Jurnal Smart Teknologi
Vol. 4, No. 5, Juli 2023, Halaman 695 – 702
ISSN: 2774-1702, http://jurnal.unmuhjember.ac.id/index.php/JST

Analisis Sentimen *Tweet* Terhadap Isu Pencalonan Presiden 2024 Menggunakan Algoritma *Multivariate Bernoulli*

Tweet Sentiment Analysis of Candidacy Issues President 2024 Using Multivariate Bernoulli Algorithm

Rega Sukmawanti [1], Deni Arifianto[2*], Reni Umilasari[3]
[1] Mahasiswa Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember
Email : regasukmawanti05@gmail.com
[2] Dosen Fakultas Teknik, Universitas Muhammadiyah Jember *Koresponden Author
Email : deniarifianto@unmuhjember.ac.id
[3] Dosen Fakultas Teknik, Universitas Muhammadiyah Jember
Email : reni.umilasari@unmuhjember.ac.id

Abstrak

Di era teknologi media sosial yang semakin berkembang, informasi dapat tersebar dengan mudah sehingga mempengaruhi cara pandang dan gaya hidup masyarakat salah satunya yaitu *Twitter.* Pada *twitter* terdapat fitur trending topik yang bertujuan untuk membagikan informasi yang sedang diperbincangkan oleh banyak pengguna salah satunya yaitu tentang isu calon pemilihan Presiden. Isu pemilihan Presiden memicu opini publik dari berbagai masyarakat. Oleh sebab itu, penting untuk menganalisis sebuah *tweet* seputar isu-isu yang diutarakan masyarakat dengan menerapkan teknik analisis sentimen. Pada penelitian ini dalam pengolahan data menggunakan metode *Multivariate Bernoulli.* Tujuan dari penelitian ini yaitu mengetahui hasil akurasi, presisi dan *recall* serta mengetahui jumlah sentimen positif dan negatif. Hasil dari metode *Multivariate Bernoulli* menunjukkan bahwa algoritma *Multivariate Bernoulli* memiliki akurasi yang cukup tinggi dalam mengklasifikasikan sebuah analisis sentimen sehingga memperoleh akurasi tertinggi pada *fold*-3 langkah uji 2 dengan nilai 93.00 %, diikuti presisi dengan nilai tertinggi 91.67 % dan *recall* dengan nilai tertinggi 93.67 %. Serta jumlah data sentimen positif terdapat 78 data dan sentimen negatif terdapat 42 data.

Kata kunci : *Twitter*, Calon Presiden, Analisis Sentimen, *Multivariate Bernoulli*.

Abstract

In the era of increasingly developing social media technology, information can be spread easily so that it influences people's perspectives and lifestyles, one of which is Twitter. On Twitter, there is a trending topic feature that aims to share information that is being discussed by many users, one of which is the issue of presidential election candidates. The issue of the presidential election sparked public opinion from various communities. Therefore, it is important to analyze a tweet about issues expressed by the public by applying sentiment analysis techniques. In this study, the data processing used the Multivariate Bernoulli method. The purpose of this study is to know the results of accuracy, precision and recall and to know the number of positive and negative sentiments. The results of the Multivariate Bernoulli method show that the Multivariate Bernoulli algorithm has a fairly high accuracy in classifying a sentiment analysis so that it obtains the highest score in the fold-8 test step 2 with a value of 93.00 %, followed by precision with the highest value of 91.67 % and recall with the highest value of 93.67 %. As well as the number of positive sentiment data there are 78 data and negative sentiment data there are 42 data.

Keywords : Twitter, Presidential Candidate, Sentiment Analysis, Multivariate Bernoulli.

Jurnal Smart Teknologi
Vol. 4, No. 5, Juli 2023, Halaman 695 – 702
ISSN: 2774-1702, http://jurnal.unmuhjember.ac.id/index.php/JST

1. PENDAHULUAN

Twitter merupakan salah satu media sosial yang cukup popular dikalangan masyarakat pada *twitter* terdapat fitur *trending* topik yang bertujuan untuk membagikan informasi yang sedang diperbincangkan oleh banyak pengguna salah satunya yaitu tentang isu calon pemilihan Presiden. Pemilihan presiden di Indonesia merupakan momen penting dimana masyarakat Indonesia dapat melalui proses demokrasi (Saputra, dkk., 2022). Pemilihan Presiden juga memicu opini publik dari berbagai masyarakat. Oleh sebab itu, penting untuk menganalisis sebuah *tweet* seputar isu-isu yang diutarakan masyarakat dengan menerapkan teknik analisis sentimen.

Analisis sentimen adalah proses meninjau opini publik terhadap suatu topik tertentu yang datanya bersumber dari internet atau media sosial (Kurniawan & Waluyo, 2022). Menurut *Classifier & Bayes* (2021) analisis sentimen itu perlu dilakukan karena penggunaan media sosial yang tumbuh untuk mempengaruhi perkembangan opini publik. Pada analisis sentimen ini, penulis menggunakan tahapan ekstraksi fitur TF-IDF. TF-IDF adalah metode yang berguna untuk menghitung nilai bobot dalam sebuah teks serta dapat mengurangi ukuran teks sehingga menghindari ruang fitur yang besar. Jadi, untuk mengetahui performa dalam analisis sentimen tersebut, selain menggunakan ekstraksi fitur TF-IDF penulis juga menerapkan algoritma *Multivariate Bernoulli.*

Menurut Suparyanto & Rosad (2020) algoritma *Multivariate Bernoulli* adalah suatu metode perkembangan dari *Naïve Nayes*. Metode ini memiliki hipotesis yang kuat dalam mengklasifikasikan suatu kejadian analisis. Metode *Multivariate Bernoulli* memiliki akurasi yang cukup baik dalam pengklasifikasiannya. Dikarenakan banyaknya opini serta adu argumen antar warganet mengenai isu pencalonan Presiden. Oleh karena itu, penelitian ini berfokus untuk menganalisis *tweet* terhadap opini publik dengan menerapkan proses klasifikasi analisis sentimen dengan hasil positif dan negatif. Dengan demikian, peneliti dapat mengetahui akurasi, presisi, dan *recall* dari metode *Multivariate Bernoulli* serta jumlah sentimen pada *tweet* yang mengacu pada opini publik atas isu pencalonan Presiden 2024.

2. TINJAUAN PUSTKA

A. *Twitter*

Twitter adalah salah satu media sosial yang cukup populer di kalangan masyarakat yang memungkinkan penggunanya untuk membagikan *tweet* yang disebut dengan *tweet* dengan batasan maksimal 280 huruf dalam satu paragraf (Kurniawan & Waluyo, 2022).

B. *Crawling* data

Crawling data merupakan proses pengambilan atau pengunduhan data dari *Twitter* dengan menggunakan *Application Program Integration* (API) *Twitter,* baik berupa data pengguna maupun data *Tweet* (Sembodo, dkk., 2016). Crawling data pada T*witter* terdapat 2 cara pencarian yaitu *by user* dan *by keyword.*

C. Analisis Sentimen

Analisis sentimen dikenal dengan *opinion mining* yang merupakan proses memahami, mengekstraksi, dan mengolah data secara otomatis untuk memperoleh informasi sentimen yang terkandung dalam sebuah kalimat (Putri & Muzakir, 2022).

D. *Text mining*

Text mining adalah proses penambangan yang dilakukan oleh komputer untuk mendapatkan sesuatu yang baru dan sebelumnya tidak diketahui atau untuk memulihkan informasi implisit. Hasil yang diperoleh berasal dari data yang diekstrak secara otomatis dari berbagai sumber data teks (Kurniawan & Waluyo, 2022)

E. *Text Pre-pocessing*

Text Pre-pocessing digunakan untuk memproses data yang telah disiapkan untuk tahapan selanjutnya. Dalam *Pre-processing*

terdapat beberapa tahapan yaitu: *cleansing, case folding, tokenizing, normalization, filtering,* dan *stemming.*

a. *Cleansing*: Pada tahap ini, bagian tertentu dari *tweet* dihapus seperti simbol, karakter, emotikon, dan tautan URL (Kurniawan & Waluyo, 2022).
b. *Case Folding*: Tahapan mengonversi semua huruf sebelumnya menjadi huruf kecil (Putri & Muzakir, 2022).
c. *Tokenizin*g: Mengubah kalimat menjadi setiap kata yang membentuk kalimat tersebut. Misalnya pada kalimat "aku pergi sebentar" maka akan menjadi "aku", "pergi", "sebentar" (Ashari, dkk., 2020).
d. *Normalization*: Proses mengubah teks dokumen dari kata atau singkatan yang tidak sesuai menjadi kata yang bermakna (Ashari, dkk., 2020).
e. *Filtering*: Proses menghilangkan kata-kata yang kurang deskriptif atau kata-kata yang tidak memiliki arti (Ashari, dkk., 2020).
f. *Stemming*: Bertujuan untuk mengubah akhiran dari setiap kata yang disaring menjadi kata dasar (Ashari, dkk., 2020).

F. Metode TF-IDF

Term Frequency (TF) adalah metode yang menghitung bobot setiap kata dalam teks. Metode ini mengasumsikan bahwa nilai kepentingan dari setiap ekspresi sebanding dengan berapa kali ekspresi tersebut muncul dalam teks. *Inverse Document Frequency* (IDF) yang mengurangi dominasi *Term Frequency* dalam beberapa data tekstual (Azizah, dkk., 2022). Berikut tahapan menghitung bobot TF-IDF (Ashari, dkk., 2020):

$$W_{dt} = tf_d \times idf_t = tf_d \times log\,(\frac{N}{df_t})$$

Keterangan:
W_{dt} = Nilai bobot *term* ke-t pada dokumen d
tf_d = Jumlah munculnya *term* t pada dokumen d
N = Jumlah dokumen yang mengandung *term* t
df_t = Jumlah dokumen secara keseluruhan

G. *Multivariate Bernoulli*

Multivariate Bernoulli merupakan salah satu model algoritma klasifikasi yang dikembangkan dari algoritma *Naive Bayes Classifier*, dimana algoritma ini cocok untuk mengklasifikasikan teks atau dokumen. Berikut merupakan persamaan pada algoritma *Multivariate Bernoulli* (Ashari, dkk., 2020):

$$P\,(c|d) = P(c) \times \prod_{i=1}^{N} P(fk|c) \times \prod_{i=1}^{M}\big(1 - P(fk|c)\big)$$

Keterangan:
$P\,(c|d)$ = Probabilitas dokumen kelas c
$P(c)$ = Probabilitas prior kelas c
$p(fk|c)$= probabilitas setiap kata

Berikut langkah perhitungan untuk menentukan prior dari kelas c:

$$P(c) = \frac{Nc}{N}$$

Keterangan:
Nc = Banyak kelas c pada seluruh dokumen
N = Banyak seluruh dokumen

Berikut langkah perhitungan probabilitas kata ke-n pada kelas c:

$$P(fk|c) = \frac{T_{ct}+1}{T_c+\sum c}$$

Keterangan:
T_{ct} = Banyaknya dokumen yang mengandung *term* pada kelas c
T_c = Jumlah data latih pada setiap kelas c
$\sum c$ = Jumlah kelas atau banyaknya kategori

H. *K-Fold Cross Validation*

Fold Cross Validation adalah salah satu metode untuk mengevaluasi kinerja sebenarnya dari model pembelajaran mesin. Pada penelitian ini, proses validasi dilakukan dengan menggunakan metode *K-Fold Cross Validation* untuk menghilangkan bias atau *noise* kata sehingga dapat meningkatkan akurasi

I. *Confusion Matrix*

Confusion Matrix adalah metode evaluasi hasil yang diperoleh pada implementasi algoritma klasifikasi menggunakan tabel (Grandis & Arumsari, 2021). Berikut tabel *Confusion Matrix*:

Jurnal Smart Teknologi
Vol. 4, No. 5, Juli 2023, Halaman 695 – 702
ISSN: 2774-1702, http://jurnal.unmuhjember.ac.id/index.php/JST

Tabel 1. *Confussion Matrix*

Kelas Prediksi	Kelas Aktual	
	Positif (+)	Negatif (-)
Positif (+)	TP	FP
Negatif (-)	FN	TN

Sumber: Jurnal

Keterangan:

a. TP (True Positive) adalah jumlah dari kelas positif yang diklasifikasikan benar.
b. FP (False Positive) adalah jumlah dari kelas negatif yang diklasifikasikan dalam kelas positif.
c. FN (False Negative) adalah jumlah dari kelas positif yang diklasifikasikan dalam kelas negatif.
d. TN *(True Negative)* adalah jumlah dari kelas negatif yang diklasifikasikan benar.

Nilai-nilai dalam tabel tersebut kemudian digunakan untuk menghitung akurasi, presisi, dan *recall.*

1. Akurasi adalah hasil membandingkan nilai prediksi yang benar dengan nilai sebenarnya (Grandis & Arumsari, 2021). Persamaan akurasi sebagai berikut:
$$\frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\%$$
2. Presisi adalah hasil dari membandingkan nilai prediksi sebenarnya dari data positif dan nilai total data positif (Grandis & Arumsari, 2021). Persamaan presisi sebagai berikut:
$$\frac{TP}{(TP+FP)} \times 100\%$$
3. *Recall* adalah hasil membandingkan nilai prediktif yang benar dari data positif dan total nilai prediktif yang benar (Grandis & Arumsari, 2021). Persamaan *recall* sebagai berikut:
$$\frac{TP}{(TP+FN)} \times 100\%$$

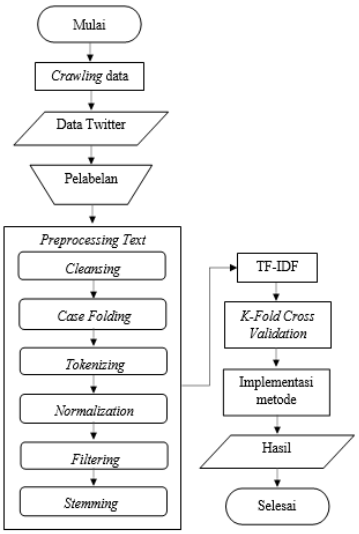
J. *Python*

Python adalah bahasa pemrograman yang dapat mengeksekusi sejumlah intruksi secara langsung dengan metode *Object Oriented Programming* (OOP) serta dapat memberikan tingkat keterbacaan *syntax. Python* dapat dipelajari dengan mudah karena telah dilengkapi dengan manajemen memori otomatis (Rahmadhika & Thantawi, 2021).

K. *Google Colab*

Google colab atau *Google Collaboratory* adalah dokumen yang dapat dieksekusi yang dapat digunakan untuk menyimpan, menulis, dan berbagi program tertulis melalui *Google Drive*. Perangkat lunak ini pada dasarnya mirip dengan *Jupyter Notebook* berbasis *cloud* gratis yang dapat diberjalan melalui *browser* (Sitio, dkk., 2022).

3. METODOLOGI PENELITIAN



Gambar 1. Diagram Alir Penelitian
Sumber: Alir Penelitian

A. Pengumpulan Data

Pada pengumpulan data ini melakukan tahapan crawling data. *crawling* data adalah proses pengambilan atau pengunduhan data dari server *twitter* dengan teknik *crawling* data menggunakan *Application Programming Integration* (API), pada data terdapat data pengguna, tanggal dan *tweet.*

B. Pelabelan Data

Setelah melakukan tahapan pengumpulan data, kemudian dilakukan proses pelabelan data yang terbagi menjadi 2 kelas yaitu positif dan negatif. Pelabelan data dilakukan secara manual dan divalidasi oleh Guru mata pelajaran Bahasa Indonesia.

Jurnal Smart Teknologi
Vol. 4, No. 5, Juli 2023, Halaman 695 – 702
ISSN: 2774-1702, http://jurnal.unmuhjember.ac.id/index.php/JST

Tabel 2. Hasil Pelabelan Data

Data	Positif	Negatif
Hasil *Crawling*	300	300
Jumlah	600	

Sumber: Hasil Pelabelan Data

4. HASIL DAN PEMBAHASAN

a) *Text Pre-pocessing*

Text Pre-pocessing digunakan untuk memproses data yang telah disiapkan untuk tahapan selanjutnya. Dalam *Pre-processing* terdapat beberapa tahapan yaitu: *cleansing, case folding, tokenizing, normalization, filtering,* dan *stemming*.

b) Pembobotan Kata

Proses pembobotan kata menggunakan fitur F-IDF. TF-IDF adalah proses mengubah kata menjadi numerik.

c) Pembagian Data

Pada pembagian data ini menggunakan 4 macam *K-Fold cross validation* yaitu 2-*fold*, 3-*fold*, 4-*fold* dan 5-*fold*. *K-Fold cross validation* tersebut akan membagi seluruh data menjadi bagian-bagian seperti data uji dan data latih.

- **Penentuan Skenario Pada Algoritma**

```
x = data['Tweet']
y = data['Sentimen']

kf=KFold(n_splits=2)
for trainingIndex,
testingIndex in kf.split(x,
y):
```

```
X_train = x[trainingIndex]
    y_train =
y[trainingIndex]
    X_test = x[testingIndex]
    y_test = y[testingIndex]
```

```
X_train = x[trainingIndex]
    y_train =
y[trainingIndex]

    X_test =x[testingIndex]
    y_test =y[testingIndex]
    y_test =y[testingIndex]
```

```
X_train = x[trainingIndex]
    y_train =
y[trainingIndex]

    X_test =x[testingIndex]
    y_test =y[testingIndex]
    y_test =y[testingIndex]
```

```
x = data['Tweet']
y = data['Sentimen']


i = 1
bernoulliAcc = 0
bernoulliPre = 0
bernoulliRec = 0

kf=KFold(n_splits=2,
shuffle=True,
random_state=25)
for trainingIndex,
testingIndex in kf.split(x,
y):
    X_train =
x[trainingIndex]
    y_train =
```

```
    X_test = x[testingIndex]
    y_test = y[testingIndex]

    Tfidf_vect =
TfidfVectorizer(max_features=5000)
    Tfidf_vect.fit(data['Tweet'])
    Train_X_Tfidf =
Tfidf_vect.transform(X_train)

    Test_X_Tfidf =
Tfidf_vect.transform(X_test)

    model_Bernoulli = BernoulliNB()

model_Bernoulli.fit(Train_X_Tfidf,y_tr
ain)

    predictions_Bernoulli =
model_Bernoulli.predict(Test_X_Tfidf)

    print('HASIL ',i)
    print()
    tes1 =
classification_report(y_test,
predictions_Bernoulli,
output_dict=True)
     print(confusion_matrix(y_test,
predictions_Bernoulli))
    print()
    print('accuracy
\t:',"{:.2f}".format((tes1['accuracy']
)*100),'%')
    print('presisi
\t:',"{:.2f}".format((tes1['0']['recal
l'])*100),'%')
    print('recall
\t\t:',"{:.2f}".format((tes1['0']['pre
cision'])*100),'%')
     bernoulliAcc =
bernoulliAcc+tes1['accuracy']
    bernoulliPre =
bernoulliPre+tes1['0']['precision']
     bernoulliRec =
bernoulliRec+tes1['0']['recall']
     print()
```

d) Rekapitulasi Hasil Akurasi, Presisi dan *Recall*

Berdasarkan pengujian "Analisis Sentimen *Tweet* terhadap Isu Pencalonan Presiden 2024 menggunakan algoritma *Multivariate Bernoulli"* serta menggunakan fitur TF-IDF maka memperoleh hasil akurasi, presisi, dan *recall* sebagai berikut:

Tabel 3. Rekapitulasi Hasil Akurasi, Presisi dan *Recall*

K-Fold	**Langkah Uji**	**Akurasi**	**Presisi**	***Recall***
2-*Fold*	Langkah Uji 1	88.33 %	84.62 %	90.30 %
	Langkah Uji 2	82.00 %	66.24 %	99.05 %
	Rata-Rata	85.17 %	75.43 %	94.67 %
3-Fold	Langkah Uji 1	88.00 %	81.44 %	92.94 %
	Langkah Uji 2	93.00 %	91.67 %	93.62 %
	Langkah Uji 3	85.00 %	73.83 %	97.53 %
	Rata-Rata	88.67 %	82.31 %	94.70 %
4-Fold	Langkah Uji 1	90.67 %	87.67 %	92.75 %
	Langkah Uji 2	90.00 %	84.29 %	93.65 %
	Langkah Uji 3	90.67 %	85.54 %	97.26 %
	Langkah Uji 4	86.00 %	72.97 %	98.18 %
	Rata-Rata	89.33 %	82.62 %	95.46 %
5-*Fold*	Langkah Uji 1	90.00 %	88.89 %	91.80 %
	Langkah Uji 2	90.83 %	79.25 %	100,00 %
	Langkah Uji 3	90,83 %	91,07 %	89,47 %
	Langkah Uji 4	90.83 %	85,29 %	98.31 %
	Langkah Uji 5	85.83 %	73.33 %	97.78 %
	Rata-Rata	89.67 %	83.57 %	95.47 %

Sumber: Hasil Perhitungan
Keterangan :

: Nilai tertinggi pada akurasi diikuti dengan presisi dan *recall*

Berdasarkan pada Tabel 3. hasil rekapitulasi memperoleh akurasi, presisi dan *recall* dengan akurasi tertinggi pada *fold* ke-3 langkah uji 2 yaitu akurasi 93.00 %, presisi 91,67 % dan *recall* 93.67 %.

Jurnal Smart Teknologi
Vol. 4, No. 5, Juli 2023, Halaman 695 – 702
ISSN: 2774-1702, http://jurnal.unmuhjember.ac.id/index.php/JST

Tabel 4. Hasil Kriteria pada *Fold*-3 Langkah Uji 2

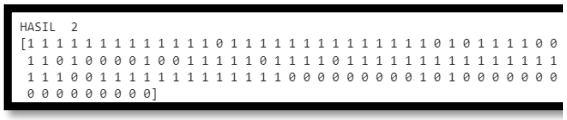
Kriteria	Jumlah
TP	42
TN	67
FP	11
FN	0

Sumber: Hasil Perhitungan

1. Akurasi = $\frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\%$
   $= \frac{(42+67)}{(42+67+11+0)} \times 100\%$
   $= 93{,}00\ \%$
2. Presisi $= \frac{TP}{(TP+FP)} \times 100\%$
   $= \frac{42}{(42+11)} \times 100\%$
   $= 91{,}67\ \%$
3. *Recall* $= \frac{TP}{(TP+FN)} \times 100\%$
   $= \frac{42}{(42+0)} \times 100\%$
   $= 93{,}67\ \%$

e) Jumlah Sentimen Positif dan Sentimen Negatif

Berdasarkan nilai akurasi tertinggi terdapat pada *fold* ke-3 langkah uji 2 diikuti presisi dan *recall*. Maka, berikut jumlah hasil sentimen positif dan sentimen negatif:



Gambar 3. Hasil Sentimen Positif dan Sentimen Negatif

Sumber: Hasil Perhitungan

Keterangan:

1 : Sentimen positif
0 : Sentimen negatif

Berdasarkan Gambar 3. jumlah data pada penelitian yang "Analisis Sentimen *Tweet* terhadap Isu Pencalonan Presiden 2024 menggunakan algoritma *Multivariate Bernoulli*" terdapat sentimen positif 78 data dan sentimen negatif terdapat 42 data yang diperoleh dari jumlah *fold* ke-3 langkah uji 2.

5. KESIMPULAN DAN SARAN

A. KESIMPULAN

Berdasarkan penelitian yang sudah dilakukan pada "Analisis Sentimen *Tweet* terhadap Isu Pencalonan Presiden 2024 menggunakan algoritma *Multivariate Bernoulli"* dapat disimpulkan sebagai berikut:

1. Nilai akurasi, presisi dan *recall* yang dihasilkan dari metode *Multivariate Bernoulli* pada "Analisis Sentimen *Tweet* terhadap Isu Pencalonan Presiden 2024 menggunakan algoritma *Multivariate Bernoulli"* menunjukkan bahwa algoritma *Multivariate Bernoulli* memiliki akurasi yang cukup tinggi dalam mengklasifikasikan sebuah analisis sentimen sehingga memperoleh akurasi tertinggi pada *fold*-3 langkah uji 2 dengan nilai 93.00 %, diikuti presisi dengan nilai tertinggi 91.67 % dan *recall* dengan nilai tertinggi 93.67 %.
2. Berdasarkan nilai akurasi tertinggi terdapat pada *fold* ke-3 langkah uji 2 diikuti presisi dan *recall* maka jumlah data sentimen positif terdapat 78 data dan sentimen negatif terdapat 42 data.

B. SARAN

Saran yang dapat penulis berikan untuk penelitian selanjutnya adalah sebagai berikut:

1. Diharapkan untuk peneliti selanjutnya menggunakan metode klasifikasi lainnya seperti *Multinomial Naive Bayes* dan *Support Vector Machine* serta menggunakan data yang lebih banyak agar mendapatkan hasil yang bervariasi.
2. Diharapkan untuk peneliti selanjutnya agar menggunakan data berita terbaru untuk mendapatkan hasil yang bervariasi.

6. DAFTAR PUSTAKA

Ashari, H., Arifianto, D., &Al Faruq, H. A.(2020). Perbandingan Kinerja Algoritma Multinominal Naive Bayes (MNB, Multivariate Bernoulli dan Rocchio Algortihm Dalam Klasifikasi Konten Berita Hoax Berbahasa Indonesia Pada Media Sosial. *Http://Repository.Unmuhjember.Ac.Id*, 1–12. Diakses pada 1 Februari 2023

Grandis, G. F., & Arumsari, Y. (2021). Seleksi Fitur Gain Ratio pada Analisis Sentimen Kebijakan Pemerintah Mengenai Pembelajaran Jarak Jauh dengan K-Nearest Neighbor. Jurnal Pengembangan Teknologi

Informasi Dan Ilmu Komputer, 5(8), 3507–3514. Diakses pada 9 Januari 2023

Kurniawan, A., & Waluyo, S. (2022). Penerapan Algoritma Naive Bayes Dalam Analisis Sentimen Pemindahan Ibukota Pada Twitter Application Of Naive Bayes Algorithm In Capital Movement Sentiment Analysis On Twitter. September, 455–461. Diakses pada 9 Januari 2023

Putri, A., & Muzakir, A. (2022). Analisis Sentimen Cyberbullying KPOP di Media Sosial Twitter Menggunakan Metode Naïve Bayes. https://jurnal.syntaxliterate.co.id/index.php/syntax-. Diakses pada 9 Januari 2023

Rahma, A. F., Agussalim, &Kartika, D. S. Y.(2021). Analisis Sentimen Hashtag Kuliner Di Indonesia Menggunakan Naive Bayes. *Jurnal Informatika Dan Sistem Informasi*, *2*(1), 19–25. https://doi.org/10.33005/jifosi.v2i1.282. Diakses pada 9 Januari 2023

Rahmadhika, M. K., & Thantawi, A. M. (2021). Rancang Bangun Aplikasi Face Recognition Pada Pendekatan CRM Menggunakan Opencv Dan Algoritma Haarcascade. *IKRA-ITH INFORMATIKA: Jurnal Komputer Dan Informatika*, *5*(1), 109–118. Diakses pada 9 Januari 2023

Rufaidha, N. F., & Irhandayaningsih, A. (2022). Perilaku Informasi Mahasiswa Fakultas Ilmu Budaya Universitas Diponegoro dalam Pemanfaatan Fitur Trending Topic Twitter Sebagai Pemenuhan Kebutuhan Informasi Abstrak. *6*(4), 493–504.*http://ejournal.undip.ac.id/index.php/anuva.* Diakses pada 9 Januari 2023

Sitio, A., Sindar, A., Marbun, M., Tiara, D., & Aswin, A. (2022). Pengenalan Data Scientist Pada Peserta PKBM AL HABIB Melalui Belajar Dasar Coding Python. *Jurnal Pengabdian Pada Masyarakat*, *7*(1), 194–200. https://doi.org/10.30653/002.202271.44. Diakses pada 9 Januari 2023

Suparyanto&Rosad. (2020). Deteksi Hoax Pada Berita Online Bahasa Inggris Menggunakan Bernoulli Naïve Bayes Dengan Ekstraksi Fitur TF-IDF.*Journal syntax admiration*,5(3), 248-253. https://jurnalsyntaxadmiration.com/index.php/jurnal/article/view/327. Diakses pada 9 Januari 2023