

Klasifikasi Komentar Publik Dalam Pemilihan Umum Presiden 2024 Dengan Menggunakan Metode Naive Bayes: Studi Kasus Identifikasi *Haters* Dan *Non-Haters*

Klasifikasi Komentar Publik Dalam Pemilihan Umum Presiden 2024 Dengan Menggunakan Metode Naive Bayes: Studi Kasus Identifikasi *Haters* Dan *Non-Haters*

Dina Imania

Program Studi S1 Sistem Informasi Fakultas Teknologi Informasi Universitas Hasyim Asy'ari Tebuireng Jombang
e-mail: Dinainmania10@gmail.com

Anita Andriani

Program Studi S1 Teknik Informatika Fakultas Teknologi Informasi Universitas Hasyim Asy'ari Tebuireng Jombang
e-mail: Anita.unhasy@gmail.com

Mahrus Ali

Program Studi S1 Sistem Informasi Fakultas Teknologi Informasi Universitas Hasyim Asy'ari Tebuireng Jombang
e-mail: mahrusali1606@gmail.com

Abstrak

Indonesia merupakan salah satu Negara yang menganut sistem demokrasi. Sistem pemerintahan demokrasi dengan mengedepankan rakyat. Sehingga saat dilakukannya pemilihan para wakil rakyat, rakyat memiliki hak tertinggi untuk melalui pemilihan yang berlangsung secara bebas. Pada tahun 2024 ini akan dilakukan pemilihan presiden Indonesia. Setiap momen pilpres atau pemilihan umum presiden banyak opini dari masyarakat untuk calon presiden yang diisukan nantinya. Namun pada praktiknya banyak terdapat komentar-komentar negatif yang bersifat "*hate speech*" yang dapat memicu konflik sosial dan merusak lingkungan politik. Penelitian ini bertujuan untuk membangun sebuah model klasifikasi komentar publik menjadi kategori "*haters*" dan "*non haters*" dalam bahasa Indonesia menggunakan algoritma Naïve Bayes. Penelitian ini menggunakan data yang berjumlah 1000 dengan rincian data latih 800 data dan data uji 200 data. Dengan tahapan pengumpulan data, pre-processing data, klasifikasi dengan Naïve Bayes, *Evaluation*, dan *Deployment*. Hasil penelitian diperoleh nilai akurasi sebesar 83,5%, nilai *precision*/presisi sebesar 82% dan nilai *recall* sebesar 90%.

Kata Kunci: Pilpres, Klasifikasi, Naive Bayes

Abstract

Indonesia is one of the countries that adheres to a democratic system. The democratic system of government prioritizes the people. So that when the election of the people's representatives is carried out, the people have the highest right to go through elections that take place freely. In 2024, the Indonesian presidential election will be held. Every moment of the presidential election or presidential general election, there are many opinions from the public about the rumored presidential candidate. However, in practice, there are many negative comments that are "hate speech," which can trigger social conflict and damage the political environment. This research aims to build a classification model of public comments into "haters" and "non-haters" categories in Indonesian using the Naïve Bayes algorithm. This research uses data totaling 1000, with details of training data of 800 and test data of 200. With the stages of data collection, data pre-processing, classification with Naïve Bayes, *Evaluation*, and *Deployment*. The results obtained an accuracy value of 83.5%, a precision value of 82%, and a recall value of 90%.

Keywords: Presidential Election; Classification; Naive Bayes.

PENDAHULUAN

Indonesia merupakan Negara yang menganut sistem demokrasi. Secara etimologis atau asal-usulnya, kata demokrasi berasal dari bahasa Yunani *demos* yang berarti rakyat dan *cratein* atau *cratos* yang berarti kekuasaan dan kedaulatan. Dari kata tersebut membentuk kata demokrasi yang dapat disimpulkan sebagai pemerintahan rakyat yang artinya rakyat memegang kekuasaan tertinggi. Sistem pemerintahan demokrasi dengan mengedepankan rakyat. Sehingga saat dilakukannya pemilihan para wakil rakyat, rakyat memiliki hak tertinggi untuk melalui pemilihan yang berlangsung secara bebas. (Rohman, 2020).

Pemilihan wakil rakyat atau disebut dengan presiden merupakan suatu proses yang penting dalam sistem demokrasi dan diselenggarakan secara periodik. Di Indonesia dilakukan pemilihan presiden sebanyak lima tahun sekali dan itu bertepatan pada tahun 2024. Tokoh politik yang ingin mencalonkan sebagai presiden Indonesia, tentunya akan melihat dan

Klasifikasi Komentar Publik Dalam Pemilihan Umum Presiden 2024 Dengan Menggunakan Metode Naive Bayes: Studi Kasus Identifikasi *Haters* Dan *Non-Haters*

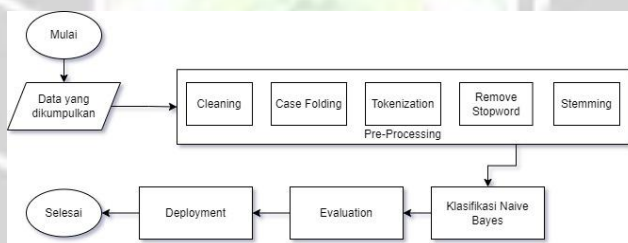
mempertimbangkan popularitas yang berdasarkan opini masyarakat, khususnya pada era yang modern ini. Salah satu opini masyarakat disampaikan pada jejaring sosial melalui Twitter dan Facebook. Dalam pemilihan umum presiden, masyarakat dapat memberikan opini-opini dan sentimen terhadap kandidat yang akan diprediksikan akan maju sebagai calon presiden. Komentar-komentar yang muncul pada media sosial khususnya Facebook dan Twitter seringkali mencerminkan pendapat dan sikap masyarakat terhadap kandidat tertentu. Namun pada praktiknya banyak terdapat komentar-komentar negatif yang bersifat “*hate speech*” yang dapat memicu konflik sosial dan merusak lingkungan politik. Para pelaku penyebar “*hate speech*” dikenal dengan sebutan *haters*. *Haters* secara harfiah berasal dari kata bahasa Inggris yang memiliki arti orang yang membenci (<http://artikata.com>).

Oleh karena itu, diperlukan suatu analisis sentimen pada komentar-komentar yang muncul pada media sosial untuk membedakan antara komentar-komentar positif dan negatif, serta untuk mengidentifikasi komentar-komentar publik yang bersifat “*hate speech*”. Metode Naive Bayes adalah metode klasifikasi dari banyaknya metode yang dapat digunakan untuk mengklasifikasikan komentar-komentar pada media sosial menjadi kategori-kategori tertentu, termasuk kategori “*haters*” dan “*non haters*”. Metode ini memanfaatkan teori peluang atau probabilitas yang ditemukan oleh ilmuwan Inggris yaitu Thomas Bayes. Metode ini mempunyai cara kerja dengan memprediksi peluang terjadinya kejadian dimasa depan berdasarkan data yang didapat sebelumnya (Nurhidayati dkk., 2023).

Dari pemaparan di atas, maka pada penelitian ini penulis menggunakan judul “Klasifikasi Komentar Publik Dalam Pemilihan Umum Presiden 2024 Dengan Menggunakan Metode Naive Bayes: Studi Kasus Identifikasi *Haters* Dan *Non-Haters*”. Penelitian ini akan mengumpulkan data berupa komentar pada media sosial Twitter dan Facebook. Data yang dikumpulkan berupa komentar publik yang berhubungan dengan calon presiden yang diprediksi akan maju pada pemilihan presiden tahun 2024. Dataset yang terbentuk akan dilakukan pelabelan secara manual dan menggunakan metode Naive Bayes dalam proses pengklasifikasi komentar ke dalam kelompok *haters* dan *non haters*.

METODE

Tahapan penelitian yang dilakukan digambarkan pada *flowchart* di bawah ini. Pengujian ini menggunakan bantuan bahasa pemrograman python.



Gambar 1. Flowchart Penelitian

Berikut penjelasan dari *flowchart* penelitian pada Gambar 1 adalah seperti berikut ini:

a. Data yang dikumpulkan

Tahapan pertama yaitu melakukan pengambilan data, data yang digunakan pada penelitian ini berjumlah 1000 data yang nantinya akan dipisah menjadi data latih (*data training*) dan data uji (*data testing*). Penelitian ini menggunakan dataset satu publik figur yang diisukan menjadi kandidat calon presiden pada pemilihan umum Presiden 2024 berdasarkan survei terhadap elektabilitas tertinggi calon presiden (*capres*) yang dilakukan oleh Lembaga survei Indekstat yaitu Ganjar Pranowo. Dapat dilihat pada website berikut ini, <https://news.detik.com/pemilu/d-6390705/survei-capres-indekstat-ganjar-165-prabowo-116-anies-99>. Berikut hasil survei lima teratas elektabilitas *capres*.

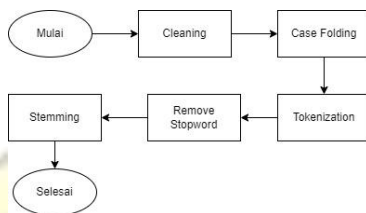
Tabel 1. Daftar 5 Publik Figur

Nama Figur	Publik	Jumlah Persentase
Ganjar Pranowo		16,5%
Prabowo Subianto		11,6%
Anies Baswedan		9,9%
Ridwan Kamil		2,3%
Puan Maharani		1,6%

Klasifikasi Komentar Publik Dalam Pemilihan Umum Presiden 2024 Dengan Menggunakan Metode Naive Bayes: Studi Kasus Identifikasi *Haters* Dan *Non-Haters*

b. Pre-Processing

Data yang dikumpulkan masih berbentuk data yang tidak terstruktur oleh karena itu, tahapan *Pre-processing* bertujuan untuk menjadikan data yang didapat tadi menjadi lebih terstruktur dan dapat meningkatkan kualitas data latih (Arum Sari & Ali Fauzi, 2019). Gambar 3.2 menjelaskan tahapan *pre-processing* data yang dilakukan pada penelitian ini



Gambar 2. Tahapan Pre-Processing Data

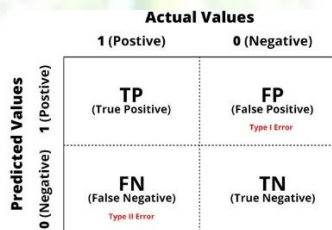
c. Klasifikasi Naive Bayes

Setelah melakukan *pre-processing* data, data melakukan *training* dengan jumlah data yang lebih besar dibandingkan dengan data uji atau data testing. Pada pembagian dataset tergantung kebutuhan. (Id, 2021:106). Salah satu pendekatan yang digunakan pada pembagian data *training* dan data uji adalah metode “*Pareto Principle*” yang dikenal sebagai hukum 80/20 (Islam, 2006:131). Sehingga jumlah data *training* atau data latih yang dipakai pada penelitian ini adalah 80% dari jumlah data yang ada dan sisanya digunakan untuk data testing 20%. Berikut Persamaan (1) Teorema Naive Bayes:

$$P(C | X) = \frac{P(X | C)P(C)}{P(X)} \quad (1)$$

d. Evaluation

Pada tahap ini, evaluasi model melakukan pengukuran performa model yang dihasilkan pada proses uji atau *training* dengan menggunakan data *testing* atau data uji dengan metode evaluasi *Confusion matrix* yang dimana metode ini dapat menghitung nilai Akurasi, nilai Presisi, dan nilai Recall. *Confusion matrix* menampilkan ringkasan dari semua hasil prediksi yang dihasilkan untuk mempermudah proses perhitungan dengan membandingkan antara hasil prediksi dengan hasil yang aktual.



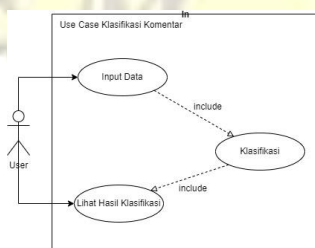
Gambar 3. Predicted Values

e. Deployment

Tahapan *deployment* adalah implementasi model *machine learning* atau mengintegrasikan model *machine learning* ke dalam lingkungan yang ada untuk membuat keputusan praktis berdasarkan data. Sederhananya, hasil dari tahapan yang dilakukan sebelumnya diimplementasikan ke sebuah aplikasi atau sistem antarmuka untuk diterapkan ke lingkungan langsung

1. Use Case

Use Case berguna untuk lebih memahami fungsionalitas sistem.



Gambar 4. Use Case Sistem

2. Activity Diagram

Diagram aktivitas berfungsi untuk menggambarkan alur implementasi *machine learning* untuk klasifikasi komentar.

Klasifikasi Komentar Publik Dalam Pemilihan Umum Presiden 2024 Dengan Menggunakan Metode Naive Bayes: Studi Kasus Identifikasi *Haters* Dan *Non-Haters*

	Komentar	Label	hasil_cleaning
0	Hanya orang bodoh yg dukung , oportunis sesat...	Haters	Hanya orang bodoh yg dukung oportunis sesat...
1	Logika bodoh	Haters	Logika bodoh
2	Dukungan untuk pak Ganjar Pranowo jadi Capres...	Non Haters	Dukungan untuk pak Ganjar Pranowo jadi Capres...
3	Sekali melangkah, pantang mundur sebelum Ganjar...	Non Haters	Sekali melangkah pantang mundur sebelum Ganjar...
4	Saya salut pada keberanian Pak Ganjar Pranowo ...	Non Haters	Saya salut pada keberanian Pak Ganjar Pranowo ...
...
995	Ganjar Pranowo tetap unggul jadi Capres 2024	Non Haters	Ganjar Pranowo tetap unggul jadi Capres
996	Tenanglanang Ganjar bim resmi dicalonkan saja ...	Non Haters	tenanglanang ganjar bim resmi dicalonkan saja ...
997	Ganjar ini prestasi nya apa? Kok bisa jadi capr...	Haters	ganjar ini prestasi nya apa kok bisa jadi capr...
998	Maafkan kami sudah tidak ada pilihan lagi di G...	Haters	maafkan kami sudah tidak ada pilihan lagi di g...
999	Kecawa dengan GANJAR	Haters	Kecawa dengan GANJAR

Gambar 8. Hasil Cleaning

b. Case Folding

Tahapan ini mengubah semua kata menjadi *lowercase* atau huruf kecil.

	hasil_cleaning	hasil_casefolding
0	Hanya orang bodoh yg dukung oportunis sesat...	hanya orang bodoh yg dukung oportunis sesat...
1	Logika bodoh	logika bodoh
2	Dukungan untuk pak Ganjar Pranowo jadi Capres...	dukungan untuk pak ganjar pranowo jadi capres...
3	Sekali melangkah, pantang mundur sebelum Ganjar...	sekali melangkah pantang mundur sebelum ganjar...
4	Saya salut pada keberanian Pak Ganjar Pranowo ...	saya salut pada keberanian pak ganjar pranowo ...
...
995	Ganjar Pranowo tetap unggul jadi Capres	ganjar pranowo tetap unggul jadi capres
996	Tenanglanang Ganjar bim resmi dicalonkan saja ...	tenanglanang ganjar bim resmi dicalonkan saja ...
997	Ganjar ini prestasi nya apa? Kok bisa jadi capr...	ganjar ini prestasi nya apa kok bisa jadi capr...
998	Maafkan kami sudah tidak ada pilihan lagi di G...	maafkan kami sudah tidak ada pilihan lagi di g...
999	Kecawa dengan GANJAR	kecawa dengan ganjar

Gambar 9. Hasil Case Folding

c. Tokenization

Tokenization merupakan tahapan memecah suatu kalimat menjadi kata-kata. Pada tahapan ini menggunakan library NLTK yang berfungsi untuk membantu dalam pekerjaan teks khususnya *tokenization*, *stemming* dan klasifikasi.

	hasil_casefolding	hasil_token
0	hanya orang bodoh yg dukung oportunis sesat...	[hanya, orang, bodoh, yg, dukung, oportunis, s...
1	logika bodoh	[logika, bodoh]
2	dukungan untuk pak ganjar pranowo jadi capres...	[dukungan, untuk, pak, ganjar, pranowo, jadi, ...]
3	sekali melangkah, pantang mundur sebelum ganjar...	[sekali, melangkah, pantang, mundur, sebelum, ...]
4	saya salut pada keberanian pak ganjar pranowo ...	[saya, salut, pada, keberanian, pak, ganjar, p...
...
995	ganjar pranowo tetap unggul jadi capres	[ganjar, pranowo, tetap, unggul, jadi, capres]
996	tenanglanang ganjar bim resmi dicalonkan...	[tenanglanang, ganjar, bim, resmi, dicalonkan...
997	ganjar ini prestasi nya apa kok bisa jadi capr...	[ganjar, ini, prestasi, nya, apa, kok, bisa, j...
998	maafkan kami sudah tidak ada pilihan lagi di g...	[maafkan, kami, sudah, tidak, ada, pilihan, la...
999	kecawa dengan ganjar	[kecawa, dengan, ganjar]

Gambar 10. Hasil Tokenization Data

d. Remove Stopword

Pada tahapan ini bertujuan untuk melakukan penghapusan kata-kata yang kurang bermakna seperti contoh kata “yang”, “dan”, “di”, “dari”, dan masih banyak lagi.

	hasil_token	hasil_stopword
0	[hanya, orang, bodoh, yg, dukung, oportunis, s...	[orang, bodoh, yg, dukung, oportunis, sesat, l...
1	[logika, bodoh]	[logika, bodoh]
2	[dukungan, untuk, pak, ganjar, pranowo, jadi, ...]	[dukungan, ganjar, pranowo, capres, berlanjut]
3	[sekali, melangkah, pantang, mundur, sebelum, ...]	[melangkah, pantang, mundur, ganjar, pranowo, ...]
4	[saya, salut, pada, keberanian, pak, ganjar, p...	[salut, keberanian, ganjar, pranowo, berani, b...
...
995	[ganjar, pranowo, tetap, unggul, jadi, capres]	[ganjar, pranowo, unggul, capres]
996	[tenanglanang, ganjar, bim, resmi, dicalonkan...	[tenanglanang, ganjar, bim, resmi, dicalonkan...
997	[ganjar, ini, prestasi, nya, apa, kok, bisa, j...	[ganjar, prestasi, nya, capres, blong, yg, ta...
998	[maafkan, kami, sudah, tidak, ada, pilihan, la...	[maafkan, pilihan, ganjar, pranowo, beralih, c...
999	[kecawa, dengan, ganjar]	[kecawa, ganjar]

Gambar 11. Hasil Remove Stopword

e. Stemming

Stemming adalah tahapan mengubah kata menjadi kata dasar atau kata aslinya. Pada tahapan ini menggunakan bantuan library swifter yang berfungsi untuk mempercepat proses *stemming* pada data. Maka dari itu dilakukan instalasi *library* dengan perintah `pip install swifter`. Setelah *library* berhasil melakukan instalasi, selanjutnya hasil dari tahapan stemming pada Gambar 12.

Klasifikasi Komentar Publik Dalam Pemilihan Umum Presiden 2024 Dengan Menggunakan Metode Naive Bayes: Studi Kasus Identifikasi *Haters* Dan *Non-Haters*

	hasil_stopword	hasil_stemming
0	[orang, bodoh, yg, dukung, oportunis, sesat, l...	[orang, bodoh, yg, dukung, oportunis, sesat, l...
1	[logika, bodoh]	[logika, bodoh]
2	[dukungan, ganjar, pranowo, capres, berlanjut]	[dukung, ganjar, pranowo, capres, lanjut]
3	[melangkah, pantang, mundur, ganjar, pranowo, ...]	[lang, pantang, mundur, ganjar, pranowo, presi...
4	[salut, keberanian, ganjar, pranowo, berani, b...	[salut, berani, ganjar, pranowo, berani, suara...
...
995	[ganjar, pranowo, unggul, capres]	[ganjar, pranowo, unggul, capres]
996	[tenangtenang, ganjar, bin, resmi, dicatankan...]	[tenangtenang, ganjar, bin, resmi, calon, sdh...
997	[ganjar, prestasi, nya, capres, tolong, yg, ta...	[ganjar, prestasi, nya, capres, tolong, yg, ta...
998	[maafkan, pilihan, ganjar, pranowo, beralih, c...	[maaf, pilih, ganjar, pranowo, alih, calon]
999	[kecewa, ganjar]	[kecewa, ganjar]

Gambar 12. Hasil Stemming

3.3 Klasifikasi Naive Bayes

Setelah melakukan pre-processing data selanjutnya akan dibuat model yang nantinya akan digunakan untuk proses klasifikasi pada data uji. Sebelum itu, dilakukan pembagian data dengan rasio 8:2 secara acak dengan bantuan Python. Sehingga didapat data latih sebanyak 80% dan data uji atau testing sebanyak 20%. Proses pembagian data dapat dilihat pada gambar dibawah ini.

```
[30] #membagi data menjadi data train dan test
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(df['hasil_stemming'],df['Label'],test_size=0.2,random_state=82)
print("Data Latih :",len(x_train))
print("Data Uji :",len(x_test))

Data Latih : 890
Data Uji : 200
```

Gambar 13. Pembagian Data Latih dan Data Uji

Langkah berikutnya adalah melakukan klasifikasi Naive Bayes pada data uji menggunakan library *Multinomial Naive Bayes*. Pada Gambar 19 merupakan implementasi dari proses klasifikasi.

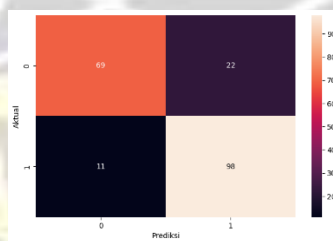
```
[ ] from sklearn.naive_bayes import MultinomialNB
#pembentukan model Naive Bayes
clf=MultinomialNB()
clf.fit(vec_latih,y_train)
```

Gambar 14. Kode Program Klasifikasi Naive Bayes

Proses klasifikasi diimplementasikan dalam Naive Bayes dilakukan dengan menghitung probabilitas antar kata atau klausa dari setiap kategori atau kelas yang bertujuan untuk menghasilkan prediksi terhadap data yang dimasukkan. Data uji yang digunakan pada penelitian ini ialah 200 data atau 20% dari data keseluruhan. Data uji digunakan untuk menguji tingkat akurasi model yang dibangun.

3.4 Evaluation

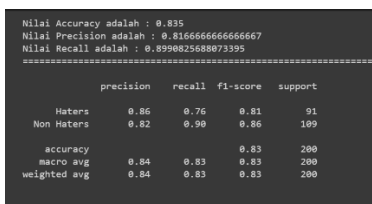
Evaluation dilakukan untuk mengetahui kinerja model dari algoritma Naive Bayes. Untuk mempermudah proses perhitungan maka pada proses ini akan ditampilkan dalam bentuk *confusion matrix*. Pengujian model klasifikasi menggunakan bantuan dari *library* python *sklearn.metrics* yang didalamnya terdapat fungsi *confusion_matrix* kemudian bentuk visualisasinya akan ditampilkan dengan menggunakan *seaborn* yang merupakan *library* yang berfungsi untuk membuat grafik atau statistik dalam python.



Gambar 15. Tampilan Confusion Matrix

Gambar 15 menampilkan hasil dari visualisasi *confusion matrix*. Setelah diketahui *confusion matrix* dari model naïve bayes yang telah dibuat, selanjutnya melakukan perhitungan nilai akurasi, presisi dan recall dari model klasifikasi Naive Bayes dengan mengimport *accuracy_score*, *precision_score*, *recall_score* yang disediakan oleh *library* *sklearn*.

Klasifikasi Komentar Publik Dalam Pemilihan Umum Presiden 2024 Dengan Menggunakan Metode Naive Bayes: Studi Kasus Identifikasi *Haters* Dan *Non-Haters*



	precision	recall	f1-score	support
Haters	0.86	0.76	0.81	91
Non Haters	0.82	0.98	0.86	109
accuracy			0.83	200
macro avg	0.84	0.83	0.83	200
weighted avg	0.84	0.83	0.83	200

Gambar 16. Hasil dari Confusion Matrix

Perhitungan akurasi secara manual dari *matrix* diatas sesuai dengan rumus adalah sebagai berikut.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Accuracy = \frac{98 + 69}{98 + 69 + 11 + 22} = \frac{167}{200} = 0,835 \times 100 = 83,5\%$$

Akurasi mengacu pada keakuratan model klasifikasi yang dibangun berdasarkan data yang diklasifikasikan dengan benar. Akurasi didapatkan dari hasil jumlah yang diprediksi dengan benar dengan total keseluruhan data uji. Dengan mengetahui nilai akurasi model, dapat dikatakan bahwa sistem dapat menentukan tingkat akurasi antara informasi yang diinginkan pengguna dengan hasil yang diprediksikan oleh sistem. Tingkat akurasi sistem pada penelitian ini adalah 83,5%.

Pada Gambar 23 nilai *precision*/presisi untuk kategori *haters* ialah 86% dan untuk kategori *non haters* ialah 82%. Nilai ini dapat diartikan bahwa kategori *haters* memiliki tingkat diprediksi dengan benar dibanding kategori *non haters*. Sedangkan tingkat keberhasilan sistem dalam menemukan kembali informasi atau recall untuk kategori *haters* ialah 76% dan kategori *non haters* ialah 90%. Hal ini dapat diartikan bahwa performa keberhasilan sistem dalam menemukan kembali informasi yang berkategori *non haters* lebih tinggi dibanding dengan menemukan informasi kembali yang berkategori *haters*. Karena *confusion matrix* didesain khusus untuk memperhatikan performa model dalam mengidentifikasi kasus positif dengan benar, maka yang difokuskan hanya pada nilai positif atau *non haters* dan didapatkan nilai presisi sebesar 82%, nilai *recall* sebesar 90%.

3.4 Deployment

Pada sistem ini, model machine learning yang telah dibangun menggunakan data training dengan format pickle. Pickle merupakan file yang digunakan pada bahasa pemrograman python untuk menyimpan model machine learning yang telah dibuat dan agar dapat diakses oleh Flask. Pada sistem ini file pickle disimpan dengan nama 'finalized_model.pkl'. Gambar 24 merupakan kode program dari deployment.

```
import pickle
with open ('finalized_model.pkl','wb') as r:
    pickle.dump(clf,r)
```

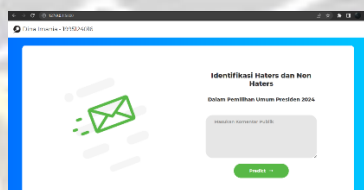
Gambar 17. Code Python Deployment

3.4.1 Implementasi Interface

Berikut Implementasi *Interface* berdasarkan rancangan *use case* dan *activity diagram*.

a. Halaman Home

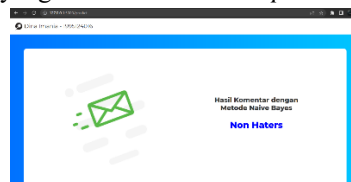
Pada halaman ini user dapat memberikan input kepada *textarea* yang nantinya inputan tersebut akan diproses oleh sistem aplikasi.



Gambar 18. Halaman Home

b. Halaman Hasil (*Non Haters*)

Pada halaman ini user dapat melihat hasil dari komentar publik dalam pemilihan umum presiden 2024 yang memiliki hasil klasifikasi *not haters*. Contoh kalimat yang dimasukkan “*Mantap Pak Ganjar, calon Presiden 2024*”.

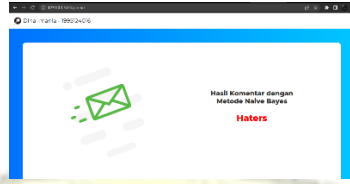


Gambar 19. Hasil Klasifikasi Non Haters

Klasifikasi Komentar Publik Dalam Pemilihan Umum Presiden 2024 Dengan Menggunakan Metode Naive Bayes: Studi Kasus Identifikasi *Haters* Dan *Non-Haters*

c. Halaman Hasil (*Haters*)

Pada halaman ini user dapat melihat hasil dari komentar publik dalam pemilihan umum presiden 2024 yang memiliki hasil klasifikasi *haters*. Contoh kalimat yang dimasukkan “*Calon presiden bego itu ya ini, si ganjar. Gak punya prestasi apa2 tiba2 nongol jadi calon presiden*”



Gambar 20. Hasil Klasifikasi *Haters*

PENUTUP

Simpulan

Berdasarkan hasil penelitian yang telah dilakukan oleh penulis pada klasifikasi komentar publik dalam pemilihan umum presiden 2024 dengan menggunakan metode naive bayes, maka penulis menarik kesimpulan, Penelitian ini berhasil menerapkan metode Naive Bayes menggunakan 1000 data, dengan rincian 800 data latih dan 200 data uji untuk klasifikasi komentar publik dalam pemilihan umum presiden 2024. Hasil penelitian yang diperoleh dengan tingkat akurasi yang baik yaitu sebesar 83,5%, nilai presisi yaitu 82% dan nilai *recall* sebesar 90%.

Saran

Adapun saran yang dapat diberikan untuk penelitian lanjutan terkait dengan penelitian ini ialah:

- Menggunakan model algoritma klasifikasi yang lain sehingga dapat membandingkan hasil uji model untuk mendapatkan algoritma dengan klasifikasi terbaik.
- Pada penelitian ini, komentar yang memiliki emoticon dan bahasa asing diabaikan, pada penelitian selanjutnya diharapkan memasukkan *emoticon* dan bahasa asing ke dalam perhitungan klasifikasi agar mendapatkan hasil yang lebih akurat.
- Pada penelitian selanjutnya diharapkan meningkatkan akurasi agar meminimalisir kesalahan dalam sistem.

DAFTAR PUSTAKA

- Arum Sari, Y., & Ali Fauzi, M. (2019). Seleksi Fitur Information Gain untuk Klasifikasi Informasi Tempat Tinggal di Kota Malang Berdasarkan Tweet Menggunakan Metode Naive Bayes dan Pembobotan TF-IDF-CF. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(5), 4907–4913. <http://j-ptiik.ub.ac.id>
- Farmadiansyah, A. Z., Hidayatullah, A. F., & Rahma, F. (2021). Deteksi Surel Spam dan Non Spam Bahasa Indonesia. *Automata*. <https://journal.uui.ac.id/AUTOMATA/article/view/19514>
- Id, I. D. (2021). *Machine learning: Teori, Studi Kasus Dan Implementasi Menggunakan Python*. Unri Press.
- Islam, K. A. (2006). *Developing and measuring training the Six Sigma way: A business approach to training and development*. John Wiley & Sons
- Nurhidayati, N. I., Yahya, Y. Y., Fathurrahman, F. F., & Samsu, L. S. S. L. (2023). Implementasi Algoritma Naive Bayes Untuk Klasifikasi Penerima Beasiswa (Studi Kasus Universitas Hamzanwadi). *Infotek: Jurnal Informatika dan Teknologi*, 6(1), 177-188
- Rohman, A. (2020). Sisi Positif Dan Negatif Demontrasi Pada Negara Demokrasi Dimasa Pandemi. *Binamulia Hukum*, 9(2), 153–170.
- Secha, K. N. (2022, November 6). *Survei Capres Indekstat: Ganjar 16,5%, Prabowo 11,6%, Anies 9,9%*. *detiknews*. <https://news.detik.com/pemilu/d-6390705/survei-capres-indekstat-ganjar-165-prabowo-116-anies-99>
- Waliyansyah, R. R., & Fitriyah, C. (2019). Perbandingan Akurasi Klasifikasi Citra Kayu Jati Menggunakan Metode Naive Bayes dan k-Nearest Neighbor (k-NN). *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(2), 157. <https://doi.org/10.26418/jp.v5i2.32473>