

ANALISIS SENTIMEN TERHADAP POPULARITAS PARTAI PENDUKUNG CAPRES 2024 DI TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER

Diterima Redaksi: 14 Oktober 2023; Revisi Akhir: 29 Oktober 2023; Diterbitkan Online: 30 November 2023

Dewi Astria Faroek¹⁾, Muhammad Yusuf^{*2)}, Grace Syatauw³⁾

^{1, 2, 3)} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Sorong
^{1, 2, 3)} Jalan Pendidikan No. 27, Kec. Sorong Timur, Kota Sorong, Papua Barat Daya, Indonesia, kode pos: 98416
e-mail: dewiastriafaroek@um-sorong.ac.id¹⁾, yusuf@um-sorong.ac.id²⁾, syatauwgrace@um-sorong.ac.id³⁾

Abstrak: Sesuai dengan sistem pemerintahan yang demokratis, Indonesia mengizinkan setiap partai politik atau gabungan partai yang mencalonkan diri dalam pemilihan umum untuk mencalonkan sendiri calon presiden dan wakil presiden, asalkan calon tersebut memenuhi persyaratan hukum. Pemilihan presiden dijadwalkan pada tahun 2024. Seorang tokoh politik yang ingin mencalonkan diri sebagai presiden pada saat itu harus sangat bergantung pada opini publik untuk mendapatkan dukungan. Menjelang pemilu 2024, partai politik semakin beralih ke media sosial untuk menyebarkan pesan mereka dan meningkatkan dukungan mereka. Tujuan dari penelitian ini adalah untuk membandingkan tingkat dukungan dua kandidat utama presiden dan wakil presiden tahun 2024 di Twitter untuk menentukan proporsi tweet yang positif, negatif, atau netral. Tujuan dari penelitian ini adalah untuk membandingkan tingkat dukungan dua kandidat utama presiden dan wakil presiden tahun 2024 di Twitter untuk menentukan proporsi tweet yang positif, negatif, atau netral. Dari hasil pengujian, menggunakan 1749 tweet didapatkan hasil Partai Demokrat akurasi 89,39%, precision 100%, recall 86,71%, dan nilai F-1 92,88%. Partai PDIP akurasi 96,60%, precision 95,10%, recall 100% dan nilai F-1 97,49%. Partai PKS akurasi 98,74%, precision 98,18%, recall 100% dan F-1 99,08%.

Kata Kunci—Analisis Sentimen, Partai Politik, Popularitas, Twitter, Naive Bayes Classifier

Abstract: In accordance with a democratic system of governance, Indonesia allows every political party or coalition of parties running in general elections to nominate their own presidential and vice-presidential candidates, provided that these candidates meet the legal requirements. The presidential election is scheduled for the year 2024. A political figure who wishes to run for president at that time must rely heavily on public opinion to gain support. As the 2024 election approaches, political parties are increasingly turning to social media to disseminate their messages and boost their support. The aim of this research is to compare the levels of support for the two main presidential and vice-presidential candidates in 2024 on Twitter in order to determine the proportion of tweets that are positive, negative, or neutral. The research objectives include comparing the levels of support for the two main presidential and vice-presidential candidates in 2024 on Twitter to determine the proportion of positive, negative, or neutral tweets. The test results, using 1,749 tweets, yielded the following outcomes: Demokrat party Accuracy of 89.39%, Precision of 100%, Recall of 86.71%, and F-1 Score of 92.88%. PDIP party Accuracy 96.60%, Precision 95.10%, Recall 100%, and F-1 Score 97.49%. PKS party: Accuracy 98.74%, Precision 98.18%, Recall 100%, and F-1 Score 99.08%.

Keywords— Sentiment Analysis, Political Parties, Popularity, Twitter, Naive Bayes Classifier

I. PENDAHULUAN

Indonesia adalah negara demokrasi, setiap partai politik atau gabungan partai yang mencalonkan diri dalam pemilihan umum dapat mencalonkan sendiri calon presiden dan wakil presiden, asalkan memenuhi persyaratan hukum. Pemilihan presiden dijadwalkan pada tahun 2024. Seorang tokoh politik yang ingin mencalonkan diri sebagai presiden pada saat itu harus sangat bergantung pada opini publik untuk mendapatkan dukungan. Banyak kelompok politik menggunakan komunitas online untuk membangun dukungan bagi kandidat mereka menjelang pemilu 2024[1]. Setelah pemerintah, legislatif, dan masyarakat, pers menempati posisi terpenting keempat. Tak dapat disangkal bahwa media berfungsi sebagai pengawas yang dapat menjaga demokrasi dan kepentingan publik tetap hidup dan sehat[2]. Beberapa peneliti telah menyimpulkan hal tersebut.

Emosi masyarakat dapat tersampaikan melalui kicauan di layanan jejaring sosial Twitter. Twitter adalah salah satu platform paling populer untuk mempromosikan komunikasi partai politik[3]. Banyak

kelompok politik yang berebut perhatian pemilih dengan menggunakan Twitter untuk menyebarkan pesan mereka menjelang pemilu 2024. Pendapat positif dan negatif orang dapat didengar dan dibagikan di jejaring sosial Twitter. Karena banyaknya informasi, mengkategorikan pendapat ini secara manual merupakan proses yang sulit dan memakan waktu. Memproses data tweet untuk membedakan antara tanggapan positif, negatif, dan netral diperlukan jika seseorang ingin mempelajari apa yang ditanggapi orang di platform Twitter. Oleh karena itu, penting untuk menentukan nada keseluruhan umpan balik. Peneliti tertarik untuk menggunakan teknik pembelajaran mesin untuk menilai kepercayaan dari tumpukan data ini. Teknik Naive Bayes Classifier digunakan untuk mengurutkan data. Teknik ini adalah pendekatan pembelajaran mesin probabilistik[4][5].

Beberapa hasil penelitian terkait analisis sentimen yaitu implementasikan algoritma DBSCAN untuk melakukan penambangan teks pada data Twitter untuk menganalisis pola voting di Pekanbaru, Indonesia[6]. Temuan studi tersebut menguatkan hasil pemilihan walikota dan wakil walikota di Pekanbaru, yang masing-masing dimenangkan oleh Firdaus dan Ayat. Analisis sentimen Twitter menggunakan metode ranking naive bayes, untuk calon presiden Indonesia 2014. Berdasarkan hasil jajak pendapat ini, 47,7% responden memiliki opini positif terhadap pasangan Prabowo Subianto dan Hatta Rajasa, sedangkan 26,4% memiliki opini negatif. Sebaliknya, tandem Joko Widodo dan Jusuf Kalla memperoleh 38% dukungan dan 34% ketidaksetujuan. dengan) menggunakan mesin vektor dukungan Twitter untuk menganalisis sentimen pemilih terhadap calon presiden 2020. Kata kunci terkait Jokowi diklasifikasikan dengan akurasi 86,82 persen oleh algoritma klasifikasi SVM dan kata kunci terkait Prabowo diklasifikasikan dengan akurasi 86,27 persen dalam penelitian ini.

Berdasarkan permasalahan diatas, peneliti tertarik untuk mengambil penelitian dengan judul “Analisis Sentimen Popularitas Partai Pendukung Capres 2024 di Twitter Menggunakan Algoritma Nave Bayes Classifier” berdasarkan uraian di atas. Bertujuan untuk membandingkan popularitas pendukung capres 2024 yang akan menjadi capres dan cawapres 2024, penelitian ini berupaya mengidentifikasi tweet mana yang positif, negatif, atau netral.

II. TINJAUAN PUSTAKA

A. Analisis Sentimen

Analisis sentiment bagian dari penambahan *teks* yang berfokus pada pendapat, perasaan, emosi, dan perilaku pengguna terhadap sesuatu seperti layanan, produk, orang, organisasi, dan atribut lainnya. Dalam era digital yang penuh dengan data teks yang diproduksi oleh pengguna melalui berbagai platform online, pemahaman sentimen menjadi semakin relevan[7].

Analisis sentiment adalah pendekatan dalam ilmu data yang bertujuan untuk mengidentifikasi, mengevaluasi, dan memahami sentiment, opini, atau emosi yang terkandung dalam teks, seperti ulasan pelanggan, tweet, artikel berita, dan lainnya. Analisis sentiment telah menjadi alat yang sangat berguna dalam berbagai aplikasi, termasuk pemahaman opini pelanggan, pemantauan media social, prediksi pasar, pemantauan merek dan banyak lagi[8].

B. Partai Politik

Dalam dinamika demokrasi, partai politik memainkan peran sentral yang krusial dalam menjembatani pemerintah dengan rakyat. Partai politik merupakan elemen inti dalam proses politik, termasuk dalam pemilihan umum, pembuatan keputusan politik dan penyampaian pandangan Masyarakat kepada pemerintah. Sistem politik pada setiap negara seringkali ditandai oleh kehadiran partai politik yang beragam[9].

C. Text Mining

Text Mining atau yang disebut dengan data mining dalam bentuk teks merupakan proses eksplorasi pengetahuan yang mengandalkan pola tertentu dalam pengambilan data secara tekstual. Konsep kerja dari proses text mining adalah mengolah kembali dokumen yang disajikan melalui predictive dan descriptive Mining[10].

Text Mining bertujuan untuk mengubah data yang tidak terstruktur menjadi data terstruktur dengan mengekstraksi data tersebut menggunakan Natural Language Processing (NLP) yang kemudian digunakan dalam proses analisis. Text Mining dapat menghasilkan informasi dan teks yang berkualitas tinggi dan benar.

D. Twitter

Twitter adalah platform media sosial yang memungkinkan pengguna untuk berbagi pemikiran, berita, gambar, video, dan pesan dalam format pendek yang disebut *tweet*. Pertumbuhan penggunaan media sosial membuka peluang baru untuk menganalisis beberapa aspek, dan pola dalam komunikasi[11]. Misalnya, data media sosial dapat dianalisis untuk mendapatkan wawasan tentang masalah-masalah, tren, partai politik, dan jenis informasi lainnya. *Twitter* adalah alat yang sangat berpengaruh dalam penyebaran berita, aktivisme, jejaring sosial, dan komunikasi daring. Ini telah digunakan dalam berbagai konteks, dari kampanye politik hingga gerakan sosial, dan menjadi platform penting dalam ekosistem media sosial[12].

E. Klasifikasi

Klasifikasi adalah landasan dasar dalam analisis data yang membuka pintu kepada pemahaman yang lebih dalam tentang pola dan karakteristik yang membedakan kelas data satu dari yang lain. Proses ini melibatkan pencarian dan identifikasi sekelompok pola atau fungsi yang menjelaskan dengan jelas perbedaan antara kelompok data, dan dengan kemampuan ini, kita dapat mengantisipasi atau memprediksi label kelas untuk data yang terkategori[13]. Pengklasifikasi membuat model algoritma berdasarkan data pelatihan yang tersedia, dan data uji diklasifikasikan berdasarkan model tertentu.

F. Naïve Bayes Classifier

Naïve Bayes Classifier adalah algoritma *machine learning* yang bersifat probabilistik yang telah menjadi pilihan yang sangat populer dalam berbagai tugas klasifikasi. Algoritma ini memiliki kemampuan untuk mengklasifikasikan beragam jenis data, termasuk sekumpulan dokumen. NBC mampu memprediksi probabilitas kejadian di masa depan berdasarkan pengalaman di masa sebelumnya[14][15]. Dalam tahap selanjutnya, kita akan menguraikan rumus dari naïve bayes classifier, yang merupakan dasar algoritma ini, untuk lebih memahami cara kerjanya dan potensi penggunaannya dalam berbagai konteks klasifikasi.

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)} \quad (1)$$

- A, B = events
- P(A|B) = probabilitas A diberikan B benar
- P(B|A) = probabilitas B diberikan A benar
- P(A), P(B) = probabilitas independen dari A dan B

G. Confusion Matrix

Confusion Matrix merupakan metode yang biasanya digunakan untuk perhitungan akurasi[16]. Saat menguji keakuratan hasil pencarian akan dievaluasi berdasarkan nilai *recall*, *precision*, *accuracy* dan *F-measure*[17]. Tabel *confusion matrix* yang berisi informasi ini dapat dilihat pada Tabel 1.

Tabel 1. Confusion Matrix

		Prediksi	
		Prediksi Benar	Prediksi Salah
Aktual	Aktual Benar	TP	FN
	Aktual Salah	FP	TN

Keterangan:

- TP (*True Positive*) : Jumlah prediksi yang benar bersifat positif
- FP (*False Positive*) : Jumlah prediksi yang salah bersifat positif
- TN (*True Negative*) : Jumlah prediksi yang benar bersifat negatif
- FN (*False Negative*) : Jumlah prediksi yang salah bersifat positif

H. Metode Agile

Agile development methods adalah tahapan kerja yang bertujuan untuk menghasilkan sistem

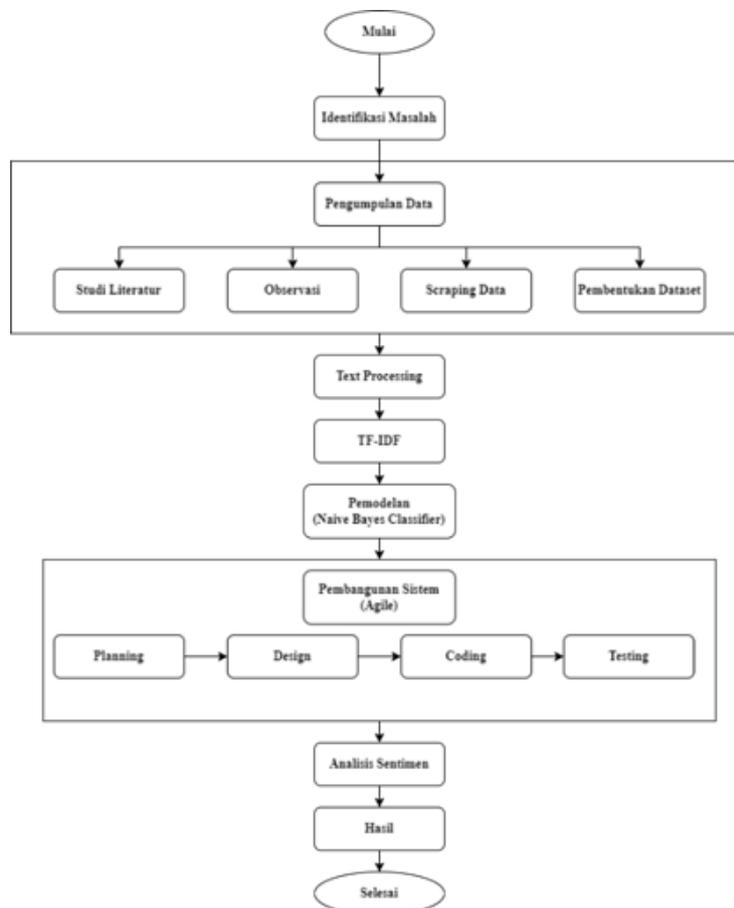
berkualitas tinggi, menjadi kerangka yang berisi langkah-langkah yang harus dilakukan untuk memproses pengembangan suatu perangkat lunak. Sistem ini berisi rencana lengkap untuk mengembangkan, memelihara, dan menggantikan perangkat lunak tertentu[18].

Tahapan dalam metode *agile* antara lain:

1. *Requirements*, membuat rencana mengenai perangkat lunak yang akan dibuat.
2. *Design* membuat rancangan mengenai perangkat lunak yang akan dibuat.
3. *Develop* merupakan proses untuk membangun perangkat lunak dengan menambahkan *coding*.
4. *Test* menguji kebenaran perangkat lunak, memastikan fungsi-fungsi perangkat lunak berjalan sesuai rancangan.
5. *Deploy* proses pemeliharaan perangkat lunak.

III. METODE PENELITIAN

Dalam perancangan sistem ini untuk mempermudah maka dijabarkan langkah-langkah yang akan dilakukan pada penelitian ini. Ada beberapa tahapan dalam perancangan sistem ini yaitu Identifikasi masalah, Pengumpulan Data, Text Processing, TF-IDF, Pemodelan Naïve Bayes Classifier, Pengembangan Sistem Model Agile, Analisis Sentimen seperti Gambar 2 berikut.



Gambar 1. Tahapan Perancangan Sistem

A. Identifikasi Masalah

Tahap ini kegiatan yang dilakukan oleh peneliti adalah mengidentifikasi masalah yang terjadi, sebagaimana yang telah dibahas pada pendahuluan yaitu perlu mengetahui popularitas partai pendukung capres 2024 di *Twitter*.

B. Pengumpulan Data

Tahap pengumpulan data terdiri atas beberapa langkah diantaranya adalah sebagai berikut:

1) Studi Literatur

Pada tahap ini peneliti menumpulkan data-data dan sumber penelitian untuk dijadikan literatur atau referensi dalam menganalisa penelitian adapun literatur tersebut berupa artikel, jurnal dan skripsi.

2) *Observasi*

Pada tahap ini peneliti melakukan observasi dengan cara mengamati dan meneliti objek yang akan diteliti, yaitu dengan mengamati platform *Twitter* terkait komentar partai politik pendukung capres 2024 yaitu partai Demokrat, PDIP dan PKS, dalam hal ini yang mana dalam komentar tersebut menuai opini atau komentar pengguna.

3) *Scrapping Data*

Scrapping data dilakukan dengan cara mengumpulkan data dari platform *Twitter* yang memuat komentar partai politik pendukung capres 2024. Data yang diambil adalah komentar partai politik pendukung capres 2024 tersebut, kemudian data yang telah diperoleh akan dijadikan sebagai data mentah.

4) *Pembentukan Dataset*

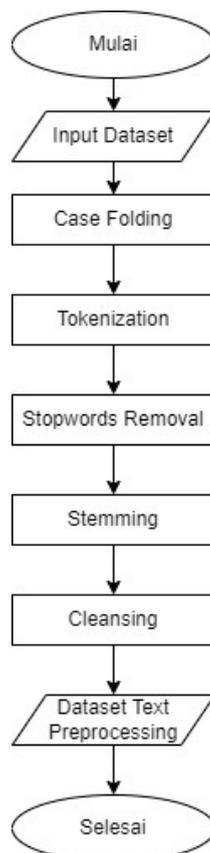
Pembentukan dataset dilakukan dengan cara mengkonversi data mentah yang telah diperoleh menjadi satu format yang sama yaitu satu *file excel* dengan format *csv* atau *xlsx*.

C. *Labelling*

Pada tahap ini kegiatan yang akan dilakukan adalah melabeli seluruh komentar atau ulasan yang terdapat pada dataset ke dalam kelas label negative, positif dan netral.

D. *Text Preprocessing*

Pada tahap ini kegiatan yang akan dilakukan adalah *Text Preprocessing* atau menyeleksi data teks agar menjadi data yang terstruktur. Terdapat beberapa metode *Text Preprocessing* yang digunakan yaitu:



Gambar 2. *Flowchart Text Preprocessing*

- 1) *Case folding* : Mengkonversi huruf data teks menjadi huruf kecil secara keseluruhan
- 2) *Tokenization* : Memecahkan kalimat pada teks menjadi terpisah setiap kata
- 3) *Stopwords Removal* : Menyeleksi data dan membuang kata yang tidak penting atau tidak bermakna
- 4) *Stemming* : Menyeleksi data dan membuang imbuhan yang terdapat pada suatu kata.
- 5) *Cleansing* : Menyeleksi data dan membuang data yang berpotensi mengurangi akurasi algoritma

E. TF-IDF

Pada tahap ini kegiatan yang akan dilakukan yaitu memboboti setiap kata pada data teks (komentar) yang terdapat pada dataset. Pembobotan kata dilakukan menggunakan metode *Term Frequency-Inverse Document Frequency*[19].

F. Pemodelan Naïve Bayes Classifier

Model algoritma Naïve Bayes Classifier dibangun menggunakan bahasa pemrograman Python[20]. Berikut alur atau langkah dari pemodelan NBC adalah sebagai berikut[21]:

1) Persiapan Dataset

Pada langkah ini yang dilakukan adalah menyiapkan dataset yang telah diperoleh pada tahap sebelumnya. Dataset disiapkan untuk digunakan sebagai bahan dalam membuat pemodelan.

2) Pembuatan Model Naïve Bayes Classifier

Pada langkah membuat model Naïve Bayes Classifier menggunakan bahasa Python. Model akan digunakan untuk klasifikasi dataset (data komentar).

3) Pengujian Model

Pada langkah melakukan pengujian terhadap model yang telah dibuat yaitu dengan melihat hasil dari klasifikasi dataset (data komentar). Hasil dalam hal ini adalah tingkat akurasi model.

4) Simpan Model

Simpan model yaitu melakukan penyimpanan model yang telah dibuat yang nantinya akan diimplementasikan pada aplikasi atau sistem.

G. Pembangunan Sistem Metode Agile

1. *Requirements* (Perancangan), pembangunan sistem diawali dengan identifikasi masalah, dari permasalahan yang didapat selanjutnya dilakukan Analisa terhadap kebutuhan-kebutuhan yang diperlukan menjadi 2 bagian yaitu[22][23] :

a) Analisa kebutuhan perangkat keras (hardware) : PC/Laptop

b) Analisa kebutuhan perangkat lunak (software)

1) *Operating system windows*

2) *Chrome/Firefox*

3) *Visual studio code*

4) *Library Twitter*

5) *Laragon*

6) *PHP*

2. *Design*, pada tahapan ini peneliti merancang sistem yang akan dibangun dengan menggunakan diagram. Diagram yang digunakan yaitu flowchart dan use case diagram dan tools yang digunakan dalam membuat perancangan sistem adalah draw.io.

3. *Develop*, pada tahapan ini peneliti akan melakukan konversi rancangan sistem yang telah dibuat pada langkah sebelumnya ke dalam code program secara mikro yang kemudian akan diintegrasikan menjadi satu kesatuan sehingga terbangunnya sebuah sistem yang dimaksud. Code program yang digunakan untuk pembangunan sistem yaitu Python.

4. *Test*, tahapan pengujian sistem menggunakan metode pengujian model algoritma Naïve Bayes Classifier yaitu *Confusion Matrix*.

5. *Deploy*, tahapan pemeliharaan sistem agar tetap berjalan dan berfungsi dengan baik.

H. Analisis Sentimen

Proses analisis sentimen menggunakan sistem pada penelitian yang akan dilakukan dimulai dari persiapan dan penginputan dataset kemudian dilakukan proses labeling atau pelabelan terhadap dataset, Text Preprocessing terhadap dataset, pembobotan TF-IDF terhadap dataset, selanjutnya pemilihan metode splitting dataset dan setelahnya adalah proses klasifikasi menggunakan algoritma Naïve Bayes Classifier kemudian menampilkan hasil dari proses klasifikasi yang telah dilakukan[24].

I. Hasil

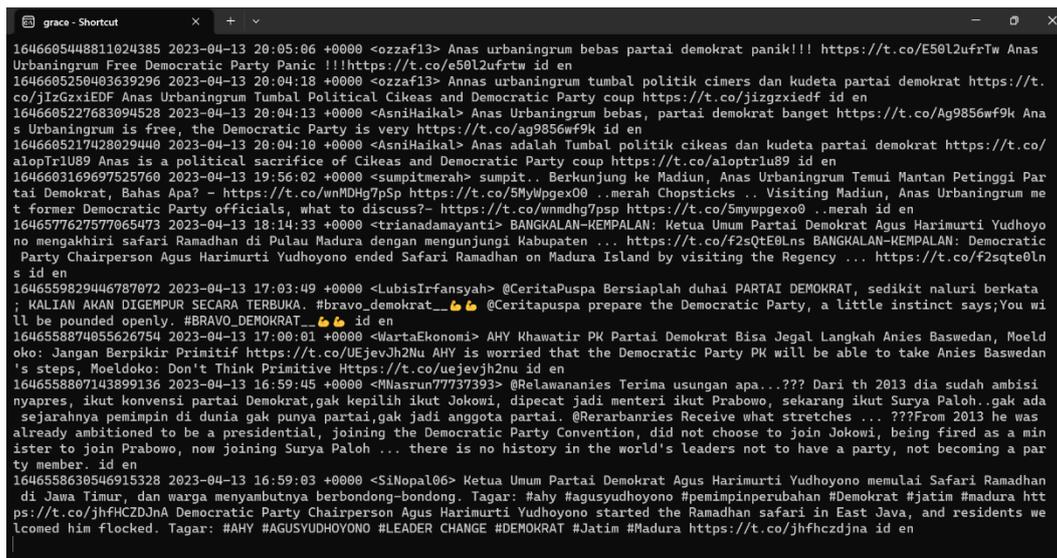
Hasil adalah tahap akhir. kegiatan yang akan dilakukan ialah membuat dan menyusun laporan penelitian berdasarkan hasil penelitian secara keseluruhan yang telah dilakukan.

IV. HASIL DAN PEMBAHASAN

Dalam penelitian ini membahas tentang analisis sentimen terhadap popularitas partai pendukung CAPRES 2024 di *Twitter* menggunakan algoritma *Naïve Bayes Classifier*.

A. Analisis Sumber Data

Data yang digunakan pada penelitian ini diambil dari kumpulan *tweets* Bahasa Indonesia yang diambil dari hastag #PartaiDemokrat, #PartaiPDIP, #PartaiPKS. Data tweets ini diperoleh dengan cara crawling yang menggunakan docker. Dalam proses crawling, secara otomatis akan mengambil data tweets yang mengandung kata “#PartaiDemokrat”, “#PartaiPDIP”, “#PartaiPKS”. Data tweets yang terkumpul nantinya akan dilakukan tahap *preprocessing teks* dan selanjutnya akan diklasifikasikan. Dalam analisis sentimen ini, tweets akan diklasifikasikan ke dalam bentuk sentimen, yaitu positif, negatif dan netral. Contoh tweets dapat dilihat pada Gambar 3 berikut ini.



Gambar 3. Contoh Tweets

B. Analisis Kebutuhan

Sebelum dilakukan perancangan sistem, maka terlebih dahulu dilakukan analisis bagaimana bentuk dan cara kerja dari sistem yang berjalan saat ini. Aplikasi yang dibangun adalah aplikasi analisis sentimen terhadap popularitas partai pendukung capres 2024 di *twitter* menggunakan algoritma *naïve bayes classifier*. Aplikasi ini menggunakan bahasa pemrograman python untuk pembangunan model dan implementasi berbasis website menggunakan visual studio code untuk pembuatan aplikasi[25][26].

1) Analisis Kebutuhan *Software*

Kebutuhan perangkat lunak yang digunakan dalam penelitian ini, antara lain:

- Operating system windows
- Chrome/Firefox
- Visual studio code
- Library Twitter
- Laragon
- PHP

2) Analisis Kebutuhan *Hardware*

Kebutuhan perangkat keras yang digunakan untuk pengembangan sistem ini antara lain:

- Laptop HP 14s
- AMD Ryzen 5 5500U with Radeon Graphics 2.10 GHz
- Memory (RAM) 12.0 GB

C. Desain Sistem

1) Implementasi *Interface*

Implementasi interface merupakan penerapan dari rancangan desain interface yang telah dirancang pada aplikasi.

a. Tampilan halaman *dashboard*

Tampilan dashboard merupakan halaman awal ketika sistem dijalankan.



Gambar 4. Tampilan Halaman Dashboard

b. Tampilan halaman dataset

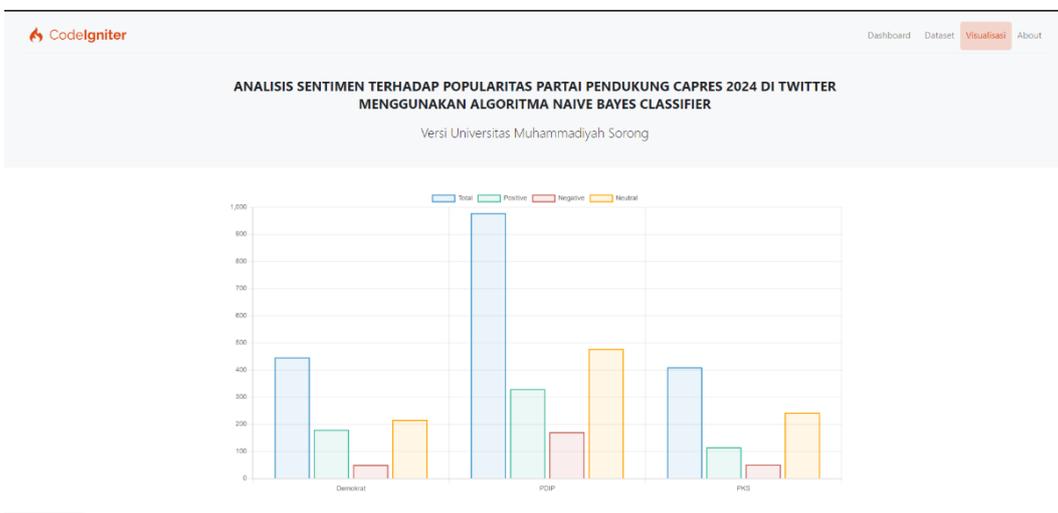
Pada halaman ini terdapat dataset yang diklasifikasi ke dalam sentimen positif, negative dan netral.



Gambar 5. Tampilan Halaman Dataset

c. Tampilan halaman visualisasi

Pada halaman ini menampilkan hasil dari pemrosesan data yang akan ditampilkan dalam bentuk visualisasi diagram.



Gambar 6. Tampilan Halaman Visualisasi

d. *Tampilan confusion matrix, akurasi dan NBC*

Pada halaman ini menampilkan tabel perhitungan dari hasil yang didapatkan pada proses data yang dilakukan.

Demokrat			PDIP			PKS		
198	Testing Positive	Testing Negative	500	Testing Positive	Testing Negative	159	Testing Positive	Testing Negative
Training Positive	TP: 137	FP: 0	Training Positive	TP: 330	FP: 17	Training Positive	TP: 108	FP: 2
Training Negative	FN: 21	TN: 40	Training Negative	FN: 0	TN: 153	Training Negative	FN: 0	TN: 49

Tabel data		Tabel data		Tabel data	
Desc	Data	Desc	Data	Desc	Data
Training Positive	137	Training Positive	347	Training Positive	110
Training Negative	61	Training Negative	153	Training Negative	49
Testing Positive	158	Testing Positive	330	Testing Positive	108
Testing Negative	40	Testing Negative	170	Testing Negative	51
Accuracy = $(TP+TN) / (TP+FP+FN+TN)$	89.39%	Accuracy = $(TP+TN) / (TP+FP+FN+TN)$	96.60%	Accuracy = $(TP+TN) / (TP+FP+FN+TN)$	98.74%
Precision = $(TP) / (TP + FP)$	100.00%	Precision = $(TP) / (TP + FP)$	95.10%	Precision = $(TP) / (TP + FP)$	98.18%
Recall = $TP / (TP + FN)$	86.71%	Recall = $TP / (TP + FN)$	100.00%	Recall = $TP / (TP + FN)$	100.00%
F-1 Score = $(2 * Recall * Precision) / (Recall + Precision)$	92.88%	F-1 Score = $(2 * Recall * Precision) / (Recall + Precision)$	97.49%	F-1 Score = $(2 * Recall * Precision) / (Recall + Precision)$	99.08%

Gambar 7. Tampilan Halaman Confusion matrix, Akurasi dan NBC

D. *Pengujian Confusion Matrix*

Pada tahap ini yang dilakukan pengujian sistem yaitu mengetahui kemampuan dari sistem. Untuk mengetahui pengaruh proporsi data latih setiap kategori dan nilai F-1 terhadap efektivitas sistem klasifikasi. Jumlah data uji yang digunakan sebanyak 2000 tweet. Dengan informasi yang dihasilkan dari pengujian ini, kami dapat mengevaluasi sejauh mana sistem ini mampu berkinerja dan memberikan wawasan yang berguna untuk perbaikan dan peningkatan sistem yang lebih lanjut.

Tabel 2. Porporasi Data Latih

Porporasi Data Latih			
Positif	Negatif	Netral	Jumlah
635	428	937	2000

Pengujian akurasi menggunakan *confusion matrix* dilakukan untuk menguji performansi metode yang digunakan dalam membangun sistem analisis sentiment berdasarkan aspel. Berikut adalah tabel dari *confusion matrix*.

Tabel 3. Hasil Confusion Matrix Partai Demokrat

Hasil Nilai Demokrat	Data
Accuracy = $(TP+TN) / (TP+FP+FN+TN)$	89.39%
Precision = $(TP) / (TP + FP)$	100.00%
Recall = $TP / (TP + FN)$	86.71%
F-1 Score = $(2 * Recall * Precision) / (Recall + Precision)$	92.88%

Didapatkan hasil dari *confusion matrix* partai demokrat yaitu *accuracy* 89,39%, *precision* 100%, *recall* 86,71% dan *F1 Score* 92,88%.

Tabel 4. Hasil Confusion Matrix Partai PDIP

Hasil Nilai PDIP	Data
Accuracy = $(TP+TN) / (TP+FP+FN+TN)$	96.60%
Precision = $(TP) / (TP + FP)$	95.10%
Recall = $TP / (TP + FN)$	100.00%
F-1 Score = $(2 * Recall * Precision) / (Recall + Precision)$	97.49%

Didapatkan hasil dari *confusion matrix* partai PDIP yaitu *accuracy* 96,60%, *precision* 95,10%, *recall* 100% dan *F1 Score* 97,49%.

Tabel 5. Hasil Confusion Matrix Partai PKS

Hasil Nilai PKS	Data
$Accuracy = (TP+TN) / (TP+FP+FN+TN)$	98.74%
$Precision = (TP) / (TP + FP)$	98.18%
$Recall = TP / (TP + FN)$	100.00%
$F-1 Score = (2 * Recall * Precision) / (Recall + Precision)$	99.08%

Didapatkan hasil dari *confusion matrix* partai PKS yaitu *accuracy* 98,74%, *precision* 98,18%, *recall* 100% dan *F1 Score* 99,08%.

V. KESIMPULAN DAN SARAN

KESIMPULAN

Berdasarkan hasil dari perancangan dan implementasi dari sistem analisis sentiment terhadap popularitas partai pendukung capres 2024 di *twitter* menggunakan algoritma *Naïve bayes classifier* berbasis *website* maka dapat ditarik kesimpulan diantaranya:

1. Algoritma *naïve bayes classifier* dapat digunakan untuk klasifikasi sentiment berbahasa Indonesia menjadi 3 kategori yaitu sentiment positif, negative, dan netral.
2. Dalam melakukan pengujian menggunakan 2000 *tweet* dapat dilihat pada hasil pengujian sistem didapatkan hasil akurasi 95,71%, *precision* 97,99%, *recall* 96,29% dan nilai F-1 97,01%.

SARAN

Untuk pengembangan sistem analisis sentimen terhadap popularitas partai pendukung capres 2024 di *twitter* menggunakan algoritma *naïve bayes classifier* berbasis *website*. Peneliti menyarankan untuk selanjutnya dapat mengatasi kesalahan penulisan sehingga mendapatkan hasil yang lebih optimal dan melakukan pengembangan sistem dengan menggunakan algoritma lainnya.

DAFTAR PUSTAKA

- [1] Z. Muda, I. Dwi, S. Arifin, and N. Manan, *Gangguan Informasi, Pemilu, dan Demokrasi*. 2023.
- [2] A. Hidayanto, “Peran Media Massa Dalam Mengakomodir Kepentingan Publik (Kebijakan Pln),” *Sintesa*, vol. 1, no. 1, pp. 18–32, 2023, doi: 10.30996/sintesa.v1i1.8177.
- [3] A. Humaidi, “Strategi Pemilihan Capres dan Cawapres Indonesia di Twitter: Kajian Wacana Kritis di Media Sosial,” vol. 1, pp. 792–793, 2019.
- [4] S. A. Hussein and M. R. Kareem, “A PROPOSED ARABIC TEXT AND TEXT IMAGE CLASSIFICATION TECHNIQUE USING A URL ADDRESS ,” *Xinan Jiaotong Daxue Xuebao/Journal Southwest Jiaotong Univ.*, vol. 54, no. 5, 2019, doi: 10.35741/issn.0258-2724.54.5.41.
- [5] A. Sumiah and N. Mirantika, “Perbandingan Metode K-Nearest Neighbor dan Naive Bayes untuk Rekomendasi Penentuan Mahasiswa Penerima Beasiswa pada Universitas Kuningan,” *Buffer Inform.*, vol. 6, no. 1, pp. 1–10, 2020.
- [6] R. Saputra *et al.*, “Menentukan Popularitas Calon Presiden dan Tren pada Pilpres 2019 menggunakan Algoritma DBSCAN,” *Puzzle Res. Data Technol. Fak. Sains dan Teknol.*, vol. 1, p. 3, 2019.
- [7] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, “Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes,” *J. Edukasi dan Penelit. Inform.*, vol. 4, no. 2, p. 113, 2018, doi: 10.26418/jp.v4i2.27526.
- [8] P. A. Permatasari, L. Linawati, and L. Jasa, “Survei Tentang Analisis Sentimen Pada Media Sosial,” *Maj. Ilm. Teknol. Elektro*, vol. 20, no. 2, p. 177, 2021, doi: 10.24843/mite.2021.v20i02.p01.
- [9] A. Annas and Z. Rusnaedy, “Journal of Government Civil Society,” *J. Gov. Civ. Soc.*, vol. 3, no. 2, pp. 117–128, 2020.

- [10] S. D. Raharjo and G. Wang, “Paper Review Text Mining Twitter,” 2020.
- [11] R. Putra Perssela, R. Mahendra, and W. Rahmadiani, “Pemanfaatan Media Sosial Untuk Efektivitas Komunikasi,” *J. Ilm. Mhs. Kuliah Kerja Nyata*, vol. 2, no. 3, pp. 650–656, 2022, doi: 10.36085/jimakukerta.v2i3.4525.
- [12] A. Rivaldy, H. A. Fedria Wowor, S. R. Maisya, and D. Safitri, “Penggunaan Twitter Dalam Meningkatkan Melek Politik Mahasiswa Ilmu Komunikasi Universitas Negeri Jakarta,” *Perspekt. Komun. J. Ilmu Komun. Polit. dan Komun. Bisnis*, vol. 5, no. 1, p. 41, 2021, doi: 10.24853/pk.5.1.41-48.
- [13] E. Fitri, “Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine,” *J. Transform.*, vol. 18, no. 1, p. 71, 2020, doi: 10.26623/transformatika.v18i1.2317.
- [14] A. Muzaki and A. Witanti, “Sentiment Analysis of the Community in the Twitter To the 2020 Election in Pandemic Covid-19 By Method Naive Bayes Classifier,” *J. Tek. Inform.*, vol. 2, no. 2, pp. 101–107, 2021, doi: 10.20884/1.jutif.2021.2.2.51.
- [15] P. B. N. Setio, D. R. S. Saputro, and Bowo Winarno, “Klasifikasi Dengan Pohon Keputusan Berbasis Algoritme C4.5,” *Prism. Pros. Semin. Nas. Mat.*, vol. 3, pp. 64–71, 2020.
- [16] W. I. Rahayu, C. Prianto, and E. A. Novia, “Perbandingan Algoritma K-Means Dan Naïve Bayes Untuk Memprediksi Prioritas Pembayaran Tagihan Rumah Sakit Berdasarkan Tingkat Kepentingan Pada Pt. Pertamina (Persero),” *J. Tek. Inform.*, vol. 13, no. 2, pp. 1–8, 2021, [Online]. Available: <https://ejurnal.poltekpos.ac.id/index.php/informatika/article/view/1383>.
- [17] H. Dalianis, “Evaluation Metrics and Evaluation,” *Clin. Text Min.*, no. 1967, pp. 45–53, 2018, doi: 10.1007/978-3-319-78503-5_6.
- [18] A. Ariesta, Y. N. Dewi, F. A. Sariasih, and F. W. Fibriany, “Penerapan Metode Agile Dalam Pengembangan Application Programming Interface System Pada Pt Xyz,” *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 7, no. 1, p. 38, 2021, doi: 10.24014/coreit.v7i1.12635.
- [19] D. W. Wibowo, M. Z. Abdullah, and J. D. Kristanto, “Penerapan Metode Tf-Idf Untuk Chatbot Pada Sistem Informasi Pelayanan Percetakan Online,” *Semin. Inform. Apl. Polinema 2020*, pp. 196–200, 2020.
- [20] F. Sodik, B. Dwi, and I. Kharisudin, “Perbandingan Metode Klasifikasi Supervised Learning pada Data Bank Customers Menggunakan Python,” *J. Mat.*, vol. 3, pp. 689–694, 2020, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/article/view/37875>.
- [21] G. K. Locarso, “Analisis Sentimen Review Aplikasi Pedulilindungi Pada Google Play Store Menggunakan NBC,” *J. Tek. Inform. Kaputama*, vol. 6, no. 2, pp. 353–361, 2022.
- [22] M. A. Firmansyah, N. Ramsari, and A. D. Rachmanto, “Rancang Bangun Sistem Informasi Penjualan Pada Toko Buku Kita Tasikmalaya Berbasis Web Menggunakan Framework Laravel 8,” *J. Teknol. Inf. dan Komun.*, vol. 12, no. 1, 2022, doi: 10.56244/fiki.v12i1.498.
- [23] A. Deviyanto and M. D. R. Wahyudi, “Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 3, no. 1, p. 1, 2018, doi: 10.14421/jiska.2018.31-01.
- [24] S. R. I. Rezeki, Y. Restiviani, and R. Zahara, “Penggunaan sosial media twitter dalam komunikasi organisasi (studi kasus pemerintah provinsi dki jakarta dalam penanganan covid-19),” *J. Islam. Law Stud.*, vol. 04, no. 02, pp. 63–78, 2020, [Online]. Available: <http://jurnal.uin-antasari.ac.id/index.php/jils/>.
- [25] R. Soekarta, “RANCANG BANGUN SISTEM INFORMASI AKADEMIK BERBASIS WEB (Studi Kasus Universitas Muhammadiyah Sorong),” 2015.
- [26] K. H. Ermin, Hanif, N. R. Muntari, and P. A. Ramadhani, “Penerapan Metode Certainty Factor untuk Mendiagnosa Penyakit Preeklamsia pada Ibu Hamil dengan Menggunakan Bahasa Pemrograman Python,” *Insect (Informatics Secur. J. Tek. Inform.*, vol. 7, no. 2, pp. 63–71, 2022, doi: 10.33506/insect.v7i2.1818.