

ANALISIS SENTIMEN PENGUMUMAN HASIL PEMILU 2024 DI SOSIAL MEDIA X MENGGUNAKAN KNN DAN NAÏVE BAYES CLASSIFIER

Theresia Aurelly Claudia Budianto, Hilman Fatoni, Maesha Ayu Syaharani, Chaerur Rozikin

Informatika, Universitas Singaperbangsa Karawang

Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361, Telp. (0267) 641177

theresiaaurelly@gmail.com

ABSTRAK

Pemilu merupakan salah satu peristiwa politik yang paling menarik perhatian publik dalam suatu negara demokratis. Hasil pemilu mencerminkan preferensi dan keinginan masyarakat terhadap calon dan partai politik yang bersaing. Di era digital saat ini, platform media sosial X telah menjadi salah satu wadah utama dimana opini dan sentimen masyarakat tercermin dengan luas. Oleh karena itu, analisis sentimen terhadap hasil pemilu 2024 di media sosial X menjadi semakin penting dalam memahami dinamika politik dan respons masyarakat. Penelitian ini membandingkan akurasi metode *KNN* dan *Naïve Bayes classifier* dalam membagi sentimen menjadi positif, negatif, dan netral. Proses pelabelan dilakukan secara otomatis menggunakan *textblob* dengan jumlah sentimen positif sebanyak 465 tweet atau 87.41%, netral sebanyak 42 tweet atau 7.89%, dan negatif sebanyak 25 tweet atau 4.70%. Dari total 532 data dengan kata kunci "Hasil Pemilu 2024", 80% digunakan sebagai data latih dan 20% sebagai data uji. *Naive Bayes* mencapai akurasi 85%, sedangkan *KNN* mencapai 68%, menandakan keunggulan *KNN* dalam mengklasifikasikan sentimen terhadap hasil pemilu 2024.

Kata kunci : Analisis Sentimen, Hasil Pemilu 2024, *KNN Classifier*, *Naïve Bayes Classifier*.

1. PENDAHULUAN

Pemilihan umum adalah salah satu momen politik yang sangat menarik perhatian masyarakat dalam konteks negara demokratis. Partisipasi masyarakat merupakan kunci utama dalam menentukan suksesnya pelaksanaan pemilihan umum [1].

Hasil dari pemilu mencerminkan aspirasi dan preferensi masyarakat terhadap calon dan partai politik yang bersaing.

Di Indonesia, pemilu 2024 diharapkan menjadi pemilu yang transparan dan adil. Akan tetapi, dengan adanya perkembangan teknologi informasi, informasi mengenai hasil pemilu dapat dengan mudah tersebar di sosial media. Media sosial merupakan platform online yang memungkinkan pengguna untuk memposting opini mereka ke media online. X seringkali menjadi pilihan utama bagi pengguna untuk melakukan hal tersebut [2].

Platform media sosial X memiliki pengaruh yang signifikan dalam membentuk sentimen dan preferensi politik publik, sehingga pesan – pesan yang di posting oleh masyarakat dapat memberikan gambaran mengenai pendapat mereka, yang menarik untuk diklasifikasikan ke dalam kategori sentimen positif, negatif, atau netral [3].

Penerapan analisis sentimen dapat mendukung dalam memahami berbagai pendapat yang terdapat dalam X. Melalui analisis sentimen, memungkinkan kita untuk menilai apakah pesan – pesan tersebut memiliki orientasi positif, negatif, atau netral [4].

Dengan demikian, kita dapat memahami lebih baik dinamika politik yang sedang berlangsung dan memperoleh wawasan yang lebih mendalam tentang preferensi dan sentimen masyarakat terhadap calon dan partai politik yang bersaing. Hal ini memberikan

kesempatan untuk meningkatkan pemahaman tentang perubahan opini publik serta memperbaiki strategi komunikasi politik di masa depan.

Analisis sentimen dapat diterapkan melalui berbagai metode klasifikasi. Beberapa metode umum dipakai diantaranya *Naïve Bayes (NB)*, *Support Vector Machine (SVM)*, dan *k-Nearest Neighbor (k-NN)* [5].

Beberapa penelitian terdahulu juga telah menggunakan analisis sentimen untuk mengambil opini dari sekumpulan dokumen teks, seperti penelitian oleh Ray, dkk. yaitu menganalisis sentimen terhadap KPU 2024 berdasarkan *tweet* media sosial *Twitter* dengan algoritma *Naïve Bayes*. Didapatkan 715 dataset *Twitter* dengan kata kunci "KPU 2024". Hasil penelitian menunjukkan akurasi algoritma *Naïve Bayes* sebesar 67,13%, dengan presisi 66,04% dan *recall* 100,00%. Evaluasi pada matriks konfusi menemukan *true positif* sebesar 457 dan *true negative* sebesar 235 [6].

Selain itu, penelitian Sarimole, dkk. (2024) mengevaluasi sentimen masyarakat terhadap isu penundaan Pemilu 2024 pada *Twitter* dengan metode *Naïve Bayes* memperoleh akurasi sebesar 98.80%, sedangkan metode *Support Vector Machine (SVM)* mencapai akurasi sebesar 91.61%. Dari hasil penelitian ini, dapat disimpulkan bahwa *Naïve Bayes* memberikan akurasi yang sedikit lebih tinggi dalam memprediksi sentimen masyarakat terhadap isu tersebut dibandingkan dengan *SVM* [7].

Penelitian lainnya oleh Jurnal Prasetyo, dkk. menganalisis sentimen relokasi Ibukota Nusantara menggunakan algoritma *Naïve Bayes* dan *K-Nearest Neighbor*. Hasilnya menunjukkan bahwa algoritma *K-Nearest Neighbor (KNN)* menghasilkan akurasi sebesar 88.12%, dengan *Precision* 93.98% dan *Recall* 81.53%. Sedangkan, algoritma *Naïve Bayes (NB)*

mencapai akurasi sebesar 82.27%, dengan *Precision* 86.36% dan *Recall* 76.93%. Dari penelitian tersebut, terlihat bahwa *KNN* memiliki kinerja yang lebih baik dalam mengukur sentimen terhadap relokasi Ibukota Nusantara [8].

Penelitian ini menggunakan algoritma *KNN* dan *Naive Bayes Classifier* untuk menganalisis sentimen hasil pemilu 2024. Pemilihan algoritma *KNN* dan *Naive Bayes* didasarkan pada penelitian sebelumnya, seperti penelitian oleh Ray, dkk. yang menunjukkan bahwa *Naive Bayes* cukup akurat dalam memprediksi sentimen masyarakat terhadap isu tertentu, seperti sentimen terhadap KPU 2024 [6].

Selain itu, penelitian oleh Prasetyo, dkk. juga menunjukkan bahwa *KNN* memiliki kinerja yang baik dalam mengukur sentimen terhadap topik tertentu, seperti relokasi Ibukota Nusantara [8].

Oleh karena itu, dengan mengacu pada penelitian-penelitian tersebut, penggunaan *KNN* dan *Naive Bayes Classifier* dalam menganalisis sentimen hasil pemilu 2024 di media sosial X diharapkan dapat memberikan hasil yang akurat dan representatif terhadap opini masyarakat secara keseluruhan.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen adalah cabang penelitian dari *data mining* dan *text mining*, atau sebuah metode yang diterapkan untuk mengolah berbagai opini yang dipublikasikan oleh konsumen atau para pakar melalui berbagai media, mengenai suatu hal tertentu agar mendapatkan sebuah sentimen (negatif, positif atau netral) yang terkandung dalam sebuah opini [9].

2.2. Pemilihan Umum (Pemilu)

Pemilu merupakan sarana kedaulatan masyarakat untuk memilih anggota Dewan Perwakilan Rakyat, anggota Dewan Perwakilan Daerah, Presiden dan Wakil Presiden, dan pemilihan anggota Dewan Perwakilan Rakyat Daerah, dilakukan secara langsung. Pemilihan Presiden atau pilpres merupakan kegiatan lima tahun sekali yang menjadi momen yang penting dalam perwujudan demokrasi di Negara Kesatuan Republik Indonesia [8].

2.3. K - Nearest Neighbor (KNN)

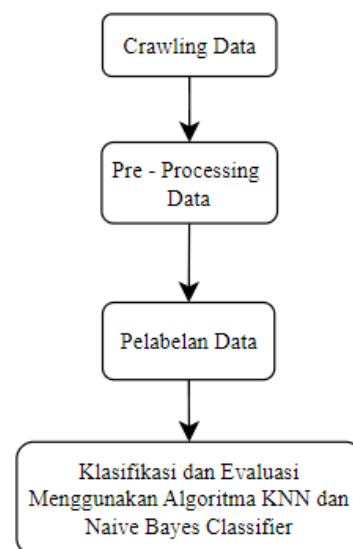
Algoritma *KNN* adalah algoritma pembelajaran terawasi yang dapat digunakan untuk tugas klasifikasi dan regresi. Ini bekerja dengan menemukan K titik data terdekat ke sampel yang diberikan, dan menggunakan label kelas atau nilai dari titik data ini untuk membuat prediksi tentang label kelas atau nilai sampel. Dalam tugas klasifikasi, *KNN* bekerja dengan mencari titik data K dalam set pelatihan yang paling dekat dengan sampel yang sedang diklasifikasikan, berdasarkan ukuran jarak seperti jarak *Euclidean*. Ini kemudian menetapkan sampel ke label kelas yang paling umum di antara K tetangga terdekat [10].

2.4. Naive Bayes Classifier

Naive Bayes Classifier adalah sebuah metode algoritma klasifikasi yang berdasarkan pada konsep probabilitas. Teorema Bayes adalah dasar konsep untuk metode *Naive Bayes*. *Naive Bayes* merupakan salah satu alat klasifikasi yang kompeten [11]. Keuntungan menggunakan metode ini adalah metode ini hanya membutuhkan sedikit data latih untuk menentukan parameter yang dibutuhkan dalam proses klasifikasi [12].

3. METODE PENELITIAN

Terdapat beberapa tahapan yang harus dilakukan dalam penelitian ini. Pertama, *Crawling data* dari X menggunakan *python*. Kedua, *Pre-Processing data* yaitu melakukan pembersihan data. Ketiga, data dilabeli menggunakan *Textblob* dengan kategori positif, negatif, dan netral. Terakhir, dilakukan klasifikasi data menggunakan algoritma *KNN* dan *Naive Bayes Classifier* (Gambar 1).



Gambar 1. Tahapan Penelitian

Berikut adalah uraian mengenai metode dari penelitian ini :

a. Crawling Data

Penelitian ini diawali dengan melakukan crawling data dari X dengan kata kunci “Hasil Pemilu 2024” menggunakan Python. Data dikumpulkan dalam rentang waktu 14 Februari 2024 hingga 23 Maret 2024, menghasilkan sebanyak 550 baris data. Hasil dari *crawling data* akan disimpan dalam bentuk CSV.

b. Pre - Processing Data

Tahapan dalam preprocessing data meliputi: *Data Cleaning*, *Case Folding*, Normalisasi, *Stopword*, Tokenisasi, dan *Stemming*. *Data cleaning* dilakukan untuk menghilangkan data yang *duplicate*, *missing value*, *mention*, *hashtag*, simbol, dan spasi yang berlebih. *Case Folding* mengubah semua kata menjadi huruf kecil agar kata-kata yang sama dengan alfabet berbeda tidak terdeteksi

sebagai kalimat yang berbeda. Normalisasi merupakan proses untuk mengubah kata yang menggunakan singkatan atau tidak baku menjadi bentuk kata baku. *Stopword* adalah proses untuk menghapus kata-kata yang tidak relevan dengan topik menggunakan *library* dari Sastrawi. Tokenisasi merupakan proses untuk memisahkan kalimat menjadi kata-kata tunggal. *Stemming* adalah proses untuk mengubah kata-kata yang memiliki imbuhan menjadi bentuk kata dasar.

c. Pelabelan Data

Proses pelabelan data dilakukan secara otomatis dengan menggunakan *Textblob*. *Textblob* menghitung nilai *polarity* dan *subjectivity* dari sebuah teks. *Polarity* berfungsi untuk menentukan kecenderungan sentimen teks, sementara *subjectivity* berfungsi untuk menilai apakah teks tersebut berupa opini atau fakta. Semakin tinggi nilai *subjectivity*, semakin besar kemungkinan teks tersebut adalah opini. Sebaliknya, semakin tinggi nilai *polarity*, semakin positif emosinya. Berdasarkan nilai *polarity*, teks diklasifikasikan ke dalam tiga kategori yaitu positif, negatif, dan netral.

d. Klasifikasi

Proses klasifikasi data menggunakan dua algoritma, yaitu *KNN* dan *Naïve Bayes Classifier*. Dari 532 data yang tersedia, 80% dijadikan data latih dan 20% data uji. Model *KNN* dan *Naïve Bayes* dibangun menggunakan data latih dan dievaluasi dengan data uji. *KNN* mengklasifikasikan sentimen teks berdasarkan kedekatan dengan K tetangga terdekat dari data latih, sedangkan *Naïve Bayes* menggunakan pendekatan probabilistik untuk menghitung kemungkinan teks termasuk dalam kategori sentimen tertentu berdasarkan distribusi kata dalam data latih.

e. Evaluasi

Evaluasi dilakukan untuk menilai seberapa bagus performa algoritma klasifikasi yang digunakan dalam penelitian. Metode yang digunakan untuk melakukan evaluasi dalam penelitian ini yaitu *confusion matrix*. *Confusion matrix* merupakan tabel yang berisi informasi perbandingan antara hasil klasifikasi prediksi dan hasil klasifikasi yang sebenarnya [13]. Berikut ini merupakan contoh *confusion matrix* 3×3 (Gambar 2) :

		Predict Class		
		A	B	C
Actual Class	A	AA	AB	AC
	B	BA	BB	BC
	C	CA	CB	CC

Gambar 2. Confusion Matrix 3×3

Nilai yang dihasilkan dari *confusion matrix*, dapat menghasilkan nilai akurasi, presisi, *recall*, dan *f1-score*. Berikut persamaan yang dapat digunakan :

$$\text{Akurasi} = \frac{AA+BB+CC}{AA+AB+AC+BA+BB+BS+CA+CB+CC} \quad (1)$$

$$\text{Presisi} = \frac{\left(\frac{AA}{AA+BA+CA} + \frac{BB}{BA+BB+BC} + \frac{CC}{CA+CB+CC}\right)}{3} \quad (2)$$

$$\text{Recall} = \frac{\left(\frac{AA}{AA+AB+AC} + \frac{BB}{BA+BB+BC} + \frac{CC}{CA+CB+CC}\right)}{3} \quad (3)$$

$$f1\text{-Score} = \frac{2 \times \text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (4)$$

Dimana :

Akurasi : Mengukur seberapa sering model membuat prediksi yang benar.

Presisi : Mengukur seberapa banyak prediksi positif yang benar-benar positif.

Recall : Mengukur seberapa banyak data positif yang berhasil terdeteksi oleh model.

f1-Score : Mengharmonisasikan presisi dan *recall*.

4. HASIL DAN PEMBAHASAN

Analisis sentimen terhadap hasil pemilu 2024 akan dilakukan menggunakan *Google Collaboratory* dengan bahasa *python*.

4.1. Crawling Data

Proses dan hasil dari *crawling* data yang telah dilakukan pada aplikasi X dengan kata kunci “Hasil Pemilu 2024” dapat dilihat pada (Gambar 3). Didapatkan data sebanyak 550 baris dan 12 kolom. Data ini masih terdapat banyak *noise* sehingga harus dilakukan *cleaning* terlebih dahulu.

```
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 550 entries, 0 to 549
Data columns (total 12 columns):
#   Column                Non-Null Count  Dtype
---  -
0   created_at            550 non-null   object
1   id_str                550 non-null   object
2   full_text             550 non-null   object
3   quote_count          550 non-null   int64
4   reply_count          550 non-null   int64
5   retweet_count         550 non-null   int64
6   favorite_count        550 non-null   int64
7   lang                  550 non-null   object
8   user_id_str           550 non-null   object
9   conversation_id_str   550 non-null   object
10  username              550 non-null   object
11  tweet_url             550 non-null   object
dtypes: int64(4), object(8)
memory usage: 51.7+ KB
```

Gambar 3. Hasil Crawling Data

4.2. Pre – Processing Data

Untuk melakukan analisis sentimen, hanya akan digunakan satu kolom, yaitu “full_text”. Pada tahap *pre-processing* ini akan dilakukan *cleaning* untuk menghilangkan data yang *duplicate*, missing value, mention, hashtag, simbol, dan spasi yang berlebih. Setelah dilakukan *cleaning* data, jumlah data yang tersisa yaitu sebanyak 532 data. Data inilah yang akan digunakan untuk menganalisis sentimen (Gambar 4).

```
#cek perubahan jumlah baris setelah di cleaning
df.shape

(532, 3)
```

Gambar 4. Jumlah Data Setelah Proses *Cleaning*

Tahap selanjutnya yaitu *case folding*, yaitu mengubah setiap kata menjadi huruf kecil. Berikut hasil dari *case folding* (Gambar 5) :

index	full_text
0	kamu kok kasar sadis pada pendukung anies dan ganjar mereka punya hak politik utk gugat hukum atas hasil pemilu pilpres 2024 selain hak politik sbg kandidat capres juga hak sbg warga indonesia yang di lindungi uu
1	momen rekapitulasi hasil penghitungan perolehan suara tingkat nasional pemilu tahun 2024 di jakarta
2	sikapi hasil pemilu 2024 pp muhammadiyah keluarkan 4 poin pernyataan
3	setelah pengumuman resmi hasil pemilu 2024 bangsa indonesia diharap bersatu kembali gempa kenceng 2 smp prabowo kerasa kate bawean sahur
4	soal hasil pemilu 2024 pp muhammadiyah beri pesan menohok ke publik

Gambar 5. Hasil *Case Folding*

Kemudian dilakukan normalisasi untuk mengubah kata yang menggunakan singkatan atau yang tidak sesuai dengan kaidah bahasa menjadi bentuk kata baku (Gambar 6).

index	full_text
0	kamu kok kasar sadis pada pendukung anies dan ganjar mereka punya hak politik untuk gugat hukum atas hasil pemilihan umum pemilihan presiden 2024 selain hak politik sebagai kandidat calon presiden juga hak sebagai warga indonesia yang di lindungi undang-undang
1	momen rekapitulasi hasil penghitungan perolehan suara tingkat nasional pemilihan umum tahun 2024 di jakarta
2	sikapi hasil pemilihan umum 2024 pimpinan pusat muhammadiyah keluarkan 4 poin pernyataan
3	setelah pengumuman resmi hasil pemilihan umum 2024 bangsa indonesia diharap bersatu kembali gempa kenceng 2 smp prabowo kerasa kate bawean sahur
4	soal hasil pemilihan umum 2024 pimpinan pusat muhammadiyah beri pesan sangat berdampak ke publik

Gambar 6. Hasil Normalisasi

Setelah dilakukan normalisasi, masih banyak kalimat yang mengandung kata yang tidak relevan dengan topik dan tidak memberikan makna signifikan dalam pemrosesan teks. Oleh karena itu, perlu dilakukan *stopword* untuk menghapusnya. Berikut hasil dari proses *stopword* (Gambar 7) :

index	full_text
0	kasar sadis pendukung anies ganjar hak politik gugat hukum pemilihan umum pemilihan presiden 2024 hak politik kandidat calon presiden hak warga indonesia lindungi undang-undang
1	momen rekapitulasi penghitungan perolehan suara nasional pemilihan umum 2024 jakarta
2	pemilihan umum 2024 pimpinan pusat muhammadiyah keluarkan 4 poin pernyataan
3	resmi pemilihan umum 2024 bangsa indonesia diharap bersatu prabowo
4	soal pemilihan umum 2024 pimpinan pusat muhammadiyah beri pesan sangat berdampak publik

Gambar 7. Hasil *Stopword*

Selanjutnya akan dilakukan pemisahan kalimat menjadi kata – kata tunggal, yang dikenal sebagai tokenisasi. Berikut hasil dari proses tokenisasi (Gambar 8).

0	[kasar, sadis, pendukung, anies, ganjar, hak, ...
1	[momen, rekapitulasi, penghitungan, perolehan, ...
2	[pemilihan, umum, 2024, pimpinan, pusat, muham...
3	[resmi, pemilihan, umum, 2024, bangsa, indones...
4	[soal, pemilihan, umum, 2024, pimpinan, pusat, ...
...	...
545	[kejahatan, pemilihan, umum, dituding, terjadi...
546	[orang, bikin, berita, apa, ibhrs, mana, mau, ...
547	[pemilihan, umum, 2024, gugat, pemilihan, umum...
548	[pemilihan, umum, 2024, nasdem, pks, terima, p...
549	[amien, rabb, nasdem, pks, terima, pemilihan, ...

Name: full_text, Length: 532, dtype: object

Gambar 8. Hasil Tokenisasi

Setelah dilakukan tokenisasi, akan dilanjutkan dengan proses *stemming* untuk mengembalikan setiap kata ke bentuk dasar atau akar katanya. Berikut hasil dari proses *stemming* (Gambar 9) :

index	full_text
0	kasar sadis dukung anies ganjar hak politik gugat hukum pemilihan umum pemilihan presiden 2024 hak politik kandidat calon presiden hak warga indonesia lindungi undang-undang
1	momen rekapitulasi hitung oleh suara nasional pemilihan umum 2024 jakarta
2	sikap pemilihan umum 2024 pimpin pusat muhammadiyah keluar 4 poin nyata
3	resmi pemilihan umum 2024 bangsa indonesia harap satu prabowo
4	soal pemilihan umum 2024 pimpin pusat muhammadiyah beri pesan sangat dampak publik

Gambar 9. Hasil *Stemming*

Setelah dilakukan *stemming*, langkah selanjutnya adalah melakukan *translate* kata. Proses *translate* setelah *stemming* membantu dalam mengembalikan makna, menghilangkan ambiguitas, dan melakukan normalisasi tambahan pada analisis sentimen. Berikut hasil dari proses *translate* (Gambar 10) :

index	full_text	tweet_english
0	kasar sadis dukung anies ganjar hak politik gugat hukum pemilihan umum pemilihan presiden 2024 hak politik kandidat calon presiden hak warga indonesia lindungi undang-undang	Roughly Sadistic Supports Anies Ganjar Political Rights Laws on General Election of Presidential Election 2024 Political Rights Candidates for Presidential Candidates for Indonesian Citizens Protect Laws
1	momen rekapitulasi hitung oleh suara nasional pemilihan umum 2024 jakarta	moment of recapitulation of calculations by the national vote of 2024 general elections in Jakarta
2	sikap pemilihan umum 2024 pimpin pusat muhammadiyah keluar 4 poin nyata	Election Attitude 2024 Leads of Muhammadiyah Center Out 4 Real Points
3	resmi pemilihan umum 2024 bangsa indonesia harap satu prabowo	Official General Election 2024 Indonesian Nation Hope Prabowo
4	soal pemilihan umum 2024 pimpin pusat muhammadiyah beri pesan sangat dampak publik	about the general election 2024 leading the center of Muhammadiyah give a message very public impact

Gambar 10. Hasil *Translate*

4.3. Pelabelan Data

Pada tahap ini, data tweet yang telah dikumpulkan diberi label berdasarkan sentimen

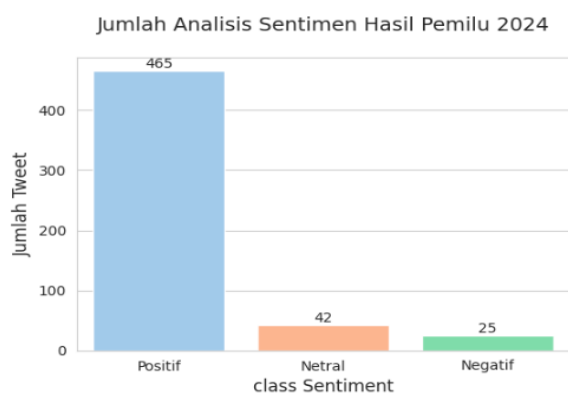
menggunakan pustaka *TextBlob*. Setiap tweet dianalisis untuk menentukan apakah sentimennya positif, netral, atau negatif. Proses ini menghasilkan kategori sentimen yang memudahkan analisis lebih lanjut dan membantu memahami distribusi sentimen dalam dataset. Berikut hasil analisis sentimen menggunakan *textblob* (Gambar 11).

```

Hasil Analisis Data:
Positif = 465 (87.41%)
Netral = 42 (7.89%)
Negatif = 25 (4.70%)
Total Data : 532
  
```

Gambar 11. Hasil Analisis Sentimen dengan *Textblob*

Berikut visualisasi analisis sentimen menggunakan *textblob* (Gambar 12).



Gambar 12. Visualisasi Hasil Analisis Sentimen menggunakan *Textblob*

4.4. Hasil Klasifikasi

Pada tahap ini, dilakukan pembuatan model *machine learning* dengan menggunakan *data training* dan *data testing* yang diambil secara acak dari keseluruhan dataset. Proses ini bertujuan untuk melakukan *cross validation* dan menghasilkan nilai prediksi untuk akurasi [14].

Klasifikasi menggunakan *Naïve Bayes Classifier* memperoleh akurasi sebesar 85%, presisi 82%, recall 85%, dan *f1-score* 80%. Proses ini mencakup *splitting data* latih dan uji, ekstraksi fitur, pelatihan model, evaluasi model, dan pembuatan *confusion matrix*. Berikut detail hasil evaluasi dengan menggunakan model *Naive Bayes Classifier* (Gambar 13).

```

Accuracy : 0.85
Classification Report:
              precision    recall  f1-score   support

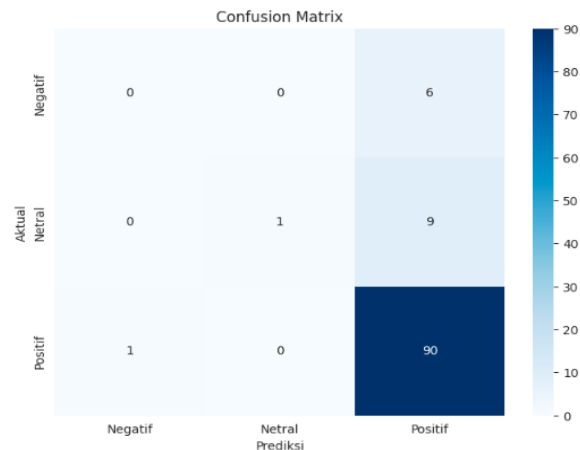
Negatif      0.00      0.00      0.00         6
Netral       1.00      0.10      0.18        10
Positif      0.86      0.99      0.92        91

accuracy      0.85
macro avg     0.62      0.36      0.37        107
weighted avg  0.82      0.85      0.80        107
  
```

Gambar 13. Hasil Evaluasi Model *Naïve Bayes*

Dan pada confusion matrix model *Naive Bayes*, hasil klasifikasinya memperoleh akurasi sebesar 85%,

presisi 82%, *recall* 85%, dan *f1-score* 80% (Gambar 14).



Gambar 14. *Confusion Matrix Naïve Bayes*

Sementara itu, klasifikasi menggunakan *KNN* memperoleh akurasi sebesar 88%, presisi 84%, *recall* 88%, dan *f1-score* 84%. Prosesnya serupa, mencakup *splitting data* latih dan uji, ekstraksi fitur, pelatihan model, evaluasi model, dan pembuatan *confusion matrix*. Berikut detail hasil evaluasi dengan menggunakan model *KNN* (Gambar 15).

```

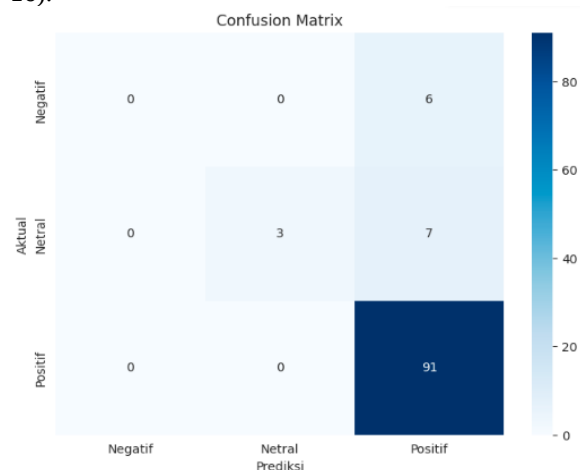
Accuracy : 0.88
Classification Report:
              precision    recall  f1-score   support

Negatif      0.00      0.00      0.00         6
Netral       1.00      0.30      0.46        10
Positif      0.88      1.00      0.93        91

accuracy      0.88
macro avg     0.62      0.43      0.46        107
weighted avg  0.84      0.88      0.84        107
  
```

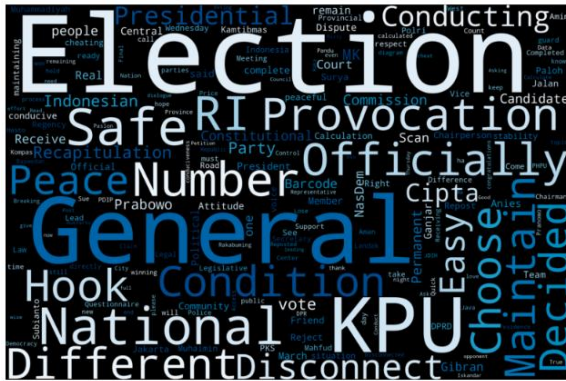
Gambar 15. Hasil Evaluasi Model *KNN*

Dan pada confusion matrix model *KNN*, hasil klasifikasinya memperoleh akurasi sebesar 88%, presisi 84%, *recall* 88%, dan *f1-score* 84% (Gambar 16).



Gambar 16. *Confusion Matrix KNN*

Hasil klasifikasi analisis sentimen terkait “Hasil Pemilu 2024” juga dapat dilihat melalui visualisasi dengan *wordcloud* seperti pada (Gambar 17).



Gambar 17. Hasil Wordcloud

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa analisis sentimen terkait hasil Pemilu 2024 dengan 532 data dari media sosial X menunjukkan 87.41% sentimen positif, 7.89% netral, dan 4.70% negatif. Ini menandakan bahwa mayoritas pengguna media sosial X memberikan tanggapan positif terhadap hasil Pemilu 2024. Algoritma KNN memiliki akurasi 88%, menunjukkan kemampuan tinggi dalam mengklasifikasikan sentimen dengan tepat, meskipun waktu komputasinya lebih tinggi dibandingkan algoritma lainnya. Di sisi lain, algoritma *Naive Bayes Classifier* mencapai akurasi 85%, menunjukkan performa yang baik meskipun sedikit lebih rendah dari KNN. Namun, kelemahan dari *Naive Bayes* terletak pada asumsi naif tentang independensi fitur, yang dapat menyebabkan estimasi kurang akurat ketika terdapat ketergantungan signifikan di antara fitur-fitur data.

Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi algoritma lain yang mungkin lebih efisien dalam hal waktu komputasi sambil tetap mempertahankan akurasi yang tinggi. Selain itu, dapat dilakukan analisis lebih mendalam mengenai ketergantungan antar fitur untuk mengatasi kelemahan pada algoritma *Naive Bayes*. Penelitian juga bisa diperluas dengan menggunakan dataset yang lebih besar dan lebih beragam untuk memastikan generalisasi hasil yang lebih baik.

DAFTAR PUSTAKA

- [1] M. R. Fahlevvi, "Analisis Sentimen Terhadap Ulasan Aplikasi Pejabat Pengelola Informasi Dan Dokumentasi Kementerian Dalam Negeri Republik Indonesia Di Google Playstore Menggunakan Metode Support Vector Machine," *J. Teknol. dan Komun. Pemerintah.*, vol. 4, no. 1, pp. 1–13, 2022, doi: 10.33701/jtkp.v4i1.2701.
- [2] I. Syahrohim, S. D. Saputra, R. W. Saputra, V. H. Pranatawijaya, and R. Priskila, "Perbandingan analisis sentimen setelah pilpres 2024 di Twitter menggunakan algoritma machine learning," *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, vol. 12, no. 2, pp. 1-9, 2024. DOI : 10.23960/jitet.v12i2.4249.
- [3] Vindua, R. and Zailani, A. U., "Analisis sentimen pemilu Indonesia tahun 2024 dari media sosial Twitter menggunakan Python," *JURIKOM (Jurnal Riset Komputer)*, vol. 10, no. 2, pp. 479-487, April 2023. DOI: 10.30865/jurikom.v10i2.5945.
- [4] L. Aji Andika And P. Amalia Nur Azizah, "Analisis Sentimen Masyarakat Terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 Pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier" *IJAS (Indonesian Journal of Applied Statistics)* 2019. DOI: <https://doi.org/10.13057/ijas.v2i1.29998>.
- [5] J. W. Iskandar dan Y. Natalian, "Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek," *J. RESTI (Rekayasa Sist. dan Teknol. Inf.)*, vol. 5, no. 6, pp. 1120–1126, 2021. DOI: <https://doi.org/10.29207/resti.v5i6.3588>.
- [6] D. P. Ray, F. N. Hasan, dan A. R. Dzirkillah, "Analisis Sentimen Terhadap KPU 2024 Berdasarkan Tweet Media Sosial Twitter Menggunakan Algoritma Naïve Bayes," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 4, hal. 2235-2243, Februari 2024. DOI: 10.30865/klik.v4i4.1587.
- [7] F. M. Sarimole and W. Septian, "Analisis Sentimen Masyarakat Terhadap Isu Penundaan Pemilu 2024 Pada Twitter Dengan Metode Naive Bayes Dan Support Vector Machine," *Jurnal Sains dan Teknologi*, vol. 5, no. 3, pp. 890-899, Februari 2024. DOI: <https://doi.org/10.55338/saintek.v5i1.2789>.
- [8] S. D. Prasetyo, S. Shofiah, dan F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *J. KomtekInfo*, vol. 10, no. 1, hal. 1-7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [9] N. T. Romadloni, I. Santoso, S. Budilaksono, "Perbandingan Metode Naive Bayes, KNN Dan Decision Tree Terhadap Analisis Sentimen Transportasi KRL Commuter Line", *Jurnal IKRA-ITH Informatika*, Vol.3, No.2, Juli 2019.
- [10] H. Yanto, "Sistem Pendukung Keputusan untuk Seleksi Usulan Pengajuan Sertifikasi Guru Menggunakan Algoritma K-Nearest Neighbor Berbasis Web" *J. KomtekInfo*, vol.5, no.2, pp. 42–50, 2018, doi: 10.35134/komtekinfo.v5i2.22.
- [11] E. W. Sholeha, S. Yunita, R. Hammad, V. C. Hardita, and Kaharuddin, "Analisis Sentimen Pada Agen Perjalanan Online Menggunakan Naïve Bayes dan K-Nearest Neighbor," *JTIM: Jurnal Teknologi Informasi dan Multimedia*, vol. 3, no. 4, pp. 203-208, 2022.
- [12] R. Syahputra, G. J. Yanris, and D. Irmayani, "SVM and Naïve Bayes Algorithm Comparison

- for User Sentiment Analysis on Twitter,” Sinkron, vol. 7, no. 2, pp. 671–678, 2022, doi: 10.33395/sinkron.v7i2.11430.
- [13] M. I. Fikri, T. S. Sabrila, dan Y. Azhar, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," SMATIKA, vol. 10, no. 02, pp. 71-76, Desember 2020. DOI: 10.32664/smatika.v10i02.455.
- [14] A. Nurian, T. N. Padilah, and Garno, "Analisis Sentimen Terhadap Pelayanan Disdukcapil Karawang Menggunakan Naive Bayes Classifier," Jurnal Informatika dan Teknik Elektro Terapan (JITET), vol. 12, no. 2, 2024. DOI : <http://dx.doi.org/10.23960/jitet.v12i2.4178>.