2, pp. 8637–8649, 2019.
- [4] P. S. Ayu, "Analisa kompetitif sosial media menggunakan metode klasifikasi naive bayes dan support vector machine," *Institut Teknologi Sepuluh Nopember*, vol. SS 141501, no. 1, pp. 1–50, 2018.
- [5] M. Furqan and A. F. A. Nasir, "Big Data Approach To Sentiment Analysis in Machine Learning-Based Microblogs: Perspectives of Religious Moderation Public Policy in Indonesia," *Journal of Applied Engineering and Technological Science*, vol. 5, no. 2, pp. 955–965, 2024, doi: 10.37385/jaets.v5i2.4498.
- [6] M. Frananda Adiezwar Ramadhan, I. Rizal Setiawan, and A. Asriyanik, "Klasifikasi Hoax Dan Fakta Menggunakan Algoritma Shallow Neural Network Pada Berita Politik Pemilihan Presiden Indonesia 2024," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 4, pp. 8006–8013, 2024, doi: 10.36040/jati.v8i4.10621.
- [7] N. Fathirachman Mahing, A. Lazuardi Gunawan, A. Foresta Azhar Zen, F. Abdurrachman Bachtar, and S. Agung Wicaksono, "Klasifikasi Tingkat Stress dari Data Berbentuk Teks dengan Menggunakan Algoritma Support Vector Machine (SVM) dan Random Forest,"

- Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 7, pp. 1527–1536, 2023, doi: 10.25126/jtiik.1078010.
- [8] N. Puspitasari, R. Rosmasari, F. W. Pratama, and H. Sulastri, “Quality Classification of Palm Oil Varieties Using Naive Bayes Classifier,” *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, vol. 13, no. 1, pp. 11–23, May 2022, doi: 10.31849/digitalzone.v13i1.9773.
- [9] A. Baita and N. Cahyono, “Analisis Sentimen Mengenai Vaksin Sinovac Menggunakan Algoritma Support Vector Machine (Svm) Dan K-Nearest Neighbor (Knn),” *Infos*, vol. 4, no. 2, pp. 42–42, 2021.
- [10] M. Furqan, S. Mayang Sari, and P. Ilmu Komputer Fakultas Sains dan Teknologi, “Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia Sentiment Analysis using K-Nearest Neighbor towards the New Normal During the Covid-19 Period in Indonesia,” 2022. [Online]. Available: [www.tripadvisor.com](http://www.tripadvisor.com)
- [11] R. Sari, “Analisis Sentimen Pada Review Objek Wisata Dunia Fantasi Menggunakan Algoritma K-Nearest Neighbor (K-Nn),” *EVOLUSI : Jurnal Sains dan Manajemen*, vol. 8, no. 1, pp. 10–17, 2020, doi: 10.31294/evolusi.v8i1.7371.
- [12] R. Azhar, A. Surahman, and C. Juliane, “Analisis Sentimen Terhadap Cryptocurrency Berbasis Python TextBlob Menggunakan Algoritma Naïve Bayes,” *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 6, no. 1, pp. 267–281, 2022.
- [13] O. Abiola, A. Abayomi-Alli, O. A. Tale, S. Misra, and O. Abayomi-Alli, “Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser,” *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, 2023, doi: 10.1186/s43067-023-00070-9.
- [14] Alfandi Safira and F. N. Hasan, “Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier,” *ZONAsi: Jurnal Sistem Informasi*, vol. 5, no. 1, pp. 59–70, 2023, doi: 10.31849/zn.v5i1.12856.
- [15] F. Novianti and K. R. N. Wardani, “Analisis Sentimen Masyarakat Terhadap Data Tweet Traveloka Selama Rapid Test Antigen Menggunakan Algoritma Naïve Bayes,” *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 3, pp. 922–933, 2023, doi: 10.29100/jipi.v8i3.3973.
- [16] 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.
- [17] S. Rabbani, D. Safitri, N. Rahmadhani, A. A. F. Sani, and M. K. Anam, “Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 153–160, 2023, doi: 10.57152/malcom.v3i2.897.
- [18] R. Soelistijadi *et al.*, “Pemodelan Prediktif Menggunakan Metode Ensemble Learning XGBoost dalam Peningkatan Akurasi Klasifikasi Penyakit Ginjal,” vol. 5, no. 4, pp. 1866–1875, 2024.