

## Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini *Twitter* Menggunakan *Naïve Bayes*

Fuad Amirullah<sup>1)</sup>, Syariful Alam<sup>2)</sup>, M.Imam Sulistyio S<sup>3)</sup>

<sup>1,2,3)</sup>Teknik Informatika, Sekolah Tinggi Teknologi Wastukencana, Purwakarta, Indonesia  
Email : <sup>1)</sup>fuadamirullah50@wastukencana.ac.id, <sup>2)</sup>syarifulalam@wastukencana.ac.id,  
<sup>3)</sup>imams@wastukencana.ac.id

---

### INFO ARTIKEL

---

**Riwayat Artikel :**

Diterima : 21 Juli 2023

Disetujui : 24 Juli 2023

---

**Kata Kunci :**

Sentimen,KPU,  
Pemilu2024,*Twitter*.

---

### ABSTRAK

Era globalisasi saat ini memengaruhi pesatnya kemajuan teknologi informasi yang beredar di masyarakat salah satu perantaranya ialah melalui media sosial *Twitter*, *Twitter* dapat dimanfaatkan sebagai sarana menyampaikan pendapat terkait saran, kritik, maupun opini-opini publik. Salah satunya perihal kinerja KPU menjelang pemilu 2024 yang muncul diakhir tahun 2022 dan awal tahun 2023, mengenai skandal manipulasi data keanggotaan parpol di website KPU, hingga penundaan pemilu. Penelitian ini menggunakan metode text mining yang terdiri dari tahapan scraping data, labeling, cleansing, preprocessing, algoritma *Naïve Bayes*, pembobotan memakai perhitungan TF-IDF, pengujian data memakai confusion matrix dengan menggunakan pemrograman python. Didapatkan data yang berjumlah 2475 data, yaitu sebanyak 321 data positif, 409 data netral, dan 1745 data negatif. Berdasarkan hasil pengujian confusion matrix dengan pemrograman python didapatkan hasil akurasi negatif sebesar 74.5% berdasarkan nilai tersebut membuktikan tingkat kepercayaan pada kinerja KPU menjelang proses pemilu 2024 mendatang tergolong kurang baik.

---

### ARTICLE INFO

---

**Article History :**

Received : July 21, 2023

Accepted : July 24, 2023

---

**Keywords:**

Sentiment,KPU,  
2024Election,*Twitter*.

---

### ABSTRACT

*The current era of globalization has greatly influenced the rapid progress of information technology circulating in society, one of the intermediaries is through the social media Twitter, Twitter can be used as a means of expressing opinions regarding criticism, suggestions, public opinions. One of them is regarding the KPU's performance ahead of the 2024 elections which will appear at the end of 2022 and early 2023, regarding the scandal of manipulation of political party membership data on the KPU website, to the postponement of the elections. This study uses the text mining method which consists of the stages of scraping data, labeling, cleansing, preprocessing, Naïve Bayes algorithm, weighting using TF-IDF calculations, data testing using a confusion matrix using python programming. There were 2475 data obtained, namely 321 positive data, 409 neutral data, and 1745 negative data. Based on the results of the confusion matrix test with python programming, a negative accuracy result of 74.5% based on this value proves the level of confidence in KPU performance ahead of the upcoming 2024 election process is classified as not good.*

## 1. PENDAHULUAN

Era globalisasi saat ini sangat memengaruhi pesatnya kemajuan teknologi informasi seperti dalam bidang ekonomi, kebudayaan, seni, pendidikan dan bahkan dunia politik. Teknologi membuat jarak tidak lagi menjadi masalah dalam berkomunikasi, dan sekarang sosial media menjadi kebutuhan wajib bagi sebagian kalangan masyarakat mengatakan komentar mereka di web media sosial semacam *Facebook*, *Twitter*, *Wechat*, ataupun sejenisnya. Dan salah satu web media sosial yang kerap digunakan untuk memperdebatkan isu - isu politik serta sosial adalah *Twitter*.(Rahayu, Fauzi and Indra, 2022)

*Twitter* merupakan website microblogging yang cukup terkenal dimana pengguna bisa mempublikasikan pembaruan status melalui "tweet", dan "Trending Topics" merupakan salah satu fitur di *Twitter* yang mempermudah pengguna dalam mencari topik yang trending saat ini di wilayah tertentu maupun seluruh dunia.(Safira *et al.*, 2023) Bertepatan di tahun 2024 merupakan tahun pesta demokrasi terbesar bagi rakyat republik Indonesia, antusiasme dan kemeriahan masyarakat terkait pasangan yang akan mencalonkan membuat Komisi Pemilihan Umum (KPU) sebagai lembaga penyelenggara pemilihan umum harus bersiap dalam menghadapi pemilihan umum yang akan diadakan serentak tersebut. KPU merupakan suatu lembaga pemerintah yang terbentuk, berdasarkan pasal 22 E Undang-Undang Dasar 1945 pada perubahan ketiga bahwa KPU adalah suatu lembaga yang bersifat nasional, tetap dan mandiri. Yang menjadikan kinerja KPU pada pemilu kali ini akan menjadi sorotan masyarakat awam maupun elite politik. Sebagai institusi yang bertanggung jawab atas pelaksanaan pemilihan umum di Indonesia baik sebelum maupun sesudah, KPU harus terus memperbaiki kinerjanya agar proses pemilihan umum berjalan lancar, demokratis, dan adil.(Metera and Mekarsari, 2022) Sehingga media sosial juga telah menjadi salah satu media yang digunakan oleh masyarakat untuk mengeluarkan pendapat, termasuk mengenai pandangan mereka terhadap kinerja KPU, seperti yang terjadi diakhir tahun 2022 dan diawal tahun 2023 mengenai desas-desus kecurangan KPU skandal manipulasi data

keanggotaan partai politik di website KPU dan dugaan penundaan pemilu.

Analisis sentimen adalah penambahan teks yang menganalisis dan mengekstraksi informasi yang bersifat subjektif dan dapat membantu pihak yang membutuhkan.(Normawati and Prayogi, 2021) penelitian tentang hal ini sebelumnya sudah pernah dilakukan oleh Amirul Haj yang berjudul Analisis Sentimen Kinerja KPU Pemilu 2019 Menggunakan Algoritma *K-Means* Dengan Algoritma *Confix Stripping Stemmer*, hasil yang didapatkan dari sentimen masyarakat terhadap kinerja KPU ini menghasilkan sentimen negatif daripada yang positif. Dan memiliki tingkat akurasi yang didapatkan dari penggunaan algoritma *K-Means* saja yaitu 84%, sedangkan kombinasi algoritma diatas yaitu 86%.(Amirul Haj, Amrizal and Arini, 2020).

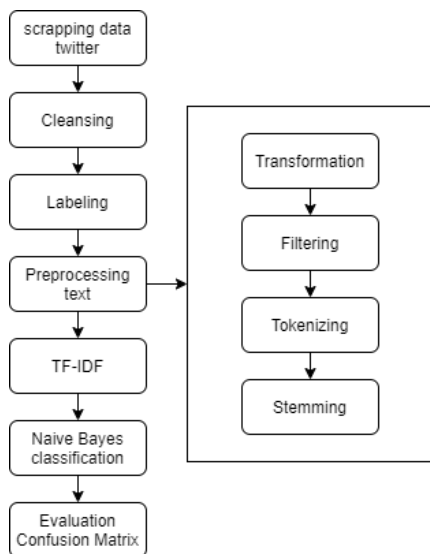
Penelitian selanjutnya menggunakan metode *Naïve Bayes Classifier* (NBC) yang merupakan model penyederhanaan Metode Bayes yang cocok dalam klasifikasi teks atau dokumen dengan banyak dataset dengan performa yang cepat dalam mengklasifikasi data dan memiliki akurasi tinggi.(Normawati and Prayogi, 2021). Dengan pembobotan tf-idf disertai penambahan konversi ikon emosi untuk mengetahui kelas sentimen yang ada dari *tweet* tentang toko JD.id. Hasil penelitian menunjukkan metode *Naïve Bayes* tanpa penambahan fitur mampu mengklasifikasi sentimen dengan nilai akurasi sebesar 96,44%, sedangkan yang disertai konversi ikon emosi mampu meningkatkan nilai akurasinya menjadi 98%.(Sari and Wibowo, 2019)

Analisis sentimen warganet Indonesia terhadap kinerja KPU melalui media sosial *Twitter* menggunakan metode sentimen analisis algoritme *Naïve Bayes* dapat dilakukan guna memperoleh gambaran yang lebih akurat tentang pandangan kepercayaan masyarakat terhadap kinerja KPU, sehingga masyarakat tidak ragu dalam mengikuti pemilu mendatang. Penulis nantinya ingin melihat tingkat kepercayaan masyarakat terhadap kinerja KPU dalam mengelola pemilu 2024 mendatang, sehingga masyarakat tidak ragu untuk mengikuti pemilu 2024 yang diselenggarakan KPU.

## 2. METODE

### 2.1. Kerangka Penelitian

Metode ini menjelaskan bagaimana prosedur pengumpulan data yang diolah penulis sampai menjadi sebuah informasi. Rancangan penelitian akan disajikan pada gambar 2.1 :



Gambar 2. 1 kerangka penelitian

### 2.2. Pengumpulan Data (Scraping Data)

Dataset yang digunakan pada penelitian adalah data yang dikumpulkan berasal dari *tweet* pada media sosial *Twitter*, untuk pengumpulan data dari *Twitter* memakai metode *scraping* menggunakan library pemrograman Python bernama *Snsrape*, (Raharjo, Sunarya and Divayana, 2022) yang menghasilkan jumlah data sebanyak 3000 data, tanggal penarikan dari 20 januari 2020 sampai dengan 31 Mei 2023. Setelah itu data *Tweet* akan disimpan dalam bentuk file *Comma Separated View(CSV)*.

Tabel 2. 1 Program python Scrapping

```

import snsrape.modules.Twitter as
snTwitter
import pandas as pd
query = "(kpu) lang:id until:2023-
05-31 since:2020-01-01"
tweets = []
limit = 3000
for tweet in
snTwitter.TwitterSearchScrapper(query).get_items():
    if len(tweets) == limit:
        break
    else:

```

```

tweets.append([tweet.date,
tweet.id, tweet.content,
tweet.user.username, ])
df = pd.DataFrame(tweets,
columns=['Datetime', 'Tweet
Id', 'Tweet', 'Username', ])
print(df)
data.to_csv("/content/kpu.csv")

```

### 2.3. Cleansing

Cleansing merupakan penghapusan kata duplikat dan kata yang sama menghasilkan jumlah data sebanyak 2477 data. cleansing data dilakukan scraping dengan menggunakan pemrograman python.

Tabel 2. 2 Program python Cleansing

```

df=df.drop_duplicates(subset=["Twee
t", "Username"])
df.head()
data.to_csv("/content/cleankpu.csv"
)

```

### 2.4. Labeling

Tahapan ini menjelaskan bagaimana kondisi ketika sesuatu mendapatkan julukan dari seseorang yang dimana julukan tersebut didasarkan pada perilakunya (Nurtikasari *et al.*, 2022). Tahapan ini membagi kategori menjadi 1 = positif, 0 = netral, -1 = negative. Dilakukan otomatis menggunakan pemrograman *python* serta manual menggunakan *Microsoft Excel*, menghasilkan 321 data positif, 409 data netral, 1745 data negatif.

Tabel 2. 3 Labeling

Tweet	User	Label
“Saya yakin dua kali pemilu kinerja KPU sangat detail akurat jadi pasti ada hal-hal seperti ini tapi teruslah dilakukan	CaturRp	1
Pengumpulan laporan bulanan dan kinerja kepada staff KPU Demak pada 4 Mei 2023. @KPU_ID @KPU_Jateng @kpu_demakkab https://t.co/PzLevwQoko	Dempet Ppk	0

Tweet	User	Label
@blank0429 Paling2 tak ditindak lanjuti oleh KPU. Hanya buat laporan dan tak ada sanksi ide keras meski itu serius berpotensi terjadinya kecurangan. Namanya negeri para bedebah ya ide2 ide, tak profesional sama sekali!	Sarahjo suaddic 1	-1

### 2.5. Preprocessing Text

Preprocessing Text membagi teks sehingga menjadi unit-unit yang lebih kecil seperti *Transformation, tokenization, Filtering, dan Stemming* (Wiguna and Rifai, 2021) menggunakan pysastrawi yaitu modul sederhana yang dimiliki oleh library Python yang memungkinkan untuk melakukan pengurangan kata-kata yang ter- infleksi dalam bahasa Indonesia ke bentuk baku-nya. (Puspa *et al.*, 2023) Proses ini menggunakan program *python*.

Tabel 2. 4 Preprocessing Text

Preprocessing Text	Text
Raw Data	@macanutara_ @aniesbaswedan @AgusYudhoyono Yg paling penting awasi kecurangan dan KPU... Kuncinya disitu.. Hanya kecurangan yg bisa kalahkan anis
Transformation	macanutara aniesbaswedan agusyudhoyono yg paling penting awasi kecurangan dan kpu kuncinya disitu hanya kecurangan yg bisa kalahkan anis
Filtering	macanutara aniesbaswedan agusyudhoyono awasi kecurangan kpu kuncinya disitu

Preprocessing Text	Text
	kecurangan kalahkan anis
Tokenizing	['macanutara_', 'aniesbaswedan', 'agusyudhoyono', 'awasi', 'kecurangan', 'kpu', 'kuncinya', 'disitu', 'kecurangan', 'kalahkan', 'anis']
Stemming	macanutara aniesbaswedan agusyudhoyono awas curang kpu kunci situ curang kalah anis

## 3. HASIL DAN PEMBAHASAN

### 3.1. Pembobotan Kata TF-IDF

Selanjutnya dilakukan proses term weighting, Proses ini dilakukan untuk menilai term atau bobot kata dari sebuah dokumen terhadap seluruh dokumen pada korpus. Proses ini dilakukan menggunakan bahasa pemrograman python, Berikut ini adalah contoh kumpulan data yang akan digunakan hingga 9 data tweet, termasuk 9 sampel yang ditandai sebagai 3 positif, 3 negatif, dan 3 netral. Data tersebut digunakan sebagai data latih, dapat dilihat pada tabel 3.1 :

Tabel 3. 1 Data Latih TF-IDF

Dokumen Tweet	Sentimen
tingkat kerja kpu teguh komitmen dedikasi profesionalisme tugas arah wima papar materi kait laksana anggaran lingkung kpu bendahara paham	Positif
percaya survey kait kerja kpu bawaslu mk	Positif
diupdate verifikasi alam kali milu kerja kpu detail akurat halhal sempurna	Positif

Dokumen Tweet	Sentimen
ridwan kamil bersama juara	
motif curang pilpres atas tunggu kpu bentuk rezim main hasil milu sekian kali	Negatif
kpu id bawaslu ri malu kecurangan main kotor ok main kasar hasil sesuai seknario aman	Negatif
curang sukses data tampil serta bukti fisik lampir serta mudah rakyat laku cross cek kerja kpu id	Negatif
catat curang kpu milu lihat video ansori oke tiktok	Netral
abu waras curang siap mari antisipasi terjadi curang indikasi kpu bawaslu perintah tindak sesuai atur viralkan	Netral
milu jurdil tidak ragu kemenangan anis antisipasi curang khususnya tingkat kpu hitung	Netral

Tabel 3. 2 Term Frequency

Kata	D1	D2	...
tingkat	1	0	...
kerja	1	0	...
kpu	1	1	...
...	...	...	...

Tabel 3. 3 Perhitungan TF-IDF

$$p_{kata} = \frac{\text{jumlah dari 1 kata di 1 kategori} + 1}{\text{jumlah kata tiap kategori} + \text{jumlah keseluruhan kata unik}}$$

$$(kpu | positif) \frac{1+1}{35+87} = \frac{2}{122} = 0.01639$$

$$(kpu | negatif) \frac{1+1}{40+87} = \frac{2}{127} = 0.00157$$

$$(kpu | netral) \frac{1+1}{33+87} = \frac{2}{120} = 0.01666$$

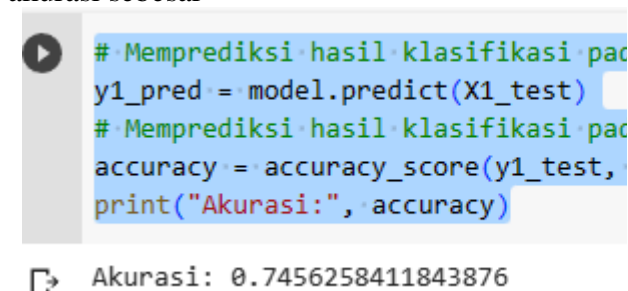
### 3.2. Klasifikasi Naïve Bayes

Setelah itu memasukan library metode *naïve bayes* menggunakan pemrograman python.

Tabel 3. 4 Program Naïve Bayes

```
from sklearn.naive_bayes import
MultinomialNB
from sklearn.metrics import
accuracy_score
from joblib import dump
from datetime import datetime
start_time = datetime.now()
# Melatih model Naive Bayes
model = MultinomialNB()
model.fit(X1_train, y1_train)
# Memprediksi hasil klasifikasi
pada data testing
y1_pred = model.predict(X1_test)
# Memprediksi hasil klasifikasi
pada data testing
accuracy =
accuracy_score(y1_test, y1_pred)
print("Akurasi:", accuracy)
```

Dengan program python menghasilkan nilai akurasi sebesar



Gambar 3. 1 Hasil Akurasi Naïve Bayes

### 3.3. Evaluation Confusion Matrix

Evaluasi bertujuan untuk mengecek kebenaran hasil klasifikasi. Pada tahap ini hasil evaluasi menggunakan metode cross validation yang nantinya akan diukur dengan menggunakan Confusion Matrix untuk mengetahui hasil akurasi.

Tabel 3. 5 Program Confusion Matrix

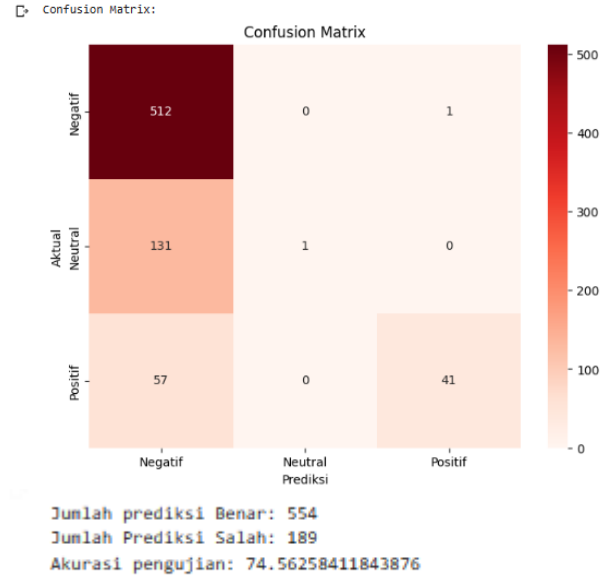
```
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import
confusion_matrix
```

```

from sklearn.metrics import
accuracy_score, precision_score,
recall_score, f1_score
from sklearn.metrics import
classification_report
from sklearn.metrics import
confusion_matrix
# Membuat confusion matrix
cm = confusion_matrix(y1_test,
y1_pred)
print("Confusion Matrix:")
labels = ['Negatif', 'Neutral',
'Positif']
# Membuat heatmap dari confusion
matrix
plt.figure(figsize=(8, 6))
sns.heatmap(cm, annot=True,
fmt="d", cmap="Reds",
xticklabels=labels,
yticklabels=labels)
plt.title("Confusion Matrix")
plt.xlabel("Prediksi")
plt.ylabel("Aktual")
plt.show()
pred_benar = (y1_test ==
y1_pred).sum()
pred_salah = (y1_test !=
y1_pred).sum()
accuracy = pred_benar/(pred_benar +
pred_salah)*100
precision =
print("Jumlah prediksi Benar:"
,
pred_benar)
print("Jumlah Prediksi Salah:"
,
pred_salah)
print("Akurasi pengujian:"
,
accuracy)
print('=====\n')
print(classification_report(y1_test
, y1_pred, zero_division=0))
print('=====\n')
print("Hasil Klasifikasi
Sentimen:")
print("Accuracy:"
,
accuracy_score(y1_test,y1_pred))
print("\nPerforma Model :\n")
    
```

```

print("Accuracy:"
,
accuracy_score(y1_test,y1_pred))
print("error_rate:"
,
1-
accuracy_score(y1_test,y1_pred))
    
```



Gambar 3. 2 Hasil Confusion Matrix dan Akurasi

### 3.3.1. Visualisasi

Pada proses ini suatu informasi akan digambarkan dengan wordcloud pada sentimen keseluruhan dokumen positif, netral, dan negatif mengenai Kinerja KPU menjelang pemilu 2024.



Gambar 3. 3 Wordcloud Sentimen Positif



Gambar 3. 4 Wordcloud Sentimen Netral





icle/view/3487/1883.

Wiguna, R.A. raffaidy and Rifai, A.I. (2021)  
'Analisis Text Clustering Masyarakat Di  
Twitter Mengenai Omnibus Law  
Menggunakan Orange Data Mining',  
*Journal of Information Systems and  
Informatics*, 3(1), pp. 1–12. Available at:  
<https://doi.org/10.33557/journalisi.v3i1.78>.